A MILP model and heuristic approach for facility location under multiple operational constraints

Mouna Kchaou Boujelben, Céline Gicquel, Michel Minoux

A MILP model and heuristic approach for facility location under multiple operational constraints

Mouna Kchaou Boujelben [1], CelineGicquel [2], MichelMinoux [3]

[1] College of Business and Economics, UAE University,
P.O. Box 15551, Al Ain, United Arab Emirates

[2] LRI, Universite Paris-Sud, 91405 Orsay, France

[3]LIP6, Universite Pierre Marie Curie, 75005 Paris, France

Abstract

In the present work, we study a multi-period facility location problem featuring many realistic constraints. In order to take into account vehicle routing from distribution centres to customers while maintaining a manageable size of the optimization problem, we develop a two-phase solution approach. In the first phase, the average distances and costs of transport from distribution centres to customers are evaluated using an exact clustering procedure based on a set-partitioning formulation. These costs serve as input to the facility location problem in the second phase, which is formulated as a mixed integer linear program and solved using a state-of-the art commercial solver. Many numerical experiments using real life data from the automotive industry are carried out in order to derive some insights related to multi-period modeling. We first show that in our case study, using static assignment decisions is better for the company as the corresponding operational benefit outweighs the additional cost to be incurred. We then compare the outputs of the multi-period model with those of its single-period counterpart. Finally, to cope with the computational difficulties encountered during the numerical experiments, we propose a linear relaxation based heuristic to solve larger instances of the problem. The heuristic method provides good quality solutions while significantly improving computation times.

Keywords: Dynamic facility location, multi-period supply chain network design, vehicle routing, linear relaxation heuristics, clustering, automotive industry

1   Introduction

Major industrial companies put considerable efforts on optimizing the planning of their supply chains in order to reduce costs, decrease lead times and improve the service offered to customers. One key question at the strategic planning level is designing the supply chain network and deciding where to locate factories and distribution centres (DCs). This question leads to studying network design and facility location problems (FLP) which aim at selecting the best locations for new facilities and assigning customers to the opened sites while optimizing a given objective function (see e.g. the reviews proposed in [29] and more recently in [21]). Although these problems have been widely studied during the last decades, diverse challenges encountered by firms nowadays raise new questions on how to model and solve complex real-life situations. Several recent works have thus focused on improving the practical relevance of facility location models by incorporating different operational features and constraints. For example, maximum capacity and single sourcing constraints have been widely considered in the literature of facility location ([30], [36]). Imposing maximum capacities aims at modeling the limitation of space and resources in each location whereas

requiring to manage all the deliveries of a given customer from a unique distribution centre is a way to simplify day-to-day logistics operations. Another type of constraint that can be included in a facility location model is the limitation of the maximum distance between a customer and a distribution centre serving him. This ensures a better proximity to customers and a higher service level ([25], [32]).

Given that the temporal aspects of real world problems can be better captured with dynamic formulations than with static models ([29]), recent works in facility location have developed a considerable interest in optimizing dynamic systems (see [6] for a recent review on facility location dynamics). Dynamic systems are related to two main features: uncertainty (e.g. in the prediction of input parameters) and time-dependency. Time-dependent or multi-period models consider parameters that vary over time-periods, such as demand, prices and costs and introduce the possibility of adapting decisions to variations in these parameters. In the present work, we formulate a multi-period facility location model taking into account time-varying customer demand. The overall network structure consists of three levels: plants in the first level, distribution centres (DCs) in the second one and customers in the third one. We assume that the number and location of the plants as well as the number and location of the customers are fixed. Given the demand of customers in each time-period and a list of potential DCs, our main concern is to locate DCs and to assign customers to them in such a way as to minimize the total distribution costs.

The model studied in this paper integrates many realistic features that have not been simultaneously taken into account in the literature of multi-period FLP. Constraints of maximum covering distance, single sourcing, minimum volume on transport links as well as minimum volume and maximum capacity on distribution centres are combined in the model. The network modeled involves two types of transport: "primary" transport from plants to DCs and "secondary" transport from DCs to customers. The main difference between these two kinds of transport is that trucks used in primary transport follow direct routes from plants to DCs while those used in secondary transport have to visit more than one customer before returning back to the DC. Such grouping of several shipments into one truck is a common practice considered by companies in order to reduce transport costs through the use of full truckloads, while ensuring frequent shipments. This introduces however more complexity to the modeling and solving of the network design problem as a Vehicle Routing Problem (VRP) should be integrated with the main facility location problem, which results in a combined location-routing problem (LRP). Although several works in the literature addressed single-period location-routing problems (see e.g. [36]), only few models were studied in the field of multi-period location routing; works like [1], [3], [17], [31] and [35] are worth mentioning in this context. However, in these papers, exact solutions could be found only for problems of small size. For instance, in [3], the MILP solver terminated without even having found a feasible solution to eight of the ten instances considered (involving 10 facilities, 40 customers and 12 time-periods).

In the present work, we propose an original way to take into account vehicle routing in a FLP while keeping a manageable size for the optimization problem. Through a two-phase solution approach, we approximately solve a multi-period LRP for real-life industrial instances: the first phase uses an exact clustering method for secondary transport distance and cost identification whereas the second phase deals with a FLP that considers the results of the first phase as input parameters. In order to validate this approach, we present several numerical experiments using real-life data from the automotive industry. We first propose an exact solution to the FLP using a state-of-the-art MILP solver, which provides us with interesting insights related to multi-period modeling. We evaluate the difference between static and dynamic assignment decisions in terms of costs and network structure and compare the outputs of the multi-period model with those of its

single-period counterpart. Finally, to cope with computational difficulties encountered during the numerical experiments, we investigate a linear relaxation based heuristic procedure to solve larger instances of the problem.

The remainder of this paper is organized as follows. In Section 2, we investigate the value of multi-period FLP and discuss some connections with the existing literature. In Section 3, we present the modeling choices considered in this paper, including problem costs and constraints as well as customer representation. In Section 4, the first phase of the 2-stage solution approach is explained and the clustering problem is formulated. The second phase of the solution approach is described in Section 5, involving the mathematical formulation of the facility location problem. Computational results and a qualitative analysis for a case study from the automotive industry are presented in Section 6. In Section 7, an efficient heuristic procedure and the related numerical results are investigated whereas conclusions and suggestions for further research are presented in the last section.

2 Literature on multi-period FLP: value and modeling considerations

In real-life problems, input parameters such as demand quantities, prices and costs are likely to vary from period to period over an extended time-horizon. Accordingly, dynamic, multi-period or time-dependent FLP, as opposed to single-period or static FLP should be considered, in this section, we do not aim at providing an exhaustive review on multi-period FLP and SCNDP. Our intent is rather to explain why multi-period models should be used and to propose a general overview of the main tactical/operational features and constraints considered in these models.

One may wonder why multi-period models should be considered for facility location if this kind of approach increases the complexity of the optimization problems. One of the reasons is that a single-period model does not consider the variation of the problem parameters from period to period and thus may lead to solutions which could be sub-optimal or even infeasible when exposed to demand or price fluctuations for example. The difference in the network structure and total cost should be evaluated by comparing the optimal solution of the dynamic problem with the optimal solution of its static counterpart (re-evaluated a posteriori using the right dynamic parameters). In their recent work [27], Melo and Saldanha Da Gama referred to this as finding "the value of the multi-period solution". To the best of our knowledge this kind of numerical comparison was only addressed in [4]. The authors provided details on how to evaluate data variation from period to period using assumptions on volume increase and yearly inflation rates. They also explained the advantages of a dynamic model as compared to a static one based on a case study in the context of reverse logistics network design for washing machines and tumble dryers.

When the planning horizon is long (typically more than 2 years), customer demand is likely to vary along time because of the growth or decline of the company sales from year to year. When considering mid-term planning (typically one year as in the case study addressed in Section 6), changes in demand from period to period are mainly explained by seasonality. In this case, an optimal solution found using multi-period modeling is likely to significantly differ from an optimal solution found using single-period modeling for a same problem. Consider, for instance, the classical problem of capacitated FLP, in which seasonal customer demand is introduced. Assume that there are four seasons, two seasons with a very low demand and two seasons with a very high demand and that opened DCs must remain open during the whole planning horizon. In such a situation, the solution obtained with a multi-period model is likely to open more DCs than the

solution obtained with a single-period model. The reason is mainly related to the constraints of maximum capacity on DCs. These constraints may be problematic during periods of peak demand that are captured in the multi-period model but not considered in the single-period model, which uses average demand values. The solution of the single-period model may even be infeasible if exposed to large fluctuations in demand. In the problem variant studied in this paper, such feasibility issues would be frequently encountered as minimum volume and maximum capacity constraints are simultaneously integrated in the model. In order to meet detailed period by period constraints, a multi-period facility location model should thus be considered.

Melo et al. pointed out in their survey [21] that literature about multi-period facility location problems for supply chains is rather scarce as more than 80% of the papers surveyed deal with single-period problems. The introduction of multiple time-periods tends to significantly increase the size of optimization problems and usually leads to computational difficulties, especially when the problem is already hard to solve in its static version. Difficulty in solving this kind of problem occurs when dealing with large-size industrial data and when taking into account several tactical/operational features and constraints. For instance, in [5], the authors investigated a dynamic distribution network design problem involving inventory and routing decisions with capacitated vehicles. In [2], Aghezzaf considered capacity expansion decisions as well as demand uncertainty through various scenarios with given probabilities of occurrence. Other types of decisions, namely production quantities [11, 20, 34] and procurement [20], were also combined with the classical location and assignment decisions in a multi-period context. Problem-specific features like budget constraints ([11, 19]), financial incentives ([19]), international factors ([34]) and bill of materials ([11, 34]) were also considered in previous models dealing with multi-period facility location. In the context of our study, we were also interested in literature on FLP and SCNDP including minimum volume constraints for distribution centres and transport links. In [15], the authors surveyed 18 works that considered this kind of constraints; only six of them studied multi-period models ([4], [9] [20], [24], [26], [33]) but, in contrast with the present work, none of them proposed a model using minimum volume constraints for both facility throughput and transport quantities.

Another feature of dynamic facility location problems is the possibility of opening and closing facilities in different time-periods ([8], [18], [23], [28]). Closing facilities is however costly and difficult at the operational level particularly if the planning horizon spans less than one or two years. Since this is the case in the present work, we assume that the decision of opening/not opening a facility cannot be reconsidered from one time-period to the next. Facilities opened at the beginning of the planning horizon remain open until the end of the planning horizon, but customer assignment decisions can be adapted to demand variations from period to period. Similar assumptions have been used e.g. in [10], [13] and [16]; these models are referred to as "implicitly dynamic models" by Arabani and Farahani [6]. This implies that the problem considered in the present work cannot be addressed by simply solving a sequence of single-period FLPs. Since each of the single-period FLPs would seek to adjust the distribution network to satisfy, as best as possible, the demand in the corresponding time-period, the obtained solution would likely differ from the ones obtained for the other periods. In this case, no indication will be provided on which of the obtained solutions should be implemented in practice, while we need to select a fixed set of facilities that is optimal over the whole planning horizon.

One of the contributions of the present paper to the existing literature is to provide a model for a multi-period facility location problem taking into account many operational constraints, namely minimum volume, maximum capacity, single sourcing and maximum covering distance constraints. We also study the impact of dynamic demand on delivery routes using a clustering procedure, which enables us to reach a good trade-off between the representation of operational routing details in

4

the model and its computational tractability. Another contribution of this paper is to provide some insights on multi-period modeling based on real-life data from the automotive industry. We compare in §6.3 the outputs of the multi-period model to those of its single-period version. In §6.2, we evaluate the difference between using static assignments and using dynamic assignments in terms of costs and location decisions. Finally, an efficient heuristic procedure based on linear relaxations is proposed to solve the multi-period problem (see Section 7): we note that few works used this kind of approach to tackle multi-period facility location problems with minimum volume constraints ([22], [33]).

3 Problem description

We study a three-level distribution network consisting of plants, distribution centres (DCs) and customers. Any product type can be manufactured at any plant. Nevertheless, it is assumed that the annual demand of each customer for the different product types have already been assigned to plants by top-level management. Demands are thus expressed as quantities to be distributed from a given source (plant) to a given destination (customer) in each time-period. Items are transported by trucks from plants to intermediate distribution centres (via "primary" transport links) before being sent by trucks to customers (via "secondary" transport links). Our aim in this work is to determine the location of the DCs to be used among a list of potential DCs and to assign customers to them in such a way as to minimize the total distribution costs. As explained in Section 2, DCs opened at the beginning of the planning horizon remain open until the end of the planning horizon. As for the customer assignments to DCs, two alternatives will be studied. The first one corresponds to dynamic assignments: customer-DC assignments and distribution routes may change from one period to the next in order to adapt to demand fluctuations. The second one corresponds to static assignments: customer-DC assignments and distribution routes are not allowed to change throughout the planning horizon. This constraint may lead to additional transport costs but enables the company to simplify its day-to-day operations. We propose to compare the two alternatives both in terms of costs and network structure.

We now provide a detailed description of the costs and constraints involved in our problem. We also discuss the aggregate customer representation used in our model.

3.1 Problem costs

In the objective function of our problem, the total cost that should be minimized involves the costs of primary transport, secondary transport and transit through DCs as well as the fixed costs related to opening DCs. It is assumed that primary transport, secondary transport and transit costs are linear functions of the corresponding quantities.

3.2 Customer representation

We assume that detailed information on the location of all final customers belonging to the studied network is available. However, in order to avoid prohibitive computation times when solving the optimization problem, we use an aggregation approach: assignment decisions are made for customers grouped into districts and not on a customer per customer basis. This means that districts are used as final customers in our MILP formulation and that distances, costs and volume constraints are expressed for DC-district transport links, which enables us to reduce the problem size.

For instance, in the case study addressed in Section 6, the number of districts (92) is significantly less than the number of end customers (448). If we consider that assignment decisions could change from period to period, the number of periods is 4, the number of plants is 16 and the number of potential DCs is 51, the number of assignment decisions (binary variables) will be reduced from 1462272 to 300288.

3.3 Problem constraints

As mentioned in the introduction, many operational constraints are considered in the proposed model:

- *Maximum covering distance constraints*: a DC cannot serve districts located farther than a given distance.

- *Single sourcing restrictions*: a given district receives all the products manufactured in the same plant through a single DC.

- *Minimum volume constraints on transport links*: the use of primary and secondary transport links in each time-period is subject to minimum volume constraints, i.e. the flow through each transport link in each time-period is either zero or greater than a given lower bound specified by the decision maker. These constraints are mainly related to the maximization of the use of truck capacities in order to reduce the transport cost per unit.

- *Minimum volume and maximum capacity constraints on DCs*: the unit transit cost mentioned in §3.1 only applies if the total throughput of the DC in each time-period is between a minimum volume and a maximum capacity.

Further details about these constraints, the underlying motivations and the approach employed to evaluate the resulting parameters could be found in [15].

We now focus on how to take into account vehicle routing in our model while keeping a manageable size for the optimization problem. We propose to develop a two-phase solution approach: the first phase uses an exact clustering method for secondary transport distance and cost identification (see Section 4) whereas the second phase deals with a FLP that considers the results of the first phase as input parameters (see Section 5).

4 First phase of the solution approach: the clustering problem and its formulation

4.1 The clustering problem

The clustering procedure described in this section is the first phase of an approximate solution approach to a facility location problem with vehicle routing from DCs to customers. The objective of this phase is the identification of secondary (DCs to demand districts) transport distances and costs, which will be an essential input to the MILP model solved in the second phase. It would be also possible to adopt a rather standard approach using distances to district barycentres. However, a numerical comparison of the two approaches presented in [14] showed that the clustering-based alternative results in a better approximation of the detailed operational constraints related to secondary transport.

We thus propose a more subtle way to evaluate these key input parameters of the facility location problem. The proposed approach aims at building the secondary transport routes from DCs to customers that might be used to deliver the products in each demand district. This information is then exploited to estimate the distances and costs from DCs to districts. Even if the considered routes will not be the ones used in practice, the resulting estimation is significantly more accurate than the one corresponding to DC-district barycentre distances.

The construction of these routes relies on a clustering algorithm which will form small groups of geographically close customers that should be served on the same route. This grouping practice aims at maximizing truckloads and thus reducing the unit transport cost. A potential cluster should meet two operational constraints: a minimum volume delivered per route and a maximum number of customers served per route. Moreover, as shown in A, the set of clusters to be formed in a given district depends on which DC is delivering the products and on the customer demand, i.e. on the time period considered. We thus iteratively apply the clustering for every district, every DC and every period. The overall procedure is illustrated in Figure. 1.
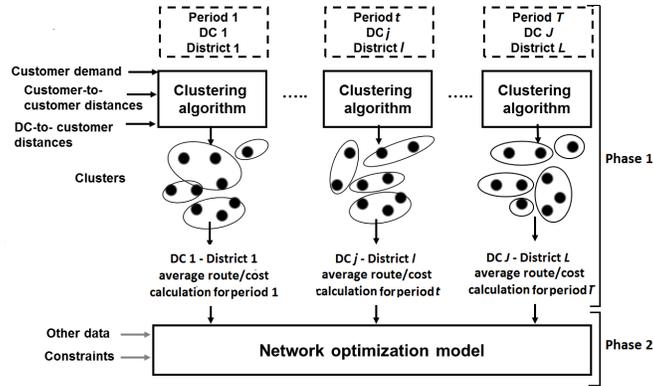

Figure 1: The overall clustering procedure

4.2   Formulation of the clustering problem

Since the clustering is carried out on limited geographical regions (districts), we propose to apply an exact clustering method based on a set-partitioning formulation. Such an approach is in contrast with the heuristic clustering methods recently proposed in [7] and [15]. The first step of the procedure consists in generating a set of potential clusters in each district. Each potential cluster has to meet the routing constraints: the total demand of the cluster in the given time-period has to be greater than a minimum volume fixed by the decision maker, and the number of customers of the cluster must not exceed a given number imposed by the decision maker. The second step of the procedure consists in selecting among the potential clusters those leading to a minimum cost function by solving to optimality the integer set-partitioning problem described in the following paragraph. The parameters of the set-partitioning formulation of the problem are as follows:

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

Evaluating the cost c_k of selecting a given cluster k in the set-partitioning solution depends on which DC is delivering the products (remember that the clustering is done for every district, every DC and every period). More precisely, c_k should be set to the total delivery cost from this DC to cluster k. To calculate this cost, the following formula is used:

$$c_k = D_k * \frac{f(Route_k)}{Truck\_Capacity}$$

- $Route_k$ is the shortest route starting at the considered DC, visiting all the customers of cluster k then coming back to the DC. It is calculated by solving a Traveling Salesman Problem (TSP) for the DC-cluster pair under consideration. If the number of customers is sufficiently small (typically less than 5, which is the case in the application presented in Section 6), this can be done by carrying out a complete enumeration of the possible solutions.

- f is a cost function with fixed and kilometric components that computes the cost of a truck on a given transport route.

- D_k is the total demand of cluster k in the time-period considered and $Truck\_Capacity$ is the maximum capacity of a truck.

- The secondary transport cost per unit from the considered DC to cluster k is approximated by the ratio of $f(Route_k)$ and $Truck\_Capacity$, assuming full truckload transport. The resulting value is then multiplied by the total demand of cluster k in the considered time-period in order to obtain the total delivery cost c_k from the DC to cluster k. Notice that in practice, the total demand of a cluster k is not necessarily delivered by a unique DC since a district could be assigned to different DCs for different product types. Nevertheless, as we are in a pre-processing step, considering the total demand of customers is a reasonable way to estimate secondary transport delivery costs.

The decision variables of the set-partitioning problem are the following:

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

Finally the set-partitioning problem can be formulated as follows:

$Minimize\, Cost = \sum_{k \in \mathbf{K}} c_k x_k$

$Subject\, to:$

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

$x_k \in \{0, 1\} \quad \forall k \in \mathbf{K}$

Through the objective function, we aim at minimizing the total delivery cost from the considered DC to the selected clusters in the considered district. Constraints stipulate that each customer n has to be assigned to exactly one cluster and that cluster selection variables x are binary variables.

4.3 Calculation of distances and costs to be used as input in the facility location problem

As stated earlier, the total demand of a cluster k is not necessarily delivered by a unique DC because cluster k could be assigned to different DCs for different product types. Calculating the total delivery cost from each DC to each district is thus not possible in the first pre-processing phase since the delivered volume is still unknown. Instead of the total delivery cost we propose to use an average cost per unit. Once the clusters obtained for each combination of district/DC/period, the average route distance R_{jlt} from each DC j to each district l in each time-period t is estimated as the weighted average of DC-cluster route distances for all clusters belonging to the district, using cluster demands as weights. An example of calculation is given in Figure. 2 for two distinct time-periods. Then, STC_{jlt} (defined in Table. 1 as the secondary transport cost per unit from DC j to district l in time-period t) is estimated using the following formula: $STC_{jlt} = \frac{f(R_{jlt})}{Truck\_Capacity}$, assuming full truckload transport and taking into account function f and $Truck\_Capacity$ that were defined in §4.2.
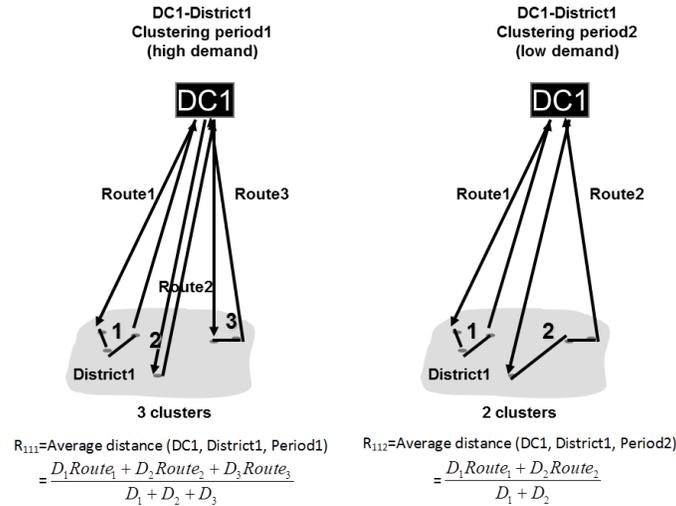

Figure 2: Example of the calculation of the average route distances from DCs to districts. D_i denotes the demand of cluster i in the considered time-period. $Route_i$ is the shortest route starting at the considered DC, visiting all the customers of cluster i then coming back to the DC in the considered time-period

5 Second phase of the solution approach: Mathematical formulation of the facility location problem

We now formulate the multi-period facility location problem as a mixed integer linear program (MILP). Two versions of this problem will be studied: the first one with dynamic assignment variables, i.e. district to DC assignment decisions may change from one period to the next, and the second one with static assignment variables. In both versions, location variables are static as earlier explained. Notice that both location and assignment variables are binary variables. In the remainder of this section, the model parameters, decision variables and MILP formulation will

be presented. Table. 1 summarizes the main parameters of the problem and provides additional comments when needed.

Decision variables are the following:

- Binary location variables $y_j = \begin{cases} 1 & \text{if DC } j \text{ is selected in the solution} \\ 0 & \text{otherwise} \end{cases}$

- Assignment variables: we study a first case with dynamic assignment variables and a second case with static assignment variables

  - *First case: dynamic assignment variables*
    In this case, the following decision variables depend on time-periods:

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

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

    x_{ijlt} and a_{jlt} are defined only if $R_{jlt} <= CD$ (maximum covering distance parameter)

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

  - *Second case: static assignment variables*
    If we choose to use static assignments then decision variables x, z and a do not depend on time-periods, we simply remove the subscript t.

The multi-period facility location problem with dynamic assignments is formulated as a MILP. To obtain the formulation of the problem with static assignments, slight changes should be made to the initial formulation, including the replacement of variables x_{ijlt}, a_{jlt} and z_{ijt} respectively with variables x_{ijl}, a_{jl} and z_{ij}.

Minimize:

$$
\begin{aligned}
Total\ cost = &\sum_{j \in \mathbf{J}} F_j y_j && (Fixed\ opening\ cost) \\
&+ \sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}} PTC_{ij} \sum_{t \in \mathbf{T}} \sum_{\substack{l \in \mathbf{L} \\ R_{jlt} \le CD}} x_{ijlt} D_{ilt} && (Primary\ transport\ cost) \\
&+ \sum_{t \in \mathbf{T}} \sum_{j \in \mathbf{J}} \sum_{\substack{l \in \mathbf{L} \\ R_{jlt} \le CD}} STC_{jlt} \sum_{i \in \mathbf{I}} x_{ijlt} D_{ilt} && (Secondary\ transport\ cost) \\
&+ \sum_{j \in \mathbf{J}} TC_j \sum_{t \in \mathbf{T}} \sum_{\substack{l \in \mathbf{L} \\ R_{jlt} \le CD}} \sum_{i \in \mathbf{I}} x_{ijlt} D_{ilt} && (Transit\ cost)
\end{aligned}
$$

Subject to:

$$\sum_{\substack{j\in\mathbf{J}\\ R_{jlt}\leq CD}} x_{ijlt} = 1 \qquad\qquad \forall t\in\mathbf{T}, i\in\mathbf{I}, l\in\mathbf{L}; D_{ilt}>0 \qquad (1)$$

$$\sum_{\substack{l\in\mathbf{L}\\ R_{jlt}\leq CD}} x_{ijlt}D_{ilt} \geq minTransport(i)z_{ijt} \qquad\qquad \forall t\in\mathbf{T}, i\in\mathbf{I}, j\in\mathbf{J} \qquad (2)$$

$$\sum_{\substack{l\in\mathbf{L}\\ R_{jlt}\leq CD}} x_{ijlt}D_{ilt} \leq M(i,j,t)z_{ijt} \qquad\qquad \forall t\in\mathbf{T}, i\in\mathbf{I}, j\in\mathbf{J} \qquad (3)$$

$$\sum_{i\in\mathbf{I}}\sum_{\substack{l\in\mathbf{L}\\ R_{jlt}\leq CD}} x_{ijlt}D_{ilt} \geq minThroughput(j)y_j \qquad\qquad \forall t\in\mathbf{T}, j\in\mathbf{J} \qquad (4)$$

$$\sum_{i\in\mathbf{I}}\sum_{\substack{l\in\mathbf{L}\\ R_{tjl}\leq CD}} x_{ijlt}D_{ilt} \leq maxThroughput(j)y_j \qquad\qquad \forall t\in\mathbf{T}, j\in\mathbf{J} \qquad (5)$$

$$\sum_{i\in\mathbf{I}} x_{ijlt}D_{ilt} \geq minTransport(j,l,t)a_{jlt} \qquad \forall t\in\mathbf{T}, j\in\mathbf{J}, l\in\mathbf{L}; R_{jlt}\leq CD \qquad (6)$$

$$\sum_{i\in\mathbf{I}} x_{ijlt} \leq |I|a_{jlt} \qquad \forall t\in\mathbf{T}, j\in\mathbf{J}, l\in\mathbf{L}; R_{jlt}\leq CD \qquad (7)$$

$$z_{ijt} \leq y_j \qquad \forall t\in\mathbf{T}, i\in\mathbf{I}, j\in\mathbf{J} \qquad (8)$$

$$y_j,\ x_{ijlt},\ a_{jlt},\ z_{ijt} \in\{0,1\} \qquad \forall t\in\mathbf{T}, i\in\mathbf{I}, j\in\mathbf{J}, l\in\mathbf{L}; R_{jlt}\leq CD \qquad (9)$$

The objective function consists in the total distribution cost (fixed cost of opening DCs, primary and secondary transport costs and transit cost).

Constraints (1) state that the demand of district l for the products of plant i has to be satisfied in each time-period t and routed through a single DC (as x are binary variables). Constraints (2) ensure that if the primary transport link between plant i and DC j is selected in time-period t ($z_{ijt} = 1$) then the corresponding total volume has to be greater than the minimum volume $minTransport(i)$. Constraints (3) stipulate that if the link (i,j) is not selected in time-period t ($z_{ijt} = 0$) then all the x_{ijlt} have to be set to 0.
Constraints (4) state that if DC j is selected ($y_j = 1$), then the flows going through j in each time-period t have to be greater than the corresponding minimum volume limit. Constraints (5) stipulate that:

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

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

Constraints (6) ensure that if the secondary transport link between DC j and district l is selected in time-period t ($a_{jlt} = 1$) then the corresponding total volume has to be greater than the minimum volume $minTransport(j,l,t)$. Constraints (7) stipulate that if the link between j and l is not selected in time-period t ($a_{jlt} = 0$) then all of the variables x_{ijlt} have to be set equal to 0. Constraints (8) state that if DC j is not opened ($y_j = 0$) then all of the variables z_{ijt} have to be set equal to 0. Constraints (9) are the integrality constraints.

The model described in this paper is highly constrained due to the combination of minimum volume, maximum capacity and maximum covering distance constraints. It is thus possible to face frequent feasibility issues due to the conflict between these constraints. To avoid this, it is possible to allow violating some of the constraints but keeping total constraint violation as low as possible via penalties. For instance, if the violation of primary transport constraints is allowed, new continuous variables v'_{ijt} and v''_{ijt} have to be introduced:

- v'_{ijt} is a positive variable that has to be greater than $minTransport(i)$.

- v''_{ijt} is a nonnegative variable used to compute the amount of violation of the minimum volume constraint on a given primary transport link (i, j) in time-period t. It has to be less than $minTransport(i)$ and will be minimized, as violation is penalized in the objective function.

Constraints (2) and (3) should be replaced with the following constraints:

$$\sum_{\substack{l \in \mathbf{L} \\ R_{jlt} \leq CD}} x_{ijlt} D_{ilt} = v'_{ijt} - v''_{ijt} \qquad \forall t \in \mathbf{T}, i \in \mathbf{I}, j \in \mathbf{J} \qquad (10)$$

$$v'_{ijt} \geq minTransport(i) z_{ijt} \qquad \forall t \in \mathbf{T}, i \in \mathbf{I}, j \in \mathbf{J} \qquad (11)$$

$$v'_{ijt} \leq M(i, j, t) z_{ijt} \qquad \forall t \in \mathbf{T}, i \in \mathbf{I}, j \in \mathbf{J} \qquad (12)$$

$$v''_{ijt} \leq minTransport(i) z_{ijt} \qquad \forall t \in \mathbf{T}, i \in \mathbf{I}, j \in \mathbf{J} \qquad (13)$$

Constraints (10) stipulate that the total volume going from plant i to DC j in time–period t is expressed as a difference between the nonnegative continuous variables v'_{ijt} and v''_{ijt}. Constraints (11)-(13) ensure that:

- If the link (i, j) is selected in time-period t ($z_{ijt} = 1$) then $v'_{ijt} \geq minTransport(i)$, $v'_{ijt} \leq M$ and $v''_{ijt} \leq minTransport(i)$.

- If the link (i, j) is not selected in time-period t ($z_{ijt} = 0$) then $v'_{ijt} = 0$ and $v''_{ijt} = 0$.

This enables us to compute the violation of minimum volume constraints on primary transport links and to penalize it in the objective function. The following term should be thus added to the objective function:

$$\sum_{i \in \mathbf{I}} \sum_{j \in \mathbf{J}} PI_{ij} \sum_{t \in \mathbf{T}} v''_{ijt}$$

6 Case study and computational experiments

In this section, we intend to test the proposed multi-period facility location model using real-life data coming from our case study. The model described in Section 5 was implemented using C++ language and the MILP was solved using the commercial solver ILOG Cplex version 12.5. Tests were carried out on a PC Intel Core (TM) i5-3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7.

We first present the case study and describe the test instances used. Then, we analyze the numerical results in order to compare the use of static assignments with the use of dynamic assignments when demand is seasonal. We also compare the output of the multi-period model with the output of its single-period counterpart. The computational difficulties encountered during these numerical experiments motivated the development of a heuristic solution approach to be presented in Section 7.

6.1 Case study

The case study addressed originates from the french car manufacturer Renault. The supply chain network under consideration is a car distribution network consisting of 16 assembly plants, 51 potential distribution centres and 448 car dealers. The country is divided into 92 districts where the company distributes its cars and each of the 448 car dealers belongs to a district. As pointed out in [12], new car sales in France show seasonal quarterly fluctuations and are typically characterized by a valley in the third quarter due to a significant drop in car registrations in August and September (see Figure. 3). Since the industrial experts involved in this study indicated the same behaviour of the company car sales, we propose to use a planning horizon of one year divided into 4 time-periods (quarter 1 covers months 1, 2, 3; quarter 2 covers months 4, 5, 6: etc.). Seasonality factors F_t for each season t are estimated using historical data of the company car sales and lead to the following values: $F_1 = 1.1$ $F_2 = 1.2$ $F_3 = 0.7$ $F_4 = 1$. Demand D_{ilt} for products of plant i, district l and time-period t is thus computed as $D_{ilt} = F_t \frac{D_{il}}{|T|}$ (D_{il} is the annual demand of district l for products of plant i as defined in Table. 1).
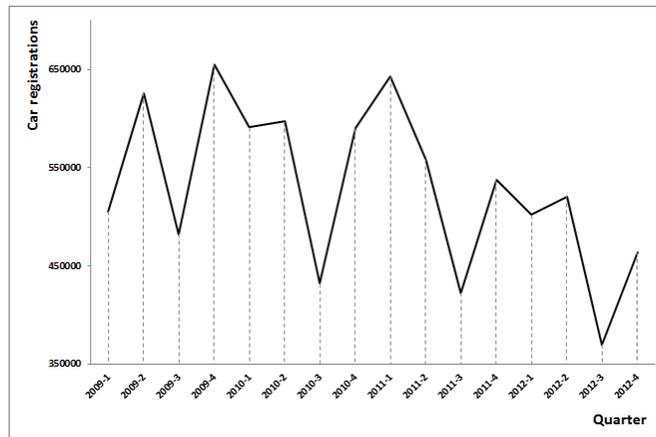

Figure 3: Historical data of passenger car registrations in France. Source: INSEE

As stated in §3.3, minimum volume constraints on transport links are introduced in our model in order to ensure high capacity utilization for trucks and thus to reduce transport costs. Moreover, preliminary computational experiments showed that solving the problem without these minimum volume constraints leads to opening many transport links with very low volume transported per year. At the operational level, this means either waiting for a long time at the factory (or the DC) until a full truckload is consolidated (which would increase the total lead time perceived by the customer) or shipping trucks that are not fully loaded (which would increase transport costs). Let us consider a simple example in which we assume that the cost of a truck from A to B is $1000, the capacity of a truck is 10 cars and products must be shipped once a week. Transporting 1 car per week means paying $1000 per car, transporting 10 cars per week means paying $100 per car while transporting 11 cars per week means paying $181.8 per car, etc. Figure. 4 shows how the cost per unit changes with the weekly transported volume. When the weekly transported volume starts to be higher than a full truckload (10 units), the actual cost per unit (dashed line) decreases significantly and stays very close to the cost per unit of a full truckload (100). Given this observation, the minimum volume in our case study is set to the volume that ensures on average the shipment of 1 full truckload per week in each time-period. In this case, a suitable approximation of the transport cost per unit is the cost of a fully loaded truck (100 per unit in

the example). Evidently, using the exact cost function showing this effect of economies of scale would be more accurate but would significantly increase the difficulty in solving the optimization problem.
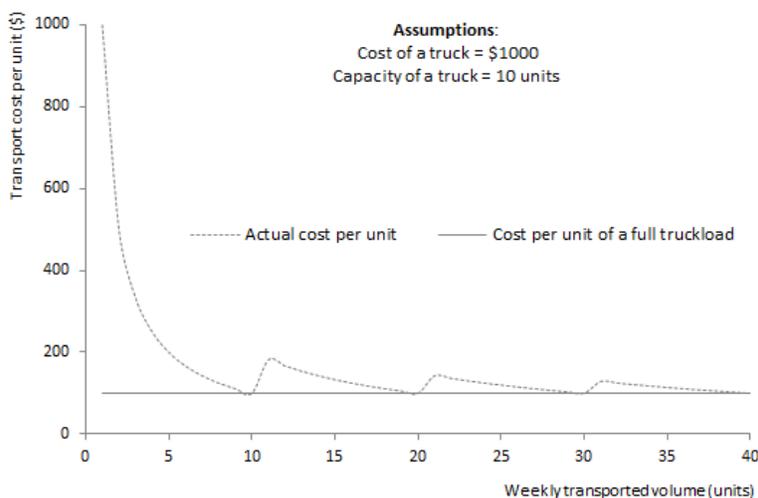

Figure 4: Example of changes in the transport cost per unit when the weekly transported volume changes

Moreover, the management of the distribution centres and the related activities is outsourced to logistics suppliers. For this reason there are no fixed opening costs to be charged to the car manufacturer but there is only a unit transit cost (considered as given in our study) to be paid to the logistics supplier each time a car goes through a DC, provided that a *minimum transit volume* (lower bound on transit flows) is handled in each time-period. Notice that this simply means that the fixed opening costs used in the model formulation (see Section 5) are set to 0. Furthermore, the other parameters of the problem are set based on the preferences of the decision maker and information from company databases: the value of the maximum covering distance is fixed at 680 kilometres for all the secondary transport routes from DCs to districts, the minimum and maximum volumes per DC per time-period are set according to some assumptions on the flexibility of each DC. Accordingly a reference dataset representing the typical values of input data and parameters is defined for the distribution network in France and used for numerical experiments.

In each district, customers are grouped into clusters according to the following constraints based on real-life conditions in car distribution:

- Minimum transport volume per cluster: value that ensures on average the shipment of 1 full truckload every week

- Maximum number of customers per cluster: 3

Then, weighted average distances from DCs to districts are calculated as explained in §4.1. Applying an optimal set-partitioning approach and using complete enumeration of all possible solutions to solve traveling salesman problems in each district was possible for this case study thanks to the fact that the maximum number of customers per cluster is limited and the demand region is small (district). The minimum volume per secondary transport link $minTransport(j, l, t)$ depends on the result of the clustering of district l in time-period t for deliveries from DC j. Its

14

value is calculated as the minimum transport volume per cluster multiplied by the number of clusters for district l, in the optimal solution of the clustering problem.

The first experiences with typical parameters and data of our case study showed that it is not always possible to comply with minimum volume constraints for primary transport links in the presence of maximum covering distance constraints. Tight maximum covering distance constraints favour the use of a large number of DCs to be close to customers but this results in splitting the global volume to be distributed from the plants in many DCs so that the constraints of minimum volume on primary transport links are likely to be violated in the corresponding solution. To avoid the resulting feasibility issues, we allow the violation of the strict minimum volume constraints on primary transport links but seek to keep these violations as low as possible via penalties. We propose to set the penalty for violating the minimum volume constraint on a primary transport link (i, j) to the unit transport cost on this link. This makes violating the constraint possible only if there is some opportunity to save costs. Figure. 5 illustrates the total cost as a function of the transported volume in this case.
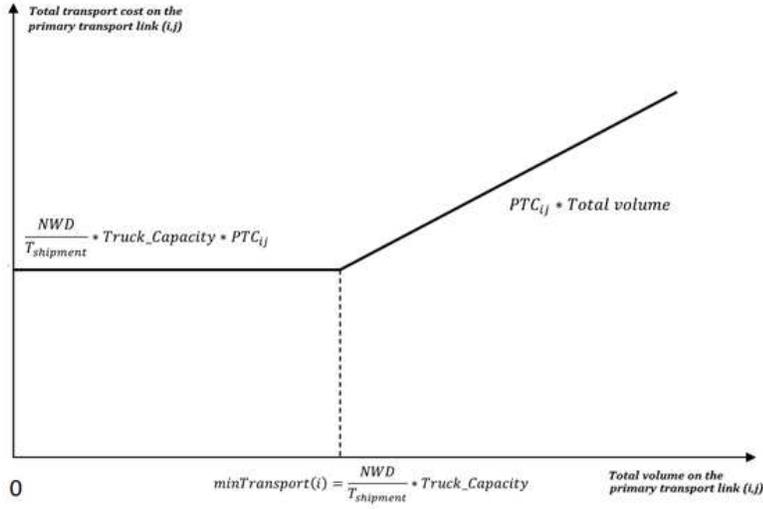

Figure 5: Total transport cost as a function of the transported volume on a given primary transport link

If the volume on the transport link is greater than the minimum required volume then all the products can be shipped in full truckloads; a unit transport cost equal to PTC_{ij} is paid. On the contrary, if the volume on the link is less than the minimum required volume, then regardless to the quantity, a total cost equal to $\frac{NWD}{T_{shipment}} * Truck\_Capacity * PTC_{ij}$ is paid. NWD is the number of working days per year, $T_{shipment}$ is the targeted shipment periodicity in days, PTC_{ij} is the cost per unit if a full truckload is used and $Truck\_Capacity$ is the average truck capacity. Observe that this amounts to keep shipping trucks with the same periodicity $T_{shipment}$ even if they are not fully loaded.

In order to test the performance of exact and heuristic solutions, other instances were also generated by varying some of the problem parameters in the reference dataset. In the first set (A), customer demand was varied to obtain 20 test instances based on the reference dataset. In each instance, a different pseudo-random coefficient in [0.5,1.5] per customer is used. The demand of each customer was multiplied by the corresponding pseudo-random coefficient. In the second set (B), the list of potential DCs was varied to obtain another 20 instances based on the reference

dataset. Each of these instances contains 46 potential DCs instead of 51. In each instance, 5 DCs were randomly selected to be removed from the list of potential DCs.

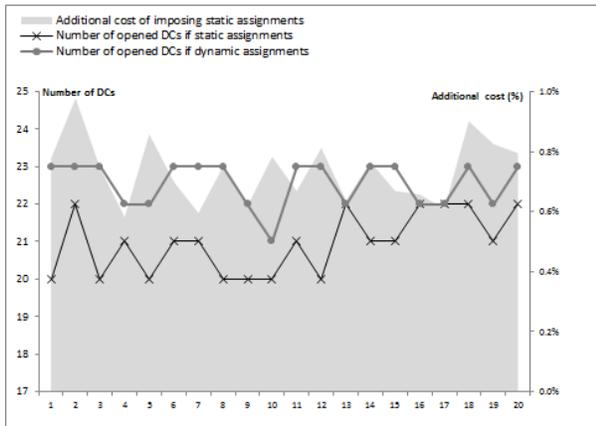## 6.2   Multi-period model: static vs. dynamic assignment decisions

In this section we propose to compare the solutions obtained using a multi-period model featuring dynamic assignment decisions with those obtained using a multi-period model featuring static assignment decisions. The solutions are obtained using the MILP formulation presented in Section 5 and then compared in terms of cost, network structure (number of opened DCs) and computation time. We first run two tests using the reference dataset: one test relates to the multi-period model with dynamic assignments and the other one relates to the multi-period model with static assignments. Note that the optimality gap of Cplex 12.5 was limited to 0.2% as proving exact optimality is very time-consuming. The number of DCs obtained in each case as well as the additional cost of imposing static assignment decisions are illustrated in Table. 2.

Table. 2 shows that when using static assignments of demand districts to DCs, there is an increase of 1% in costs while reducing the number of opened DCs from 25 to 21. During periods of low demand (period 3 in our case), splitting flows in too many locations is not possible because of the strict minimum throughput constraints on DCs. The number of DCs is thus reduced in order to meet these constraints in all periods, including periods of low demand. However, when using dynamic assignments, it is possible to adapt flows to demand variation from period to period and thus to open more DCs in order to reduce total costs.

To confirm these results, further tests were carried out using instances A and B defined in §6.1. As heavy computation times were needed to carry out each test, especially for the version with dynamic assignments (more than 3 days as shown in Table. 2), the tolerance of Cplex was limited to 0.5%. Moreover, the optimal solution of the problem with static assignments was used as an initial feasible solution to the problem with dynamic assignments in order to reduce computation times in the latter case. Computational results are summarized in Figure. 6. The values shown in Figures. 6(a) and 6(b) confirm the previous observations obtained using the reference dataset, i.e imposing static assignments leads to opening fewer DCs and to an increase in cost that stays below 1%.

The reduction in the number of opened DCs and the cost increase observed while using the static assignment model are mainly explained by the need to comply with strict minimum throughput constraints on DCs in low demand periods. Relaxing these constraints or modifying the related numerical parameters might thus have an impact on the conclusion of our computational study. In order to better assess this possible impact, we carried out a sensitivity analysis on the minimum throughput constraint violation. We thus investigated a variant of our problem in which violation of constraints (4) is allowed but penalized in the objective function. We conducted additional computational experiments on two sets of instances based on (1) our reference dataset, (2) our reference dataset in which the minimum throughput values $minThrougput(j)$ are doubled. Moreover, we considered a unit penalty cost for violating the DC minimum throughput constraints varying between $0.1TC_j$ and $10TC_j$. The corresponding results are displayed in Figures 7(a) and 7(b). They show that in all test instances, the number of DCs opened in the case of static assignments is lower than the number opened in the case of dynamic assignments and that the cost increase varies between 0.47% and 0.9%. Observe that the difference in DC number is more noticeable and the additional cost is higher when the unit penalty cost is above $2TC_j$. However, for all instances, the cost increase stays below 1%. It thus seems that for our case study, the selection of the parameters on the DC minimum throughput constraints has a limited impact on the cost

increase resulting from the use of static assignments. Incurring an additional cost not exceeding 1% when imposing static assignments of demand districts to DCs was considered as acceptable by our case study company especially in view of the corresponding operational benefit: having the same distribution flows over the whole planning horizon and managing fewer distribution centers considerably simplifies day-to-day operations. Accordingly, a strategy with static assignments appears to be the best alternative in our case study, even if demand is seasonal. In the next steps of our computational investigations, we will thus consider only the multi-period model with static assignments. Nevertheless, we note that this conclusion pertains mainly to our case study and that in other applicative contexts, exploiting the flexibility allowed by the use of dynamic assignments might prove of practical interest.
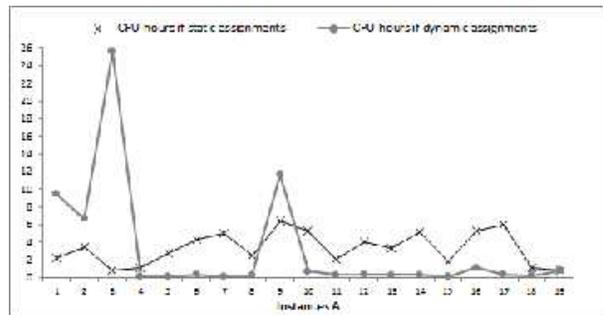

(a) Number of opened DCs and additional cost (Instances A)
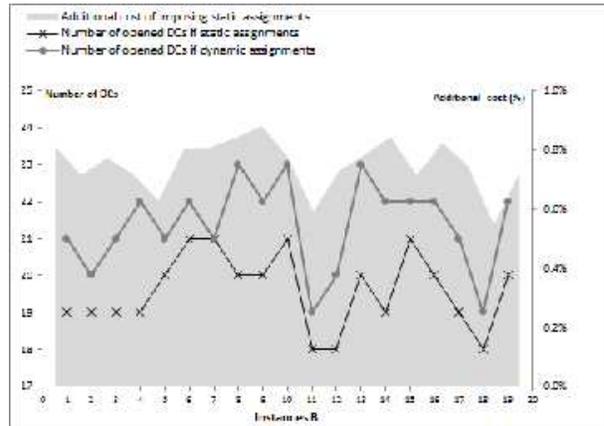
(b) Number of opened DCs and additional cost (Instances B)
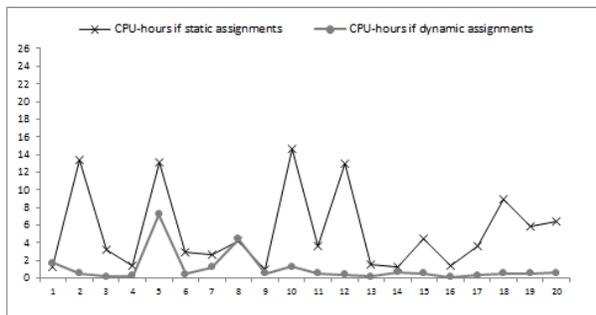
(c) Computation time (Instances A)

(d) Computation time (Instances B)

Figure 6: Multi-period model using dynamic assignments vs. multi-period model using static assignments.

Moreover, results from Figures. 6(c) and 6(d) show that long computation times (more than 13 hours for some instances) are required by Cplex to find solutions within 0.5% of exact optimality. Out-of-memory problems were also encountered for a few instances. Note that the reduced computation times obtained when using dynamic assignments are explained by the fact that an initial feasible solution (the one obtained with static assignments) is provided to Cplex. The computational difficulties faced during these tests motivated the development of a heuristic procedure capable of obtaining good quality solutions within shorter computation times. The heuristic solution approach based on linear relaxation is explained in Section 7. In the next subsection, we
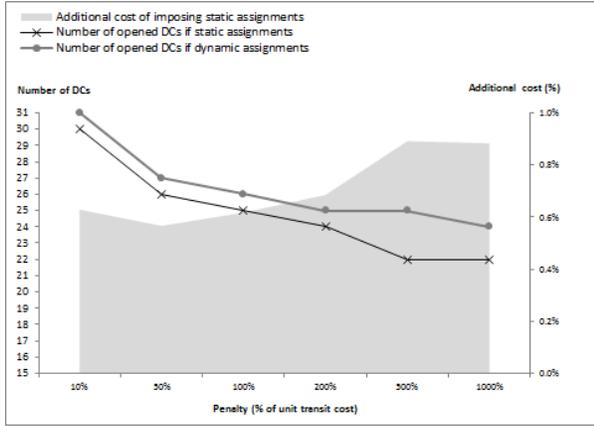
(a) Reference dataset
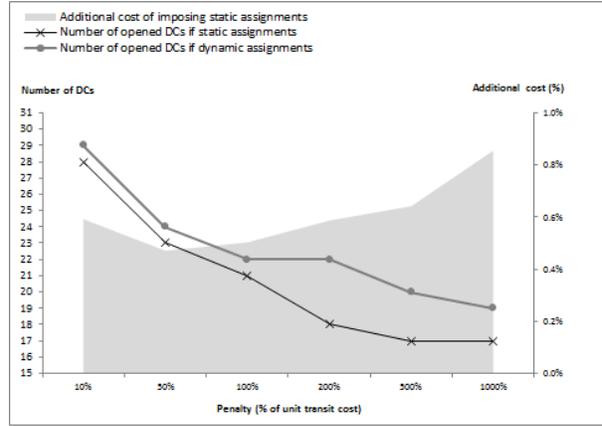
(b) Reference dataset with minimum throughput values doubled

Figure 7: Using dynamic assignments vs. using static assignments when violation of minimum throughput constraints is allowed.

compare the output of the multi-period model using static assignments with the output of its single-period counterpart.

6.3 Multi-period model vs. single-period model

We focus in this section on a comparison between the multi-period model using static assignments and its single-period counterpart. The two models tested here thus consider static location and assignment decisions but the only difference is that the multi-period model takes into account seasonal demand and constraint satisfaction in each time-period, whereas the single-period model uses average estimates over the whole planning horizon. The main objective of our experiments is to gain a better understanding of the impact of demand seasonality on the network structure and to identify whether it is worth using a more complex multi-period model. The solutions are obtained using the MILP formulation presented in Section 5. The obtained cost cannot however be used to compare solutions in a fair way. While the solution of the multi-period model meets the detailed period by period constraints, the solution of the single-period model is likely to violate these constraints in some time-periods. A fair comparison would thus require to include an estimation of the financial penalties to be incurred whenever constraints are violated. Since obtaining such an estimation was not possible in our case, we chose to compare the models not in terms of cost but rather in terms of network structure (number of DCs).

We first run two tests using the reference dataset and Cplex 12.5 with an optimality gap limited to 0.2%. The number of DCs opened in the case of a multi-period model is 21 while it is 26 in the case of a single-period model. The difference in the number of DCs is significant as the multi-period version has to meet strict minimum volume constraints for all opened DCs in all time-periods while keeping the same network configuration and routes over the whole planning horizon. In order to meet minimum volume constraints in low-demand periods, it is thus necessary to reduce the number of opened DCs. This might explain why only 21 facilities are opened in this case. The single-period model does not capture demand variation and thus is not required to satisfy operational constraints period by period. This makes possible the use of a higher number of DCs (26) in order to reduce distribution costs. These observations are confirmed by the tests done with instances A and B, the results of which are summarized in Figure. 8. Accordingly, we

18

can conclude that the main benefit of using a multi-period model in our case is to ensure that the operational constraints required by the decision maker are met in each time-period, even in periods of low demand. The single-period model uses average estimates when expressing costs and constraints and thus its optimal solution is likely to violate detailed period by period constraints.
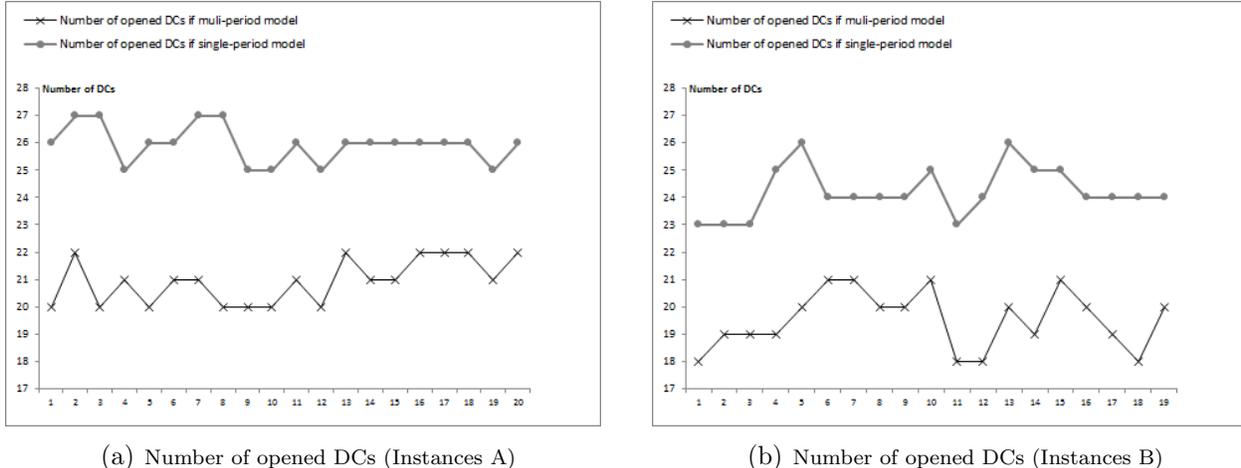

(a) Number of opened DCs (Instances A)

(b) Number of opened DCs (Instances B)

Figure 8: Multi-period model vs. single-period model.

7 Heuristic solution approach

As shown in Section 6, long computation times, sometimes exceeding 13 hours, are required by Cplex to find solutions within 0.5% of exact optimality for the multi-period facility location problem with static assignments. Developing a heuristic method capable of obtaining good quality solutions within reduced computation times is thus needed. We focus here on the resolution of the problem with static assignments as the use of dynamic assignments was found to be less relevant for our case study company. We first describe the heuristic method. We then study its performance through a comparison with the straight application of Cplex to the MILP formulation.

7.1 Description of the heuristic solution approach

In order to improve computation times, a heuristic method based on a linear relaxation approach is now investigated (see Figure. 9). An argument supporting this idea is the tightness of the lower bounds provided by the linear relaxation of the formulation presented in Section 5 (typical deviations are less than 4.5%). The heuristic method comprises two stages: the first stage focuses on location decisions whereas the second one deals with assignment decisions. In the first stage, our objective is to decide which DCs should be opened, i.e. to determine the value of each location variable y. We propose to solve a partial linear relaxation of the original MILP where the 0-1 constraints on variables x, z and a have been removed whereas they have been kept for y variables. Our preliminary computational experiments showed that this is an efficient way to quickly identify which DCs to open. After fixing the binary location variables in the first stage, the resulting assignment problem is easier to solve than the original MILP as a large fraction of the binary variables x are set to 0 (namely those corresponding to unselected DCs). We propose, in the second stage, to fix as many assignment variables as possible before reintroducing the 0-1

19

constraints removed in the first stage and solving the resulting MILP. To this end, we develop a dynamic version of the fixing strategy introduced in [15] for solving heuristically a single-period distribution network design problem. This fixing strategy should enable us to decrease the number of binary assignment variables before solving the final MILP and thus to reduce computation times in the second stage of the heuristic procedure.

The primary objective of the fixing strategy is to close any secondary transport link (j, l) where the minimum volume constraint is violated (this is one of the main issues noticed in the optimal solution of the first stage of the heuristic), provided that the problem remains feasible. Thus, we make sure that after closing (j, l), the throughput of DC j in each time-period t would be kept above its minimum volume $minThroughput(j)$ and that there would be another DC k capable of delivering the demand of l in each time-period t without exceeding its maximum volume of throughput $maxThroughput(k)$. Notice here that checking the constraint satisfaction before setting variables x to 0 has to be made for all time-periods. This procedure is repeated for all the secondary transport links then the resulting linear problem is solved. In case the problem becomes infeasible, the fixing constraints added in the previous iteration are removed and the loop is stopped. Otherwise, the procedure is iterated until there are no further flows that can be fixed at 0 or the problem becomes infeasible. The second step of the fixing strategy is to reintroduce the 0-1 constraints for all the free assignment variables and to solve the resulting MILP with a commercial solver.
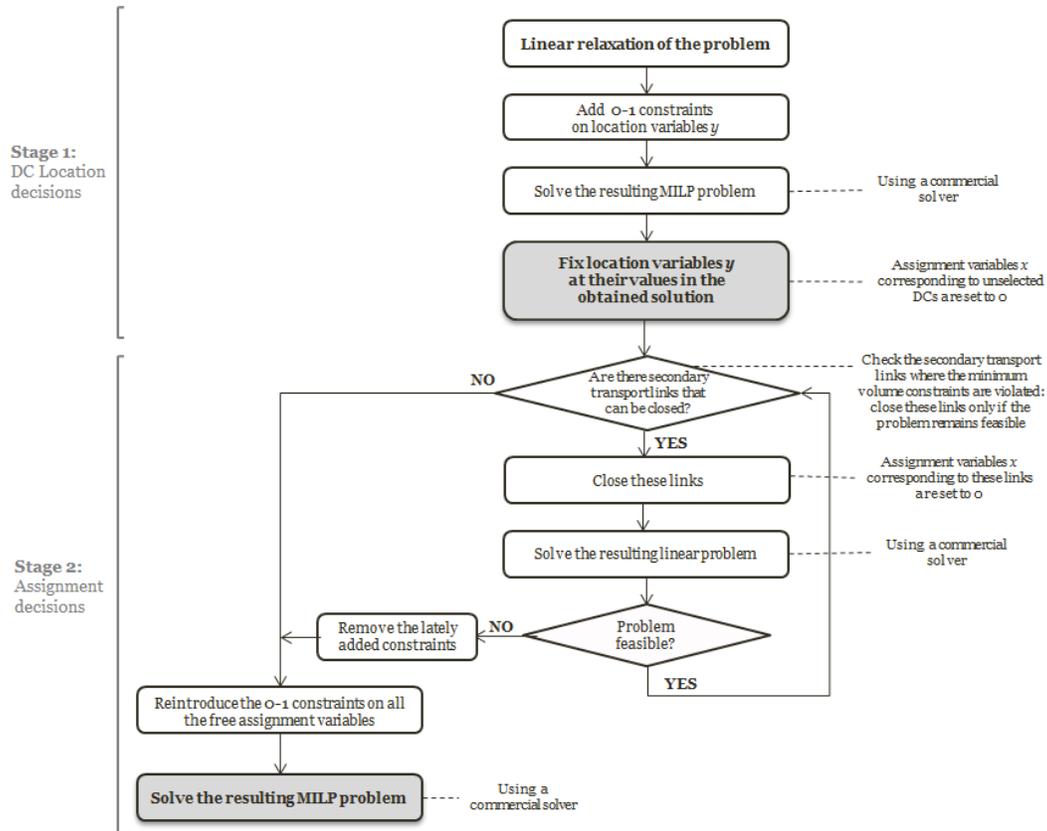

Figure 9: The linear-relaxation based heuristic

7.2 Computational experiments

We propose to study the performance of the heuristic approach described in §7.1 by comparing it with the performance of an exact solution approach based on the resolution of the MILP formulation using CPLEX solver. We carry out 40 tests using instances A and B defined in §6.1. All the linear and mixed integer linear programs (including those used in the heuristic) are solved with Cplex, limiting the MILP optimality gap to 0.2%.A PC Intel Core(TM) i5–3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7, was used. The obtained results are summarized in Table. 3 and Table. 4. For each test, the computation time required by Cplex to find a solution within 0.2% of optimality is provided as well as the computation time of the heuristic method and its quality (columns 2-4).

The numerical results show that the heuristic approach significantly reduces computation times from several hours to less than two minutes while providing good quality solutions. The heuristic quality varies between 0.82% and 2.48% with an average of 1.64%. To further confirm the performance of the heuristic method, we carried out additional experiments and solved the MILP formulation using Cplex with an optimality gap set to 2% (average quality of our heuristic) instead of 0.2%. In Table. 3 and Table. 4, the resulting computation times and quality measures (calculated each as the relative difference between the Cplex solution within 2% of optimality and the Cplex solution within 0.2% of optimality) are displayed in columns 5 and 6. These additional tests show that the heuristic method is capable in 12 out of 40 instances of providing better quality solutions than Cplex-2% within faster computation times. For the other instances, Cplex-2% provides solutions of better quality but requires longer computation times.

8   Conclusion and future research

In this paper, a multi-period facility location problem involving several operational features has been investigated. We proposed an original way of evaluating transport routes and costs from distribution centres to customers based on a dynamic clustering procedure. We formulated our problem as a MILP in which location decisions are static whereas DC-district assignments can be static or dynamic.

We carried out various tests based on our case study dealing with car distribution. The qualitative analysis showed that multi-period modeling allows building a distribution network which captures demand seasonality and meets strict operational constraints in all time-periods. This results in opening fewer distribution centres than when using a single-period model. Furthermore, it was observed that the additional cost to be incurred by our case study company when deciding to keep static assignment decisions in a multi-period context stays within an acceptable limit. Using static assignments thus appeared to be the best alternative at the operational level as dealing with the same distribution routes over the whole planning horizon significantly simplifies day-to-day operations. Moreover, in that case, fewer distribution centres are opened, which also facilitates the work of operational managers. As far as computation times are considered, we proposed a heuristic procedure based on linear relaxation to reduce the time required to find good quality solutions. The heuristic output was compared with the output of a straight application of a state-of-the–art MILP solver. The heuristic provided good quality solutions (about 1.6%) within short computation times (less than two minutes).

Finally, the proposed model and solution approach were applied to a facility location problem with time-varying demand but could also be used in a context of uncertain demand with discrete scenarios having each a fixed probability of occurrence. Our approach is indeed very similar to the one employed in two-stage stochastic location problems where location decisions are made in the first stage and assignment decisions occur after random parameters become known. One of the challenging options for future research would be thus to study a stochastic version of the problem. Another outlook that may be interesting for future research is to apply the model to case-studies in other fields than car distribution. This can present new challenges in the pre-processing clustering step if the number of customers per demand region is large or the maximum cardinality of a cluster is greater than 3. In this case, applying a MILP solver to the clustering problem based on a set-partitioning formulation will lead to extensive computation times and using complete enumeration to compute optimal traveling salesman (TSP) tours in secondary transport will no longer be possible. Thus, one interesting research direction could be to implement efficient

heuristics for these two problems.

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

We aim here at analyzing the results of the pre-processing clustering approach applied to each combination of district/DC/period, using the case study introduced in §6.1. First, we study cluster variation according to the DC delivering products, then, we examine this variation from period to period when customer demand changes.

## A.1   Variation of the clustering results according to the DC delivering products

We consider here the solution of the clustering for one of the districts (namely district 51), comprising six customers, when served from different DCs in time-period 3. Figure. 10 shows the clusters obtained in this district when served by DCs in Auxerre, Batilly and Chaligny. The DC in Auxerre situated at the south-west of the district leads to 3 clusters of 1 customer and 1 cluster of 3 customers. The DC in Batilly situated at the east of the district leads to 2 clusters of 1 customer and 2 clusters of 2 customers. The DC in Chaligny situated at the south-east of the district leads to 1 cluster of 1 customer, 1 cluster of 2 customers and 1 cluster of 3 customers. The obtained results show thus that clusters of a given district could vary according to the DC serving it. Therefore we chose to iteratively apply the clustering procedure for each DC.

## A.2   Variation of the clustering results according to time-periods

We study here the clustering output of district 1, comprising five customers, when served from DC Bourges in different time-periods (period 2 with $F_2 = 1.2$, period 3 with $F_3 = 0.7$ and period 4 with $F_4 = 1$). Figure. 11 shows the resulting clusters according to time-periods. Period 2 corresponding to the highest volume leads to 3 clusters of 1 customer and 1 cluster of 2 customers. Period 4 with seasonality factor set to 1 leads to 2 clusters of 2 customers and 1 cluster of 1 customer. Period 3 corresponding to the lowest volume leads to 1 cluster of 3 customers and 1 cluster of 2 customers. This means that when demand gets lower, the average size of clusters gets higher in order to ensure full truckload transport every week. The obtained results show thus that clusters in a given district could vary according to time-periods. For this reason, we chose to iterate applying the clustering procedure for each time-period.

A.3   Conclusion

Based on several experiments using the clustering algorithm described in Section 4, we analyzed the optimal clustering solutions (number and size of clusters) in variety of situations. As shown in some examples, the results in a given district could change when the DC delivering the products changes or the time-period (i.e. customer demand) is different. Accordingly, we should carry out a separate clustering procedure for every district, every DC and every time period, as proposed in Section 4.
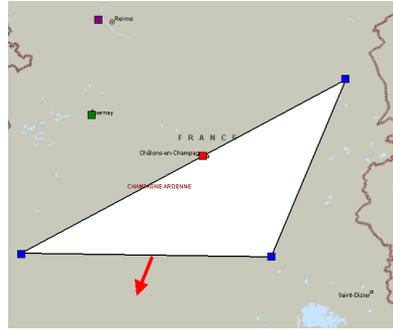
(a) Clusters of district 51 when delivered from DC Auxerre in period 3
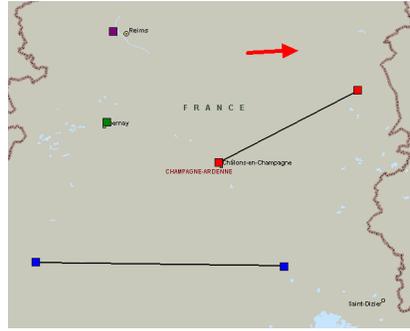
(b) Clusters of district 51 when delivered from DC Batilly in period 3
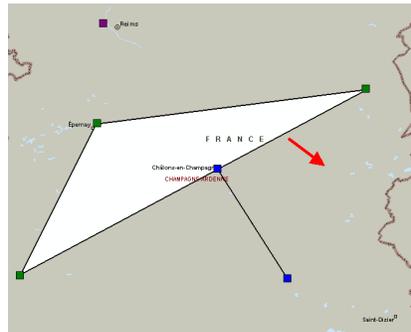
(c) Clusters of district 51 when delivered from DC Chaligny in period 3

Figure 10: Cluster variation in district 51 when delivered from DCs Auxerre, Batilly and Chaligny. The arrow indicates the direction situating the DC.

References

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

[2] E. Aghezzaf. Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society 56 pp. 453—462*, 2012.

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

[4] S. A. Alumur, S. Nickel, F. Saldanha da Gama, and V. Verter. Multi-period reverse logistics network design. *European Journal of Operational Research 220 pp. 67—78*, 2012.

[5] D. Ambrosino and M. G. Scutellá. Distribution network design: New problems and related models. *European Journal of Operational Research 165 pp. 610—624*, 2005.

[6] A. B. Arabani and R. Zanjirani Farahani. Facility location dynamics: An overview of classifications and applications. *Computers and Industrial Engineering 62 pp. 408—420*, 2012.

[7] S. Barreto, C. Ferreira, J. Paixão, and B. Sousa Santos. Using clustering analysis in a capacitated location-routing problem. *European Journal of Operational Research 179 pp. 968—977*, 2007.

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

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

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

Figure 11: Cluster variation in district 1 when served from DC Bourges in different time-periods. Customers belonging to a same cluster are linked together.

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

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

[10] K. Dogan and M. Goetschalckx. A primal decomposition method for the integrated design of multi-period production-distribution systems. *IIE Transactions 31 pp. 1027–1036*, 1999.

[11] B. Fleischmann, S. Ferber, and P. Henrich. Strategic planning of BMW's global production network. *Interfaces 36 pp. 194—208*, 2006.

[12] B. Gautier. Le marché automobile: Un cycle spécifique? *Economie et statistique 282*, 1995.

[13] H. Gunnarsson, M. Rönnqvist, and J. Lundgren. Supply chain modelling of forest fuel. *European Journal of Operational Research 158 pp. 103—123*, 2004.

[14] M. Kchaou Boujelben. Modeling and solving a distribution network design problem with multiple operational constraints. Application to a case-study in the automotive industry. *Ph.D. Dissertation, Department of Industrial Engineering, Ecole Centrale Paris*, 2013.

[15] M. Kchaou Boujelben, C. Gicquel, and M. Minoux. A distribution network design problem in the automotive industry: MIP formulation and heuristics. *Computers and Operations Research 52 pp. 16—28*, 2014.

[16] P. Kouvelis and M. Rosenblatt. A mathematical programming model for global supply chain management: Conceptual approach and managerial insights. *In Supply Chain Man-*

agement: Models, Applications, and Research directions, Applied Optimization, Kluwer, Dordrecht (Chapter 10) pp. 245—277, 2002.

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

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

[19] E. Melachrinoudis and H. Min. The dynamic relocation and phase-out of a hybrid, two-echelon plant/warehousing facility: A multiple objective approach. *European Journal of Operational Research 123 pp. 1—15*, 2000.

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

[21] M. T. Melo, S. Nickel, and F. Saldanha da Gama. Facility location and supply chain management: A review. *European Journal of Operational Research 196 pp. 401—412*, 2009.

[22] M. T. Melo, S. Nickel, and F. Saldanha da Gama. An efficient heuristic approach for a multi-period logistics network redesign problem. *TOP pp. 1–29*, 2011.

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

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

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

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

[27] S. Nickel and F. Saldanha da Gama. Multi-period facility location. *In Location Science. Springer International Publishing*, 2015.

[28] S. Nickel, F. Saldanha da Gama, and H.-P. Ziegler. Multi-stage stochastic supply network design problem with financial decisions and risk management. *Omega 40 pp. 511–524*, 2012.

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

[30] S. Park, T.-E. Lee, and C. S. Sung. A three-level supply chain network design model with risk-pooling and lead times. *Transportation Research Part E 46 pp. 563—5819*, 2010.

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

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

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

[34] D. Vila, A. Martel, and R. Beauregard. Designing logistics networks in divergent process industries: A methodology and its application to the lumber industry. *International Journal of Production Economics 102 pp. 358—378*, 2006.

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

[36] V. F. Yu, S.-W. Lin, W. Lee, and C.-J. Ting. A simulated annealing heuristic for the capacitated location routing problem. *Computers and Industrial Engineering 58 pp. 288-299*, 2010.

Parameter	Description	Additional comments
I	Set of plant indices ($i = 1..|I|$)	
J	Set of DC indices ($j = 1..|J|$)	
L	Set of district indices ($l = 1..|L|$)	
T	Set of time-period indices ($t = 1..|T|$)	
D_{il}	Total demand of district l for the products manufactured in plant i during the whole planning horizon	
D_{ilt}	Demand of district l for the products manufactured in plant i in time-period t	
$minThroughput(j)$	Minimum volume that has to go through DC j in each time-period if it is opened	
$maxThroughput(j)$	Maximum volume that can go through DC j in each time-period if it is selected in the solution	
$minTransport(i)$	Minimum volume that has to go through any opened primary transport link starting at plant i in each time-period	In our case it is evaluated as the volume ensuring on average full truckload transport within a given shipment periodicity
$minTransport(j, l, t)$	Minimum volume that has to go through any opened secondary transport link from DC j to district l in time-period t	See §6.1 for information on how to evaluate this parameter
R_{jlt}	Average length of the delivery route from DC j to district l in time-period t	This is an output of the clustering phase
STC_{jlt}	Secondary transport cost per unit from DC j to district l in time-period t	
PTC_{ij}	Primary transport cost per unit from plant i to DC j	
TC_j	Transit cost per unit for a product going through DC j	
F_j	Fixed cost to be paid if potential DC j is opened	
CD	Maximum covering distance parameter	The length of a delivery route R_{jlt} from DC j to any district l that it serves in any time-period t should not exceed CD)
$totProd_{it}$	Total volume produced by plant i in time-period t	
$M(i, j, t)$	Big value involved in the constraints related to the variables v'_{ijt} and v''_{ijt} defined later in this section	Set to $min\left(maxThroughput(j), totProd_{it}\right)$
PI_{ij}	Penalty parameter for violating the minimum volume constraint on a primary transport link from plant i to DC j	

Instance	Optimal solution	CPU (hours)
Multi-period model with dynamic assignments	25 opened DCs	84.6
Multi-period model with static assignments	21 opened DCs, 1% increase in cost as compared to the instance using dynamic assignments	9.4

Table 2: Multi-period model: using dynamic assignments vs. using static assignments (reference dataset)

Instance	CPU(hours) of Cplex solution with optimality limit 0.2%	CPU(hours) of the heuristic solution	Quality of the heuristic solution (%)	CPU(hours) of Cplex solution with optimality limit 2%	Quality of Cplex solution with optimality limit 2%
A.1	1.26	0.04	1.26	0.52	1.77
A.2	13.36	0.01	1.55	5.80	1.77
A.3	3.26	0.01	1.52	2.33	0.18
A.4	1.47	0.02	1.44	0.93	0.32
A.5	13.12	0.02	1.58	1.98	1.08
A.6	2.97	0.01	2.2	1.63	0.14
A.7	2.71	0.01	2.1	1.32	0.91
A.8	4.15	0.01	1.65	1.67	1.63
A.9	1.07	0.01	1.76	1.07	0.44
A.10	14.6	0.05	1.85	2.25	1.04
A.11	3.67	0.01	1.65	3.08	0.55
A.12	12.92	0.03	1.45	1.72	1.20
A.13	1.49	0.02	1.42	0.83	1.59
A.14	1.29	0.06	1.66	0.89	0.26
A.15	4.41	0.01	1.79	3.21	1.84
A.16	1.44	0.02	1.89	1.40	0.80
A.17	3.61	0.01	1.56	1.44	1.78
A.18	8.89	0.01	1.88	3.15	1.59
A.19	5.88	0.01	2.48	3.05	0.36
A.20	6.43	0.02	1.54	3.71	0.68

Table 3: CPU and solution quality of the heuristic procedure applied to the multi-period problem with static assignments when varying customer demand (test instances A)

Instance	CPU(hours) of Cplex solution with optimality limit 0.2%	CPU(hours) of the heuristic solution	Quality of the heuristic solution (%)	CPU(hours) of Cplex solution with optimality limit 2%	Quality of Cplex solution with optimality limit 2%
B.1	1.44	0.02	1.13	0.80	1.23
B.2	3.46	0.02	1.92	2.31	0.98
B.3	0.86	<0.01	1.49	0.65	0.35
B.4	1.13	0.02	1.84	0.82	0.62
B.5	2.85	<0.01	1.58	2.17	1.36
B.6	4.40	0.01	1.27	2.00	0.46
B.7	5.08	0.01	0.82	0.91	0.87
B.8	2.58	0.01	1.83	1.04	1.73
B.9	6.43	0.01	1.78	1.65	1.67
B.10	OOM	0.01	-	1.27	-
B.11	5.33	0.01	1.82	1.40	1.46
B.12	2.06	0.01	1.55	1.71	0.56
B.13	4.01	0.01	1.14	2.68	1.65
B.14	3.42	0.01	1.73	2.82	0.75
B.15	5.19	0.02	1.38	11.51	1.71
B.16	1.83	0.01	1.89	0.93	0.81
B.17	5.29	0.02	1.59	3.47	1.00
B.18	6.05	0.01	1.46	3.07	1.61
B.19	1.19	0.02	1.59	0.74	1.82
B.20	0.89	0.02	1.75	1.04	0.17

Table 4: CPU and solution quality of the heuristic procedure applied to the multi-period problem with static assignments when varying the list of potential DCs (test instances B). "OOM" denotes an out of memory after several hours of calculation.