



Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Capacitated p -center problem with failure foresightInmaculada Espejo^a, Alfredo Marín^b, Antonio M. Rodríguez-Chía^{a,*}^a Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, E-11510 Puerto Real (Cádiz), Spain^b Departamento de Estadística e Investigación Operativa, Universidad de Murcia, Spain

ARTICLE INFO

Article history:

Received 4 December 2014

Accepted 26 May 2015

Available online xxx

Keywords:

 p -center

Capacities

Integer programming

ABSTRACT

This paper deals with a generalized version of the capacitated p -center problem. The model takes into account the possibility that a center might suffer a disruption (being unavailable to serve some demand) and assumes that every site will be covered by its closest available center. The problem is of interest when the goal is to locate emergency centers while, at the same time, taking precautions against an unforeseen incident (natural disaster, labor strike, accident...) which can cause failure of a center. We consider different formulations for the problem and extensive computational tests are reported, showing the potentials and limits of each formulation in several types of instances. Finally, a preprocessing phase for fixing variables has been developed and different families of valid inequalities have been proposed to strengthen the most promising formulations, obtaining in some cases much better resolution times.

© 2015 Elsevier B.V. and Association of European Operational Research Societies (EURO) within the International Federation of Operational Research Societies (IFORS). All rights reserved.

1. Introduction

Facility location models have been extensively studied in the literature. Various kinds of facilities have been modeled, such as routers or servers in communication networks, warehouses or distribution centers in a supply chain, hubs or transshipment nodes in passenger transport, and hospitals or emergency facilities in a public service system, among many others. In general, the goal of these problems is to locate the facilities among a set of candidate sites and to assign customers to the facilities, optimizing some effectiveness measure that usually depends of the distances between the facilities and the customers, see for instance Daskin (1995) and Drezner and Hamacher (2002) and the references therein. The p -center problem (pCP) is a well-known discrete optimization location problem which consists of locating p centers out of n sites and assigning (allocating) the remaining $n - p$ sites to the centers so as to minimize the maximum distance (cost) between a site and the corresponding center. It was shown in Kariv and Hakimi (1979) that pCP is NP-hard. A straight application of pCP is the location of emergency services like ambulances, hospitals or fire stations, since the whole population should be inside a small radius around some emergency center. pCP has been extensively studied, and both exact and heuristic algorithms have been proposed. Recent articles on this issue are Mladenović, Labbé, and Hansen (2003), Elloumi, Labbé, and Pochet (2004), Daskin (2000)

and Calik and Tansel (2014). We also refer the reader to Chapter 5 of Daskin (1995). A recent survey on location of emergency services can be consulted in Basar, Aatay, and Unluuyurt (2012).

Emergency services have, in practice, a limited capacity. Consequently, capacitated versions of pCP have also received attention in the literature. In the capacitated p -center problem (CpCP), each site has a demand and a capacity, and the total demand of the sites assigned to a center cannot exceed its capacity. Thus, CpCP is the problem of finding the set of p locations and the assignment pattern that satisfies the capacity constraints while minimizing the maximum distance site-closest center. A local search heuristic for CpCP has been developed in Scaparra, Pallotino, and Scutellà (2004). The special case where all demands are equal has been studied in Bar-Ilan, Kortsarz, and Peleg (1993) and Khuller and Sussmann (2000). To the best of our knowledge, only in three papers, CpCP has been solved exactly. Jaeger and Goldberg (1994), where the special case of an underlying tree network is approached; Özsoy and Pinar (2006), where an algorithm developed in Ilhan and Pinar (2001) for the p -center problem is modified and extended, succeeding in providing optimal solutions in short times for small instances of the problem; and Albareda-Sambola, Díaz, and Fernández (2010) where the authors propose two different Lagrangian duals and an exact algorithm able to solve larger instances. In Kalcsics, Nickel, Puerto, and Rodríguez-Chía (2010), the capacitated ordered discrete location problem is studied; among many others, CpCP is a particular case obtained by fixing a parameterized vector used to get the objective function of the model to $(0, \dots, 0, 1)$.

In addition to the capacity constraint of the facilities, there is an important aspect that deserves attention, namely, the possibility of

* Corresponding author. Tel.: +34 956012787; fax: +34 956016288.

E-mail address: antonio.rodriguezchia@uca.es (A.M. Rodríguez-Chía).

disruption of the facilities. Although most models in the literature on Location Theory have treated facilities as if they could never fail, there is a wide variety of reasons for which facilities can fail in practice (industrial accidents, natural disaster, labor strikes, *et cetera*). Some illustrative examples of real situations of the damage caused in facilities can be seen in the introduction of O'Hanley, Scaparra, and García (2013). There is some research on discrete location dealing with facility disruptions, although it is mainly focused on stochastic assumptions and reliability issues in supply chain design. In those models, some known probability of failure is considered for each of the facilities and the goal is to minimize the expected cost; if a facility fails, customers originally assigned to it have to be reassigned to other facilities. In Ball and Lin (1993) the authors attempt to increase the system availability through redundant coverage. In Snyder and Daskin (2005) an implicit formulation of the stochastic p -median problem as a linear integer program is developed, where the facilities are subject to failure with the same probability. In Zhan, Shen, and Daskin (2007) the authors provide several heuristics for the stochastic fixed-charged problem, which is also formulated as a nonlinear mixed integer program. Heuristics with bounds on the worst-case performance and the study of asymptotic properties of the solutions of the stochastic p -median problem are carried out in Berman, Krass, and Menezes (2007). In Cui, Ouyang, and Shen (2010), the uniform failure probability assumption in Snyder and Daskin (2005) is relaxed and the failure probabilities are facility-specific. In addition, the authors propose a mixed-integer formulation and a continuum approximation formulation to solve large-scale problems. In Shen, Zhan, and Zhang (2011), the problem is formulated as a two-stage stochastic program and then as a nonlinear integer program. Several heuristics and a 4-approximation algorithm are provided. Recently, O'Hanley et al. (2013) extend Berman et al. (2007) and Cui et al. (2010) by proposing an exact linear reformulation of the unreliable p -median problem given unequal facility disruption probabilities. Unlike the previous references, some recent papers deal with facility disruption without considering probabilities. In O'Hanley and Church (2011) the authors maximize a combination of initial coverage and the minimum coverage level after the loss of the most critical facilities. From a bilevel point of view, Aksent, Aras, and Piyade (2013) consider the location of p facilities and, in a second step, the protection of some of these under budget constraints. A related model that also considers capacity constraints is given in Aksent and Aras (2012). There is even one more paper in this line that considers the possibility of disruption in a hub location environment (Parvaresh, Hussein, Golpayegany, & Karimi, 2014), where some heuristics are developed with the aim of minimizing the effects of an attacker after designing the hub network.

Other related research fields are the fortification of reliability of the service to cover a set of existing facilities and the backup set covering problems. The interdiction-fortification p -median problem is generally formulated as a bilevel programming problem, where the idea is to protect the existing critical facilities under the events of disruption, see Snyder, Scaparra, Daskin, and Church (2006) and Scaparra and Church (2008). In Lei and Church (2011), the authors use explicit closest assignment constraints for a novel multilevel facility interdiction problem that optimize worst levels of facility disruptions. A recent paper in this field that takes capacities into consideration (see also references therein) is Liberatore, Scaparra, and Daskin (2012), where the authors consider the protection of a capacitated median system with a limited amount of protective resources subject to partial or complete disruption of the facilities involved. On the other hand, several backup set covering problems have also appeared in the literature under the common idea of covering the demand points with several centers in order to guarantee the coverage in case of either failure or an overflow into one or some centers (in this sense, the model proposed in this paper can be considered a backup problem). In these models there are two natural objectives:

Minimization of the number of open centers and maximization of the backup coverage. Several times the model is approached from the point of view of multiobjective optimization (Storbeck & Vohra, 1988) and model BACOP1 in Hogan and ReVelle (1986) are two examples of such an approach. At other times, both objectives are combined in a unique function as in model BACOP2 of Hogan and ReVelle (1986). In Weaver and Church (1985) the authors take a different approach with the Vector Assignment p -Median Model. In their model, a fixed fraction of each node's demand is allocated among the facilities on the basis of their proximity, e.g., 70 percent of a node's demand is assigned to the closest facility, 10 percent to the next closest facility, and so forth. Very recently, Curtin, Hayslett-McCall, and Qiu (2010) present a new backup coverage model that is appropriate for patrol area design, see also the references therein.

This paper seeks to generalize the capacitated p -center problem by considering the possibility that a facility might suffer a disruption, being unavailable to serve some demand. These two joint aspects (capacity constraint and failure possibility) make the model we consider in this paper more fitting for some real situations. What we are going to consider is that sites must still be assigned to their closest centers and so any center j must have enough capacity to satisfy the demands of the sites which are closer to j than to any other center. In addition, an unpredictable incident may occur in any of the p centers, forcing it to be closed. Then each customer allocated to it is re-assigned to another center, which must be its second closest center. Consequently, the capacity of any center must be enough to receive also the sites re-assigned to it in case of failure of any other center. The goal then is to minimize the second lowest distance from a site to the set of centers, since in case of an accident which produces a damage in one of the emergency services, all the sites should still be inside the smallest possible radius around an available center. Since we consider emergency services which could be essential for saving lives, increasing the operation cost (with respect to not considering failure possibility) to prevent facility disruptions is fully justified. Observe that this problem is NP-hard because its uncapacitated $2p$ -facility version can be reduced to a pCP. We call our model *Capacitated Second p -Center Problem* (CSpCP, in short).

The paper is organized as follows. In the next section, we formalize the problem and illustrate it with an example. In Section 3, different integer programming formulations for CSpCP are introduced. A comparison of CSpCP with CpCP with closest assignment and preliminary computational results are given in Section 4. In Section 5, a heuristic approach is developed and then, in Section 6, we use this heuristic to improve the formulations using different variable fixing strategies and strengthening formulations with new valid inequalities which are computationally compared. We end with some conclusions.

2. The problem

Let $N = \{1, \dots, n\}$ be the given set of sites. Throughout the paper we assume, without loss of generality, that the set of candidate sites for centers is identical to N . Let h_i denote the demand of site $i \in N$, b_i the capacity of a center located at site $i \in N$, and $p \geq 2$ be the number of centers to be located. For each pair (i, j) , $i, j \in N$, let d_{ij} be the distance (cost, travel time) from i to j . We assume $d_{ii} = 0 \forall i \in N$ and $d_{ij} > 0 \forall i, j \in N: i \neq j$. We do not assume other special properties like satisfaction of triangle inequality, that is to say, strictly speaking d is not a distance. But we need to make an additional assumption to deal with the case of ties among several distances from the same site. If this is the case, in order to break ties we suppose that there are preferences on the centers such that sites will undoubtedly choose one center before the others. In practice, ties can be broken by slightly perturbing the tied distances. Summarizing, we will also assume $d_{ia} \neq d_{ib} \forall i, a, b \in N: a \neq b$. Then CSpCP is to choose p centers $\{i_1, \dots, i_p\}$ among the n sites to minimize the distance from any site to its second closest

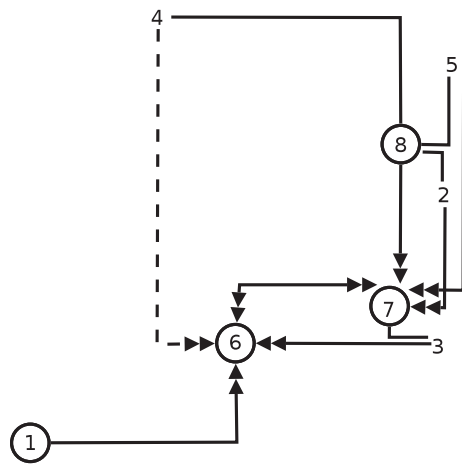


Fig. 1. Example of how points are allocated to the two closest centers.

center while satisfying the capacity limitations. This means that for any $i \in \{i_1, \dots, i_p\}$,

$$b_i \geq \max_{j \in \{i_1, \dots, i_p\} \setminus \{i\}} \sum_{k: i = \arg \min_{\ell \in \{i_1, \dots, i_p\} \setminus \{j\}} d_{k\ell}} h_k$$

must hold. We present now an example of the problem for the sake of clarity.

Example 2.1. Consider eight points in the plane with coordinates (1.1,3.3), (6.4,5.6), (6.3,3.7), (2.7,9.1), (7.0,8.3), (3.7,3.8), (5.9,3.9) and (6.1,6.5), demands 5, 2, 2, 4, 1, 5, 4 and 4 and capacities 6, 7, 13, 11, 15, 13, 15 and 15, respectively. The ℓ_1 norm is used to compute distances between points. Taking $p = 4$, the optimal solution of the problem is given in Fig. 1. Centers are drawn inside circles; the closest center to a non-center point is indicated by a segment, and the second closest center with respect to a point is depicted with a two tip arrow. For instance, point 5 is not a center, its closest center is 8 and its second closest center is 7.

To see that the solution satisfies the capacity constraints, in the following table we give for each center j (column 1), the total demand of points having j as their closest center (column 2), the additional demand assigned to j in case of failure of either 1, 6, 7 or 8 (columns 3–6), the maximum of the last four amounts (column 7), the sum of columns 2 and 7 (column 8) and the capacity of center j (column 9).

Center	Demand	1 fails	6 fails	7 fails	8 fails	Max	Total	Capacity
1	5	-	0	0	0	0	5	6
6	5	5	-	6	4	6	11	13
7	6	0	5	-	7	7	13	15
8	11	0	0	0	-	0	11	15

For instance, center 7 is the closest center with respect to points 3 and itself, with a total demand of $2 + 4 = 6$. In case of failure of center 1, 7 would not receive additional demand. In case of failure of 6, 7 would receive the demand of point 6, which is 5. In case of failure of center 8, 7 would receive the demand of points 2, 5 and 8, i.e., $2 + 1 + 4 = 7$. This is the worst case, and then in case of failure of center 8, center 7 would need enough capacity to cover the demand of $6 + 7 = 13$. Since the capacity of 7 is 15, the corresponding capacity constraint is satisfied.

The objective value is given by the maximum distance from a point (in this case point 4) to its second closest center (in this case 6), as indicated in the figure by a dashed arrow.

3. Integer programming formulations

The model proposed in this work can be formulated as an integer linear programming problem. In this section we present several formulations.

3.1. Formulation using variables with three indices

For the first formulation we will define the following two usual sets of binary variables, which will be used to determine the location of the centers and the allocation of the sites to the centers (the closest and the second-closest ones, respectively).

$$y_j = \begin{cases} 1, & \text{if a center is located at site } j, \\ 0, & \text{otherwise,} \end{cases} \quad j \in N,$$

$$x_{ijk} = \begin{cases} 1, & \text{if the closest and the second-closest} \\ & \text{centers to site } i \\ & \text{are } j \text{ and } k, \text{ respectively,} \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j, k \in N.$$

The formulation is

$$(P3) \min z \quad (1)$$

$$\text{s.t. } \sum_{j \in N} \sum_{\substack{k \in N \\ k \neq i}} x_{ijk} = 1 \quad \forall i \in N, \quad (2)$$

$$\sum_{j \in N} y_j = p \quad (3)$$

$$\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} \sum_{\substack{k \in N \\ k \neq i, \ell}} x_{i\ell k} + \sum_{\substack{\ell \in N \\ d_{i\ell} < d_{ij}}} \sum_{\substack{k \in N \\ d_{ik} > d_{ij} \\ k < i}} x_{i\ell k} + y_j \leq 1 \quad \forall i, j \in N, \quad (4)$$

$$\sum_{\ell \in N} \sum_{\substack{k \in N \\ k \neq \ell, j}} h_{\ell} x_{\ell j k} + \sum_{\substack{\ell \in N \\ \ell \neq i}} h_{\ell} x_{\ell i j} \leq b_j y_j \quad \forall i, j \in N : i \neq j, \quad (5)$$

$$z \geq \sum_{j \in N} \sum_{k \in N} d_{ik} x_{ijk} \quad \forall i \in N, \quad (6)$$

$$y_j \in \{0, 1\} \quad \forall j \in N, \quad (7)$$

$$x_{ijk} \in \{0, 1\} \quad \forall i, j, k \in N. \quad (8)$$

The objective function (1) together with constraints (6) are used to get the maximum distance of any site with respect to its second closest center by means of an additional decision variable z which is minimized in the objective. Constraints (2) force each site to be allocated to a pair of different centers. Constraint (3) fixes the number of centers to p .

In order to ensure that a site is not assigned to centers other than the two closest ones, constraints (4) are incorporated into the formulation. Given a site i and a center j (i.e., $y_j = 1$), x -variables in the first term (forced to take value 0) correspond with routes whose first center is further from i than j . Similarly, x -variables in the second term are fixed to 0 if the corresponding route goes first to a center closer to i than j but afterward to a center further to i than j . Both sets of variables can be aggregated (so producing a stronger formulation) since in case $y_i = 0$, the resulting constraint is implied by (2).

Constraints (5) have a double mission: preventing sites being assigned to sites which are not centers and keeping the demand of sites assigned to the same center under the capacity of this center. The demand added up on the left hand side splits into $\sum_{\ell \in N} \sum_{\substack{k \in N \\ k \neq \ell, j}} h_{\ell} x_{\ell j k}$, which is the demand of sites assigned to j as their first center, and $\sum_{\substack{\ell \in N \\ \ell \neq j}} h_{\ell} x_{\ell i j}$, which is the demand of sites having j as their second center and i as their first center. Finally, binarity of all variables is fixed in constraints (7) and (8).

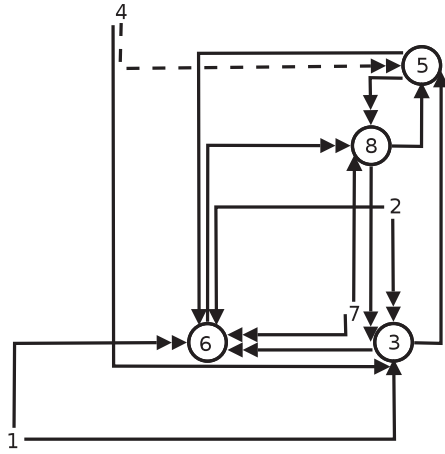


Fig. 2. Illustration of the need for (4).

In what follows, we show the need for the families of constraints describing formulation (P3). It is clear that constraints (2), (3), (5) and (6) are necessary. Regarding family of constraints (4), we again consider the instance given in Example 2.1. After removing constraints (4), the optimal solution we obtain is depicted in Fig. 2. Here, for instance, site 1 is not allocated first to its closest center. The reason is to exploit the capacity of further centers in order to reduce the objective value. Also, site 5 is allocated first to 6 (and then to 8), not to itself. Similarly, sites 8 and 3 are allocated first to 5 (and then to 3 and 6, respectively).

Note that x_{ijk} -variables with $i = k$ or $j = k$ will never take value 1 at any optimal solution and can be fixed at zero. Several other variables can be fixed at zero in (P3). The second closest center w.r.t. i will never be closer to i than the closest center, then

$$x_{ijk} = 0 \quad \forall i, j, k \in N : d_{ij} \geq d_{ik}.$$

Furthermore, y -variables can be removed from the formulation by taking into account that

$$y_i = \sum_{k \in N} x_{iik} \quad \forall i \in N.$$

3.2. Formulation using variables with two indices A

In this formulation we will use the following sets of binary variables:

$$x_{jj} = \begin{cases} 1, & \text{if a center is located at site } j, \\ 0, & \text{otherwise,} \end{cases} \quad j \in N,$$

$$x_{ij} = \begin{cases} 1, & \text{if the closest center to site } i \text{ is } j, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j \in N.$$

Note that $x_{jj} = 1$ implies self allocation of site j . Additionally, we incorporate into the formulation the set of binary variables

$$w_{ij} = \begin{cases} 1, & \text{if the second closest center to site } i \text{ is } j, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j \in N,$$

necessary to determine the allocation of sites to their second closest centers.

Then the formulation we propose is

$$(P2A) \min z$$

$$\text{s.t. } \sum_{j \in N} x_{jj} = p \quad (9)$$

$$\sum_{j \in N} x_{ij} = 1 \quad \forall i \in N, \quad (10)$$

$$\sum_{\substack{j \in N \\ j \neq i}} w_{ij} = 1 \quad \forall i \in N, \quad (11)$$

$$x_{ij} + w_{ij} \leq x_{jj} \quad \forall i, j \in N, \quad (12)$$

$$\sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} x_{j\ell} + \sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} x_{i\ell} \leq 1 \quad \forall i, j \in N, \quad (13)$$

$$x_{jj} + \sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} w_{i\ell} \leq 1 + x_{ij} \quad \forall i, j \in N : i \neq j, \quad (14)$$

$$\sum_{\substack{\ell \in N \\ \ell \neq j}} h_{\ell} x_{\ell j} + \sum_{\substack{\ell \in N \\ d_{i\ell} < d_{ij}}} h_{\ell} (w_{\ell j} + x_{ii} - 1) \leq (b_j - h_j) x_{jj} \quad \forall i, j \in N : i \neq j, \quad (15)$$

$$z \geq \sum_{\ell \in N} d_{i\ell} w_{i\ell} \quad \forall i \in N, \quad (16)$$

$$x_{ij}, w_{ij} \in \{0, 1\} \quad \forall i, j \in N.$$

The goal is to minimize the maximum of the distances to the second center assigned to each site, given in constraints (16). Constraints (9)–(12) resemble the classical constraints in discrete location problems. In our formulation they ensure allocation of a site to two different centers and fix the number of centers at p . Due to the capacity constraints we need to add closest assignment constraints (CAC) to the formulation (a recent and complete study of CAC can be consulted in Espejo, Marín, and Rodríguez-Chía, 2012). Constraints (13) improve on those proposed in Wagner and Falkson (1975), and force a variable x_{ij} to take value one if j is the closest center with respect to site i .

Similar constraints are needed to force allocation (by means of w -variables) to the second closest center. To this end, constraints (14) act as follows. In the cases $x_{jj} = 0$ and $x_{jj} = x_{ij} = 1$, (14) reduces to a trivial constraint due to (11). The case of interest is $x_{jj} = 1$ and $x_{ij} = 0$, i.e., j is a center but i is not allocated to j . Then, due to (13), there must exist a center closer to i than j , and (14) reads $\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} w_{i\ell} \leq 0 \Rightarrow w_{i\ell} = 0 \quad \forall \ell \in N : d_{i\ell} > d_{ij}$. The effect is that i is not allocated to a second closest center further than j . A version of (14) has been also developed in Lei (2010), see also Lei and Church (2011).

Constraints (15) are the capacity constraints. If $x_{jj} = 0$, it follows $x_{\ell j} = w_{\ell j} = 0 \quad \forall \ell \in N$ and then $\sum_{\substack{\ell \in N \\ d_{i\ell} < d_{ij}}} h_{\ell} (x_{ii} - 1) \leq 0$, which holds trivially. If $x_{jj} = 1$, b_j bounds the sum of $\sum_{\ell \in N} h_{\ell} x_{\ell j}$ and $\sum_{\substack{\ell \in N \\ d_{i\ell} < d_{ij}}} h_{\ell} (w_{\ell j} + x_{ii} - 1)$. The first addend measures the demand of sites which have j as the closest center. The second addend, for a given value of i , is only of interest if $x_{ii} = 1$ (otherwise it is non-positive), in which case it counts the demand of the sites with j as the second closest center which are closer to i than to j . That is to say, the second addend measures the demand which will go to j if center i fails.

All variables must be constrained to be 0–1 valued due to the capacity constraints (unsplittable demand). Note that taking $i = j$ in (12) it follows that $w_{jj} = 0 \quad \forall j \in N$. Throughout the rest of the paper, we will use these identities when required.

In order to show the need for several families of constraints in formulation (P2A) we consider again the instance given in Example 2.1. After removing constraints (13), the solution is depicted in Fig. 3. Note that, in order to reduce the objective value, allocation is made first to centers which are further than the closest center with respect to a given site; for instance, site 1 is allocated to 5 instead of 6. After removing, instead, constraints (14), the obtained solution is depicted in Fig. 4. Here, for instance, site 3 is not allocated to its second closest center. The reason is to exploit the capacity of further centers in order to reduce the objective value. After removing constraints (12), all points would be allocated to themselves by means

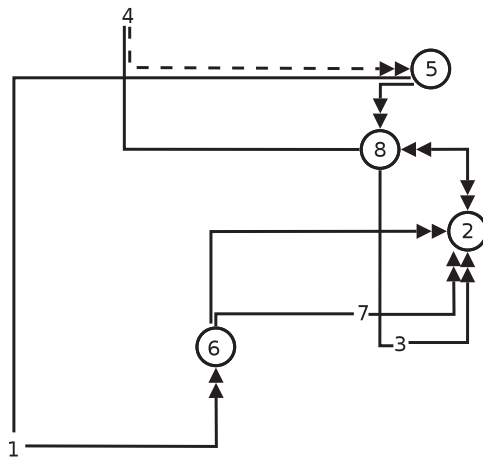


Fig. 3. Illustration of the need for (13).

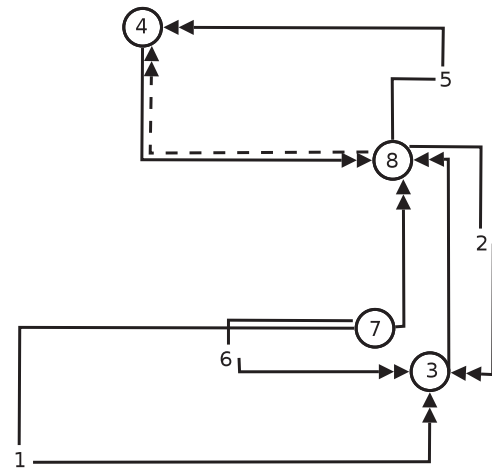


Fig. 4. Illustration of the necessity of (14).

of w -variables, with an objective value equal to zero. It is clear that constraints (9)–(11), (15) and (16) are necessary.

Analogously to (P3), several variables can be fixed at zero in (P2A). Sorting the distances from a fixed site i to the rest of the sites by

$$d_1^i \leq d_2^i \leq \dots \leq d_n^i,$$

since there are p centers, i will never be assigned to the $p - 1$ furthest sites. Then

$$x_{ij} = 0 \quad \forall i, j \in N : d_{ij} > d_{n-p+1}^i.$$

Analogously, the second closest center w.r.t. i will never be one of the $p - 2$ furthest sites and

$$w_{ij} = 0 \quad \forall i, j \in N : d_{ij} > d_{n-p+2}^i.$$

Also, since i is always the unique closest site with respect to itself,

$$w_{ij} = 0 \quad \forall i \in N.$$

3.3. Formulation using variables with two indices B

It is observed in formulation (P2A) that the x -variables play a secondary role. It is possible, with a complete knowledge of $x_{ij} \forall j \in N$ plus the w -variables, to build capacity constraints and CAC, thus reducing the size of the formulation. The result of this reduction is called formulation (P2B). Whether this change entails a quality reduction of the formulation and, in such a case, if the size reduction makes up for it, will be a matter of forthcoming sections. Formulation (P2B) is as follows:

$$\begin{aligned} \text{(P2B) } \min \quad & z \\ \text{s.t. } & (9), (11), (16) \\ & w_{ij} \leq x_{jj} \quad \forall i, j \in N, \end{aligned} \quad (17)$$

$$\sum_{\substack{\ell \in N \\ d_{j\ell} > d_{ji}}} 2w_{i\ell} + \sum_{\substack{\ell \in N \\ d_{j\ell} \leq d_{ji}}} x_{\ell\ell} \geq 2 \quad \forall i, j \in N : i \neq j, \quad (18)$$

$$\sum_{\substack{\ell \in N \\ d_{i\ell} \geq d_{ij}}} (p-1)w_{i\ell} \leq \sum_{\substack{\ell \in N \\ d_{i\ell} \geq d_{ij}}} x_{\ell\ell} \quad \forall i, j \in N: i \neq j, \quad (19)$$

$$\begin{aligned} & \sum_{\ell \in N} \sum_{\substack{q \in N \\ d_{\ell j} < d_{\ell q}}} h_{\ell}(w_{\ell q} + x_{jj} - 1) + \sum_{\substack{\ell \in N \\ d_{\ell i} < d_{\ell j}}} h_{\ell}(w_{\ell j} + x_{ii} - 1) \\ & \leq b_i x_{jj} \quad \forall i, j \in N : i \neq j, \end{aligned} \quad (20)$$

$$\begin{aligned} x_{jj} &\in \{0, 1\} \quad \forall j \in N, \\ w_{ij} &\in \{0, 1\} \quad \forall i, j \in N. \end{aligned} \quad (21)$$

Given two different sites i and j , the corresponding constraint (18) is trivially satisfied when $\sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} x_{\ell\ell} \geq 2$, i.e., when there are two or more centers inside a radius d_{ij} around i . Otherwise, if $\sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} x_{\ell\ell} \in \{0, 1\}$, constraint (18) forces some of the variables in the first term to take value 1, that is to say, it forces the distance in the objective function to be greater than d_{ij} .

To complement (18), and due to the capacity limitations, it is necessary to add some other constraints preventing i for being allocated to a center further than the second closest center. Constraints (19) have this effect. The trivial case is $\sum_{d_{i\ell} \geq d_{ij}}^{\ell \in N} w_{i\ell} = 0$. Otherwise, if $\sum_{d_{i\ell} \geq d_{ij}}^{\ell \in N} w_{i\ell} = 1$, it means that the reference center of i is not inside radius d_{ij} . Since the total number of center is p , this would only occur when the number of centers at least at a distance d_{ij} , given by the second term of (19), is $p - 1$.

The combination of constraints (18) and (19) produces the desired allocation to the second closest center. Similar constraints were used in (Belotti, Labbé, Maffioli, & Ndiaye, 2007) in the context of obnoxious location.

In order to establish the limit on the capacity of the centers using this reduced set of variables, constraints (20) are incorporated into the formulation. Each constraint in this family takes care of the capacity of center j if center i fails. The first case is when $x_{jj} = 0$, i.e., j is not going to be a center. In such a case, the addends of the first term will be $w_{\ell q} - 1$, which are less than or equal to 0, and the addend of the second term will again be $x_{ii} - 1$ less than or equal to 0, and the constraint is trivially satisfied. The interesting case is then $x_{jj} = 1$. Replacing x_{jj} by 1 in the first term, it becomes $\sum_{\ell \in N} \sum_{\substack{q \in N \\ d_{\ell q} < d_{\ell i}}} h_{\ell} w_{\ell q}$.

Since j is now a center, a variable $w_{\ell q} = 1$ when q is outside a radius $d_{\ell j}$ around a site ℓ means that j must be the closest center with respect to ℓ . Then the first term sums up the demand of the sites with j as the closest center, to be taken always into account when limiting the capacity of j . But the capacity of j must also be enough to cover the demand of some sites after the failure of another center, represented in (20) by i . When i is actually a center, variable x_{ji} will take value 1 (otherwise the second term will be negative). Note that the second term only considers those sites ℓ closer to i than to j , and variables $w_{\ell j}$ taking value 1. The combination of both situations implies that, in case of failure of i , the demand of ℓ will be satisfied by j , and this demand given by the second term must be added to that of the first term to be bounded by b_j .

Although we have introduced (P2B) with all constraints that can be considered intuitively necessary, not all of them are needed to obtain a correct formulation, as we show in the following result.

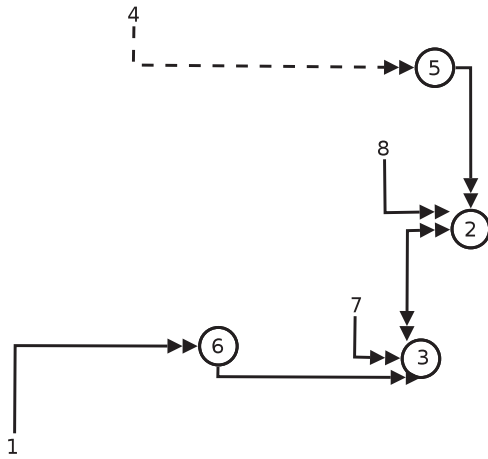


Fig. 5. Illustration of the need for (18).

Proposition 3.1. Constraints (17) are implied by constraints (9), (11), (18), (19) and (21).

Proof. Assume that, for $i \in N$ and $j \in N \setminus \{i\}$, $w_{ij} = 1$ holds.

Replacing w_{ij} by 1 in (11) and using (21) it follows that $w_{ii} = 0 \forall i$ such that $d_{ii} > d_{ij}$. Now, replacing these values in (18),

$$\sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} x_{\ell\ell} \geq 2.$$

Taking (9) into consideration, the last inequality is equivalent to

$$\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} x_{\ell\ell} \leq p - 2. \quad (22)$$

On the other hand, replacing w_{ij} by 1 in (19) and using (11) again, it follows that

$$p - 1 \leq \sum_{\substack{\ell \in N \\ d_{i\ell} \geq d_{ij}}} x_{\ell\ell}$$

which, by the assumption of no ties in the distances from the same point, can be re-written as

$$p - 1 \leq x_{jj} + \sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} x_{\ell\ell}. \quad (23)$$

Finally, bringing together (22) and (23),

$$p - 1 \leq x_{jj} + \sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} x_{\ell\ell} \leq x_{jj} + p - 2,$$

and $x_{jj} \geq 1$ trivially holds. \square

The remaining constraints are not implied by the others. Consider the instance given in Example 2.1. After removing constraints (18), the solution obtained is depicted in Fig. 5. Notice that, in order to reduce the objective value, second allocation is made to centers which are closest with respect to a given site. For instance, site 7 is allocated to center 3. Removing families of constraints (19), we obtain solutions with wrong allocations. To show this, consider $n = 4$, $p = 3$, distance ℓ_1 , points in coordinates (2, 9), (1, 1), (7, 6), (10, 7), demands 2, 2, 1, 1 and capacities 8, 6, 6, 7. The solution has opened centers 1, 2 and 3 and $w_{12} = w_{21} = w_{31} = w_{41} = 1$, where we can see that site 1 is not allocated to its second closest center, site 3 (see Fig. 6).

Several variables can be fixed at zero in (P2B). The second closest center w.r.t. i will never be one of the $p - 2$ furthest sites nor the

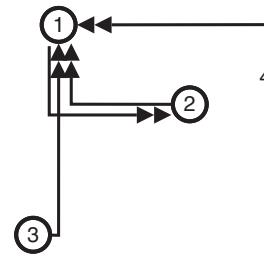


Fig. 6. Illustration of the need for (19).

closest one (itself), then

$$w_{ij} = 0, \quad \forall i, j \in N : d_{ij} > d_{n-p+2}^i.$$

$$w_{ii} = 0, \quad \forall i \in N.$$

Formulations (P2A) and (P2B) are both of interest from a theoretical point of view due to the following result.

Proposition 3.2. The linear relaxations of (P3), (P2A) and (P2B) do not dominate one another.

Proof. To prove the thesis it suffices to show two examples of instances in which the linear relaxation of each formulation gives a larger value than the linear relaxation of the other. For both examples we take a set of six points in the plane, $p = 3$ and the ℓ_1 norm is used to compute distances between points.

In the first example, the points are (1, 8), (5, 0), (9, 9), (5, 2), (6, 2) and (4, 9), with demands 3, 2, 3, 3, 2, 5 and capacities 16, 3, 11, 14, 4, 15 respectively. Centers in the optimal solution are 3, 4 and 6. The two closest centers with respect to each site are (6, 3), (4, 6), (3, 6), (4, 6), (4, 6) and (6, 3), respectively. The optimal value is 10 (distance from 2 to 6, see Fig. 7, left hand side). The optimal values of the linear relaxation of (P3), (P2A) and (P2B) are 8, 6.04 and 5.77, respectively.

In the second example, the points are (6, 4), (0, 9), (1, 6), (4, 0), (4, 4) and (1, 7), with demands 3, 1, 2, 3, 2, 2 and capacities 11, 3, 13, 15, 14, 5 respectively. Centers in the optimal solution are 1, 5 and 6. The two closest centers with respect to each site are (1, 5), (6, 5), (6, 5), (5, 1), (5, 1) and (6, 5), respectively. The optimal value is 9 (distance from 2 to 5, see Fig. 7, right hand side). The optimal values of the linear relaxation of (P3), (P2A) and (P2B) are 4, 4.59 and 5, respectively. \square

4. Analysis of the model and formulations

The goal of this section is twofold. On the one hand, to provide the readers a better understanding of the model considering a comparison between the solutions of CpCP and CSpCP which highlights that there are significant benefits in using the proposed model in case of disruption. On the other hand, to compare the three proposed formulations for CSpCP in order to identify which of them is more promising from the computational point of view.

4.1. Comparison between versions of CpCP with and without failure foresight

It could be argued that the prevention of a disaster that rarely occurs will imply a change in the solution to the problem that leads to some sites being allocated to centers further than the reference centers they would have in the solution of CpCP in order to guarantee sufficient capacity. But this argument would rapidly lead to the invalidation of all kinds of p -center problems, which are based on the improvement of the worst case situation at the expense of making the average situation worse than in the median models. In contrast, when the decision maker decides to use the solution to the CpCP instead of its counterpart with median objective, the idea is to provide an opportunity to everybody (in particular to the furthest one, although it

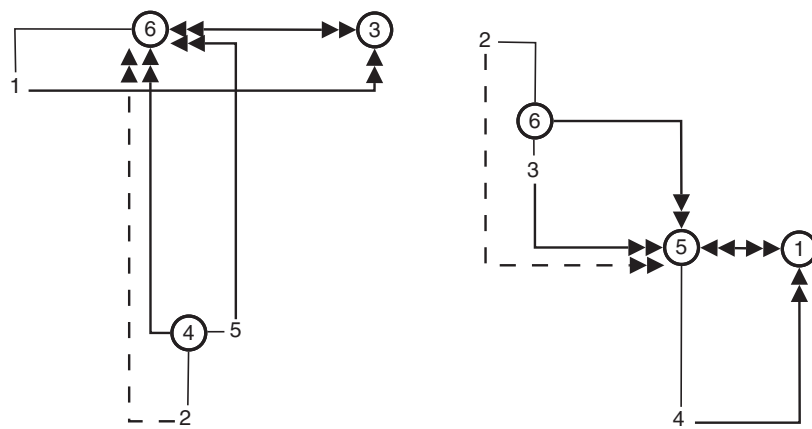


Fig. 7. Examples of Proposition 3.2.

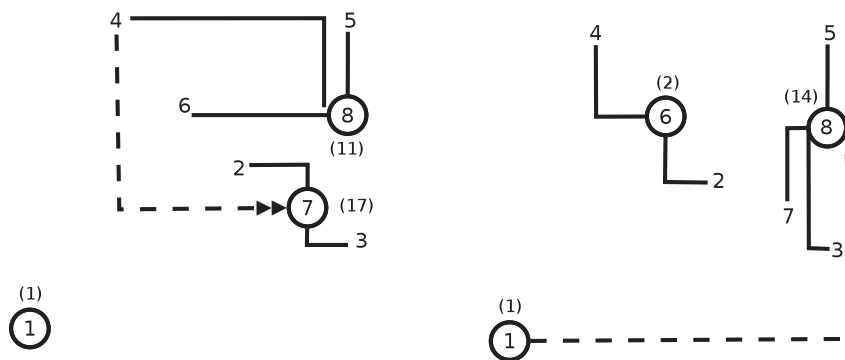


Fig. 8. Example 4.1.

could be a rare case) to reach an emergency center when it is needed. Model CSpCP is nothing but the widening of this humanitarian point of view to the case of disruption.

In order to highlight the strengths of taking into account the failure possibility, in the following we provide a comparison of the optimal solutions obtained for a set of instances by using both CSpCP and CpCP with closest assignment (i.e., when sites are served by their closest centers). Besides the optimal values of both problems, we have computed some illustrative measures of goodness of these solutions as follows. In the case of CSpCP, we compute the maximum distance between each site and its closest center, that is to say, the objective value that the optimal solution of CSpCP gives in the CpCP model. In the case of the optimal solution to CpCP, we consider the failure of one of the centers and distinguish between the cases (i) where each site previously allocated to the center that failed is allocated to its second closest center and (ii) it is allocated to any center with enough capacity (this is done by solving an assignment problem between these sites and the remaining centers taking into account the available capacity of these centers). The maximum distance from every site to its corresponding new center is then calculated.

We start with an example which shows how inadequate it may be to simply solve CpCP and then re-assign sites when the failure of one of the centers occurs. It is worth mentioning that when an emergency situation occurs, injured people, following a natural and intuitive pattern, try to go to the second closest center, since in this type of situations it is implausible that the information as to which center has enough capacity to cover their demands is available.

Example 4.1. We are given eight points in the plane with coordinates (0.6, 2.3), (4.4, 4.6), (6.3, 3.7), (2.7, 9.1), (7.0, 8.3), (3.7, 6.8), (5.9, 3.9) and (6.9, 6.5), demands 5, 2, 2, 4, 1, 5, 4 and 4 and capacities 6, 7, 13, 11, 15, 13, 25 and 25, respectively. Taking ℓ_1 to measure the distances

and $p = 3$, the optimal solution to CSpCP is depicted at the left hand side of Fig. 8. The free capacity in each center is indicated between brackets. The closest center to a non-center point is indicated by a segment. The optimal value (the furthest second closest center with respect to a point in the case of failure) is depicted with a dashed line between site 4 and center 7.

Consider now the possibility of using the optimal solution to CpCP and re-assigning sites in case of failure of some of the centers. The optimal solution to CpCP is depicted at the right hand side of the figure. Now, after a disaster center 1 fails, and the demand of 1 must be allocated to another center with enough free capacity, the only feasible solution is to choose center 8, which is not the second closest center with respect to 1. The situation is still worse if the disaster affects 8, since the free capacity of the remaining centers is not enough to satisfy the demand of sites originally allocated to the unavailable center.

Observing the left hand side of Fig. 8, the reader can realize that the optimal solution of CSpSp requires the system to have a spare capacity of 29 to accommodate the customers reallocation in case of failure. Although the fact that a real system can afford more than 100 percent extra capacity to deal with potential disruptions which in general are rare could be considered as excessive, this extra capacity does not necessarily have to be too expensive. Nowadays, some hospitals have operating rooms or even corridors (areas of rooms for the patients) that usually are closed and then opened when there is an increase of the demand (epidemics, accidents, etc.). Therefore, the hospital only has to cover the expenses of keeping them in good shape to be used in case of necessity. In the following we give some alternative types of centers where including extra capacity does not necessarily imply significant additional costs:

1. The governments have some centers to store oil to address stock-outs or to fight against the price fluctuations of the oil. These cen-

Table 1

Comparison of different optimal solutions for CSpCP and CpCP with closest assignment.

	CpCP					CSpCP		Imp_F (in percent)	Wors_F (in percent)	Imp_NF (in percent)	Wors_NF (in percent)
	OV	OV_F	F_N2C	Pos	WPos	OV	OV_NF				
L & S	340.4	INF (5)	828.3	2nd	4th	708.6	646.9	14.5	–16.9	47.4	–90.1
Orlib_50	59.4	INF (4)	675.0	5th	5th	593.0	571.2	12.1	–13.8	89.6	–861.0
Orlib_50 ^b	44.4	INF (5)	135.0	3rd	4th	119.7	81.7	11.3	–12.8	45.7	–84.2
Orlib_80	51.9	INF (6)	132.1	3rd	3rd	99.2	74.2	24.9	–33.2	30.1	–43.2
45_5	36.2	INF (3)	80.0	2nd	4th	58.5	49.0	26.8	–36.6	26.0	–35.1
45_10	36.2	INF (3)	80.0	2nd	4th	58.5	43.3	26.8	–36.6	16.2	–19.4
75_5	36.5	INF (4)	95.3	4th	5th	62.1	51.7	34.8	–53.4	29.4	–41.7
75_10	36.5	INF (1)	95.3	4th	5th	57.1	50.5	40.0	–66.8	27.7	–38.3
100_5	33.1	INF (4)	76.3	3rd	4th	53.6	40.3	29.8	–42.4	17.8	–21.7
100_10	33.1	63.8	63.8	2nd	4th	51.9	38.1	18.7	–23.0	13.1	–15.0

ters could be used to cover the regular demand of a region (gas stations, airports, etc.), but in addition have an extra capacity to cover the aforementioned special situations.

- Centers of staple products, blankets, tents, machines to get potable water, non perishable food, snowplows etc. that are used to cover small incidents (like storms) can have an extra capacity to replace the damaged material of other center in case of a bigger disaster (like an earthquake or a hurricane).
- A reservoir to store water for consumption to cover the demand of a region could also be used to cover the demand of additional cities if, suddenly, they have supply problems by an unexpected reason like contamination of the water in the reservoir that initially cover those cities.

Summarizing, our model could become a useful tool for the decision-maker or system-planner to evaluate whether it is recommendable to establish a strategy to prevent possible incidents in one of the centers or on the contrary, it is not feasible from a financial point of view. Therefore, this model provides the decision-maker with a quantitative rule to select the best option relative to foresee emergency situations.

Table 1 reports the results of the comparison between CpCP and CSpCP for a set of instances. Columns “OV” give the corresponding optimal values. Given the optimal solution of CpCP, column “OV_F” provides the maximum distance from each site to their second closest center. The number in brackets shows the number of times in which it is not possible to re-assign sites to their second closest centers (because they do not have enough capacity) when one of the centers in the optimal solution to CpCP fails. Column “F_N2C” provides the maximum distance from sites to a center with enough capacity (not necessarily the second closest one) when the closest one fails. Column “Pos” reports the position of “F_N2C” in the ordered sequence of distances between the site for which the maximum distance is achieved and the centers, whereas “WPos” gives the worst position of the distance between each site and the center that would cover this site if its closest one were to fail in the ordered sequence of distances between this site and the centers. Column “OV_NF” reports the maximum distance between the sites and their closest center for the optimal solution of CSpCP (when no center fails).

Column Imp_F gives $\frac{F_N2C - CSpCP_OV}{F_N2C} \times 100$, i.e., the improvement percentage, with respect to the maximum distance from each client to its center if one of these centers fails, by considering the optimal solution of CSpCP (values in column CSpCP_OV) instead of the optimal solution of CpCP (values in column F_N2C; observe that in this case is not a fair comparison because we have not assumed the closest assignment, otherwise in all the cases except the last one this solution would not be feasible). Conversely, column Wors_F reports the worsening percentage, with respect to the maximum distance from each client to its center if one of these centers fails, by considering the optimal solution of CpCP instead of the optimal solution of CSpCP, i.e., $\frac{CSpCP_OV - F_N2C}{CSpCP_OV} \times 100$.

Analogously, column Imp_NF gives $\frac{CSpCP_OV_NF - CpCP_OV}{CSpCP_OV_NF} \times 100$, i.e., the improvement percentage, with respect to the maximum distance from each client to its closest center, by considering the optimal solution of CpCP (values in column CpCP_OV) instead of the optimal solution of CSpCP (values in column CSpCP_OV_NF). Conversely, column Wors_NF reports the worsening percentage, with respect to the maximum distance from each client to its closest center, by considering the optimal solution of CSpCP instead of the optimal solution of CpCP, i.e., $\frac{CpCP_OV - CSpCP_OV_NF}{CpCP_OV} \times 100$.

The instances considered in the comparison are L & S ($n = 90$, $p = 10$, where the 10 sites with the largest demand have been removed, see Lorena and Senne, 2004), Orlib_50 ($n = 50$, $p = 5$, with all capacities equal to 150, see Osman and Christofides, 1994), Orlib_50^b (this instance is Orlib_50 where one of its points has been removed, i.e., $n = 49$, $p = 5$) and Orlib_80 ($n = 80$, $p = 10$ where the 10 sites with the largest demand have been removed, see Osman and Christofides, 1994) and some instances with points randomly generated in the plane (using ℓ_1 to measure the distances); demands and capacities are randomly generated as $U[6, \dots, 15]$ and $[n/p][h_1 + \max_i h_i]/2 + [U(0, 10 \cdot type)]$, respectively, with $type = 5, 10$: 45_5 ($n = 45$, $p = 5$, $type=5$), 45_10 ($n = 45$, $p = 5$, $type=10$), 75_5 ($n = 75$, $p = 5$, $type=5$), 75_10 ($n = 75$, $p = 5$, $type=10$), 100_5 ($n = 100$, $p = 7$, $type=5$) and 100_10 ($n = 100$, $p = 7$, $type=10$).

The difference between n_5 and n_{10} is simply how the capacity of the centers ($type$ 5 or 10) is generated. It can be observed in column OV_F of CpCP that all the instances except the last one are infeasible when sites are assigned to their second closest center if the closest one fails). Columns Pos and WPos show that there are sites that should be allocated to centers sorted between the third and the fifth positions.

From column Imp_F we see the clear advantage of considering the failure foresight which provides a significant improvement (25 percent on average) of the objective value with respect to the CpCP when a center fails.

Although the comparison between the solutions of CpCP and CSpCP highlights that there are significant benefits in using the proposed model in case of disruption, sometimes this comes at the cost of a deterioration of the objective if failure does not occur. This could be measured by comparing columns Imp_F and Wors_NF, but it is worth mentioning that this is not completely fair, since closest assignment in case of failure was not assumed to obtain the values in column F_N2C (otherwise, this solution would not be feasible). Ignoring this point, we can see that the improvement percentages in case of disruption (Imp_F) are smaller than the worsening percentages (Wors_NF) in the first five problems and the seventh one, but quite the opposite in the other cases. Therefore, depending on the cases, the difference of the maximum distance from the clients to their centers for the optimal solution of CpCP and CSpCP when a disruption occurs could be bigger or smaller than the difference of these values when no disruption occurs. Just taking into account these

percentages, the prevention against failure of some of the centers might not be advisable in the first five examples and the seventh one. However, in the other cases, to foresee the failure of one center is highly recommendable.

An extreme case of the above comparison could be the second instance in Table 1, Orlib_50. In this case, the largest site-center distance drops from 675 to 593 when using our model as compared to a standard p -center model, i.e., an improvement of 12.1 percent (see column Imp_F). However, the maximum distance in normal conditions increases from 59.44 to 571.22, i.e., a worsening of 861 percent (see column Wors_NF). A deeper analysis of this data set allows us to recognize that there is an isolated point, a , which is quite far away from the remaining points. Indeed, the coordinates of this point are $a = (70, 719)$ and the 49 remaining points are in the rectangle $[4, 100] \times [4, 234]$. The optimal solution of the CpCP contains a as a center; for this reason the optimal objective value of this problem is relatively small (59.44). However, if we consider the possibility that a center fails, due to the particular configuration of this example, the maximum distance from a center to its closest available center will increase extremely (we can see that this maximum distance for the optimal solution of CpCP increases to 675, even assuming that the sites are not assigned to the closest available center because otherwise this solution would not be feasible). Therefore, the reason for these large differences of percentages in the improvement and worsening of the objective is due to the particular configuration of this problem. Indeed, we have solved the same instance deleting this isolated point, Orlib_50^b, and the results are more reasonable, see the third row of Table 1.

In addition, from the comparison of columns Imp_NF and Wors_F, we can see that in the first three problems, the improvement percentage (considering the optimal solution of CpCP instead of the one of CSpCP when no disruptions occurs) is bigger than the worsening percentage (considering the optimal solution of CpCP instead of the one of CSpCP when a failure occurs). However, the behavior is quite the opposite in the remaining cases.

In conclusion, taking into account both comparisons of these percentages, it seems to be recommendable not to prevent the failure of one of the centers for the first three problems. On the other hand, it is highly recommendable to foresee it in the last five problems except the seventh one. For the fourth, fifth and seventh problems, depending on the utility function of the decision-maker for the differences of these percentages, it may be advisable one or the other choice. Moreover, from the comparison we can conclude that considering the possibility of failure in this kind of model may result in a significant improvement with respect to not considering it. In particular, if failure foresight is not considered, after a disruption in a center, sites may not be allocated to their closest centers (with a high social and political cost, in the case of emergency centers after a natural disaster) resulting in a hardly implementable solution (in an emergency situation, people demanding attention would not know which center would be available to attend them). Moreover, from all these reflexions, we can conclude that independently on whether or not the solution of the CSpCP is implemented by the system-planner/decision-maker, it is indubitable that this model provides an objective and quantitative rule to support his/her decision.

4.2. Comparing formulations

Table 2 gives a comparison of the number of binary variables and constraints of the three previous formulations. We can observe that formulation (P2B) is the one with the lowest number of variables and constraints while (P3) has the highest number. Note that, in what follows, all possibilities of fixing variables at zero in all formulations have been carried out.

Before trying to improve the performance of the previously studied formulations, we will compare them by means of a simple com-

Table 2

Number of variables and constraints of formulations (P3), (P2A) and (P2B).

Formulation	Number of binary variables	Number of constraints
(P3)	n^3	$2n^2 + n + 1$
(P2A)	$2n^2$	$4n^2 + n + 1$
(P2B)	$n^2 + n$	$4n^2 - n + 1$

putational study. The formulations were implemented, as they have been presented in the previous sections, in the commercial solver Xpress IVE 1.22, running on a 3.40 gigahertz PC with 16 gigabytes of RAM memory. The cut generation option of Xpress was disabled in order to compare the relative performance of the formulations cleanly.

In order to produce a set of test instances, we generated random points in the plane. Demands and capacities were randomly generated as $U\{6, \dots, 15\}$ and $\lceil n/p \rceil [(h_1 + \max_i h_i)/2] + [U(0, 10 \cdot \text{type})]$, respectively, where $\text{type} = 3, 4, 5, 10$. The distance considered was ℓ_1 . We compared the three formulations on a testbed of five instances for each combination of (i) type of capacity, (ii) n in $\{15, 20, 25, 30, 35, 40, 45\}$ and (iii) different values of p in $\{3, 5, 8, 10\}$.

Tables 3 and 4 report the results of formulations (P3), (P2A) and (P2B). For each size $n \in \{15, 20, 25, 30, 35, 40, 45\}$, the first column of these tables stands for the different values of p and the second for each one of these three formulations. The remaining columns are grouped in five blocks of four columns. The first four blocks provide the results for the different types of capacity generated (3, 4, 5 and 10) and the last one for the averaged results of these four blocks. In each block, the column u provides the number of unfeasible instances, t gives the average time in seconds of the overall solution process, gap reports the average gap in the root node, and nod represents the average number of thousands of nodes (the above averages refer to the five instances analyzed). The number of instances that exceed the time limit (two hours of CPU) is indicated with a superscript and the averaged times in these cases were calculated by fixing the times of these instances to 7200. As expected because of a lower overall capacity, almost all the instances of type 3, 4 and 5 for $p = 3$ were unfeasible. The symbol “-” in the tables represents that either all the instances are unfeasible or, if some of them are feasible, all these instances need more than two hours of CPU to be solved.

In Tables 3 and 4, we can observe that the running times of formulation (P2A) are much shorter than those of formulation (P2B) and the times of formulation (P2B) are much shorter than the ones of (P3). Actually, the running times of formulation (P3) exceed the time limit for some instances of size greater than or equal to 25 and for almost all instances of size greater than or equal to 35. Regarding the influence of the capacity on the running times, as a general trend the running times for the instances with $\text{type} = 10$ are smaller than the running times of the rest of the values of type , in the cases of small and medium values of p . On the other hand, for large values of p the running times among the instances with different types of capacities are quite similar.

Concerning the gaps of these three different formulations, we can see that the average gaps for (P3) (in the cases where all the instances are solved before reaching the time limit) are greater than those of (P2A). Moreover, the gaps provided by formulation (P2A) are always greater than those provided by (P2B). In addition, the average gaps for each of these formulations over the instances with different types of capacities are almost identical.

Observe that, although the gaps of formulation (P2A) are worse than those of (P2B), its running times are smaller. A possible explanation of this is given by column nod which shows that the number of nodes checked by (P2A), in most of the instances, is much lower than those of (P2B) and (P3).

Table 3
Comparing formulations I.

		type = 3				type = 4				type = 5				type = 10				mean			
		u	t	gap	nod	u	t	gap	nod	u	t	gap	nod	u	t	gap	nod	u	t	gap	nod
<i>n</i> = 15																					
3	(P3)	4	31	47	35.4	1	12	58	12.3	1	14	57	14.4	0	58	46	91.3	6	30	52	42.8
	(P2A)	4	2	39	0.3	1	2	49	0.4	1	2	48	0.3	0	2	36	0.4	6	2	43	0.3
	(P2B)	4	1	31	1.0	1	2	44	2.4	1	1	43	1.9	0	1	31	0.9	6	1	38	1.6
5	(P3)	0	8	30	15.2	0	8	30	15.4	0	11	31	20.4	0	12	30	27.0	0	10	30	19.5
	(P2A)	0	1	29	0.7	0	1	29	0.7	0	1	30	0.7	0	1	30	0.6	0	1	30	0.7
	(P2B)	0	1	29	1.2	0	1	29	1.0	0	1	30	1.8	0	1	29	1.2	0	1	29	1.3
8	(P3)	0	1	3	0.7	0	1	3	0.6	0	1	3	0.7	0	0	3	0.6	0	1	3	0.6
	(P2A)	0	0	3	0.1	0	0	3	0.1	0	0	3	0.1	0	0	3	0.1	0	0	3	0.1
	(P2B)	0	0	3	0.3	0	0	3	0.3	0	0	3	0.3	0	0	3	0.3	0	0	3	0.3
10	(P3)	0	0	0	0.1	0	0	0	0.1	0	0	0	0.0	0	0	0	0.0	0	0	0	0.1
	(P2A)	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0
	(P2B)	0	0	0	0.1	0	0	0	0.1	0	0	0	0.1	0	0	0	0.1	0	0	0	0.1
<i>n</i> = 20																					
3	(P3)	5	–	–	–	5	–	–	–	4	13	65	1.3	0	2893 ²	60	1398.1	14	2413	61	1165.3
	(P2A)	5	–	–	–	5	–	–	–	4	7	46	0.5	0	6	43	0.6	14	7	44	0.6
	(P2B)	5	–	–	–	5	–	–	–	4	8	41	5.7	0	5	36	2.7	14	5	37	3.2
5	(P3)	0	750	47	474.6	0	1544	44	1072.6	0	1183	45	833.7	0	1107	44	911.1	0	1146	45	823.0
	(P2A)	0	6	44	2.5	0	5	41	2.1	0	4	42	1.5	0	4	40	1.8	0	5	42	2.0
	(P2B)	0	11	40	20.2	0	6	37	8.6	0	5	38	6.8	0	4	36	5.5	0	7	38	10.3
8	(P3)	0	22	13	20.2	0	62	13	52.4	0	46	13	44.5	0	39	13	36.7	0	42	13	38.5
	(P2A)	0	2	13	1.0	0	2	13	1.1	0	2	13	1.1	0	2	13	0.8	0	2	13	1.0
	(P2B)	0	2	13	2.5	0	2	13	2.8	0	2	13	2.6	0	2	13	1.8	0	2	13	2.4
10	(P3)	0	4	5	5.2	0	2	5	2.2	0	3	5	3.4	0	2	5	1.8	0	3	5	3.1
	(P2A)	0	1	5	0.2	0	1	5	0.3	0	1	5	0.4	0	1	5	0.3	0	1	5	0.3
	(P2B)	0	1	5	1.2	0	1	5	0.9	0	1	5	1.0	0	1	5	0.8	0	1	5	1.0
<i>n</i> = 25																					
3	(P3)	5	–	–	–	5	–	–	–	3	380	69	31.9	0	1583 ¹	64	292.4	13	1239 ¹	65	218.0
	(P2A)	5	–	–	–	5	–	–	–	3	26	52	1.5	0	16	40	1.1	13	19	44	1.2
	(P2B)	5	–	–	–	5	–	–	–	3	55	45	25.0	0	21	31	5.3	13	31	35	10.9
5	(P3)	0	5813 ⁴	58	1302.3	0	5744 ³	56	1426.6	0	5785 ⁴	56	1408.4	0	5782 ⁴	57	1407.1	0	5781 ¹⁵	57	1386.1
	(P2A)	0	20	42	3.9	0	21	42	4.5	0	20	42	4.8	0	17	42	10004.2	0	19	42	4.3
	(P2B)	0	31	35	17.8	0	26	35	13.3	0	31	35	16.6	0	22	35	9.2	0	28	35	14.2
8	(P3)	0	5537 ³	28	2253.2	0	5650 ³	29	2366.5	0	4208 ¹	28	1577.9	0	3695 ³	28	1630.3	0	4773 ¹⁰	28	1957.0
	(P2A)	0	11	27	6.2	0	13	27	5.9	0	14	27	7.4	0	12	27	5.9	0	13	27	6.4
	(P2B)	0	31	26	31.6	0	32	26	30.2	0	21	26	21.4	0	26	26	24.6	0	27	26	26.9
10	(P3)	0	677 ¹	16	359.0	0	1795 ¹	16	683.1	0	868	16	380.8	0	1450 ¹	16	693.8	0	1198 ³	16	529.2
	(P2A)	0	8	16	6.3	0	7	16	4.0	0	6	16	3.9	0	7	16	5.7	0	7	16	5.0
	(P2B)	0	14	16	20.6	0	8	16	11.0	0	18	16	27.4	0	12	16	16.1	0	13	16	18.8
<i>n</i> = 30																					
3	(P3)	5	–	–	–	3	6504 ²	71	355.8	3	4786 ²	53	281.7	0	3413 ²	66	217.3	11	4405 ⁶	64	262.4
	(P2A)	5	–	–	–	3	73	45	2.9	3	73	47	2.5	0	58	45	2.1	11	65	46	2.3
	(P2B)	5	–	–	–	3	160	36	19.5	3	127	38	22.2	0	119	37	15.1	11	130	37	17.7
5	(P3)	0	7200 ⁵	51	628.3	0	7200 ⁵	57	632.7	0	7200 ⁵	51	778.1	0	7200 ⁵	64	822.6	0	7200 ²⁰	56	715.4
	(P2A)	0	62	42	8.5	0	52	42	8.0	0	56	42	6.8	0	54	42	7.5	0	56	42	7.7
	(P2B)	0	135	34	32.5	0	104	33	19.2	0	95	33	17.7	0	79	33	14.0	0	103	33	20.9
8	(P3)	0	4907 ³	38	529.4	0	4970 ³	41	714.1	0	6035 ⁴	38	866.5	0	4832 ³	36	731.5	0	5186 ¹³	38	710.4
	(P2A)	0	41	25	19.8	0	47	25	26.1	0	44	25	27.0	0	56	25	20.2	0	47	25	23.3
	(P2B)	0	168	23	65.4	0	300	23	115.0	0	168	23	65.2	0	132	23	46.8	0	192	23	73.1
10	(P3)	0	4756 ³	24	883.2	0	4826 ³	25	826.4	0	4695 ³	25	888.9	0	4717 ³	25	964.0	0	4748 ¹²	25	890.6
	(P2A)	0	44	18	38.2	0	34	17	13.2	0	81	17	44.3	0	42	17	40.1	0	50	17	33.9
	(P2B)	0	130	18	73.6	0	83	17	40.7	0	157	17	76.6	0	113	17	65.8	0	121	17	64.2

Table 4
Comparing formulations II.

		type = 3				type = 4				type = 5				type = 10				mean			
		<i>u</i>	<i>t</i>	gap	nod	<i>u</i>	<i>t</i>	gap	nod	<i>u</i>	<i>t</i>	gap	nod	<i>u</i>	<i>t</i>	gap	nod	<i>u</i>	<i>t</i>	gap	nod
<i>n</i> = 35																					
3	(P3)	5	–	–	–	5	–	–	–	4	7200 ¹	13	296.0	0	6650 ⁴	67	167.0	14	6742 ⁵	58	188.5
	(P2A)	5	–	–	–	5	–	–	–	4	179	45	3.1	0	196	49	3.1	14	193	48	3.1
	(P2B)	5	–	–	–	5	–	–	–	4	347	35	19.1	0	440	39	35.2	14	425	39	32.5
5	(P3)	0	–	47	–	0	–	64	–	0	–	46	–	0	–	63	–	0	–	55	–
	(P2A)	0	355	43	21.4	0	309	43	20.3	0	322	42	22.5	0	236	41	19.9	0	306	42	21.0
	(P2B)	0	1453	36	194.4	0	1057	36	110.6	0	772	34	69.5	0	500	33	38.8	0	945	35	103.3
8	(P3)	0	–	46	–	0	–	42	–	0	–	46	–	0	–	40	–	0	–	44	–
	(P2A)	0	303	30	119.7	0	294	29	115.9	0	285	29	107.6	0	277	29	111.0	0	290	29	113.5
	(P2B)	0	862	25	162.5	0	627	24	139.3	0	587	24	107.2	0	611	24	115.2	0	672	24	131.0
10	(P3)	0	–	36	–	0	–	30	–	0	–	34	–	0	–	31	416.4	0	–	33	–
	(P2A)	0	273	22	65.2	0	379	22	101.0	0	536	22	171.6	0	655	22	314.8	0	461	22	163.1
	(P2B)	0	393	19	116.7	0	617	19	149.2	0	580	19	134.7	0	568	19	128.5	0	539	19	132.3
<i>n</i> = 40																					
3	(P3)	5	–	–	–	5	–	–	–	5	–	–	–	0	–	71	–	15	–	71	–
	(P2A)	5	–	–	–	5	–	–	–	5	–	–	–	0	582	51	4.6	15	582	51	4.6
	(P2B)	5	–	–	–	5	–	–	–	5	–	–	–	0	1486	42	85.3	15	1486	42	85.3
5	(P3)	0	–	58	–	0	–	65	–	0	–	58	–	0	–	66	–	0	–	62	–
	(P2A)	0	1011	46	28.4	0	878	46	32.7	0	837	45	26.7	0	637	45	24.3	0	841	46	28.0
	(P2B)	0	1677	35	76.3	0	1755	36	84.6	0	1214	34	51.0	0	1261	34	52.6	0	1477	35	66.1
8	(P3)	0	–	48	–	0	–	55	–	0	–	57	–	0	–	57	–	0	–	54	–
	(P2A)	0	924	37	223.6	0	940	37	225.1	0	927	37	230.3	0	816	37	215.4	0	902	37	223.6
	(P2B)	0	1241	30	84.0	0	1398	30	95.0	0	1757	30	111.9	0	1445 ¹	30	91.2	0	1460 ¹	30	95.5
10	(P3)	0	–	52	–	0	–	49	–	0	–	46	–	0	–	48	–	0	–	49	–
	(P2A)	0	2475	33	997.2	0	2215	33	1006.0	0	2390	33	916.4	0	2145	33	926.9	0	2306	33	961.6
	(P2B)	0	3410 ¹	30	319.3	0	4091 ²	30	386.6	0	3577 ²	30	386.1	0	3629 ¹	30	351.1	0	3677 ⁶	30	360.8
<i>n</i> = 45																					
3	(P3)	5	–	–	–	5	–	–	–	5	–	–	–	0	–	62	–	15	–	62	–
	(P2A)	5	–	–	–	5	–	–	–	5	–	–	–	0	1282	49	7.6	15	1282	49	7.6
	(P2B)	5	–	–	–	5	–	–	–	5	–	–	–	0	3002	39	103.6	15	3002	39	103.6
5	(P3)	0	–	59	–	0	–	70	–	0	–	68	–	0	–	67	–	0	–	66	–
	(P2A)	0	1989	46	41.0	0	1988	46	49.8	0	1879	46	42.0	0	1933	46	53.2	0	1947	46	46.5
	(P2B)	0	3179	34	78.4	0	3253	34	68.7	0	2744	34	60.0	0	3195	34	59.4	0	3093	34	66.6
8	(P3)	0	–	64	–	0	–	47	–	0	–	52	–	0	–	54	–	0	–	54	–
	(P2A)	0	2562	35	300.8	0	2491	35	304.5	0	2373	35	293.5	0	2162	35	299.7	0	2397	35	299.6
	(P2B)	0	3023 ¹	28	242.8	0	2937 ¹	29	246.6	0	3599 ¹	29	259.8	0	3119 ¹	29	249.0	0	3170 ⁴	29	249.6
10	(P3)	0	–	49	–	0	–	43	–	0	–	45	–	0	–	50	–	0	–	47	–
	(P2A)	0	3997	31	860.1	0	4018	31	619.1	0	4105	31	882.4	0	4669	31	894.3	0	4197	31	814.0
	(P2B)	0	5279 ³	28	486.1	0	6586 ⁴	29	565.0	0	6114 ³	29	493.3	0	6183 ³	30	511.6	0	6040 ¹³	29	514.0

Algorithm 1: Heuristic approach.

```

Initialize  $\mathcal{L} := \emptyset$  (the list of already tested feasible solutions),
 $cont := 0, iteration := 0, MAXIT, obj^* = +\infty$ ;
Generate a random solution,  $CS = \{i_1, \dots, i_p\}$  (current solution);
3 Check feasibility of  $CS$ ;
while  $iteration < MAXIT$  do
    iteration := iteration + 1;
    if  $CS$  is unfeasible then
        cont := 1;
        DO a FEASIBILITY MOVEMENT.
    else
        cont := 0;
        Compute the objective value  $f(CS)$ ;
        if  $f(CS) < obj^*$  then
             $obj^* := f(CS)$  and  $S^* := CS$ 
14 DO an IMPROVEMENT MOVEMENT.
    Let  $CS'$  be the new generated solution;
    if  $CS' \notin \mathcal{L}$  then
         $\mathcal{L} := \mathcal{L} \cup CS'$ ;
        GO TO LINE 3
    else
        if  $cont = 1$  then
            GO TO LINE 8
        else
            GO TO LINE 14
Return  $S^*$  and  $obj^*$ ; the best solution found and the
corresponding objective value;

```

5. A heuristic approach

In this section we develop a heuristic algorithm to address the problem under consideration. The main idea of this procedure is to generate new solutions iteratively through local movements. Depending on whether the current solution is feasible or not, we make an improvement movement or a feasibility movement, respectively. This means that we move from a current solution to another one in its neighborhood, in such a way that, hopefully, it will have a lower objective value when the current solution is feasible, or it could, otherwise, be feasible. The detailed pseudo-code is depicted in Algorithm 1.

In the improvement movement, given a current feasible solution with open centers $CS = \{j_1, \dots, j_p\}$ and objective value $f(CS)$, we choose $j_{\min} \in CS$ and $i_{\min} \in N$ such that $d(i_{\min}, j_{\min}) = f(CS)$. Then, we look for j'_{\min} , the closest node to j_{\min} such that (i) $j'_{\min} \notin CS$, and (ii) $d(i_{\min}, j'_{\min}) < d(i_{\min}, j_{\min})$. Set $CS' = (CS \setminus \{j_{\min}\}) \cup \{j'_{\min}\}$. Hopefully, this movement would decrease the current objective value whenever this solution was feasible and would not imply an excessive increase in remaining distances between sites and their corresponding second closest centers.

In the feasibility movement, given a current solution, for instance (\hat{x}, \hat{w}) for (P2A), with open centers $CS = \{j_1, \dots, j_p\}$, we choose $j_{\min} \in CS$ maximizing the gap of the capacity constraints, i.e.

$$j_{\min} := \arg \max_{j \in N} \left(\max_{\substack{s \in N \\ j_s \neq j}} \sum_{\substack{t \in N \\ d_{tj_s} < d_{tj}}} h_t (\hat{w}_{tj} + \hat{x}_{j_s j_s} - 1) + \sum_{\substack{t \in N \\ t \neq j}} h_t \hat{x}_{tj} - b_j \hat{x}_{jj} \right).$$

Then we look for j'_{\min} , the closest node to j_{\min} such that (i) $j'_{\min} \notin CS$, and (ii) $b_{j'_{\min}} > b_{j_{\min}}$, i.e., the closest node with larger capacity. Set $CS' = (CS \setminus \{j_{\min}\}) \cup \{j'_{\min}\}$. Hopefully this movement would have generated a new solution where the maximum gap of the capacity constraints have decreased and it could be feasible if this gap is less than or equal to 0 for each pair of centers.

If one of the previous movements is not possible, we randomly choose the new place to locate a center j'_{\min} . Whenever these local movements cannot generate a new solution which has not been generated previously, a completely new random solution is generated.

We checked the heuristic on the same testbed as the initial computational study. The algorithm needs around 8, 25, 59 and 125 seconds for sizes 45, 75, 100 and 125, respectively, to carry out 1000 iterations. Fig. 9 illustrates the efficiency of this heuristic approach. The three graphs on the left report the instances which are optimality solved (gap=0) and the numbers at the top of the bars providing the percentage of instances where the heuristic does not solve optimally the problem (gap > 0) report the average gap. The three on the right part show the iteration where the heuristic algorithm reaches the best solution (using a limit of 5000 iterations). In both cases, the graphics represent the percentage of instances as a function of n (top), types of capacity (center) and p (bottom), respectively. In the first case, we can see that the percentage of instances solved optimally decreases when n increases. However, the types of capacity and p do not have a large influence. In the right part, we can observe that the percentage of instances that find the best result before the first 100 iterations is quite high, especially for small values of n . Again, we can see that this behavior is quite similar for different types of capacity and p . Therefore, we can conclude that the most influential aspect in the results of the algorithm is the size of the instances analyzed.

6. Improvements

This section is devoted to improving the preliminary computational results obtained in Section 4.2. Since (P3) is not a very efficient formulation from the computational time point of view, we concentrate on improving the performance of formulations (P2A) and (P2B). Different techniques were studied but we only show those which provide the best running times. After these improvements, we will see that formulation (P2A) again reports the best computational times.

6.1. Improving formulation (P2A)

In order to solve our problem using (P2A), we tested several families of valid inequalities and a variable fixing strategy based on the best value found by the heuristic given in Section 5. This will be a successful approach which, for most of the instances, will significantly reduce the computational times given in Tables 3 and 4.

6.1.1. Valid inequalities

The first sets of valid inequalities that we try to use to reinforce formulation (P2A) are those coming from formulation (P2B), i.e., (18)–(20). Our first result shows that (20) will not improve formulation (P2A).

Proposition 6.1. Constraints (20) are implied by constraints $x_{jj} \leq 1 \forall j \in N$, (14) and (15).

Proof. We consider two cases for $j, \ell \in N$. First, we assume that there exists $q \in N$ such that $d_{\ell q} > d_{ej}$. In such a case, from $x_{jj} - 1 \leq 0 \forall j \in N$ it follows that

$$\sum_{\substack{q \in N \\ d_{ej} < d_{\ell q}}} (x_{jj} - 1) \leq x_{jj} - 1 \Rightarrow 1 + x_{\ell j} - x_{jj} + \sum_{\substack{q \in N \\ d_{ej} < d_{\ell q}}} (x_{jj} - 1) \leq x_{\ell j}.$$

Now from (14)

$$\sum_{\substack{q \in N \\ d_{ej} < d_{\ell q}}} w_{\ell q} + \sum_{\substack{q \in N \\ d_{ej} < d_{\ell q}}} (x_{jj} - 1) \leq x_{\ell j}.$$

This inequality trivially holds in the second case, i.e., $d_{\ell q} \leq d_{ej}, \forall q \in N$.

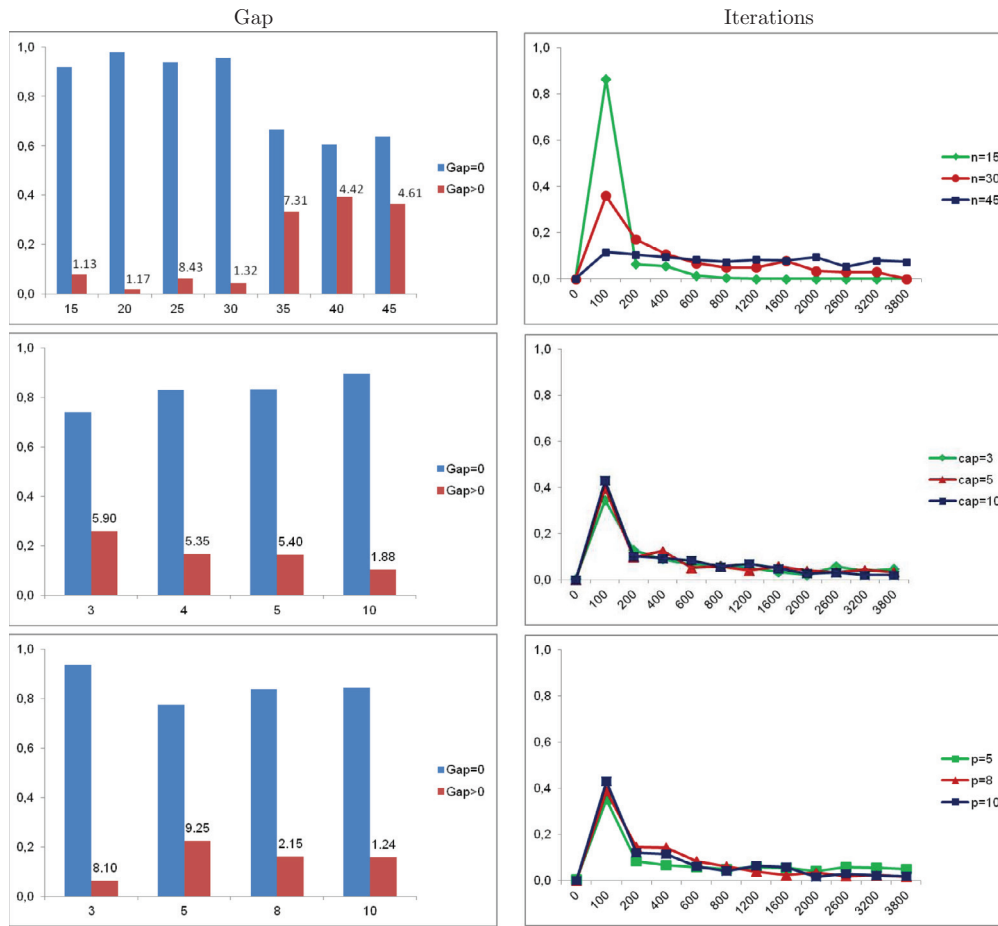


Fig. 9. Results obtained with the heuristic.

Thus, for all $j, \ell \in N$ it holds that

$$\sum_{\substack{q \in N \\ d_{\ell j} < d_{\ell q}}} (w_{\ell q} + x_{jj} - 1) \leq x_{\ell j}.$$

For a given $j \in N$ we can linearly combine these inequalities with non negative weights h_ℓ to obtain

$$\sum_{\ell \in N} \sum_{\substack{q \in N \\ d_{\ell j} < d_{\ell q}}} h_\ell (w_{\ell q} + x_{jj} - 1) \leq \sum_{\ell \in N} h_\ell x_{\ell j} \leq$$

[from (15)]

$$\begin{aligned} h_j x_{jj} - \sum_{\substack{\ell \in N \\ d_{\ell j} < d_{\ell j}}} h_\ell (w_{\ell j} + x_{ii} - 1) + (b_j - h_j) x_{jj} \\ = - \sum_{\substack{\ell \in N \\ d_{\ell j} < d_{\ell j}}} h_\ell (w_{\ell j} + x_{ii} - 1) + b_j x_{jj}. \end{aligned}$$

The inequality between the first and last terms matches (20). \square

Consequently, among the constraints of (P2B) we will only consider (18) and (19) as valid inequalities for (P2A). Additionally, we will incorporate the following two families of inequalities:

$$\sum_{\substack{\ell \in N \\ d_{\ell i} > d_{ij}}} x_{i\ell} + \sum_{\substack{\ell \in N \\ d_{\ell i} \leq d_{ij}}} w_{i\ell} \leq 1 \quad \forall i, j \in N, \quad (24)$$

$$z \geq \sum_{\substack{\ell \in N \\ \ell \neq i}} \hat{d}_{i\ell} x_{i\ell} + \hat{d}_{ji} x_{ji} + d_{ji} w_{ji} \quad \forall i, j \in N. \quad (25)$$

Here $\hat{d}_{i\ell} := \min_{j \in N \{d_{ij} : d_{ij} > d_{i\ell}\}}$, i.e., $\hat{d}_{i\ell}$ is the distance immediately larger than $d_{i\ell}$ from site i to another site.

Constraints (24) act as follows. If $\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} x_{i\ell} = 1$, it means that i is allocated to a closest center further than j . Thus, the second closest center of i is further than j , or equivalently, $w_{i\ell} = 0 \quad \forall \ell \in N : d_{i\ell} \leq d_{ij}$. If $\sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} w_{i\ell} = 1$ it means that i is allocated to a second closest center no further than j ; then, i is not allocated to a center further than j , or equivalently, $x_{i\ell} = 0 \quad \forall \ell \in N : d_{i\ell} > d_{ij}$.

The meaning of inequalities (25) is the following. If $x_{ji} = 1$, then $x_{i\ell} = 0, \forall \ell \neq i$ and $w_{ji} = 0, \forall j \in N$. Thus, the second closest center from site j is at least \hat{d}_{ji} . If $w_{ji} = 1$ then $x_{ji} = 0$ and $x_{i\ell} = 0, \forall \ell \neq i$. Thus, the objective function is at least d_{ji} . If $x_{i\ell} = 1$ for some $\ell (\neq i) \in N$ then $x_{ji} = w_{ji} = 0$. Moreover, the second closest center from site i is at least $\hat{d}_{i\ell}$.

In what follows we see that valid inequalities (24) are stronger than valid inequalities (18) when we consider them to enforce formulation (P2A).

Proposition 6.2. Constraints (18) are implied by (11), (12) and (24).

Proof. Using (11) and taking into account that $w_{jj} = 0 \quad \forall j \in N$, constraints (24) can be rewritten as:

$$\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} w_{i\ell} + \sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} x_{i\ell} \geq 1.$$

Hence, applying (12) we obtain

$$\sum_{\substack{\ell \in N \\ d_{i\ell} > d_{ij}}} w_{i\ell} + \sum_{\substack{\ell \in N \\ d_{i\ell} \leq d_{ij}}} (x_{i\ell} - w_{i\ell}) \geq 1,$$

and this is equivalent to (18) \square

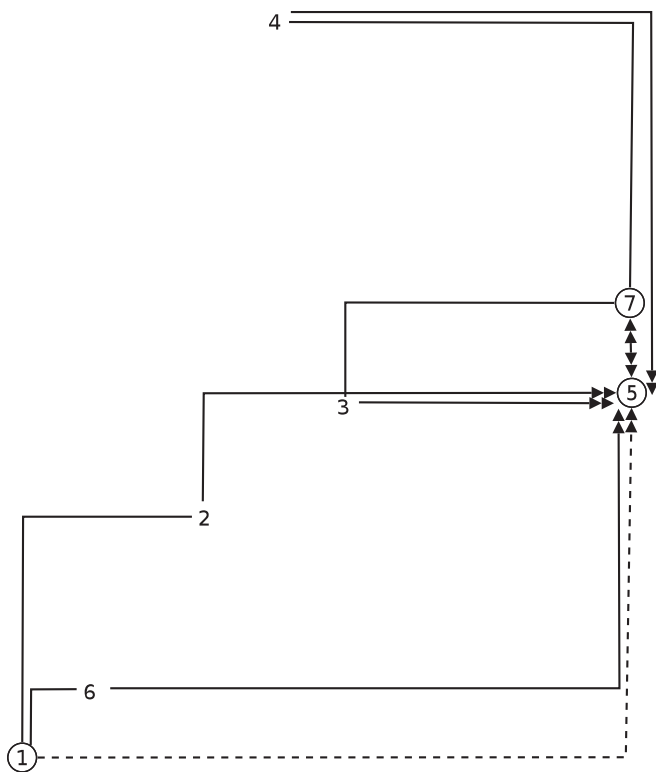


Fig. 10. Example of Proposition 6.3.

Let then (P2AI) be the formulation resulting from adding constraints (19), (24) and (25) to (P2A). We see now that all these constraints make a contribution to the lower bound given by the linear relaxation of (P2AI).

Proposition 6.3. *None of the sets of constraints (13)–(15), (19), (24) and (25) are implied by other constraints in Formulation (P2AI).*

Proof. To prove the thesis it suffices to show an instance in which the linear relaxation gives a larger value without any one of the sets of constraints (13)–(15), (19), (24) and (25) than the linear relaxation of the whole set. We take a set of seven points in the plane, $p = 3$ and the ℓ_1 norm is used to compute distances between points. The points are (0, 1), (3, 5), (5, 7), (4, 9), (10, 6), (1, 1), (10, 7) with demands 3, 3, 2, 3, 0, 3, 2 and capacities 9, 1, 6, 1, 9, 2, 9, respectively.

Centers in the optimal solution are 1, 5 and 7. The two closest centers with respect to each site are (1, 5), (1, 5), (7, 5), (7, 5), (5, 7), (1, 5), (7, 5), respectively. The optimal value is 15 (distance from 1 to 5, see Fig. 10), and the linear relaxation of (P2AI) takes value 7.648. The optimal values of the linear relaxations of (P2AI) when relaxing (13), (15), (19), (24) and (25) are 7.217, 7.63, 6.604, 7.618, 7.39 and 7.026, respectively. \square

6.1.2. Variable fixing

In order to fix w_{ij} -variables we will take advantage of our heuristic approach presented in Section 5. Indeed, since this approach gave very good computational times either to solve the problem or to provide good solutions, in this section we will use it to obtain feasible solution and consequently, the upper bounds on the optimal value of our problem. Hence, it allows us to fix w_{ij} -variables as follows. If vo_h^* is the best objective value obtained by the heuristic approach, we can fix at 0 w_{ij} -variables and x_{ij} -variables such that $d_{ij} > vo_h^*$. In addition to this preprocessing phase, in the methodology that we have developed, every time that a feasible solution is found in the branching tree, we have again applied a process of fixing variables, such that if this feasible solution provides a better objective value than those

obtained so far, we can fix at 0 all the w_{ij} - and x_{ij} -variables such that d_{ij} are greater than this objective value. Thus, the dimension of the problem decreases iteratively and, consequently, the process is sped up.

6.1.3. Computational analysis

Table 5 reports the results of the methodology described above when some of the families, subfamilies or combinations of valid inequalities (19), (24) and (25) are added to formulation (P2A). We have attempted cut-and-branch and branch-and-cut methods with all the inequalities. Only the families of constraints which provided the best results are shown. We have fixed the parameter in the Heuristic at $it = 1000$ and since the different ways of generating the capacity do not make a significant influence in the running times, as shown in Tables 3 and 4 and Fig. 9, in this table we only consider the type of capacity 5. For each size $n \in \{45, 75, 100, 110, 125\}$, the first column of these tables stands for the different values of p . Thus, nine blocks of two columns provide the average times and the gaps of the five instances solved by adding a family of valid inequalities to formulation (P2A). Namely, block OUR reports the results of solving our problem with formulation (P2A) using the above methodology without adding an additional family of valid inequalities and blocks “19”, “24” and “25” mean that inequalities (19), (24) and (25) are added, respectively. On the other hand, blocks “19_{ij}”, “24_{ij}” and “25_{ij}” report the results of solving formulation (P2A) using a Cut & Branch procedure including the most violated constraint of families (19), (24) and (25) for each $i, j \in N$, respectively. Finally, blocks “19_{ij}&24_{ij}” and “19_{ij}&24_{ij}&25_{ij}” report the results of the respective combinations of valid inequalities in a Cut & Branch procedure. The best running times have been bold-faced for each n and p . In the particular case of size 125, the number of instances where the CPU times exceeds two hours is indicated as superscripts. As a general rule, to compute the average times of the five instances analyzed for each n and p , we have taken 7200 seconds as the running times of those instances that exceed the time limit.

In Table 5, we observe that the use of this methodology to solve our problem has reduced the computational times needed to solve the instances with $n = 45$ by more than two orders of magnitude and we have been able to solve all the checked instances of size lower than or equal to 110 in less than two hours. Moreover, none of the families of valid inequalities analyzed dominates the others. For small-sized problems, i.e. $n = 45$ and 75, “19_{ij}&24_{ij}&25_{ij}” and “19_{ij}&24_{ij}” provide the best average computational times, respectively. However, for larger instances it seems that “24” and “24_{ij}” present a better trend. We can observe that all the approaches provide similar gaps. Finally, in order to check the limit of our solution approach we have repeated this analysis, taking $it = 10,000$ and we have been able to solve the five instances tested for problems of size 170 in less than two hours.

6.2. Improving formulation (P2B)

Following the same ideas developed for (P2A), we propose several improvements for (P2B). In this case, although the improvements in formulation (P2B) are worthwhile from a theoretical point of view, we have not reported the computational times because the results obtained with (P2A) are much better than those obtained using (P2B) (for instance, for $n = 100$ these improvements of (P2B) are not able to solve most of the instances studied in less than two hours, whereas with the ones in (P2A) we were able to solve these problems in approximately 12 minutes).

6.2.1. Valid inequalities

We have studied different families of valid inequalities as follows.

1. Consider $i, a, t \in N (i \neq a, t)$ such that (i) $d_{it} < d_{ia}$ and (ii) $d_{tr} \geq \max\{d_{ti}, d_{ta}\}$ for all $r \in N$ such that $d_{ir} > d_{ia}$. We have the following

Table 5
Improving formulation (P2A).

p	OUR		19		19 _{ij}		24		24 _{ij}		25		25 _{ij}		19 _{ij} &24 _{ij}		19 _{ij} &24 _{ij} &25 _{ij}	
	t	gap	t	gap	t	gap	t	gap	t	gap	t	gap	t	gap	t	gap	t	gap
$n = 45$																		
5	31	27	41	27	33	27	30	23	36	27	39	19	31	27	31	27	25	27
8	15	5	17	5	14	5	18	4	15	5	15	5	15	5	17	5	13	5
10	16	19	19	19	16	19	15	18	15	19	14	9	14	19	16	19	16	19
mean	21	17	26	17	21	17	21	15	22	17	23	11	20	17	21	17	18	17
$n = 75$																		
5	641	27	337	27	267	27	281	23	859	27	451	20	722	27	143	27	252	27
8	111	18	137	18	94	18	98	17	101	18	90	16	88	18	84	18	102	18
12	170	28	356	28	174	28	195	25	176	28	226	22	195	28	177	28	188	28
15	168	27	394	27	170	27	196	26	158	27	154	22	197	27	161	27	191	27
mean	273	25	306	25	176	25	193	23	324	25	230	20	301	25	141	25	183	25
$n = 100$																		
7	1368	33	1070	33	1254	33	709	28	672	33	411	21	513	33	955	33	719	33
11	1084	32	1791	32	603	32	928	27	870	32	1097	24	518	32	611	32	749	32
16	1021	31	2614	31	1183	31	1152	27	810	31	1372	25	1377	31	1165	31	1433	31
20	675	26	2033	26	712	26	841	24	693	26	1075	21	670	26	583	26	776	26
mean	1037	31	1877	30	938	30	908	27	761	31	989	23	770	31	829	30	919	30
$n = 110$																		
7	763	33	1651	33	882	33	841	33	1124	33	751	33	606	33	797	33	827	33
12	1136	36	3425	36	1516	36	1120	36	1571	36	1741	36	1696	36	1504	36	1818	36
18	1278	33	3436	33	1189	33	1015	33	1180	33	1673	33	1509	33	1308	33	1101	33
22	2425	31	5625	31	2361	31	2216	31	2191	31	3444	31	3035	31	2186	31	2577	31
mean	1401	33	3534	33	1487	33	1298	33	1517	33	1902	33	1712	33	1449	33	1581	33
$n = 125$																		
8	2381 ¹	23	3914 ²	23	2604	23	2587 ¹	23	2457 ¹	23	3419 ¹	23	2570 ¹	23	2356¹	23	2562 ¹	23
14	4271 ²	38	5915 ²	38	4131 ²	38	3872 ¹	38	4077 ¹	38	4688 ²	38	4141 ¹	38	3650¹	38	4176 ²	38
20	4788 ²	37	6599 ⁴	37	4777¹	37	5067 ²	37	5055 ²	37	5227 ³	37	5551 ³	37	5232 ³	37	4890 ²	37
25	3662	32	6929 ⁴	32	3619	32	3672	32	3407	32	5715 ²	32	5458 ²	32	3833	32	4875 ²	32
mean	3776	33	5839	33	3783	33	3800	33	3749	33	4762	33	4430	33	3768	32	4126	33

two sets of constraints

$$\begin{aligned} w_{ia} + x_{ji} - 1 &\leq w_{ta}, & \text{if } d_{ti} < d_{ta} \\ w_{ia} + x_{ji} - 1 &\leq w_{ti}, & \text{if } d_{ti} \geq d_{ta}. \end{aligned} \quad (26)$$

Constraints (26) have an effect only if $x_{ji} = w_{ia} = 1$. In this case, the closest center to site i is itself and the second closest one is a ; this means that $x_{jj} = 0$ for any $j \neq i$ such that $d_{ij} < d_{ia}$. Thus, $x_{ti} = 0$. Moreover, since $d_{tr} > \max\{d_{ti}, d_{ta}\}$, for all $r \in N$ such that $d_{ir} > d_{ia}$ the closest and second closest centers of site t will be, respectively, (i) i and a if $d_{ti} < d_{ta}$ and (ii) a and i if $d_{ti} \geq d_{ta}$. Therefore, in the first case we will have that $w_{ta} = 1$ and in the second case that $w_{ti} = 1$.

2. Following the same arguments, we can reinforce the second family of valid inequalities of (26) as follows:

$$\sum_{\substack{a \in N: d_{ji} \geq d_{ta} \\ d_{ja} < d_{ia} \\ d_{tr} \geq \max\{d_{ti}, d_{ta}\}, \forall r: d_{ir} > d_{ia}}} w_{ia} + x_{ji} - 1 \leq w_{ti}, \quad \forall t, i. \quad (27)$$

3. Constraints (26) have been obtained for any $i, a, t \in N$ with $i \neq a, t$ and $d_{it} < d_{ia}$. Now, following similar arguments, we extend these constraints to the case where $d_{it} \geq d_{ia}$. Consider $i, a, t \in N$ with $i \neq a$ such that (i) $d_{it} \geq d_{ia}$ and (ii) $d_{tr} \geq \max\{d_{ti}, d_{ta}\}$ for all $r \in N$ such that $d_{ir} > d_{ia}$. We have the following set of constraints:

$$w_{ia} + x_{ji} - 1 \leq w_{ta} + w_{ti}. \quad (28)$$

4. The last set of constraints of this family is

$$x_{ji} + w_{ia} - 1 \leq w_{ai} + \sum_{\substack{k \in N \\ d_{ak} < d_{ai}, d_{ik} > d_{ia}}} x_{kk} \quad \forall i, a \in N. \quad (29)$$

Constraints (29) act as follows. If $x_{ji} = w_{ia} = 1$, this means that $\sum_{d_{ij} < d_{ia}} x_{jj} = x_{ji} = 1$. Then, the second closest center to a should be either i or a center such that the distance from i is larger than d_{ia} . In the latter case, a should be closer to that center than i .

7. Conclusions

Emergency services are vulnerable to disruptions caused by large natural disasters, terrorist attacks or sabotage. The consequences of a failure in this type of services are often disastrous despite their rare occurrence. For this reason, it is worth taking into account this aspect in the design phase for locating these types of facilities. Moreover, although most of the existing papers in the literature considering reliability issues in supply chain design have avoided the capacity constraints, because they increase the complexity of the models, in order to be more realistic it is relevant to consider these constraints also. Therefore, in this paper we have presented several linear integer formulations for the problem of locating emergency services with capacity constraints taking into account the possibility of a failure. Some have been improved by means of valid inequalities and preprocessing techniques for fixing variables. A heuristic has been developed which can also be used in the preprocessing phase of a methodology that we have provided to obtain an optimal solution. This latter approach provided the best results, reducing by more than two orders of magnitude the running times for $n = 45$ and allowing all the instances tested with $n \leq 110$ to be solved in under two hours.

Acknowledgements

The authors would like to thank to the anonymous referees for their careful reading of the manuscript and their valuable comments. This research has been partially supported by the Spanish Ministerio de Economía y Competitividad, grants nos. MTM2012-36163-C06-04, MTM2010-19576-C02-02 and MTM2013-46962-C2-2-P, ERDF funds, Fundación Séneca, grant no. 08716/PI/08 and Junta de Andalucía, grant no. FQM-5849.

References

- Aksen, D., & Aras, N. (2012). A bilevel fixed charge location model for facilities under imminent attack. *Computers & Operations Research*, 39, 1364–1381.
- Aksen, D., Aras, N., & Piyade, N. (2013). A bilevel p -median model for the planning and protection of critical facilities. *Journal of Heuristics*, 19, 373–398.

- Albareda-Sambola, M., Díaz, J., & Fernández, E. (2010). Lagrangean duals and exact solution to the capacitated p -center problem. *European Journal Operational Research*, 201, 71–81.
- Ball, M. O., & Lin, F. L. (1993). A reliability model applied to emergency service vehicle location. *Operations Research*, 41(1), 18–36.
- Bar-Ilan, J., Kortsarz, G., & Peleg, D. (1993). How to allocate network centers. *Journal of Algorithms*, 15, 385–415.
- Basar, A., Aatay, B., & Unluyurt, T. (2012). A taxonomy for emergency service station location problem. *Optimization Letters*, 6, 1147–1160.
- Belotti, P., Labbé, M., Maffioli, F., & Ndiaye, M. (2007). A branch-and-cut method for the obnoxious p -median problem. *4OR*, 5, 299–314.
- Berman, O., Krass, D., & Menezes, M. (2007). Facility reliability issues in network p -median problems: Strategic centralization and co-location effects. *Operations Research*, 55, 332–350.
- Calik, H., & Tansel, B. C. (2014). Double bound method for solving the p -center location problem. *Computers and Operations Research*, 40, 2991–2999.
- Cui, T., Ouyang, Y., & Shen, Z. (2010). Reliable facility location design under the risk of disruptions. *Operations Research*, 58, 996–1011.
- Curtin, K. M., Hayslett-McCall, K., & Qiu, F. (2010). Determining optimal police patrol areas with maximal covering and backup covering location models. *Networks and Spatial Economics*, 10(1), 125–145.
- Daskin, M. (1995). *Network and discrete location: Models, algorithms, and applications*. New York: Wiley.
- Daskin, M. (2000). A new approach to solving the vertex p -center problem to optimality: Algorithm and computational results. *Communications of the Operations Research Society of Japan*, 45, 428–436.
- Drezner, Z., & Hamacher, H. (2002). *Facility location: Applications and theory*. New York: Springer Verlag.
- Elloumi, S., Labbé, M., & Pochet, Y. (2004). New formulation and resolution method for the p -center problem. *INFORMS Journal on Computing*, 16, 84–94.
- Espejo, I., Marín, A., & Rodríguez-Chía, A. (2012). Closest assignment constraints in discrete location problems. *European Journal Operational Research*, 219, 49–58.
- Hogan, K., & ReVelle, C. (1986). Concepts and applications of backup coverage. *Management Science*, 32, 1434–1444.
- Ilhan, T., & Pinar, M. (2001). An efficient exact algorithm for the vertex p -center problem. http://www.optimization-online.org/DB_HTML/2001/09/376.html.
- Jaeger, M., & Goldberg, J. (1994). A polynomial algorithm for the equal capacity p -center problem on trees. *Transportation Science*, 28, 167–175.
- Kalcsics, J., Nickel, S., Puerto, J., & Rodríguez-Chía, A. (2010). The ordered capacitated facility location problem. *Top*, 18, 203–222.
- Kariv, O., & Hakimi, S. (1979). An algorithmic approach to network location problems I: The p -centers. *SIAM Journal on Applied Mathematics*, 37(3), 513–538.
- Khuller, S., & Sussmann, Y. (2000). The capacitated k -center problem. *SIAM Journal on Discrete Mathematics*, 13, 403–418.
- Lei, T. (2010). *Location modelling utilizing closest and generalized closest transport/interaction assignment constructs*. Santa Barbara, CA: University of California Ph.D. thesis.
- Lei, T. L., & Church, R. L. (2011). Constructs for multilevel closest assignment in location modeling. *International Regional Science Review*, 34, 339–367.
- Liberatore, F., Scaparra, M. P., & Daskin, M. S. (2012). Hedging against disruptions with ripple effects in location analysis. *Omega*, 40, 21–30.
- Lorena, L. A. N., & Senne, E. L. F. (2004). A column generation approach to capacitated p -median problems. *Computers and Operations Research*, 31(6), 863–876. <http://www.lac.inpe.br/~loreana/instancias.html>.
- Mladenović, N., Labbé, M., & Hansen, P. (2003). Solving the p -center problem with tabu search and variable neighborhood search. *Networks*, 42, 48–64.
- O'Hanley, J. R., & Church, R. L. (2011). Designing robust coverage networks to hedge against worst-case facility losses. *European Journal of Operational Research*, 209, 23–36.
- O'Hanley, J. R., Scaparra, M., & García, S. (2013). Probability chains: A general linearization technique for modeling reliability in facility location and related problems. *European Journal of Operational Research*, 230, 63–75.
- Osman, I. H., & Christofides, N. (1994). Capacitated clustering problems by hybrid simulated annealing and tabu search. *International Transactions in Operational Research*, 1(3), 317–336. <http://people.brunel.ac.uk/~mastijb/jeb/orlib/pmedcapinfo.html>.
- Özsoy, F., & Pinar, M. (2006). An exact algorithm for the capacitated vertex p -center problem. *Computers and Operations Research*, 33, 1420–1436.
- Parvaresh, F., Hussein, S. M. M., Golpayegany, S. A. H., & Karimi, B. (2014). Hub network design problem in the presence of disruptions. *Journal of Intelligent Manufacturing*, 25, 755–774.
- Scaparra, M., Pallotino, S., & Scutellá, M. (2004). Large-scale local search heuristics for the capacitated vertex p -center problem. *Networks*, 32, 241–255.
- Scaparra, P., & Church, R. (2008). A bilevel mixed integer program for critical infrastructure protection planning. *Computers and Operations Research*, 35, 1905–1923.
- Shen, Z., Zhan, R., & Zhang, J. (2011). The reliable facility location problem: Formulations, heuristics, and approximation algorithms. *INFORMS Journal on Computing*, 23, 470–482.
- Snyder, L., & Daskin, M. (2005). Reliability models for facility location: The expected failure cost case. *Transportation Science*, 39, 400–416.
- Snyder, L., Scaparra, M., Daskin, M., & Church, R. (2006). Planning for disruptions in supply chain networks. *INFORMS Tutorial in operations research* (pp. 234–257) Baltimore, MD. <http://dx.doi.org/10.1287/educ.1063.0025>.
- Storbeck, J. E., & Vohra, R. V. (1988). A simple trade-off model for maximal and multiple coverage. *Geographical Analysis*, 20, 220–230.
- Wagner, J., & Falkson, L. (1975). The optimal nodal location of public facilities with price-sensitive demand. *Geographical Analysis*, 7, 69–83.
- Weaver, J. R., & Church, R. L. (1985). A median location model with non-closest facility service. *Transportation Science*, 19, 58–74.
- Zhan, R., Shen, Z.-J., & Daskin, M. (2007). *System reliability with location-specific failure probabilities*. Technical Report. Berkeley, CA: University of California.