Computers & Operations Research 36 (2009) 1966 – 1981

Contents lists available at ScienceDirect

Computers & Operations Research

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

# A comparison of formulations and solution methods for the minimum-envy location problem

Inmaculada Espejo[a], Alfredo Marín[b], Justo Puerto[c], Antonio M. Rodríguez-Chía[a,*]

[a]Facultad de Ciencias, Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, Pol Río San Pedro, 11510 Puerto Real (Cádiz), Spain
[b]Departamento de Estadística e Investigación Operativa, Universidad de Murcia, Spain
[c]Departamento de Estadística e Investigación Operativa, Universidad de Sevilla, Spain

ARTICLE INFO

Available online 28 June 2008

*Keywords:*
Discrete location
Equity
Integer programming

ABSTRACT

We consider a discrete facility location problem with a new form of equity criterion. The model discussed in the paper analyzes the case where demand points only have strict preference order on the sites where the plants can be located. The goal is to find the location of the facilities minimizing the *total envy* felt by the entire set of demand points. We define this new total envy criterion and provide several integer linear programming formulations that reflect and model this approach. The formulations are illustrated by examples. Extensive computational tests are reported, showing the potentials and limits of each formulation on several types of instances. Finally, some improvements for all the formulations previously presented are developed, obtaining in some cases much better resolution times.



## 1. Introduction

Over the last decade, the increasing interest in equity issues has resulted in new methodologies in the area of Operations Research. Continuing this trend, this paper examines a discrete location problem that consists of establishing a fixed number of $p$ plants to cover $M$ demand points based solely on their preference orders on the sites where these plants can be located. The goal is to minimize the dissatisfaction of the demand points arising from pairwise comparisons among them. Equitable location is a matter which has been addressed previously in the literature but, whereas in other papers preferences over sites were represented by numerical scales, our approach requires only ordinal information, giving rise to a new paradigm that deals with the concept of *envy*.

The concept of envy has been widely studied in the literature of Decision Theory. In particular, one of the most well-known criteria to judge fairness and satisfaction is envy-freeness, i.e., a solution of a decision process such that every agent of this process likes its own solution at least as much as any other agent's. In particular, the envy-free criterion has been used in problems of fair division of indivisible items among people (see e.g. [1–5]), problems of allocating heterogeneous indivisible objects (see e.g. [6,7]), queueing problems (see e.g. [8]), auction problems (see e.g. [9,10]) and implementation of game solution concepts (see e.g. [11]), among others. To the best of our knowledge, addressing location problems from this perspective has not been previously considered in the literature.

According to the Cambridge Advanced Learner's English dictionary the first meaning of the term *envy* is "to wish that you had something that another person has". In our framework envy is defined with respect to the revealed preference of each demand point for the sites of the potential serving facilities. We assume that demand points only have from a most-preferred to a least-preferred strict preference order for the sites where the plants can be located. The goal is to find the location of the facilities minimizing the total envy felt by the entire set of demand points. A limitation of our model is that it is assumed that the decision-maker has a previous complete knowledge of the preferences of all customers at the demand points or, alternatively, that all customers do not lie when they are asked about their preferences.

During the last two decades there has been a major effort to develop location models which capture more features of real problems. One of the features that has attracted more attention has been the issue of equity in cost or distance-to-travel distribution. Nevertheless, while for efficiency and effectiveness there is almost a consensus that median and center, respectively, are the most representative objective functions, this is not the case when modelers look for an equity criteria. Many different proposals for such a criteria can be found in the literature. Several authors have considered different general aspects such as: how to define equality measures, how to measure equality, what properties equality functions have and what

* Corresponding author. Tel.: +34 9560 16087; fax: +34 9560 16050.
*E-mail addresses:* inmaculada.espejo@uca.es (I. Espejo), amarin@um.es (A. Marín), puerto@us.es (J. Puerto), antonio.rodriguezchia@uca.es (A.M. Rodríguez-Chía).

0305-0548/$ - see front matter 
doi:10.1016/j.cor.2008.06.013

they should have, and how to compare the solutions provided by the different models.

Thus for instance, Savas [12] pointed out the insufficiency of efficiency and effectiveness measures in location models for public facilities. Halpern and Maimon [13] considered a large number of tree networks in order to determine the agreement and disagreement of the solutions to location problems using the median, center, variance and Lorenz measure. Mulligan [14] designed a simple experiment consisting of locating a facility in an interval of the real straight line regarding three demand points. Apart from a comparative analysis of the optimal solutions for nine equality measures: the median, center, range, Gini coefficient, mean absolute deviation, Hoover's concentration index, variance, standard deviation and Theil's index, Mulligan provided the standardized travel distance curves for them. Erkut [15] proposed a general framework for quantifying inequality and presented some axioms for the appropriateness of the inequality measures. He also showed that only two of his considered measures—the coefficient of variation and the Gini coefficient—hold both the scale-invariance property and the principle of transfers or Pigou–Dalton property. Berman and Kaplan [16] address the equity question using taxes. Recently, Mesa et al. [17] and Garfinkel et al. [18] addressed some algorithmic aspects of equity measures on network location and routing.

A review of the existing literature on equity measurement in Location Theory and a discussion on how to select an appropriate measure of equality is contained in the paper by Marsh and Schilling [19]. Also the equality objectives literature is reviewed in the paper by Eiselt and Laporte [20] within a general discussion of objectives in Location Theory based on the physics concepts of pulling, pushing and balancing forces.

In this paper, we study a discrete location problem and try to find sites minimizing the overall envy felt by demand points according to their own revealed preference scales. Thus, we introduce a new element into the literature of equity methods in location analysis. Since this is the first paper dealing with this problem, our aim is to find an adequate formulation for the problem and to implement an efficient solution method based on this formulation. Thus, we provide three different formulations for the problem of minimizing the total envy, either based on known ideas for other discrete location problems or created *ad hoc* for this case. These formulations are compared in a preliminary computational study to test their performance. Our second formulation seems to outperform the others when the number of plants is not very small. The rest of the formulations have a quite similar performance. Then we have derived several improvements of these formulations. In particular, we try some valid inequalities of the first formulation, we develop an *ad hoc* cut-and-branch algorithm (a cut-and-branch algorithm consists of adding valid inequalities to the formulation of the problem and then going with the enforced formulation to a branch-and-bound scheme) for the second formulation, based on some families of valid inequalities and two variable fixing strategies, and we test a slight modification of the objective function of our third formulation. The improved formulations and/or solution methods are again computationally compared on the same testbed.

The paper is organized as follows. In the next section, we formalize the concept of equity and pose the problem we aim to formulate. In Sections 3–5, three mixed integer programming formulations for the problem are introduced. Preliminary computational results are given in Section 6. Then, in Sections 7–9 we try to improve the formulations using different strategies which are computationally compared in Section 10. We close the paper with some conclusions.

## 2. Formalization of the minimum envy location problem

Let $A=\{1,\dots,M\}$ the given set of $M$ sites. We assume without loss of generality that the set of candidate sites for new plants is identical to the set of clients.

Let $O=(O_{ij})_{i,j=1,\dots,M}$ be the $M\times M$ preferences matrix. Each row of matrix $O$ is given by a permutation of $A$, in such a way that the smaller $O_{ij}$, the most preferred site $j$ is for client $i$. Let $1\leqslant p\leqslant M-1$ be the number of plants to be located. A solution to the minimum envy location problem (MELP) is given by a set of $p$ sites $X\subset A$, $|X|=p$.

We assume that each client will be served by its favourite plant, i.e., given a solution $X$, we assume that each client $i$ will be supplied from a site $j\in X$ such that

$$O_{ij}=O_i(X):=\min_{k\in X}\{O_{ik}\}.$$

$P_i(X)$ denotes the plant assigned to client $i\in A$ in a solution $X$. The definition of the envy of client $i_1$ for client $i_2$, that will be used throughout the paper, is given by

$$e_{i_1i_2}(X)=\begin{cases}0 & \text{if } O_{i_1,P_{i_1}(X)}\leqslant O_{i_2,P_{i_2}(X)},\\ O_{i_1,P_{i_1}(X)}-O_{i_2,P_{i_2}(X)} & \text{otherwise.}\end{cases}$$

Our goal is to minimize the total envy induced by the choice of a set of locations, that is

$$\min_{X\subset A,|X|=p} F(X):=\sum_{i_1}\sum_{i_2\neq i_1} e_{i_1i_2}(X).$$

Taking into account that

$$e_{i_1i_2}(X)+e_{i_2i_1}(X)=|O_{i_1,P_{i_1}(X)}-O_{i_2,P_{i_2}(X)}|,$$

the objective function can be expressed as

$$F(X)=\sum_{i_1=1}^{M-1}\sum_{i_2=i_1+1}^{M}|O_{i_1,P_{i_1}(X)}-O_{i_2,P_{i_2}(X)}|. \tag{1}$$

Notice that this objective function is related to the sum of absolute differences, as studied in [17,21]. However, in these two papers the preferences are expressed, rather than as a permutation of $A$ (i.e., in an ordinal scale), as the distances from a demand point to the remainders (preferences of demand point $i$ are given by $(d(i,1),\dots,d(i,M))$). Thus, it is clear that our approach cannot distinguish between slight and large preferences. For more research on equity objectives for location problems where these weighted preferences are taken into account, we refer the reader to [22].

**Example 1.** In Fig. 1 we show a small example. There are six points on a line, each one representing a customer and also a potential plant, sorted from left to right. We must locate two plants. Customers prefer closer points and, in case of ties, they prefer points to the right-hand side. Then, the preferences matrix is

$$O=\begin{pmatrix}1&2&3&4&5&6\\3&1&2&4&5&6\\4&2&1&3&5&6\\5&4&2&1&3&6\\6&4&3&2&1&5\\6&5&4&3&2&1\end{pmatrix}.$$

The interpretation of this matrix is as follows. Each row is associated with a point. For instance, row 5 means that the fifth client (located at point 7) prefers itself, then site 4 (located at point 4), then site 3 (located at point 2), then site 2 (located at point 1), then site 6 (located at point 14) and, finally, site 1 (located at point 0).

The optimal solution of the $p$-median (resp. $p$-center) problem on this instance, where the aim is to minimize the total (resp. maximum) distance between plants and their allocated customers, is obtained locating plants at 2 and 14 (4 and 14) and the optimal allocation pattern can be seen in the first (second) picture of Fig. 1.

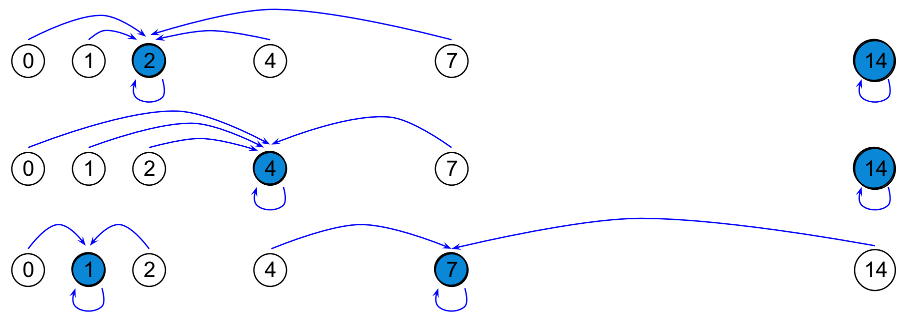


**Fig. 1.** Different optimal solutions of $p$-median, $p$-center and minimum-envy problems on a small instance.

In the optimal solution of the $p$-median problem, customer at 1 is allocated to its second favourite point (its favourite point is itself), customer at 0 is allocated to its third favourite point and so on. We summarize the allocation of all points, with respect to their preferences, in vector (3,2,1,2,3,1). Then, customer at 0 feels that it may have been better allocated to 0 or 1 whereas, for instance, customer at 4 is allocated to its second favourite point, and we quantify this part of customer at 0's envy by 1. In total, the envy felt by the customers in this solution, calculated by means of (1), is $(1+2+1+0+2)+(1+0+1+1)+(1+2+0)+(1+1)+(2)=16$.

Despite the fact that $p$-center is in the family of equitable objectives, the envy produced by the associated optimal solution is still larger, 21, corresponding with the vector $(4,3,2,1,2,1)$. Notice that, in order to reduce the maximum allocation distance, a plant is located at 14 thus disappointing customer at 0 which has to go to a plant not satisfying its expectations.

An optimal solution of the minimum envy problem for this instance is shown in the last picture of Fig. 1. By locating plants at 1 and 7, the corresponding vector is $(2,1,2,3,1,2)$ and the total envy is 13. The essence of the envy criterion is explained by the fact that the unsatisfaction of customer at 14 is not proportional to the distance to the allocated plant, since it "knew" beforehand that all but one (itself) of the possible locations for plants were far from it. Moreover, customer at 4, allocated to a point which is close to it but is in the third position of its preferences, envies the rest of the customers which were allocated to their first or second favourite points.

The reader may note that using ordinal ranking in location problems is not new and there is an important body of literature in competitive location that deals with this issue. Examples are the use of Condorcet and Simpson rules or the Borda count in location problems or in the spatial theory of voting (see e.g. [23–26]). In all these cases, facilities (users) are seen as voters that have revealed preferences induced by the distances to the servers (plants). Then, facilities are only allowed to use their ordinal preference to make their decisions (one person one vote) regardless of the actual values of the distance inducing the preferences. Under the same framework, one may use a different criterion and rather than, for instance, using the simple Borda count one may like to minimize the total envy produced by the preferences (as it is proposed in this paper).

There are many possible applications even not directly inspired in the Location field. Consider, for example, the following examination problem. Let $A=\{1,\dots,M\}$ be the set of people that apply for a given set of positions. The exam consists of $p$ lessons (questions) chosen out of $T$ and the applicants (students) must answer just one out of these $p$ possible choices. The examiner, rather than making a random choice of the lessons (questions), tries to do a fair selection of them. To this end, he asks each one of the applicants to sort the lessons (questions) from the most preferred one to the less preferred one and he (the examiner) chooses those $p$ lessons minimizing the overall envy felt by the applicants according to their own revealed preference scales. The reader may note that this procedure avoids the potential bad luck that some applicants may suffer due to a random selection of questions.

In the following, we present several mixed integer linear programming formulations for the MELP.

## 3. First formulation

The first formulation of the MELP uses variables which are *natural* with respect to the definition of the problem. There will be three sets of variables. The first one is the usual set of location variables:

$$y_j = \begin{cases} 1 & \text{if a plant is located at site } j, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall j \in A.$$

The second set of variables, $z_i$ with $i \in A$, is used to measure $O_i(X)$, the minimum of the entries of row $i$ in matrix $O$ which are located in columns which are plants. Then, $z_i$ will take value 1 if there is a plant in the site most preferred by $i$, value 2 if this is not the case but there is a plant in the second most preferred site, and so on.

Finally, making abuse of notation, $e_{ij}$ will represent the envy that either client $i$ feels for client $j$ or vice versa, for every $i,j \in A, i \neq j$.

Using the latter set of variables, the objective function can be easily expressed as

$$\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} e_{ij}.$$

Variables $e_{ij}$ are obtained from variables $z_i$ by adding the following sets of constraints to the formulation for every $i=1,\dots,M-1$, $j=i+1,\dots,M$:

$$e_{ij} \geqslant z_i - z_j,$$

$$e_{ij} \geqslant z_j - z_i.$$

In turn, variables $z_i$ are going to be calculated from the location variables $y_j$. On the one hand, variable $z_i$ must be forced to be greater than or equal to $k$ if none of the $k-1$ most-preferred-by-$i$ sites receives a plant:

$$z_i + \sum_{\ell : O_{i\ell} \leqslant k-1} (k - O_{i\ell}) y_\ell \geqslant k, \quad \forall i,k \in A. \tag{2}$$

In the above constraints, if $y_\ell = 0$ for all the $k-1$ most-preferred-by-$i$ sites, variable $z_i$ will assume a value of, at least, $k$. The coefficients of variables $y_\ell$ are as small as possible, in such a way that if $y_\ell = 1$ for some $\ell$, its coefficient $k - O_{i\ell}$ goes to the right-hand side of the inequality, reducing its value to $O_{i\ell}$.

On the other hand, since every client must be assigned to its most preferred plant, variables $z_i$ must be forced to take a maximum value of $O_{ik}$ if there is a plant at site $k$. The reader may note that by breaking this condition, smaller values of the objective could be obtained by assigning a client to a plant which is not the most preferred one available. Closest allocation can be enforced by means of the (not yet optimized) constraints

$$z_i + (M - O_{ik})y_k \leqslant M, \quad \forall i, k \in A. \tag{3}$$

Here, if $y_k = 0$ then $z_i$ will be bounded by $M$, i.e., not really bounded, whereas if $y_k = 1$ then $z_i$ will be bounded by $O_{ik}$.

Putting all these constraints together, fixing the number of plants to $p$ and taking into account in constraints (2) and (3) that no client will be assigned to any of its $p - 1$ less preferred sites, the first formulation is

$$\begin{aligned} (\text{F1}) \quad \min \quad & \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} e_{ij} \\ \text{s.t.} \quad & e_{ij} \geqslant z_i - z_j, \quad \forall i = 1, \ldots, M-1, \ j = i+1, \ldots, M, \\ & e_{ij} \geqslant z_j - z_i, \quad \forall i = 1, \ldots, M-1, \ j = i+1, \ldots, M, \\ & z_i + \sum_{\ell : O_{i\ell} \leqslant k-1} (k - O_{i\ell}) y_\ell \geqslant k, \quad \forall i \in A, \ k = 1, \ldots, M-p+1, \\ & z_i + (M-p+1-O_{ik})y_k \leqslant M-p+1, \quad \forall i, k \in A : O_{ik} \leqslant M-p, \qquad (4) \\ & \sum_{j=1}^{M} y_j = p, \\ & y_j \in \{0,1\}, \quad \forall j \in A. \qquad (5) \end{aligned}$$

**Example 2.** Consider $M = 5$, $p = 2$ and the preferences matrix

$$O = \begin{pmatrix} 1 & 4 & 3 & 2 & 5 \\ 2 & 1 & 5 & 3 & 4 \\ 4 & 2 & 1 & 5 & 3 \\ 5 & 4 & 3 & 1 & 2 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}.$$

Consider a solution with plants in sites 2 and 5, i.e., $X = \{2, 5\}$ and $y = (0, 1, 0, 0, 1)$. Then, the values of the $z$-variables are given by

$z = (4, 1, 2, 2, 1)$.

Therefore client 1 goes to its fourth favourite site and $z_1 = 4$. Then, the $e$-variables take the values

$$e = \begin{pmatrix} 0 & 3 & 2 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The total envy (the sum of the elements in matrix $e$) is 14.

## 4. Second formulation

We will build our second formulation through an intermediate formulation obtained from (F1) adding unrestricted variables $x_i$, $i \in A$, such that the vector $(x_i)$ contains the values of variables $(z_i)$ sorted in non-increasing order. Thus, the objective function can be obtained from the values of the $x$-variables simply by doing

$$F(X) = \sum_{i_1} \sum_{i_2 > i_1} (x_{i_2} - x_{i_1}) = \sum_{i=1}^{M} (2i - M - 1)x_i.$$

The constraints needed to obtain the values of the $z$-variables from the values of the $y$-variables are the same as in (F1). In order to obtain the values of the $x$-variables, we add two families of constraints.

The first family enforces the sorted shape of the $x$-vector:

$x_i \leqslant x_{i+1}, \quad \forall i = 1, \ldots, M-1.$

The second family, containing $O(2^M)$ constraints, ensures that each value in the $x$-vector matches a value of the $z$-vector (we extract one of the constraints of this family, the corresponding to the entire set $A$, which is in fact an equality):

$$\sum_{i=k}^{M} x_i \geqslant \sum_{i \in S} z_i, \quad \forall k = 2, \ldots, M, \ \forall S \subset A : |S| = M - k + 1, \tag{6}$$

$$\sum_{i=1}^{M} x_i = \sum_{i=1}^{M} z_i.$$

When $k = M$ in (6), it is enforced that $x_M$ will be greater than or equal to the maximum value of variables $z_i$. Due to the shape of the objective function, $x_M$ will be equal to this maximum. When $k = M-1$, the second largest value of $z$ is assigned to $x_{M-1}$, and so on.

Therefore, the intermediate formulation is as follows:

$$\begin{aligned} \min \quad & \sum_{i=1}^{M} (2i - M - 1)x_i \\ \text{s.t.} \quad & z_i + \sum_{\ell : O_{i\ell} \leqslant k-1} (k - O_{i\ell}) y_\ell \geqslant k, \quad \forall i \in A, \ k = 1, \ldots, M-p+1, \\ & z_i + (M-p+1-O_{ik})y_k \leqslant M-p+1, \quad \forall i, k \in A : O_{ik} \leqslant M-p+1, \\ & \sum_{j=1}^{M} y_j = p, \\ & x_i \leqslant x_{i+1}, \quad \forall i = 1, \ldots, M-1, \\ & \sum_{i=1}^{M} x_i = \sum_{i=1}^{M} z_i, \\ & \sum_{i=k}^{M} x_i \geqslant \sum_{i \in S} z_i, \quad \forall k = 2, \ldots, M, \quad \forall S \subset A : |S| = M-k+1, \\ & y_j \in \{0,1\}, \quad \forall j \in A. \qquad (7) \end{aligned}$$

**Example 3.** Consider again the data and solution of Example 2. The values of the $z$-variables were given by

$z = (4, 1, 2, 2, 1)$.

Constraints (7) for $k = 5$ are $x_5 \geqslant 4$, $x_5 \geqslant 1$, $x_5 \geqslant 2$, $x_5 \geqslant 2$, $x_5 \geqslant 1$. Then, $x_5$ will be at least 4. Constraints (7) for $k = 4$ can be summarized as $x_4 + x_5 \geqslant \max\{5, 6, 3, 2, 4\}$. Then, $x_4 + x_5$ will be at least 6. In turn, we get $x_3 + x_4 + x_5 \geqslant 8$, $x_2 + x_3 + x_4 + x_5 \geqslant 9$ and $x_1 + x_2 + x_3 + x_4 + x_5 = 10$. The objective function is $-4x_1 - 2x_2 + 2x_4 + 4x_5$, which can be rewritten as

$$\begin{aligned} & -6(x_1 + \cdots + x_5) + 2[x_5 + (x_4 + x_5) + (x_3 + x_4 + x_5) \\ & \quad + (x_2 + x_3 + x_4 + x_5) + (x_1 + x_2 + x_3 + x_4 + x_5)] \\ & \quad = -60 + 2[x_5 + (x_4 + x_5) + (x_3 + x_4 + x_5) + (x_2 + x_3 + x_4 + x_5) \\ & \quad + (x_1 + x_2 + x_3 + x_4 + x_5)]. \end{aligned}$$

This function is minimized when each constraint from above is satisfied as an equality, i.e., when each $x$-variable takes its lowest possible value. Then, $x = (1, 1, 2, 2, 4)$, with objective value 14.

The high number of constraints makes it impossible to use this formulation without a tailored scheme. The problem comes with constraints (7), which can be summarized by writing, for all $k$ in $\{2, \ldots, M\}$,

$$\sum_{i=k}^{M} x_i \geqslant \max \left\{ \sum_{i \in S} z_i : S \subset A, \ |S| = M - k + 1 \right\}. \tag{8}$$

In other words, $\sum_{i=k}^{M} x_i$ must be greater than or equal to the sum of the $M-k+1$ maximum values in vector $z$. We designed a separation procedure for (8) (and consequently for (7)) in which only the $M-k+1$ maximum values in $z$ were compared with $\sum_{i=k}^{M} x_i$ and, if this last sum was the smallest one among these two, the corresponding constraint was added. The separation procedure was iterated at every node of a branching tree, until no more constraints (7) were violated by the optimal solution of the linear subproblem. The results obtained with this method were not satisfactory. For this reason, this formulation is not further considered in our computational experiments.

In [27], the authors introduced some constraints whose effect was to sort the values of a set of variables which can be used here to avoid the exponential number of constraints of the previous formulations. To build our second formulation we define

$$z_{ik} = \begin{cases} 1 & \text{if there are no plants located at the } k-1 \\ & \text{favourite sites of client } i, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall i,k \in A.$$

and

$$x_{ik} = \begin{cases} 1 & \text{if there are no plants located at the } k-1 \\ & \text{favourite sites of the } i\text{-th luckiest client,} \\ 0 & \text{otherwise,} \end{cases} \qquad \forall i,k \in A.$$

Observe that $z_i = \sum_{k\in A} z_{ik}\ \forall i$ and $x_i = \sum_{k\in A} x_{ik}\ \forall i$. As it is illustrated in the following example, each column of the matrix $(x_{ik})$ is obtained by sorting the entries in the corresponding column of the matrix $(z_{ik})$.

**Example 4.** Consider again the data and solution of Example 2. The values of the $z_{ij}$ variables will be given by

$$z = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For instance, in row 1 corresponding to client 1 there are four 1's since its three most preferred sites do not contain any plant, but the fourth most preferred site does. Moving the ones to the bottom of the columns, the following values of the $x$-variables are obtained:

$$x = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

At the same time, the rows of the latter matrix match the rows of the former matrix, but they are now sorted in lexicographic order. The total envy of 14 is obtained when summing up the number of 0's over the 1's of every row. In particular, client 1 (in the last row on the second matrix since it is the unluckiest client) gives 10 to the total envy because there are 2 zeros in the second column and 4 zeros in the third and fourth columns, whereas clients 1 and 4, tied in the third and fourth positions of the luckiness ranking, each give 2 to the total envy. Customers 2 and 5 are tied in the luckiest positions and have zero envy.

Before developing the constraints needed in the formulation, we observe that, following the definition of the above variables,

$$z_{i1} = x_{i1} = 1, \quad \forall i \in A,$$

since all the clients must be allocated to some plant. Thus, we will no longer need these variables. In the same way, taking into account that no client will be assigned to its $p-1$ least preferred sites, we fix the following variables:

$$z_{i,M-p+2} = \cdots = z_{iM} = 0, \quad \forall i \in A,$$
$$x_{i,M-p+2} = \cdots = x_{iM} = 0, \quad \forall i \in A.$$

The first three sets of constraints of this formulation are used to obtain the values of the $z$-variables from the solution $X$ ($|X|=p$), that is defined by the $y$-variables. Note that, for each $i \in A$, variables $z_{ik}$ must be sorted in non-increasing order (i.e., they will take value one for some $k=1,\dots,q$ and then value zero for $k=q+1,\dots,M$):

$$z_{ik} \geqslant z_{i,k+1}, \quad \forall i \in A, \ k=2,\dots,M-p.$$

To ensure that the change from one to zero in the $i$-th row of the matrix of $z$-variables is made in the right position, two conditions must be imposed. The change must be made in correspondence with a site where a plant was installed:

$$z_{iO_{ij}} - z_{i,O_{ij}+1} \leqslant y_j, \quad \forall i,j \in A : O_{ij} \leqslant M-p+1, \tag{9}$$

and it must be the favourite plant of client $i$:

$$z_{i,O_{ij}+1} + y_j \leqslant 1, \quad \forall i,j \in A : O_{ij} \leqslant M-p.$$

Once the $z$-values are obtained, two more families of constraints are needed to get the (sorted) variables $x$ from them. The idea is to copy the matrix $z$ into the matrix $x$ but moving, in each column, the ones to the bottom and the zeros to the top

$$\sum_{i=1}^{M} x_{ik} = \sum_{i=1}^{M} z_{ik}, \quad \forall k=2,\dots,M-p+1,$$

$$x_{ik} \geqslant x_{i-1,k}, \quad \forall i=2,\dots,M, \ k=2,\dots,M-p+1.$$

Putting all these constraints together, fixing the number of plants to $p$, and taking into account the variables we have fixed above, the resulting formulation is

$$\begin{aligned}
\text{(F2)} \quad \min \quad & \sum_{i=1}^{M}(2i-M-1)\sum_{k=2}^{M-p+1} x_{ik} \\
\text{s.t.} \quad & z_{ik} \geqslant z_{i,k+1}, \quad \forall i \in A, \ k=2,\dots,M-p, \\
& z_{iO_{ij}} - z_{i,O_{ij}+1} \leqslant y_j, \quad \forall i,j \in A : 2 \leqslant O_{ij} \leqslant M-p, \\
& 1 - z_{i2} \leqslant y_j, \quad \forall i,j \in A : O_{ij} = 1, \qquad (10) \\
& z_{i,M-p+1} \leqslant y_j, \quad \forall i,j \in A : O_{ij} = M-p+1, \qquad (11) \\
& z_{i,O_{ij}+1} + y_j \leqslant 1, \quad \forall i,j \in A : O_{ij} \leqslant M-p, \\
& \sum_{i=1}^{M} x_{ik} = \sum_{i=1}^{M} z_{ik}, \quad \forall k=2,\dots,M-p+1, \\
& x_{ik} \geqslant x_{i-1,k}, \quad \forall i=2,\dots,M, \ k=2,\dots,M-p+1, \\
& \sum_{j=1}^{M} y_j = p, \\
& y_j \in \{0,1\}, \quad \forall j \in A, \\
& x_{ik} \in \{0,1\}, \quad \forall i \in A, \ k=2,\dots,M-p+1.
\end{aligned}$$

Constraint families (10) and (11) are obtained from (9) when $O_{ij} = 1$ and $O_{ij} = M-p+1$, respectively, after fixing $z_{i1}$ to 1 and $z_{i,M-p+2}$ to 0.

## 5. Third formulation

The next formulation for the MELP we are going to consider is based on the work by Ogryczak and Tamir [28]. We have adapted their formulation for the ordered median problem to our case as follows.

Variables $z_i$ are defined again as in the former formulations. Following the idea of Ogryczak and Tamir, we look for the sum of the $q$ largest $z$-values. $z_{(i)}$ represents the values $z_i$ sorted in non-increasing order.

For any integer $q$, $1 \leqslant q \leqslant M$, consider the following function defined in $[0, +\infty)$:

$$f_q(t) := qt + \sum_{i=1}^{M} \max\{0, z_i - t\}.$$

This is a convex piecewise linear function with slopes moving from $q - M$ to $q$ in integer steps, whose minimum is reached either when the slope is 0 or, if this is not the case, when the slope changes from negative to positive, i.e., when $t$ equals the $q$-th maximum value of the vector $z$, namely $z_{(q)}$. Thus, the minimum value of $f$ is

$$f_q(z_{(q)}) = qz_{(q)} + \sum_{i=1}^{M} \max\{0, z_i - z_{(q)}\} = qz_{(q)} + \sum_{i=1}^{q} (z_{(i)} - z_{(q)}) = \sum_{i=1}^{q} z_{(i)},$$

i.e., the sum of the $q$ maximum values. Therefore, minimizing the sum of the $q$-largest $z$-values can be linearized as

$$\min qt_q + \sum_{i=1}^{M} d_{iq} \quad \text{s.t. } d_{iq} \geqslant 0,\ \forall i, \quad d_{iq} \geqslant z_i - t_q,\ \forall i.$$

We overcharge our notation and use the index $q$ in variables $d$ and $t$ because we are interested in solving this problem for each value $q$ from 1 to $M - 1$ and combine the latter linearization with the constraints used in the former formulations to get the values of the $z_i$ variables in our third formulation:

$$\begin{aligned}
\text{(F3)} \quad \min \quad & \sum_{i=1}^{M} \sum_{q=1}^{M-1} 2d_{iq} + \sum_{q=1}^{M-1} 2qt_q - (M-1) \sum_{i=1}^{M} z_i, \\
\text{s.t.} \quad & z_i + \sum_{\ell : O_{i\ell} \leqslant k-1} (k - O_{i\ell}) y_\ell \geqslant k, \ \forall i \in A, \ k = 1, \ldots, M-p+1, \\
& z_i + (M-p+1-O_{ik}) y_k \leqslant M-p+1, \quad \forall i, k \in A : O_{ik} \leqslant M-p+1, \\
& \sum_{j=1}^{M} y_j = p, \\
& d_{iq} \geqslant -t_q + z_i, \quad \forall q = 1, \ldots, M-1, \ i \in A, \qquad (12) \\
& d_{iq} \geqslant 0, \quad \forall q = 1, \ldots, M-1, \ i \in A, \\
& y_j \in \{0, 1\}, \quad \forall j \in A.
\end{aligned}$$

**Example 5.** We continue using Example 2, where $z = (4, 1, 2, 2, 1)$. The constraints (12), for $q = 3$, are

$$d_{13} \geqslant 4 - t_3, \ d_{23} \geqslant 1 - t_3, \ d_{33} \geqslant 2 - t_3, \ d_{43} \geqslant 2 - t_3, \ d_{53} \geqslant 1 - t_3.$$

The addends of the objective function corresponding with these variables are

$$\sum_{i=1}^{5} 2d_{i3} + 6t_3.$$

This part of the objective function reaches its minimum when $t_3 \in [1, 2]$. In this region, $1 - t_3 \leqslant 0$ (and then $d_{23} = d_{53} = 0$), but $4 - t_3 > 0$ and $2 - t_3 \geqslant 0$, and then $d_{13} = 4 - t_3$, $d_{33} = d_{43} = 2 - t_3$, giving a partial objective function of $2(4 - t_3 + 2 - t_3 + 2 - t_3) + 6t_3$ with slope 0. The optimal value of this part of the objective is 16. Solving, in similar way, the cases $q = 1, 2, 4$, the values to be added to the objective function are 8, 12 and 18, respectively. The final value is $8 + 12 + 16 + 18 - (5 - 1) \cdot 10 = 14$.

A similar formulation has been obtained by combining Ogryczak and Tamir's idea with the second formulation (i.e., using $z_{ik}$-variables instead of $z_i$-variables). In this case, the results obtained were not so satisfactory and for this reason we have omitted this formulation.

## 6. Comparing formulations

Before trying to improve the performance of the previously studied formulations, we will compare them by means of a simple computational study. The formulations were implemented, as they have been presented in the previous sections, in the commercial solver Xpress IVE 1.17.12, running on a 2.40 GHz PC with 2.00 GB of RAM memory. The cut generation option of Xpress was disabled in order to compare the relative performance of the formulations cleanly. The reader should note that by means of a good heuristic algorithm the running times of all the formulations and solutions methods presented throughout the paper could be reduced. Moreover, the step of the branching process in which a good heuristic solution is produced by Xpress can produce small perturbations in the comparison between formulations.

In order to produce a set of test instances, we generated the preferences matrices in three different ways. First we generated points in the plane at random, and considered the closest point, the most preferred (so every point is its own favorite), getting the results of Table 1. The second way was like the first one, but considering that every point was the least preferred by itself, producing Table 2. For the third set of instances, we generated the rows of the preferences matrix as random permutations of $A$, and the results are given in Table 3. We tested the three formulations on a testbed of five instances for each combination of (i) preferences, (ii) $M$ in $\{20, 30, 40\}$ and (iii) different values of $p$ depending on the case. In Tables 1–3, $LP$ is the average time in seconds needed to solve the linear relaxation, $\bar{t}$ is the average time in seconds of the overall solution process and $n$ is the average number of nodes of the branching tree. The time limit was fixed to one hour of CPU (in the tables, $>1H$ indicates this time has been exceeded, at least, in one of the instances). Throughout the paper, the minimum average time and the minimum average number of nodes have been boldfaced in the tables for each set of instances.

We can see that the time needed to solve the linear relaxations is fairly low. However, in all the instances studied the linear relaxation optimal value was 0. Observe that the three choices for the preferences have a different impact on the running times. When the value of $p$ is small, the times increase from Tables 1 to 3, but these times decrease for medium and large values of $p$. However, when comparing the efficiency of the formulations, the results were independent of the data generator.

The computational times largely depend on the size of $p$. Indeed, for the small values of $p$, (F2) is the formulation which reports largest times and (F1) is the most efficient formulation. Notice that the large figures associated to (F2) when $M = 40$ and $p = 2, 4$ are not due to any outlier, but the results are similar for the five instances of each size. However, for medium and large values of $p$, (F2) drastically changes and becomes the most efficient formulation.

For a more comprehensive computational analysis the reader is referred to [29], where alternative formulations, which did not provide better running times, are also studied.

## 7. Improving the first formulation

The lower bound on the optimal value of (F1) obtained by relaxing the integrality constraints (5) (LP bound) is usually equal to zero. MELP is particularly difficult in this aspect. The LP bound given by different formulations is zero even after adding some valid inequalities. Nevertheless, the effectiveness of the inequalities must

**Table 1**
Customers prefer closer sites and self-service is allowed.

| | $M=20$ | | | | $M=30$ | | | | $M=40$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ |
| (F1) | 2 | 0.0 | **1.2** | **144.2** | 3 | 0.1 | **21.1** | **641.4** | 2 | 0.2 | **47.0** | **532.6** |
| (F2) | | 0.1 | 10.8 | 420.2 | | 0.6 | 63.4 | 948.2 | | 2.3 | 2379.9 | 14480.0 |
| (F3) | | 0.0 | 1.5 | 177.0 | | 0.1 | 22.7 | 718.2 | | 0.5 | 50.7 | 1077.0 |
| (F1) | 3 | 0.0 | **2.4** | 349.4 | 6 | 0.1 | 83.0 | 5006.6 | 4 | 0.2 | 281.5 | 3955.4 |
| (F2) | | 0.1 | 7.3 | **317.8** | | 0.4 | **4.5** | **113.4** | | 1.9 | 265.8 | **2142.6** |
| (F3) | | 0.0 | 2.5 | 327.8 | | 0.1 | 81.9 | 4357.8 | | 0.2 | **263.6** | 3373.8 |
| (F1) | 5 | 0.0 | 3.9 | 847.8 | 10 | 0.0 | 177.0 | 19892.2 | 8 | 0.2 | 2269.1 | 55702.6 |
| (F2) | | 0.1 | **2.6** | **63.0** | | 0.3 | **1.3** | **35.4** | | 1.3 | **16.6** | **207.8** |
| (F3) | | 0.0 | 4.4 | 917.0 | | 0.1 | 174.5 | 17987.0 | | 0.2 | 2443.9 | 57277.8 |
| (F1) | 7 | 0.0 | 4.0 | 1153.0 | 12 | 0.0 | 219.7 | 30517.8 | 10 | 0.1 | $>1H$ | – |
| (F2) | | 0.1 | **1.4** | **27.4** | | 0.3 | **0.7** | **7.8** | | 1.2 | **13.5** | **80.6** |
| (F3) | | 0.0 | 4.4 | 1255.8 | | 0.1 | 425.6 | 51375.0 | | 0.2 | $>1H$ | – |
| (F1) | 10 | 0.0 | 8.5 | 5382.6 | 15 | 0.0 | 680.2 | 205131.4 | 16 | 0.1 | 3097.7 | 117527.7 |
| (F2) | | 0.0 | **0.2** | **3.0** | | 0.2 | **1.9** | **13.0** | | 0.8 | **2.0** | **12.2** |
| (F3) | | 0.0 | 9.5 | 5579.8 | | 0.1 | 996.8 | 249648.0 | | 0.2 | $>1H$ | – |
| (F1) | 12 | 0.0 | 16.5 | 16316.2 | 22 | 0.0 | 2453.0 | $2\times10^6$ | 20 | 0.1 | $>1H$ | – |
| (F2) | | 0.0 | **0.1** | **1.0** | | 0.0 | **0.2** | **1.0** | | 0.6 | **3.8** | **1.8** |
| (F3) | | 0.0 | 16.4 | 14263.0 | | 0.1 | 3298.1 | $2\times10^6$ | | 0.2 | $>1H$ | – |

**Table 2**
Customers prefer closer sites but self-service is forbidden.

| | $M=20$ | | | | $M=30$ | | | | $M=40$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ |
| (F1) | 2 | 0.0 | **1.1** | **120.6** | 3 | 0.1 | **23.9** | **799.0** | 2 | 0.2 | **44.9** | **437.0** |
| (F2) | | 0.1 | 28.2 | 2117.8 | | 0.5 | 814.2 | 21899.4 | | 2.0 | 2717.5 | 17868.5 |
| (F3) | | 0.0 | 1.4 | 167.8 | | 0.1 | 26.5 | 861.9 | | 0.3 | 53.1 | 482.2 |
| (F1) | 3 | 0.0 | **0.4** | **129.4** | 6 | 0.1 | 52.1 | **2672.6** | 4 | 0.2 | **312.0** | 4173.4 |
| (F2) | | 0.1 | 16.6 | 1791.0 | | 0.4 | **52.0** | 2749.8 | | 1.8 | $>1H$ | – |
| (F3) | | 0.0 | 2.5 | 315.4 | | 0.1 | 66.9 | 3258.6 | | 0.2 | 326.3 | **4029.4** |
| (F1) | 5 | 0.0 | 2.9 | 584.6 | 10 | 0.1 | 87.7 | 7987.8 | 8 | 0.2 | 2000.2 | 43647.4 |
| (F2) | | 0.1 | **2.5** | **364.6** | | 0.4 | **4.5** | **420.4** | | 1.4 | **846.7** | **28928.2** |
| (F3) | | 0.1 | 3.2 | 597.0 | | 0.1 | 82.3 | 6807.0 | | 0.2 | 2105.2 | 38412.2 |
| (F1) | 7 | 0.0 | 3.0 | 722.2 | 12 | 0.0 | 49.6 | 5036.2 | 10 | 0.2 | 2624.0 | 62600.0 |
| (F2) | | 0.1 | **1.0** | **101.4** | | 0.3 | **1.7** | **97.4** | | 1.3 | **139.6** | **4933.4** |
| (F3) | | 0.0 | 2.5 | 595.4 | | 0.1 | 48.8 | 4261.0 | | 0.2 | 2763.5 | 67043.0 |
| (F1) | 10 | 0.0 | 1.0 | 315.4 | 15 | 0.0 | 16.1 | 2396.6 | 16 | 0.1 | 1555.1 | 41557.0 |
| (F2) | | 0.1 | **0.4** | **28.6** | | 0.2 | **2.2** | **21.8** | | 0.8 | **4.2** | **136.0** |
| (F3) | | 0.0 | 1.3 | 411.0 | | 0.1 | 15.8 | 2054.0 | | 0.2 | 1677.0 | 43331.5 |
| (F1) | 12 | 0.0 | **0.4** | 129.4 | 22 | 0.0 | **0.1** | 2.6 | 20 | 0.1 | 315.3 | 19865.8 |
| (F2) | | 0.0 | **0.4** | **10.6** | | 0.1 | 0.2 | **1.8** | | 0.6 | **2.0** | **53.8** |
| (F3) | | 0.0 | **0.4** | 107.4 | | 0.0 | 0.2 | 3.0 | | 0.2 | 127.5 | 6387.0 |

be tested in the overall process, since the linear relaxation of a subproblem obtained during the branch-and-bound solution procedure after fixing some $y$-variables to 0 or 1 can be improved with these inequalities.

In order to solve the MELP using the first formulation, we tightened constraints (4). These constraints, for a given $i \in A$, force $z_i$ to take a value of at most $O_{ik}$ if $y_k = 1$. If, for some other $\ell \in A$ with $O_{i\ell} < O_{ik}$, the variable $y_\ell$ is also equal to one, the value of $z_i$ will be at most $O_{i\ell}$. This can be written as $z_i + (M-p+1-O_{ik})y_k + (O_{ik} - O_{i\ell})y_\ell \leqslant M-p+1$, an enforcement of (4). Additional $y$-variables corresponding with sites which are most preferred can be incorporated into the constraint, leading to the general case

$$z_i + \sum_{\ell=1}^{q-1}(O_{ik_{\ell+1}} - O_{ik_\ell})y_{k_\ell} + (M-p+1-O_{ik_q})y_{k_q} \leqslant M-p+1 \quad (13)$$

$\forall i \in A, S := \{k_1, \ldots, k_q\} \subset A : O_{ik_1} < \cdots < O_{ik_q}$.

There is an exponential number of valid inequalities in family (13). We implemented a branch-and-cut algorithm designed as follows. In (F1), we replaced constraints (4) for all $i \in A$ and $k \in A$ such that $O_{ik} \leqslant M-p+1$, by the tighter constraint

$$z_i + \sum_{\ell : O_{i\ell} < O_{ik}} y_\ell + (M-p+1-O_{ik})y_k \leqslant M-p+1,$$

obtained from (13) when $k_q$ is replaced by $k$ and $S = \{\ell \in A : O_{i\ell} < O_{ik}\}$. Then, at every node of the branching tree and for every value of $i$ we proceeded as follows in order to build the most violated constraint of type (13) for the given index $i$ and, if the constraint was violated by the fractional optimal solution, add it to the formulation.

- Let $y^*$ be the optimal solution to the linear relaxation of the problem and let $y_t^*$ be the largest value in $y^*$. Add variable $y_t$ to the constraint with coefficient $M-p+1-O_{it}$.
- Repeat.
  Let $y_s^*$ be the largest value in $y^*$ such that $O_{is} < O_{it}$. If such a value does not exist, stop.

**Table 3**
Random preferences.

| | $M=20$ | | | | $M=30$ | | | | $M=40$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ |
| (F1) | 2 | 0.0 | **1.6** | **201.8** | 3 | 0.1 | **28.1** | **869.8** | 2 | 0.3 | **90.2** | **1082.6** |
| (F2) | | 0.1 | 62.7 | 4650.6 | | 0.6 | > 1*H* | – | | 2.1 | > 1*H* | – |
| (F3) | | 0.0 | 2.0 | 242.6 | | 0.1 | 34.0 | 1033.4 | | 0.0 | 96.6 | 1150.2 |
| (F1) | 3 | 0.0 | **2.2** | **273.0** | 6 | 0.1 | 66.4 | 2903.0 | 4 | 0.2 | **459.6** | **5656.6** |
| (F2) | | 0.1 | 15.5 | 1262.6 | | 0.4 | 113.5 | 5463.8 | | 2.0 | > 1*H* | – |
| (F3) | | 0.0 | 2.5 | 293.4 | | 0.1 | **63.2** | **2436.6** | | 0.3 | 507.0 | 5660.2 |
| (F1) | 5 | 0.0 | 2.5 | 389.8 | 10 | 0.1 | 32.0 | 2085.8 | 8 | 0.2 | 977.1 | 15782.6 |
| (F2) | | 0.1 | 2.4 | **329.4** | | 0.3 | **2.7** | **137.8** | | 1.3 | **212.8** | **4813.8** |
| (F3) | | 0.0 | **2.3** | 346.2 | | 0.1 | 27.8 | 1443.0 | | 0.2 | 862.9 | 12009.4 |
| (F1) | 7 | 0.0 | 1.4 | 270.2 | 12 | 0.0 | 18.6 | 1384.6 | 10 | 0.2 | 985.2 | 19158.2 |
| (F2) | | 0.1 | **1.1** | **44.6** | | 0.3 | **1.1** | **39.4** | | 1.2 | **28.4** | **645.4** |
| (F3) | | 0.0 | 1.7 | 329.4 | | 0.1 | 20.5 | 1495.4 | | 0.2 | 651.4 | 11759.0 |
| (F1) | 10 | 0.0 | **0.5** | 145.0 | 15 | 0.0 | 4.7 | 50.0 | 16 | 0.1 | 280.2 | 9982.2 |
| (F2) | | 0.0 | **0.5** | **16.2** | | 0.2 | **0.8** | **19.0** | | 0.8 | **2.7** | **31.8** |
| (F3) | | 0.0 | 0.6 | 166.2 | | 0.1 | 6.4 | 648.0 | | 0.2 | 336.1 | 12150.0 |
| (F1) | 12 | 0.0 | **0.2** | 33.0 | 22 | 0.0 | **0.0** | 2.0 | 20 | 0.0 | 75.3 | 3708.2 |
| (F2) | | 0.0 | 0.3 | **6.2** | | 0.1 | 0.1 | **1.0** | | 0.6 | **2.1** | **39.4** |
| (F3) | | 0.0 | **0.2** | 51.0 | | 0.0 | 0.2 | 3.0 | | 0.2 | 71.1 | 3689.0 |

Add variable $y_s$ to the constraint with coefficient $O_{it} - O_{is}$. Let $t := s$.

Table 4 reports a comparative analysis of the results provided when solving formulation (F1) with and without the previous separation method (lines (F1R) and (F1), respectively). In general, the inequalities reduce significantly the number of nodes of the branching tree but the total time needed to solve the problems is higher. In the remaining tables we have written the computational times inside a box whether they improved the previous times.

Another family of valid inequalities we produced in order to avoid the lower bound taking value 0 is the following:

$$\sum_{i \in S} \sum_{j \in S: j > i} e_{ij} \geqslant p y_a, \tag{14}$$

where $S \subseteq A : |S| = p + 1$ is a set of clients whose favorite sites are all different and $a$ is the favorite site of one of the clients in set $S$. Then, if $y_a = 1$, the favorite site of one client in $S$, say $i_a$, contains a plant. Since there are $p$ plants and the favorite sites of the clients in $S$ are different, there will be another client, say $i_b$, whose favorite site does not contain a plant. Those other clients in $S$ whose favorite site does not contain a plant envy client $i_a$, whereas those whose favorite site contains a plant are envied by $i_b$, and the total envy is, at least, $p$. Again, the number of inequalities in family (14) is exponential. Several attempts to use these constraints to speed up the resolution of MELP with formulation (F1) led to better lower bounds but at the expense of much larger computational times.

## 8. Improving the second formulation

In order to solve the MELP using the second formulation, we tested several families of valid inequalities and two variable fixing strategies. This is going to be a successful approach which, for most of the instances, will give the best computational times.

### 8.1. Valid inequalities

We present the following groups of inequalities:

$$x_{ik} \geqslant x_{i,k+1}, \quad \forall i \in A, \ k = 2, \dots, M - p, \tag{15}$$

$$p z_{ik} \leqslant \sum_{j: O_{ij} \geqslant k} y_j, \quad \forall i \in A, \ k = 2, \dots, M - p, \tag{16}$$

$$(p-1)(z_{ik} - z_{i,k+1}) \leqslant \sum_{j: O_{ij} \geqslant k+1} y_j, \quad \forall i \in A, \ k = 2, \dots, M - p, \tag{17}$$

$$\sum_{i=1}^{s} x_{M+1-i,k} \geqslant \sum_{i \in S} z_{ik}, \quad \forall s \in A, S \subset A : |S| = s, \ k = 2, \dots, M - p + 1, \tag{18}$$

$$z_{ik} + \sum_{j: O_{ij} < k} y_j \geqslant 1, \quad \forall i \in A, \ k = 2, \dots, M - p + 1. \tag{19}$$

Inequalities (15) are natural, since the rows of the $x$-matrix match the rows of the $z$-matrix, and the rows of the latter matrix are sorted. Nevertheless, these constraints are not always satisfied by the optimal solution of the linear relaxation of (F2), and these can be used as valid inequalities for (F2). This family of inequalities were considered in [27] to tighten the formulation of the discrete ordered median problem.

Inequalities (16) are an adaptation of a set of constraints introduced in [30] for the obnoxious $p$-median problem, whereas (17) were an improvement of (16) developed in [31] for the simple plant location problem with order. The meaning of (16) and (17) is the following. Assume $z_{ik} = 1$ for some $i \in A$ and some $k \in \{2, \dots, M - p\}$. Then, there cannot be any plant in the $k - 1$ favourite sites of client $i$. Consequently, the $p$ plants must be in the rest of the sites, and it follows $p \leqslant \sum_{j: O_{ij} \geqslant k} y_j$, supporting (16). Now, assume $z_{ik} - z_{i,k+1} = 1$ for some $i \in A$ and some $k \in \{2, \dots, M - p\}$. Then, there cannot be any plant in the $k - 1$ favourite sites of client $i$ and there is a plant in its $k$-th favourite site. Consequently, the $p - 1$ remaining plants must be in the rest of the sites, and $p - 1 \leqslant \sum_{j: O_{ij} \geqslant k+1} y_j$ follows, supporting (17).

Inequalities (18) were also used in [27]. Columns of the $x$-matrix contain as many 1's as their corresponding columns in the $z$-matrix, but the ones of the $x$-matrix are at the bottom of the columns. Consequently, if there are $s$ ones in some part of the $z$-column, the sum of the $s$ last values of the $x$-column is at least equal to $s$, leading to (18). Note that there is an exponential number of constraints in this family.

Finally, inequalities (19) say in a different way that $z_{ik}$ must be equal to one if there are no plants in the $k - 1$ sites preferred by

**Table 4**
Comparison of (F1) with and without additional constraints.

| | $M=20$ | | | | $M=30$ | | | | $M=40$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ |
| *Customers prefer closer sites and self-service is allowed* | | | | | | | | | | | | |
| (F1) | 2 | 0.0 | $\boxed{1.2}$ | **144.2** | 3 | 0.1 | $\boxed{21.1}$ | 641.4 | 2 | 0.2 | $\boxed{47.0}$ | 532.6 |
| (F1R) | | 0.0 | 1.7 | 155.0 | | 0.1 | 24.8 | **557.0** | | 0.2 | 74.4 | **496.6** |
| (F1) | 3 | 0.0 | **2.4** | 349.4 | 6 | 0.1 | **83.0** | 5006.6 | 4 | 0.2 | **281.5** | 3955.4 |
| (F1R) | | 0.0 | 2.8 | **285.0** | | 0.1 | 95.7 | **3910.2** | | 0.2 | 359.1 | **3355.8** |
| (F1) | 5 | 0.0 | **3.9** | 847.8 | 10 | 0.0 | **177.0** | 19892.2 | 8 | 0.2 | **2269.1** | 55702.6 |
| (F1R) | | 0.0 | 4.5 | **622.6** | | 0.1 | 180.1 | **11339.8** | | 0.1 | 2655.5 | **42317.8** |
| (F1) | 7 | 0.0 | **4.0** | 1153.0 | 12 | 0.0 | 219.7 | 30517.8 | 10 | 0.1 | $>1H$ | – |
| (F1R) | | 0.0 | 4.9 | **786.2** | | 0.0 | **188.5** | **13592.6** | | 0.1 | $>1H$ | – |
| (F1) | 10 | 0.0 | **8.5** | 5382.6 | 15 | 0.0 | 680.2 | 205131.4 | 16 | 0.1 | **3097.7** | **117527.7** |
| (F1R) | | 0.0 | 12.9 | **4125.8** | | 0.0 | **635.6** | **116827.0** | | 0.1 | $>1H$ | – |
| (F1) | 12 | 0.0 | **16.5** | 16316.2 | 22 | 0.0 | **2453.0** | $2\times10^6$ | 20 | 0.1 | $>1H$ | – |
| (F1R) | | 0.0 | 25.3 | **14071.8** | | 0.0 | $>1H$ | – | | 0.1 | $>1H$ | – |
| *Customers prefer closer sites but self-service is forbidden* | | | | | | | | | | | | |
| (F1) | 2 | 0.0 | $\boxed{1.1}$ | **120.6** | 3 | 0.1 | $\boxed{23.9}$ | 799.0 | 2 | 0.2 | $\boxed{44.9}$ | 437.0 |
| (F1R) | | 0.0 | 1.6 | **120.6** | | 0.1 | 30.4 | **655.8** | | 0.2 | 64.9 | **353.8** |
| (F1) | 3 | 0.0 | $\boxed{0.4}$ | **129.4** | 6 | 0.1 | **52.1** | 2672.6 | 4 | 0.2 | $\boxed{312.0}$ | 4173.4 |
| (F1R) | | 0.0 | 2.7 | 245.8 | | 0.1 | 75.6 | **2389.0** | | 0.2 | 416.1 | **3450.2** |
| (F1) | 5 | 0.0 | **2.9** | 584.6 | 10 | 0.1 | 87.7 | 7987.8 | 8 | 0.2 | **2000.2** | 43647.4 |
| (F1R) | | 0.0 | 4.1 | **499.8** | | 0.1 | **87.5** | **4195.0** | | 0.2 | 2405.5 | **29669.0** |
| (F1) | 7 | 0.0 | **3.0** | 722.2 | 12 | 0.0 | **49.6** | 5036.2 | 10 | 0.2 | **2624.0** | 62600.0 |
| (F1R) | | 0.0 | 3.8 | **519.0** | | 0.1 | 61.4 | **2890.2** | | 0.1 | 3020.4 | **35502.0** |
| (F1) | 10 | 0.0 | **1.0** | **315.4** | 15 | 0.0 | **16.1** | 2396.6 | 16 | 0.1 | 1555.1 | 41557.0 |
| (F1R) | | 0.0 | 2.1 | 337.8 | | 0.1 | 23.4 | **1117.8** | | 0.2 | **1261.4** | **24426.3** |
| (F1) | 12 | 0.0 | $\boxed{0.4}$ | 129.4 | 22 | 0.0 | $\boxed{0.1}$ | **2.6** | 20 | 0.1 | **315.3** | 19865.8 |
| (F1R) | | 0.0 | 0.7 | **90.6** | | 0.0 | 0.7 | **2.6** | | 0.1 | 523.3 | **11281.8** |
| *Random preferences* | | | | | | | | | | | | |
| (F1) | 2 | 0.0 | $\boxed{1.6}$ | 201.8 | 3 | 0.1 | $\boxed{28.1}$ | 869.8 | 2 | 0.3 | $\boxed{90.2}$ | 1082.6 |
| (F1R) | | 0.0 | 2.1 | **183.4** | | 0.1 | 40.4 | **728.6** | | 0.2 | 147.9 | **1052.3** |
| (F1) | 3 | 0.0 | $\boxed{2.2}$ | 273.0 | 6 | 0.1 | **66.4** | 2903.0 | 4 | 0.2 | $\boxed{459.6}$ | 5656.6 |
| (F1R) | | 0.0 | 2.8 | **218.2** | | 0.1 | 84.4 | **2320.6** | | 0.2 | 767.3 | **5431.0** |
| (F1) | 5 | 0.0 | **2.5** | 389.8 | 10 | 0.1 | **32.0** | 2085.8 | 8 | 0.2 | **977.1** | 15782.6 |
| (F1R) | | 0.0 | 3.5 | **381.0** | | 0.1 | 43.2 | **1470.2** | | 0.3 | 1033.7 | **9091.4** |
| (F1) | 7 | 0.0 | **1.4** | 270.2 | 12 | 0.0 | **18.6** | 1384.6 | 10 | 0.2 | **985.2** | 19158.2 |
| (F1R) | | 0.0 | 2.2 | **248.6** | | 0.1 | 30.4 | **1037.0** | | 0.2 | 1118.1 | **14588.3** |
| (F1) | 10 | 0.0 | **0.5** | 145.0 | 15 | 0.0 | **4.7** | **50.0** | 16 | 0.1 | **280.2** | 9982.2 |
| (F1R) | | 0.0 | 1.0 | **128.2** | | 0.1 | 8.7 | 372.6 | | 0.2 | 390.0 | **4941.0** |
| (F1) | 12 | 0.0 | $\boxed{0.2}$ | 33.0 | 22 | 0.0 | $\boxed{0.0}$ | **2.0** | 20 | 0.0 | **75.3** | 3708.2 |
| (F1R) | | 0.0 | 0.3 | **32.2** | | 0.0 | 0.2 | 3.8 | | 0.2 | 124.6 | **2098.2** |

client $i$. These inequalities are set covering constraints, and following the ideas of [32], a paper in which the facets with coefficients in $\{0,1,2\}$ for a set covering problem were characterized, we get the following new set of valid inequalities for (F2):

$$\sum_{(i,k)\in Q} z_{ik} + \sum_{j\in\bigcup_{\ell=1}^{q-1} J_\ell} y_j + \sum_{j\in J_q} 2y_j \geqslant 2, \tag{20}$$

for all $i\in A$, $q=3,\dots,M(M-p+1)$, $Q\subset A\times\{2,\dots,M-p+1\}:|Q|=q$, where

$$J_\ell := \{j\in A : |\{(i,k)\in Q : O_{ij}<k\}| = \ell\}.$$

These constraints are obtained by choosing any set $Q$ of different pairs $(i,k)$, adding the corresponding constraints (19) up, replacing by 2 all the coefficients equal to $|Q|$ and the right-hand side and replacing the remaining coefficients by 1. Therefore (i) if all $y$-variables in the constraint take value zero, clearly there will be at least 2 $z$-variables taking value 1, (ii) if only one $y$-variable in the second addend (i.e., which is not shared by all the added constraints (19)) takes value one, there will at least one $z$-variable, coming from a constraint (19) with all the $y$-variables equal to zero, which will take value 1.

As seen, each of these families were (successfully, in general) used for solving some kind of discrete optimization problem. We considered each family of valid inequalities separately and in combination with the others in order to test their usefulness in our own context.

Nevertheless, the LP bound given by (F2) strengthened with the constraints is not going to be better, as can be seen in the following result.

**Proposition 1.** *The optimal value of the linear relaxation of* (F2), *even after adding constraints* (15)–(20), *is* $z^2_{LP}=0$.

**Proof.** $z^2_{LP}\geqslant 0$, since the objective function is $\sum_{i_1}\sum_{i_2>i_1}\sum_{k=2}^{M-p+1}(x_{i_2k}-x_{i_1k})$ and (F2) includes the constraints $x_{ik}\geqslant x_{i-1,k}$.

On the other hand, $z^2_{LP}\leqslant 0$, since the fractional solution given by

$$y_j = \frac{p}{M}, \quad \forall j\in A, \tag{21}$$

**Table 5**
Performance of several families of valid inequalities added to (F2).

| Ineq. | $\bar{t}$ | $n$ | $\bar{t}$ | $n$ | $\bar{t}$ | $n$ | $\bar{t}$ | $n$ | $\bar{t}$ | $n$ | $\bar{t}$ | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $M=20$ | | | | | | | | | | | | |
| | $p=2$ | | $p=3$ | | $p=5$ | | $p=7$ | | $p=10$ | | $p=12$ | |
| None | 10.8 | 420.2 | 7.3 | 317.8 | 2.6 | 63.0 | 1.4 | 27.4 | 0.2 | 3.0 | $\boxed{0.1}$ | **1.0** |
| (15) | **4.6** | 96.2 | 2.9 | 71.0 | 1.2 | 42.2 | $\boxed{0.4}$ | 10.0 | 0.2 | 5.0 | **0.1** | **1.0** |
| (17) | 14.1 | 415.0 | 9.5 | 370.6 | 3.1 | 68.6 | 1.7 | 15.4 | $\boxed{0.1}$ | **1.0** | 0.2 | 4.2 |
| (15)(16) | 4.7 | 89.8 | **2.6** | 81.8 | 1.5 | 31.0 | 0.7 | 9.4 | 0.3 | 5.0 | **0.1** | **1.0** |
| (15)(17) | 5.5 | **71.4** | 3.1 | 91.8 | 1.6 | 29.4 | 0.9 | 8.2 | 0.6 | 10.6 | **0.1** | **1.0** |
| (15)(19) | 5.1 | 101.0 | 2.9 | **69.8** | 1.1 | **23.4** | 1.1 | 14.2 | 0.4 | 8.6 | **0.1** | **1.0** |
| (15)(16)(17) | 5.0 | 111.4 | 3.2 | 75.8 | 1.0 | **23.4** | 0.9 | 11.8 | **0.1** | **1.0** | **0.1** | **1.0** |
| (15)(16)(19) | 6.1 | 99.0 | 3.6 | 75.0 | $\boxed{0.9}$ | 34.2 | 0.7 | **6.2** | 0.3 | 4.2 | **0.1** | **1.0** |
| $M=30$ | | | | | | | | | | | | |
| | $p=3$ | | $p=6$ | | $p=10$ | | $p=12$ | | $p=15$ | | $p=22$ | |
| None | 63.4 | 948.2 | 4.5 | 113.4 | 1.3 | 35.4 | $\boxed{0.7}$ | 7.8 | 1.9 | 13.0 | $\boxed{0.2}$ | **1.0** |
| (15) | $\boxed{10.0}$ | 159.0 | $\boxed{3.0}$ | **8.7** | $\boxed{1.1}$ | **15.0** | **0.7** | 2.3 | 1.4 | 11.4 | **0.2** | **1.0** |
| (16) | 55.8 | 630.6 | 8.2 | 24.1 | 1.4 | 25.4 | 0.9 | 10.8 | $\boxed{0.3}$ | **1.0** | **0.2** | **1.0** |
| (16)(19) | 75.7 | 675.0 | 8.7 | 48.3 | 2.3 | 30.6 | 1.2 | **1.7** | **0.3** | **1.0** | 0.3 | **1.0** |
| (15)(17)(19) | 15.6 | **105.4** | 4.7 | 38.1 | 1.7 | 15.8 | 1.2 | 4.3 | 0.8 | 5.0 | 0.3 | **1.0** |
| (16)(17)(19) | 97.9 | 777.4 | 12.0 | 34.4 | 2.1 | 22.6 | 1.2 | **1.7** | 0.4 | **1.0** | **0.2** | **1.0** |
| $M=40$ | | | | | | | | | | | | |
| | $p=2$ | | $p=4$ | | $p=8$ | | $p=10$ | | $p=16$ | | $p=20$ | |
| None | 2379.9 | 14480.0 | 265.8 | 2142.6 | 16.6 | 207.8 | 13.5 | 80.6 | 2.0 | 12.2 | 3.8 | 1.8 |
| (15) | 176.9 | 838.2 | $\boxed{51.6}$ | 413.8 | $\boxed{13.1}$ | 143.8 | 7.2 | 86.5 | 2.3 | 18.5 | 1.0 | **1.0** |
| (16) | 2106.9 | 7885.4 | 283.1 | 1449.0 | 32.6 | 257.8 | 9.0 | 93.5 | 2.0 | 7.0 | $\boxed{0.6}$ | **1.0** |
| (19) | $>1H$ | – | 338.4 | 1865.8 | 18.8 | 159.4 | 10.6 | 108.0 | $\boxed{1.7}$ | 7.5 | $\boxed{0.6}$ | **1.0** |
| (15)(17) | **165.4** | 562.2 | 58.8 | 374.2 | 25.6 | 110.2 | $\boxed{7.0}$ | 66.5 | 2.1 | **4.0** | 1.0 | **1.0** |
| (15)(19) | 319.1 | 642.6 | 66.4 | 423.8 | 26.0 | 120.2 | $\boxed{7.0}$ | 73.0 | 2.2 | 8.5 | $\boxed{0.6}$ | **1.0** |
| (15)(16)(17) | 219.4 | **532.6** | 78.4 | **333.8** | 46.6 | 155.0 | 18.5 | 72.5 | 9.2 | 13.5 | 1.5 | **1.0** |
| (15)(16)(19) | 230.2 | 579.4 | 115.7 | 509.8 | 34.3 | **87.0** | 23.2 | 85.0 | 6.3 | 4.5 | 1.6 | **1.0** |
| (15)(16)(17)(19) | 353.3 | 718.0 | 155.2 | 473.4 | 73.6 | 119.0 | 29.7 | **72.0** | 8.6 | **4.0** | 5.8 | **1.0** |

$$z_{ik} = x_{ik} = \begin{cases} \dfrac{M-p}{M} & \text{if } 2 \leqslant k \leqslant p+1, \\ \dfrac{M-k+1}{M} & \text{if } k \geqslant p+2, \end{cases} \quad \forall i \in A, \tag{22}$$

with objective value 0, satisfies all the constraints. It can be easily checked in constraints (15)–(19). Then, let $a$ be the number of different $y$-variables in an inequality of the family (20) and let $q \geqslant 3$ be the cardinality of $Q$. Thus, if we denote $K_{\max} = \max\{k : (i,k) \in Q\}$ we have that $K_{\max} \leqslant a+1$. Hence, since the values of $z_{ik}$ decrease when $k$ increases, the value of $z_{ik}$ defined in (22) is, at least, equal to $(M-(a+1)+1)/M$ (the number of $y_j$ satisfying $O_{ij} < k$ is $k-1$). On the other hand, there are $a$ $y$-variables in (20) with coefficient 1 or 2, and the value of $y_j$ defined in (21) is $p/M$, thus the left-hand side minus the right-hand side of (20) after replacing (21) and (22) will be greater than or equal to $\frac{q(M-(a+1)+1)}{M} + \frac{ap}{M} - 2$. Moreover, since $a \leqslant M$ and $q > 2$:

$$\frac{q(M-(a+1)+1)}{M} + \frac{ap}{M} - 2 = \frac{(q-2)M + a(p-q)}{M} \geqslant \frac{(q-2)a + a(p-q)}{M} = \frac{a(p-2)}{M}.$$

When $p \geqslant 2$, the last term is non-negative, and (20) is satisfied.

Consider now the case $p = 1$, i.e,

$$y_j = \frac{1}{M}, \quad \forall j \in A, \tag{23}$$

$$z_{ik} = \frac{M-k+1}{M}, \quad \forall i \in A, \quad k \geqslant 2. \tag{24}$$

In order to prove that this solution satisfies constraint (20), we will analyze the incremental contribution of the elements in $Q$ to the value of the variables $z$ and $y$ in the left-hand-side of this constraint. Consider $Q = \{(i_t, k_t)$ with $t = 1, \ldots, q\}$; the first element $(i_1, k_1)$, after replacing (23) and (24), contributes with $(M-k_1+1)/M + (k_1-1)(1/M) = 1$ and the second one, $(i_2, k_2)$, with $(M-k_2+1)/M + (k_2-1)(1/M) = 1$. Thus, the overall contribution of $(i_1, k_1)$ and $(i_2, k_2)$ is 2. Every new element of $Q$ that we consider will not decrease the value of this total. To see this, note that this total decreases only when a $y$-variable with coefficient 2 changes to coefficient 1. This change may happen only when there is an element $(i_r, k_r) \in Q$ and columns $j$ such that $O_{i_r j} \geqslant k_r$. Each of these columns $j$ is associated with a variable $y_j$ that will change its coefficient in (20) from 2 to 1. The number of those variables is clearly bounded from above by $M - (k_r - 1)$; and thus the overall decrease is, at most, $(M-(k_r-1))/M$. Nevertheless, this decrease on the $y$-variables is compensated by the value added by $z_{i_r k_r}$ which, according to (24), equals $(M-(k_r-1))/M$. $\square$

Table 5 shows the running times when solving formulation (F2) with several sets of additional valid inequalities. The preferences matrices used here were those based on distances between random points in the plane. In the first column we indicate which subset of valid inequalities is added to (F2). We only considered families with $O(M^2)$ inequalities, i.e., (15)–(17) and (19). Again we solved five instances of each size and showed the average times and number of nodes. Only sets of constraints which give the best results for at least one combination of parameters are shown. Based on the number of nodes of the branching tree and the times obtained in the experiment, we can conclude that the family of valid inequalities whose inclusion in the formulation is well-grounded is (15). However, in some cases the combination of (15) with other inequalities is the best option. With these improvements, the formulation becomes competitive in those instances with $p$ small.

We tested the inequalities in family (20) in several ways, with no success. Regarding the inequalities in family (18), we implemented a separation procedure similar to that developed in [27], getting the results given in Table 6. The idea of this separation method is (i) sorting for each $i$, in non-increasing order, the $z$-values of the optimal solution of the linear relaxation $(\tilde{x}_{ik}, \tilde{z}_{ik})$ into a vector $\tilde{z}_{ik}^{\geqslant}$, and (ii) checking the inequalities

$$\sum_{i=1}^{s} \tilde{x}_{M+1-i,k} \geqslant \sum_{i=1}^{s} \tilde{z}_{ik}^{\geqslant}, \quad \forall s \in A, \quad k = 2, \ldots, M-p+1.$$

**Table 6**
Comparison between (F2) with and without separation of constraints (18).

| | $M=20$ | | | $M=30$ | | | $M=40$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $\bar{t}$ | $n$ | $p$ | $\bar{t}$ | $n$ | $p$ | $\bar{t}$ | $n$ |
| *Customers prefer closer sites and self-service is allowed* | | | | | | | | | |
| (F2) | 2 | 10.7 | 420.2 | 3 | 63.4 | 948.2 | 2 | 379.9 | 14480.0 |
| (F2)(15) | | **4.6** | 96.2 | | $\boxed{10.0}$ | 159.0 | | 176.9 | 838.2 |
| (F2)(15)(18) | | 9.0 | **67.8** | | 16.7 | **110.6** | | **101.6** | **602.2** |
| (F2) | 3 | 7.3 | 317.8 | 6 | 4.5 | 113.4 | 4 | 265.8 | 2142.6 |
| (F2)(15) | | **2.9** | 71.0 | | $\boxed{3.0}$ | **8.7** | | 51.6 | 413.8 |
| (F2)(15)(18) | | 4.5 | **55.8** | | 5.6 | 53.0 | | $\boxed{35.8}$ | **237.0** |
| (F2) | 5 | 2.6 | 63.0 | 10 | 1.3 | 35.4 | 8 | 16.6 | 207.8 |
| (F2)(15) | | **1.2** | 42.2 | | $\boxed{1.1}$ | 15.0 | | $\boxed{13.1}$ | 143.8 |
| (F2)(15)(18) | | 1.5 | **17.4** | | 2.7 | **13.0** | | 19.2 | **119.8** |
| (F2) | 7 | 1.4 | 27.4 | 12 | $\boxed{0.7}$ | 7.8 | 10 | 13.5 | 80.6 |
| (F2)(15) | | $\boxed{0.4}$ | 10.0 | | $\boxed{0.7}$ | **2.3** | | **7.2** | 86.5 |
| (F2)(15)(18) | | 0.8 | **8.6** | | 1.4 | 4.2 | | 10.5 | **56.2** |
| (F2) | 10 | **0.2** | **3.0** | 15 | 1.9 | 13.0 | 16 | **2.0** | **12.2** |
| (F2)(15) | | **0.2** | 5.0 | | 1.4 | 11.4 | | 2.3 | 18.5 |
| (F2)(15)(18) | | 0.6 | 3.8 | | **1.2** | **6.2** | | 5.0 | 15.8 |
| (F2) | 12 | $\boxed{0.1}$ | **1.0** | 22 | $\boxed{0.2}$ | **1.0** | 20 | 3.8 | 1.8 |
| (F2)(15) | | $\boxed{0.1}$ | **1.0** | | $\boxed{0.2}$ | **1.0** | | **1.0** | **1.0** |
| (F2)(15)(18) | | 0.3 | **1.0** | | 0.5 | **1.0** | | 3.1 | 1.4 |
| *Customers prefer closer sites but self-service is forbidden* | | | | | | | | | |
| (F2) | 2 | 28.2 | 2117.8 | 3 | 814.2 | 21899.4 | 2 | 2717.5 | 17868.5 |
| (F2)(15) | | 4.4 | 193.4 | | **28.2** | 825.4 | | 125.2 | 876.6 |
| (F2)(15)(18) | | **4.0** | **140.6** | | 29.9 | **665.4** | | **108.0** | **719.0** |
| (F2) | 3 | 16.6 | 1791.0 | 6 | 52.0 | 2749.8 | 4 | > 1*H* | – |
| (F2)(15) | | 3.8 | 275.4 | | $\boxed{6.7}$ | 304.6 | | $\boxed{114.6}$ | 1815.0 |
| (F2)(15)(18) | | **3.6** | **154.2** | | 9.3 | **187.4** | | 137.8 | **1382.7** |
| (F2) | 5 | $\boxed{2.5}$ | 364.6 | 10 | 4.5 | 420.4 | 8 | 846.7 | 28928.2 |
| (F2)(15) | | 3.8 | 275.4 | | $\boxed{3.1}$ | **171.6** | | $\boxed{23.1}$ | 632.2 |
| (F2)(15)(18) | | 2.9 | **103.8** | | 6.6 | 173.0 | | 88.6 | **479.2** |
| (F2) | 7 | 1.0 | 101.4 | 12 | **1.7** | 97.4 | 10 | 139.6 | 4933.4 |
| (F2)(15) | | $\boxed{0.7}$ | 67.4 | | **1.7** | 72.2 | | 21.7 | 740.6 |
| (F2)(15)(18) | | 1.5 | **45.4** | | 3.7 | **65.0** | | $\boxed{15.7}$ | **231.7** |
| (F2) | 10 | **0.4** | 28.6 | 15 | 2.2 | 21.8 | 16 | **4.2** | 136.0 |
| (F2)(15) | | 1.3 | **17.4** | | 1.5 | 15.0 | | 4.6 | 75.0 |
| (F2)(15)(18) | | 0.8 | 20.2 | | $\boxed{1.4}$ | **13.0** | | 7.4 | **70.6** |
| (F2) | 12 | **0.4** | 10.6 | 22 | 0.2 | 1.8 | 20 | **2.0** | 53.8 |
| (F2)(15) | | 0.6 | 6.6 | | $\boxed{0.1}$ | **1.0** | | 2.6 | 19.0 |
| (F2)(15)(18) | | 0.9 | **4.6** | | 1.3 | **1.0** | | 4.1 | **12.0** |
| *Random preferences* | | | | | | | | | |
| (F2) | 2 | 62.7 | 4650.6 | 3 | > 1*H* | – | 2 | > 1*H* | – |
| (F2)(15) | | **6.2** | 372.6 | | **93.5** | **2489.8** | | > 1*H* | – |
| (F2)(15)(18) | | 6.7 | **245.5** | | 127.3 | 3204.6 | | > 1*H* | – |
| (F2) | 3 | 15.5 | 1262.6 | 6 | 113.5 | 5463.8 | 4 | > 1*H* | – |
| (F2)(15) | | **3.1** | 287.4 | | $\boxed{6.6}$ | 245.4 | | 1342.9 | **16320.0** |
| (F2)(15)(18) | | 5.8 | **179.4** | | 7.8 | **180.2** | | **1253.0** | 16697.7 |
| (F2) | 5 | 2.4 | 329.4 | 10 | 2.7 | 137.8 | 8 | 212.8 | 4813.8 |
| (F2)(15) | | $\boxed{1.2}$ | 83.0 | | $\boxed{2.2}$ | 64.6 | | 32.3 | **182.2** |
| (F2)(15)(18) | | 2.0 | **75.4** | | 2.9 | **28.2** | | $\boxed{21.5}$ | 269.0 |
| (F2) | 7 | **1.1** | 44.6 | 12 | **1.1** | 39.4 | 10 | 28.4 | 645.4 |
| (F2)(15) | | 2.1 | 36.2 | | 1.4 | 23.4 | | $\boxed{10.3}$ | 166.0 |
| (F2)(15)(18) | | 1.3 | **31.0** | | 2.6 | **20.2** | | 11.2 | **119.0** |
| (F2) | 10 | **0.5** | 16.2 | 15 | **0.8** | 19.0 | 16 | **2.7** | 31.8 |
| (F2)(15) | | 1.0 | 11.0 | | 1.1 | **12.0** | | 32.0 | **4.7** |
| (F2)(15)(18) | | 0.7 | **9.0** | | 1.9 | 18.0 | | 5.9 | 15.0 |
| (F2) | 12 | **0.3** | 6.2 | 22 | **0.1** | **1.0** | 20 | **2.1** | 39.4 |
| (F2)(15) | | **0.3** | 3.4 | | **0.1** | **1.0** | | 2.8 | **16.0** |
| (F2)(15)(18) | | 0.8 | **2.6** | | 0.3 | 1.5 | | 3.6 | 21.0 |

If these inequalities are not violated, no other inequality of family (18) will be violated (see the referred paper for details).

In Table 6, (F2) corresponds with the original formulation without any valid inequality, and (F2)(15) refers to the formulation (F2) plus the set of valid inequalities with best performance, (15), and (F2)(15)(18) adds to the latter case the separation of a subset of inequalities in family (18), those with small values of $k$. The number of nodes in the branching tree experiments almost always a significant reduction. In general, the computational times are slightly reduced in the most difficult instances, whereas in the best solved cases, implementing the separation scheme has no positive effect.

### 8.2. Variable fixing

Due to the definition of the variables in (F2), one can expect that many variables in the right-hand part of the matrices will take value 0 in the optimal solution. The size of the formulation could be reduced if some (hopefully many) of these variables were fixed beforehand. In this subsection we describe a number of variable fixing possibilities for the set of $x$-variables which are useful in the overall solution process.

#### 8.2.1. Fixing x-variables to 0

In order to fix $x_{ik}$-variables to 0 for a given $k \in A$, we will deal with an auxiliary problem that maximizes the number of uncovered clients. Here we consider that a client is covered if it is supplied by one of its $k-1$ favourite plants. In conclusion, this auxiliary problem provides the maximum number of ones in the $k$-th column of the $x$-matrix.

Variables $\delta_i$ in the auxiliary problem are defined as follows:

$$\delta_i = \begin{cases} 1 & \text{if there are no plants in the } k-1 \\ & \text{favourite sites of client } i, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall i \in A.$$

Then, we formulate the following problem (for a given value of $k$ in $\{2, \dots, M-p+1\}$):

$$\begin{aligned} \text{(Fd)} \quad \max \quad & \sum_{i=1}^{M} \delta_i \\ \text{s.t.} \quad & \delta_i + y_j \leqslant 1, \quad \forall i,j \in A : O_{ij} \leqslant k-1, \\ & \sum_{j=1}^{M} y_j = p, \qquad\qquad\qquad\qquad (25) \\ & y_j \in \{0,1\}, \quad \forall j \in A. \end{aligned}$$

If $I_k$ is the optimal value of problem (Fd), we have that, in any feasible solution of (F2),

$$x_{ik} = 0, \quad \forall i = 1, \dots, M - I_k.$$

Problem (Fd) is usually quickly solved, as could be expected considering that, by neglecting constraint (25), the coefficient matrix of (Fd) is totally unimodular. The trade-off between the effort devoted to solving several problems (Fd) (for different values of $k$) and the advantage of fixing $x$-variables to zero when solving (F2) will determine the usefulness of this preprocessing.

#### 8.2.2. Fixing x-variables to 1

In a similar way, in this section we fix $x_{ik}$ variables to 1. We consider now an auxiliary problem that minimizes the number of uncovered clients (again a client is covered, if the closest plant is in one of its $k-1$ favourite sites). Thus this problem provides the minimum number of ones in the $k$th column of the $x$-matrix.

Let $\beta_i$ be defined as follows:

$$\beta_i = \begin{cases} 1 & \text{if there are no plants in the } k-1 \\ & \text{favourite sites of client } i, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall i \in A.$$

We formulate the following problem (for every $k=2, \dots, M-p+1$):

$$\begin{aligned} \text{(Fb)} \quad \min \quad & \sum_{i=1}^{M} \beta_i \\ \text{s.t.} \quad & \beta_i + \sum_{j:O_{ij}<k} y_j \geqslant 1, \quad \forall i \in A, \\ & \sum_{j=1}^{M} y_j = p, \\ & y_j \in \{0,1\}, \quad \forall j \in A, \\ & \beta_i \geqslant 0, \quad \forall i \in A. \end{aligned}$$

If $I'_k$ is the optimal value of problem (Fb), we have that

$$x_{ik} = 1, \quad \forall i = M - I'_k + 1, \dots, M.$$

**Example 6.** Let us go back to Example 2. The optimal solution of the LP relaxation is

$$z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.6 & 0.4 & 0 \\ 0.6 & 0.4 & 0 \\ 0.6 & 0.4 & 0.5 \\ 0.6 & 0.4 & 0.5 \end{pmatrix}.$$

Note that the columns presented here are 2, 3 and 4, since the first column was fixed to one and the last column was fixed to 0 ($p=2$). The optimal value of the LP relaxation is 0. If we consider (Fd) for $k=3$, i.e.,

$$\begin{aligned} \max \quad & \sum_{i=1}^{5} \delta_i \\ \text{s.t.} \quad & \delta_1 + y_1 \leqslant 1, \quad \delta_1 + y_4 \leqslant 1, \quad \delta_2 + y_1 \leqslant 1, \quad \delta_2 + y_2 \leqslant 1, \\ & \delta_3 + y_2 \leqslant 1, \quad \delta_3 + y_3 \leqslant 1, \quad \delta_4 + y_4 \leqslant 1, \quad \delta_4 + y_5 \leqslant 1, \\ & \delta_5 + y_3 \leqslant 1, \quad \delta_5 + y_5 \leqslant 1, \quad \sum_{j=1}^{5} y_j = 2, \\ & y_j \in \{0,1\}, \quad \forall j = 1, \dots, 5, \end{aligned}$$

the optimal solution of this problem is $y=(0,1,1,0,0)$, $\delta=(1,0,0,1,0)$. This means that the solution given by plants in sites 2 and 3 is the one with largest number of clients not covered by their two favourite sites. There are two uncovered clients, namely clients 1 and 4. Therefore, no feasible solution can contain more than two ones in the third column of the $x$-matrix and we can fix $x_{13}=x_{23}=x_{33}=0$. Analogously, we fix $x_{12}=x_{22}=0$ and $x_{14}=x_{24}=x_{34}=0$. If we consider (Fb) for $k=3$, i.e.,

$$\begin{aligned} \min \quad & \sum_{i=1}^{5} \beta_i \\ \text{s.t.} \quad & \beta_1 + y_1 + y_4 \geqslant 1, \quad \beta_2 + y_1 + y_2 \geqslant 1, \\ & \beta_3 + y_2 + y_3 \geqslant 1, \quad \beta_4 + y_4 + y_5 \geqslant 1, \\ & \beta_5 + y_3 + y_5 \geqslant 1, \quad \sum_{j=1}^{5} y_j = 2, \\ & \beta_i \geqslant 0, \quad \forall i = 1, \dots, 5, \\ & y_j \in \{0,1\}, \quad \forall j = 1, \dots, 5, \end{aligned}$$

**Table 7**
Comparison between (F2) with and without preprocessing.

| | $M=20$ | | | | | $M=30$ | | | | | $M=40$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $\bar{t}_P$ | $\%v_0$ | $\%v_1$ | $\bar{t}$ | $p$ | $\bar{t}_P$ | $\%v_0$ | $\%v_1$ | $\bar{t}$ | $p$ | $\bar{t}_P$ | $\%v_0$ | $\%v_1$ | $\bar{t}$ |
| *Customers prefer closer sites and self-service is allowed* | | | | | | | | | | | | | | | |
| (F2) | 2 | – | – | – | 10.8 | 3 | – | – | – | 63.4 | 2 | – | – | – | 2379.9 |
| (F2)(15) | | – | – | – | 4.6 | | – | – | – | 10.0 | | – | – | – | 176.9 |
| (F2)(15)Pre | | 0.4 | 58.6 | 15.9 | **1.5** | | 0.7 | 71.5 | 10.5 | **5.2** | | 0.9 | 78.1 | 14.3 | **59.0** |
| (F2) | 3 | – | – | – | 7.3 | 6 | – | – | – | 4.5 | 4 | – | – | – | 265.8 |
| (F2)(15) | | – | – | – | 3.0 | | – | – | – | 3.0 | | – | – | – | 51.6 |
| (F2)(15)Pre | | 0.5 | 59.1 | 11.2 | **1.2** | | 0.7 | 72.5 | 5.5 | **1.7** | | 1.0 | 78.4 | 8.2 | **16.6** |
| (F2) | 5 | – | – | – | 2.6 | 10 | – | – | – | 1.3 | 8 | – | – | – | 16.6 |
| (F2)(15) | | – | – | – | 1.2 | | – | – | – | 1.1 | | – | – | – | 13.1 |
| (F2)(15)Pre | | 0.4 | 61.4 | 6.9 | **0.6** | | 0.8 | 72.5 | 3.9 | **0.9** | | 1.2 | 78.9 | 4.3 | **4.6** |
| (F2) | 7 | – | – | – | 1.4 | 12 | – | – | – | **0.7** | 10 | – | – | – | 13.5 |
| (F2)(15) | | – | – | – | **0.4** | | – | – | – | **0.7** | | – | – | – | 7.2 |
| (F2)(15)Pre | | 0.4 | 61.7 | 5.5 | 0.5 | | 0.8 | 73.8 | 3.5 | 0.8 | | 1.3 | 79.1 | 3.6 | **2.9** |
| (F2) | 10 | – | – | – | **0.2** | 15 | – | – | – | 1.9 | 16 | – | – | – | 2.0 |
| (F2)(15) | | – | – | – | **0.2** | | – | – | – | 1.4 | | – | – | – | 2.3 |
| (F2)(15)Pre | | 0.4 | 62.8 | 5.0 | 0.4 | | 0.7 | 73.7 | 3.3 | **0.8** | | 1.1 | 79.4 | 2.6 | **1.2** |
| (F2) | 12 | – | – | – | **0.1** | 22 | – | – | – | **0.2** | 20 | – | – | – | 3.8 |
| (F2)(15) | | – | – | – | **0.1** | | – | – | – | **0.2** | | – | – | – | **1.0** |
| (F2)(15)Pre | | 2.2 | 70.5 | 5.0 | 2.5 | | 1.6 | 79.6 | 3.3 | 1.7 | | 1.7 | 79.9 | 3.3 | 2.0 |
| *Customers prefer closer sites but self-service is forbidden* | | | | | | | | | | | | | | | |
| (F2) | 2 | – | – | – | 28.1 | 3 | – | – | – | 814.1 | 2 | – | – | – | 2717.5 |
| (F2)(15) | | – | – | – | 4.4 | | – | – | – | 28.1 | | – | – | – | 125.2 |
| (F2)(15)Pre | | 0.5 | 57.5 | 14.4 | **2.7** | | 0.8 | 70.9 | 9.8 | **20.5** | | 1.0 | 77.8 | 13.6 | **74.6** |
| (F2) | 3 | – | – | – | 16.6 | 6 | – | – | – | 52.0 | 4 | – | – | – | > 1*H* |
| (F2)(15) | | – | – | – | 3.8 | | – | – | – | 6.7 | | – | – | – | 114.6 |
| (F2)(15)Pre | | 0.5 | 58.2 | 9.8 | **2.4** | | 1.0 | 71.3 | 4.4 | **4.2** | | 1.3 | 77.8 | 7.7 | **90.8** |
| (F2) | 5 | – | – | – | 2.5 | 10 | – | – | – | 4.5 | 8 | – | – | – | 846.7 |
| (F2)(15) | | – | – | – | 3.8 | | – | – | – | 3.1 | | – | – | – | 23.1 |
| (F2)(15)Pre | | 0.5 | 59.5 | 5.4 | **1.2** | | 0.9 | 71.3 | 2.6 | **1.8** | | 1.7 | 78.0 | 3.5 | **22.1** |
| (F2) | 7 | – | – | – | 1.0 | 12 | – | – | – | 1.7 | 10 | – | – | – | 139.6 |
| (F2)(15) | | – | – | – | **0.7** | | – | – | – | 1.7 | | – | – | – | 21.7 |
| (F2)(15)Pre | | 0.5 | 59.3 | 3.4 | **0.7** | | 0.9 | 71.9 | 1.8 | **1.2** | | 1.9 | 78.1 | 2.7 | **14.3** |
| (F2) | 10 | – | – | – | **0.4** | 15 | – | – | – | 2.2 | 16 | – | – | – | 4.2 |
| (F2)(15) | | – | – | – | 1.3 | | – | – | – | 1.5 | | – | – | – | 4.6 |
| (F2)(15)Pre | | 0.5 | 60.9 | 2.0 | 0.5 | | 0.8 | 71.6 | 1.4 | **1.0** | | 1.5 | 78.0 | 1.5 | **2.1** |
| (F2) | 12 | – | – | – | **0.4** | 22 | – | – | – | 0.2 | 20 | – | – | – | **2.0** |
| (F2)(15) | | – | – | – | 0.6 | | – | – | – | **0.1** | | – | – | – | 2.6 |
| (F2)(15)Pre | | 2.0 | 67.4 | 1.1 | 2.1 | | 1.5 | 76.0 | 0.0 | 1.7 | | 1.7 | 78.1 | 1.1 | **2.0** |
| *Random preferences* | | | | | | | | | | | | | | | |
| (F2) | 2 | – | – | – | 62.7 | 3 | – | – | – | > 1*H* | 2 | – | – | – | > 1*H* |
| (F2)(15) | | – | – | – | 6.2 | | – | – | – | 93.5 | | – | – | – | > 1*H* |
| (F2)(15)Pre | | 1.0 | 65.8 | 14.6 | **6.1** | | 1.8 | 72.0 | 10.6 | **65.2** | | 2.0 | 78.2 | 13.7 | **2005.5** |
| (F2) | 3 | – | – | – | 15.5 | 6 | – | – | – | 113.5 | 4 | – | – | – | > 1*H* |
| (F2)(15) | | – | – | – | **3.1** | | – | – | – | **6.6** | | – | – | – | 1342.9 |
| (F2)(15)Pre | | 1.1 | 66.6 | 9.0 | 4.3 | | 3.1 | 74.5 | 4.1 | 6.9 | | 4.3 | 79.6 | 7.8 | **397.5** |
| (F2) | 5 | – | – | – | 2.4 | 10 | – | – | – | 2.7 | 8 | – | – | – | 212.8 |
| (F2)(15) | | – | – | – | **1.2** | | – | – | – | **2.2** | | – | – | – | 32.3 |
| (F2)(15)Pre | | 1.0 | 68.5 | 4.3 | 1.7 | | 3.4 | 77.2 | 1.5 | 4.0 | | 11.3 | 80.8 | 3.1 | **20.7** |
| (F2) | 7 | – | – | – | **1.1** | 12 | – | – | – | **1.1** | 10 | – | – | – | 28.4 |
| (F2)(15) | | – | – | – | 2.1 | | – | – | – | 1.4 | | – | – | – | **10.3** |
| (F2)(15)Pre | | 1.0 | 69.8 | 2.4 | 1.3 | | 3.2 | 77.8 | 1.2 | 3.6 | | 16.8 | 82.1 | 2.1 | 19.3 |
| (F2) | 10 | – | – | – | **0.5** | 15 | – | – | – | **0.8** | 16 | – | – | – | **2.7** |
| (F2)(15) | | – | – | – | 1.0 | | – | – | – | 1.1 | | – | – | – | 32.0 |
| (F2)(15)Pre | | 0.9 | 71.8 | 1.2 | 0.9 | | 2.5 | 78.7 | 0.6 | 2.6 | | 17.2 | 83.4 | 1.1 | 17.7 |
| (F2) | 12 | – | – | – | **0.3** | 22 | – | – | – | **0.1** | 20 | – | – | – | **2.1** |
| (F2)(15) | | – | – | – | **0.3** | | – | – | – | **0.1** | | – | – | – | 2.8 |
| (F2)(15)Pre | | 1.6 | 73.2 | 0.5 | 1.8 | | 1.5 | 81.7 | 0.0 | 1.8 | | 11.5 | 84.4 | 0.9 | 11.8 |

the optimal solution of this problem is $y=(0,1,0,1,0)$, $\beta=(0,0,0,0,1)$. Then, the solution given by plants in sites 2 and 4 is the one with smallest number of clients not covered by their two favourite sites. There is one uncovered client, namely client 5. Therefore, no feasible solution can contain less than one 1 in the third column of the $x$-matrix and we can fix $x_{15}=1$. Analogously, we fix $x_{32}=x_{42}=x_{52}=1$. In a set of 15 variables, 12 have been fixed. The solution of the linear relaxation of the subproblem obtained after fixing the variables is

$$z=\begin{pmatrix} 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \\ 1 & 0.5 & 0 \\ 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 \end{pmatrix}, \quad x=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

with an optimal value of 10.

**Table 8**
Comparison between (F3) with and without a modification in the objective function.

| | $M=20$ | | | | $M=30$ | | | | $M=40$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ | $p$ | $LP$ | $\bar{t}$ | $n$ |
| *Customers prefer closer sites and self-service is allowed* | | | | | | | | | | | | |
| (F3) | 2 | 0.0 | 1.5 | 177.0 | 3 | 0.1 | 22.7 | 718.2 | 2 | 0.5 | 50.7 | 1077.0 |
| (F3R) | | 0.0 | **1.4** | **153.4** | | 0.1 | **19.2** | **617.0** | | 0.2 | $\boxed{42.3}$ | **577.8** |
| (F3) | 3 | 0.0 | 2.5 | 327.8 | 6 | 0.1 | 81.9 | 4357.8 | 4 | 0.2 | 263.6 | **3373.8** |
| (F3R) | | 0.0 | **2.2** | **291.4** | | 0.1 | **65.8** | **3955.0** | | 0.2 | **218.1** | 3568.6 |
| (F3) | 5 | 0.0 | 4.4 | 917.0 | 10 | 0.1 | 174.5 | 17987.0 | 8 | 0.2 | 2443.9 | 57277.8 |
| (F3R) | | 0.0 | **3.6** | **718.6** | | 0.1 | **141.5** | **10941.0** | | 0.1 | **1568.4** | **37586.2** |
| (F3) | 7 | 0.0 | 4.4 | 1255.8 | 12 | 0.1 | 425.6 | 51375.0 | 10 | 0.2 | $>1H$ | – |
| (F3R) | | 0.0 | **4.0** | **874.6** | | 0.0 | **270.5** | **23519.8** | | 0.1 | **3200.6** | **81377.5** |
| (F3) | 10 | 0.0 | 9.5 | 5579.8 | 15 | 0.1 | 996.8 | 249648.0 | 16 | 0.2 | $>1H$ | – |
| (F3R) | | 0.0 | **7.6** | **2937.0** | | 0.0 | **686.0** | **107674.0** | | 0.2 | $>1H$ | – |
| (F3) | 12 | 0.0 | **16.4** | 14263.0 | 22 | 0.1 | 3298.1 | **1999284.0** | 20 | 0.2 | $>1H$ | – |
| (F3R) | | 0.0 | 17.8 | **10633.4** | | OOM | – | – | | 0.2 | $>1H$ | – |
| *Customers prefer closer sites but self-service is forbidden* | | | | | | | | | | | | |
| (F3) | 2 | 0.0 | **1.4** | 167.8 | 3 | 0.1 | 26.5 | 861.9 | 2 | 0.3 | 53.1 | 482.2 |
| (F3R) | | 0.0 | **1.4** | **137.4** | | 0.1 | **22.2** | **786.2** | | 0.2 | $\boxed{39.3}$ | **409.4** |
| (F3) | 3 | 0.0 | 2.5 | 315.4 | 6 | 0.1 | 66.9 | 3258.6 | 4 | 0.2 | 326.3 | 4029.4 |
| (F3R) | | 0.0 | **2.2** | **279.4** | | 0.1 | **57.8** | **2917.8** | | 0.2 | **214.4** | **3193.8** |
| (F3) | 5 | 0.0 | 3.2 | 597.0 | 10 | 0.1 | 82.3 | 6807.0 | 8 | 0.2 | 2105.2 | 38412.2 |
| (F3R) | | 0.0 | **2.7** | **466.6** | | 0.1 | **64.4** | **4516.0** | | 0.1 | **1281.3** | **28625.8** |
| (F3) | 7 | 0.0 | 2.5 | 595.4 | 12 | 0.1 | 48.8 | 4261.0 | 10 | 0.2 | 2763.5 | 67043.0 |
| (F3R) | | 0.0 | **2.3** | **474.2** | | 0.0 | **31.6** | **2208.2** | | 0.1 | **1829.3** | **45384.0** |
| (F3) | 10 | 0.0 | 1.3 | 411.0 | 15 | 0.1 | 15.8 | 2054.0 | 16 | 0.2 | 1677.0 | 43331.5 |
| (F3R) | | 0.0 | **1.1** | **281.4** | | 0.0 | **8.6** | **782.0** | | 0.1 | **1346.5** | **40464.6** |
| (F3) | 12 | 0.0 | **0.4** | 107.4 | 22 | 0.0 | **0.2** | **3.0** | 20 | 0.2 | 127.5 | 6387.0 |
| (F3R) | | 0.0 | **0.4** | **96.2** | | 0.0 | 0.4 | 30.0 | | 0.1 | **96.2** | **3906.0** |
| *Random preferences* | | | | | | | | | | | | |
| (F3) | 2 | 0.0 | 2.0 | 242.6 | 3 | 0.1 | 34.0 | 1033.4 | 2 | 0.0 | 96.6 | 1150.2 |
| (F3R) | | 0.0 | **1.7** | **175.4** | | 0.1 | $\boxed{23.9}$ | **699.4** | | 0.2 | $\boxed{70.4}$ | **1028.2** |
| (F3) | 3 | 0.0 | 2.5 | 293.4 | 6 | 0.1 | 63.2 | 2436.6 | 4 | 0.3 | 507.0 | 5660.2 |
| (F3R) | | 0.0 | $\boxed{2.2}$ | **257.4** | | 0.1 | **53.5** | **2183.8** | | 0.2 | $\boxed{319.4}$ | **4102.5** |
| (F3) | 5 | 0.0 | 2.3 | 346.2 | 10 | 0.1 | **27.8** | **1443.0** | 8 | 0.2 | 862.9 | 12009.4 |
| (F3R) | | 0.0 | **2.1** | **317.8** | | 0.1 | 29.8 | 1504.0 | | 0.2 | **447.3** | **6487.8** |
| (F3) | 7 | 0.0 | 1.7 | 329.4 | 12 | 0.1 | 20.5 | 1495.4 | 10 | 0.2 | 651.4 | 11759.0 |
| (F3R) | | 0.0 | **1.2** | **204.6** | | 0.0 | **15.3** | **919.0** | | 0.2 | **423.3** | **7875.0** |
| (F3) | 10 | 0.0 | **0.6** | 166.2 | 15 | 0.1 | 6.4 | 648.0 | 16 | 0.2 | **336.1** | 12150.0 |
| (F3R) | | 0.0 | 0.7 | **160.6** | | 0.1 | **3.8** | **326.0** | | 0.1 | 388.1 | **10534.0** |
| (F3) | 12 | 0.0 | $\boxed{0.2}$ | 51.0 | 22 | 0.0 | **0.2** | **3.0** | 20 | 0.2 | 71.1 | 3689.0 |
| (F3R) | | 0.0 | $\boxed{0.2}$ | **41.4** | | 0.0 | 0.3 | 30.0 | | 0.1 | **59.6** | **2340.0** |

In the computational experiment, in order to avoid excessive running time, we obtained $I_k$ and $I'_k$ for $k \in \{2,\dots,10\}$ (in the cases $M{=}20$ with $p=12$ and $M=30$ with $p=22$, we considered $k \in \{2,\dots,9\}$). The results of solving the preprocessed formulation are presented in Table 7, where columns $\%v_0$ and $\%v_1$ denote the average percentage of variables fixed to 0 and 1, respectively, in the preprocessing phase. It can be seen that the computational times $\bar{t}$ needed for solving the MELP using the combination of the second formulation, the valid inequalities (15) and the preprocessing phase (carried out using time $\bar{t}_P$) are extremely good (notice that the total time includes the preprocessing time).

## 9. Improving the third formulation

The improvement we have considered for the third formulation is quite simple: we have subtracted a small number from some coefficients of the objective function. The idea is to avoid the multiplicity of optimal solutions prompted by the use of Ogryczak and Tamir's formulation. As seen in Section 5, these authors considered the function

$$f_q(t) := qt + \sum_{i=1}^{M} \max\{0, z_i - t\}.$$

The $q$-th $z$-value, $z_{(q)}$, is always a minimum of this function, with $f_q(z_{(q)}) = \sum_{i=1}^{q} z_{(i)}$. But, for any $b$ in the interval $(z_{(q-1)}, z_{(q)})$ (which has sense only if $z_{(q-1)} \neq z_{(q)}$) we get

$$f_q(b) = qb + \sum_{i=1}^{M} \max\{0, z_i - b\} = qb + \sum_{i=1}^{q} (z_{(i)} - b) = \sum_{i=1}^{q} z_{(i)},$$

still the optimal value. On the contrary, if the objective is changed to

$$F_q(t) := (q-\varepsilon)t + \sum_{i=1}^{M} \max\{0, z_i - t\},$$

**Table 9**
Best computational times (seconds) obtained with each formulation for $M=40$.

| | $p=2$ | $p=4$ | $p=8$ | $p=10$ | $p=16$ | $p=20$ |
|---|---|---|---|---|---|---|
| *Customers prefer closer sites and self-service is allowed* | | | | | | |
| First formulation | 47 | 281 | 2269 | $>1H$ | 3098 | $>1H$ |
| Second formulation | 59 | **17** | **5** | **3** | **1** | **1** |
| Third formulation | **42** | 218 | 1568 | 3201 | $>1H$ | $>1H$ |
| *Customers prefer closer sites but self-service is forbidden* | | | | | | |
| First formulation | 45 | 312 | 2000 | 2624 | 1261 | 315 |
| Second formulation | 75 | **91** | **22** | **14** | **2** | **2** |
| Third formulation | **39** | 214 | 1281 | 1829 | 1347 | 96 |
| *Random preferences* | | | | | | |
| First formulation | 90 | 460 | 977 | 985 | 280 | 75 |
| Second formulation | 2005 | 398 | **21** | **10** | **3** | **2** |
| Third formulation | **70** | **319** | 447 | 423 | 336 | 60 |

then

$$f_q(z_{(q)}) = (q-\varepsilon)z_{(q)} + \sum_{i=1}^{M}\max\{0, z_i - z_{(q)}\} = (q-\varepsilon)z_{(q)} + \sum_{i=1}^{q}(z_{(i)} - z_{(q)}) = -\varepsilon z_{(q)} + \sum_{i=1}^{q} z_{(i)},$$

while, for $b \in (z_{(q-1)}, z_{(q)})$, we get

$$f_q(b) = (q-\varepsilon)b + \sum_{i=1}^{M}\max\{0, z_i - b\} = (q-\varepsilon)b + \sum_{i=1}^{q}(z_{(i)} - b) = -\varepsilon b + \sum_{i=1}^{q} z_{(i)}.$$

Since $b < z_{(q)}$, then $f_q(b) > f_q(z_{(q)})$, and $b$ is not an optimal solution of the problem anymore. If $\varepsilon$ is small enough, the perturbation of the objective function will not affect the correctness of the formulation.

In Table 8 the results obtained for formulation (F3) with and without the modified objective function have been compared. We can see that with the modification the amount of nodes of the branching tree is significantly reduced and the computational times are smaller. In this table, OOM means an "out of memory" error, possibly due to the huge number of nodes in the branching tree.

## 10. Conclusions

The minimum envy location problem, like many other equity problems in the Mathematical Programming literature, has the disadvantage of having an objective function in which the absolute values of several differences must be calculated. Whereas the classical strategy in the field of Statistics has been to replace the absolute values by squares with the aim of obtaining a differentiable function, when building an Integer Programming formulation one can deal with these values using different strategies. We have considered a wide repertoire of formulations (each with its own flavor), have improved them in most cases by means of valid inequalities, strengthening of the constraints, preprocessing and other different techniques, and compared them in a computational framework. Some of the approaches are clearly outperformed by the others, and there is an improved formulation (the so-called second formulation) which seems to give better results in terms of computational times and size of the branching tree, except for very small values of $p$, which are best solved with the third formulation. A summary of the best computational times obtained with each formulation for the instances with 40 points is presented in Table 9.

Further research on this topic could include more tailored solution schemes, in particular those which allow to obtain better lower bounds than the simple linear relaxation, like Lagrangean relaxation, as well as further investigation into the polyhedral characteristics of the formulations. Another matter of future research is the addition of capacities to the plants, i.e., considering that a maximum number of customers can be allocated to each plant or that every plant has a limited capacity and every customer has a known demand.

## Acknowledgments

This research has been partially supported by Spanish Ministry of Education and Science Grant numbers MTM2004-0909, MTM2006-14961-C05-04, MTM2007-67433-C02-(01,02), HI2006-0123 and HA2003-0121, RDEF funds, Fundación Séneca, Grant number 02911/PI/05 and Junta de Andalucía, Grant numbers P06-FQM-01364 and P06-FQM-01366.

## References

[1] Brams SJ, Fishburn PC. Fair division of indivisible items between two people with identical preferences: envy-freeness. Pareto-optimality and equity. Social Choice and Welfare 2000;17:247–67.
[2] Dall'Aglio M, Hill TP. Maximin share and minimax envy in fair-division problems. Journal of Mathematical Analysis and Applications 2003;281(1): 346–61.
[3] Pápai S. Groves sealed bid auctions of heterogeneous objects with fair prices. Social Choice and Welfare 2003;20:371–85.
[4] Reijnierse JH, Potters JAM. On finding an envy-free Pareto-optimal division. Mathematical Programming 1998;83(2):291–311.
[5] Webb WA. An algorithm for super envy-free cake division. Journal of Mathematical Analysis and Applications 1999;239(1):175–9.
[6] Ohseto S. Strategy-proof assignment with fair compensation. Mathematical Social Sciences 2005;50:215–26.
[7] Shioura A, Sun N, Yang Z. Efficient strategy proof fair allocation algorithms. Journal of the Operations Research Society of Japan 2006;49(2):144–50.
[8] Chun Y. No-envy in queueing problems. Economic Theory 2006;29(1):151–62.
[9] Haake CJ, Raith MG, Su FE. Bidding for envy-freeness: a procedural approach to $n$-player fair-division problems. Social Choice and Welfare 2002;19(4):723–49.
[10] Sung SC, Vlach M. Competitive envy-free division. Social Choice and Welfare 2004;23(1):103–11.
[11] Moulin H. Cooperative microeconomis: a game theoretic introduction. Princeton, NJ: Princeton University Press; 1995.
[12] Savas ES. On equity in providing public services. Management Science 1978;24:800–8.
[13] Halpern J, Maimon O. Accord and conflict among several objectives in locational decisions on tree networks. In: Thisse JF, Zoller HG, editors. Locational analysis of public facilities. Amsterdam: North-Holland; 1983.
[14] Mulligan GF. Equality measures and facility location. Regional Science 1991;70:345–65.
[15] Erkut E. Inequality measures for location problems. Location Science 1993;1: 199–217.
[16] Berman O, Kaplan EH. Equity maximizing facility location schemes. Transportation Science 1990;24:137–44.
[17] Mesa JA, Puerto J, Tamir A. Improved algorithm for several network location problems with equality measures. Discrete Applied Mathematics 2003;130: 437–48.
[18] Garfinkel R, Fernández E, Lowe T. The $k$-centrum shortest path problem. TOP 2006;14(2):279–92.
[19] Marsh MT, Schilling DA. Equity measurement in facility location analysis: a review and framework. European Journal of Operational Research 1994;74: 1–17.

[20] Eiselt HA, Laporte G. Objectives in location problems. In: Drezner Z, editor. Facility location: a survey of applications and methods. Berlin: Springer; 1995.
[21] López-de-los-Mozos MC, Mesa JA. The sum of absolute differences on a network: algorithm and comparison with other equality measures. INFOR: Information System and Operational Research 2003;41:195–210.
[22] Marín A, Nickel S, Velten S. A reformulated and extended covering model for flexible discrete location problems. In: Proceedings of the EURO Winter Institute on Location and Logistics, University of Lisbon, 2007. Submitted for publication.
[23] Hansen P, Thisse JF. Outcomes of voting and planning. Journal of Public Economics 1981;16:1–15.
[24] Carrizosa E, Conde E, Muñoz-Márquez M, Puerto J. Simpson points in planar problems with locational constraints. The polyhedral-gauge case. Mathematic of Operations Research 1997;22:291–300.
[25] Hakimi SL. Locations with spatial interactions: competitive locations and games. In: Mirchandani PB, Francis RL, editors. Discrete location theory. New York: Wiley; 1990.
[26] Labbé M. Outcomes of voting and planning in single facility location problems. European Journal of Operational Research 1985;20:299–313.
[27] Marín A, Nickel S, Puerto J, Velten S. A flexible model and efficient solution strategies for discrete location problems. Discrete Applied Mathematics, 2008 (forthcoming). doi: 10.1016/j.dam.2008.03.013.
[28] Ogryczak W, Tamir A. Minimizing the sum of the k largest functions in linear time. Information Processing Letters 2003;85:117–22.
[29] Espejo M, Marín A, Puerto J, Rodríguez-Chía A. A comparison of formulations and solution methods for the minimum-envy location problem. Additional results. Research report available at ⟨www.arxiv.org⟩, code arXiv:0806.3134.
[30] Belotti P, Labbé M, Maffioli F, Ndiaye M. A branch-and-cut method for the obnoxious p-median problem. 4OR 2007;5(4):299–314.
[31] Cánovas L, García S, Labbé M, Marín A. A strengthened formulation for the simple plant location problem with order. Operations Research Letters 2007;35(2):141–50.
[32] Balas E, Ng SM. On the set covering polytope: I. All the facets with coefficients in $\{0,1,2\}$. Mathematical Programming 1989;43:57–69.