The Facility Location Problem with Bernoulli Demands

Maria Albareda-Sambola[a], Elena Fernández[b*], Francisco Saldanha-da-Gama[c]

[a] *Universitat Politècnica de Catalunya, Dept. d'Estadística i Investigació Operativa, Carrer Colom, 11, 08022 Terrassa, Spain*
[b] *Universitat Politècnica de Catalunya, Dept. d'Estadística i Investigació Operativa, Campus Nord, C5-208, Jordi Girona 1–3, 08034 Barcelona, Spain*
[c] *Universidade de Lisboa, Faculdade de Ciências, Dept. Estatística e Investigação Operacional e Centro de Investigação Operacional,*
Bloco C6, Piso 4, 1749-016 Lisboa, Portugal

July 30, 2008

Abstract

In this paper we address a discrete capacitated facility location problem in which customers have Bernoulli demands. The problem is formulated as a two-stage stochastic program. The goal is to define an *a priori* solution for the location of the facilities and for the allocation of customers to the operating facilities that minimize the expected value of the recourse function. A closed form is presented for the recourse function and an approximation is proposed for situations in which it can not be evaluated exactly. The stochastic assignment problem resulting from setting the operating facilities is studied and a procedure is proposed to approximate its optimal solution. We formulate the capacitated facility location problem with Bernoulli demands under perfect information and, finally, we propose a heuristic to find an *a priori* solution. Computational results are presented for the homogeneous case.

1 Introduction

A Facility Location Problem looks for the best location for a set of facilities that must satisfy requests of service coming from a given set of customers. Often it is assumed that customers demands of service are part of the input data of the problem and, thus, are known in advance. However, it is well-known that in practice this very seldom occurs, since usually there is a high level of uncertainty associated with the demand. Typical examples of location problems with non-deterministic demands include not only the location of emergency facilities, but also any location problem where the demand levels might change over different time periods (postal services, super-markets, warehouses to distribute goods with seasonal-dependent demand, airports, etc.) Brandeau and Chiu [9], Louveaux [23] and Snyder [28] have surveyed different aspects of Stochastic Location Problems.

There are several alternatives for handling problems that exhibit uncertainty in service demands. If no assumption is made on the probability distribution of the demands, one possibility is to guarantee the robustness of the obtained solutions by taking into account several possible scenarios. Depending on the context, robustness can be evaluated by means of different measures like, for instance, the minimization of the cost associated with the most adverse scenario or the minimization of the cost associated with the most likely scenario. Some related references are, for instance, Averbakh and Berman [3, 4, 5], Carrizosa and Nickel [11], or Conde [12, 13]. Other possibilities arise when the stochastic nature of the demands is captured by means of some assumption on their probability distribution. For instance, when service requests that arrive when the servers are busy are rejected, some authors have considered probabilistic constraints to guarantee that the probability that customers with demand receive the requested service is above a certain threshold (e.g. ReVelle [26] and Toregas *et al.* [29]). From a different point of view, when it is accepted that requests that arrive when servers are busy can wait until some server is free, queueing-location models address the minimization of the expected waiting time in the queue or the expected length of the queue. Some references along this line are Marianov and ReVelle [25], Carrizosa

*Corresponding author. *e-mail address*: e.fernandez@upc.edu

et al. [10], or Fernández *et al.* [15]. A different methodology to handle these problems comes from Stochastic Programming. One possibility is to define a recourse function that, given an *a priori* solution, indicates the best decision to make for each possible realization of the random vector, and then to find the *a priori* solution that minimizes the expected cost of this recourse function. Louveaux and Peeters [24] and Laporte *et al.* [22] explore this possibility for the Uncapacitated Facility Location Problem. In both papers, a set of scenarios is considered each of which occurring with some specified probability. Riis and Andersen [27] study a capacity expansion problem arising in telecommunication networks that involves no location decisions. The authors consider the realistic situation in which demand is uncertain and follows some known probability distribution. A stochastic integer programming formulation is proposed for the problem.

Frequently, the request of service of a given customer is expressed in terms of a "quantity" that represents the amount of resources of a service center that will be consumed if the customer is served. Nevertheless, in many situations requests for service are unitary in the sense that some service is requested that requires one unit of resource (e.g. one worker) from the service center. In such situation, each possible scenario is characterized by the set of customers with demand. This is, for instance, the case when servers are repairmen and requests of service from customers correspond to requests for repairing. Now, requests can be modeled by means of binary vectors. In this context, if demands are stochastic, the components of such a vector are 0-1 random variables. In this paper we consider the particular case in which the 0-1 random variables are independent and follow a Bernoulli probability distribution.

The facility location problem that we address in this paper is stated in a discrete setting. Accordingly, we assume that there is a finite set of potential locations for the facilities. Moreover, each customer must be assigned to one single facility. The decision regarding where to locate the facilities and how to allocate the customers to the operating facilities must be made before knowing which customers will in fact request the service. Actually, depending on the context, locational and assignment decisions are often made on a strategic or tactical level due to the resources that may have to be allocated to the implementation of these decisions (e.g. financial resources or human workforce). Therefore, not only it is not desirable to change these decisions too often but also, such decisions are often made and implemented before knowing exactly how the future will be, for instance, in terms of the potential demand. In this case, the planning must be ready before all requests for the service are received (see Snyder [28] for a deeper discussion on these aspects).

We assume that each facility has a capacity that represents the maximum number of customers that can be served from it (in case it is operating). Thus, in some realizations, there might be more customers allocated to one operating facility than the capacity of the facility. In such situation, if there are more customers requesting the service than the capacity of the facility to which they are allocated, some customers can not be served and a penalty cost is incurred for each request declined. We also consider the situation in which a facility can only be installed if a pre-specified minimum number of demand points is assigned to it. It should be noted that often, this imposition is not drawn by the costs but by preliminary decisions made to assure a somehow balanced system. This requirement is quite general, since the situation in which no such limits exist becomes a particular case of the problem we are studying. The goal is to decide where to locate the facilities and how to allocate all the customers so as to minimize the overall expected cost which, in addition to the penalty cost above mentioned also includes the costs for installing the facilities and the demand satisfaction costs. This is a Stochastic Combinatorial Optimization Problem, in which the stochasticity is associated with the customers demands, which we assume to follow a Bernoulli distribution.

Several Stochastic Combinatorial Optimization Problems with Bernoulli demands have been addressed in the literature in the context of the minimization of a recourse function. This is the case of the probabilistic travelling salesman problem. Jaillet [19] introduces this problem and presents a closed form expression for computing efficiently the expected value of the length of any given tour. Laporte *et al.* [21] propose a linear stochastic program model that is solved using a branch-and-cut approach. More recently, Bianchi and Campbell [7] propose a heuristic approach for the problem. The reader is referred to this latter work for additional references on this problem. Berman and Simchi-Levi [6] study a single-vehicle location-routing problem with Bernoulli demands. The authors formulate the problem and develop a lower bound on the value of the optimal *a priori* tour. Quite recently, Albareda-Sambola *et al.* [1] addressed a study of location-routing problems with Bernoulli demands. The authors propose heuristics and lower bounds to minimize the expected value of the defined recourse function. Albareda-Sambola *et al.* [2] propose an exact algorithm for minimizing the expectation of a recourse function for the Stochastic Generalized Assignment Problem, where it is assumed that the customers demands follow a Bernoulli distribution. To the best of our knowledge, no work exists in the literature that uses a similar methodology for discrete location problems with Bernoulli demands.

In this paper the optimization of the above described Facility Location Problem with Bernoulli Demands (FLPBD) is studied. In the next section the problem is formally described and a recourse function is presented that allows the FLPBD to be formulated as a two-stage stochastic program. The goal is to find an *a priori* solution that minimizes the expected cost of the recourse function. A closed expression for this function is derived, which is simple to evaluate in the homogeneous situation (all customers have demand with the same probability). Since its evaluation becomes unaffordable as the number of customers increases, the general case is addressed in Subsection 2.1, in which an approximation of the recourse function is proposed. Like in most location problems, the assignment subproblem to be solved when the set of facilities to open is known is of interest on its own. In Section 3 a formulation is presented for this stochastic problem and a solution algorithm is proposed which is based on the resolution of a minimum cost flow problem in an auxiliary network. The problem under perfect information that results for a particular a realization of the random vector is addressed in Section 4. This problem is indeed closely related to the FLPBD and is also of great interest, since its resolution for a large sample of realizations of the random vector gives us valuable information for evaluating the solutions obtained for the overall stochastic problem. Feasible solutions are obtained with a heuristic procedure that is proposed in Section 5. Section 6 is devoted to the computational experience performed to evaluate the proposed methodologies. The paper ends with some conclusions and suggestions for further research.

2 Definition of the problem

Let I and J denote the set of indices for the potential locations of facilities and for customers, respectively. Let $n = |J|$. We assume that the demand of service of each customer $j \in J$ is given by a binary random variable, that we denote by ξ_j, indicating whether or not customer j has demand. We further assume that ξ_j follows a Bernoulli probability distribution of parameter p_j. For $i \in I$, we consider, in addition, the following notation:

f_i	Fixed set-up cost for opening facility i;
ℓ_i	Minimum number of customers that have to be assigned to facility i if it is opened;
K_i	Maximum number of customers that can be served from facility i when it is opened;
c_{ij}	Cost for serving customer j from facility i ($j \in J$).

It should be noted that for a realization of the random vector not all the customers need to have demand. For this reason, we distinguish between the assignment of customers to plants, which is done *a priori* and is independent of the potential realizations, and the service of customers from open plants, which is decided *a posteriori*, once the realization is known. Thus, an *a priori* solution is given by a set of operating facilities together with an assignment of all the customers to these facilities, such that for any open plant the set of customers that are assigned to it is at least ℓ_i. Observe that K_i is an upper bound on the number of customers that can be served from an opened plant and, thus, this is not taken into account for the feasibility of *a priori* solutions. Let J_i denote the set of customers assigned to facility i in the *a priori* solution, with $z_i = |J_i|$. Throughout the text the customers with demand in a potential realization are referred to as the *demand customers*. Also, for a potential realization, η_i denotes the number of demand customers assigned to facility i.

Given an *a priori* solution, the *a posteriori* solution indicates the decisions to make, once demand customers are known. If the number of customers assigned to an open facility i does not exceed its upper bound, i.e. $z_i \leq K_i$, then in the *a posteriori* solution all the demand customers indexed in J_i receive service from plant i. However, when the number z_i exceeds K_i, if η_i is also greater than K_i then the *a posteriori* solution consists of randomly selecting K_i demand customers from J_i for being actually served from plant i. Each of the remaining $(\eta_i - K_i)$ demand customers remain unserved. A service cost c_{ij} is incurred for each served customer j. A penalty cost that we denote by g is incurred for every unserved customer. The recourse function is the expected cost, over all possible realizations of the demand vector and all realizations of the random choices for resolving the excess demand in plants, of the *a posteriori* solution.

The FLPBD consists of finding a set of facilities to open and an allocation of the customers to the opened facilities, in such a way that the sum of the fixed cost associated with the open facilities and the recourse function is minimized. In order to give a mathematical programming formulation for this problem, we consider the following sets of decision variables:

$$y_i = \begin{cases} 1 & \text{if a facility is established at } i \\ 0 & \text{otherwise} \end{cases} \qquad (i \in I)$$

$$x_{ij} = \begin{cases} 1 & \text{if customer } j \text{ is allocated to facility } i \\ 0 & \text{otherwise} \end{cases} \qquad (i \in I,\ j \in J)$$

The FLPBD can be formulated as follows:

$$(P) \quad min \quad \sum_{i \in I} f_i y_i + \mathbb{E}_\xi(\text{Service cost} + \text{Penalty cost}) \tag{1}$$

$$s.\,t. \quad \sum_{i \in I} x_{ij} = 1 \qquad j \in J \tag{2}$$

$$x_{ij} \le y_i \qquad i \in I,\ j \in J \tag{3}$$

$$\ell_i y_i \le \sum_{j \in J} x_{ij} \qquad i \in I,\ j \in J \tag{4}$$

$$y_i \in \{0,1\} \qquad i \in I \tag{5}$$

$$x_{ij} \in \{0,1\} \qquad i \in I,\ j \in J \tag{6}$$

As mentioned above, the objective function (1) includes the fixed costs for opening the facilities and the recourse function. Constraints (2) assure that all customers will be assigned to (exactly) one facility while constraints (3) assure that this assignments can only be done to operating facilities. Constraints (4) state the minimum number of customers that should be assigned to one operating facility. Finally, constraints (5) and (6) are domain constraints.

We next analyze in more detail the recourse function. To do so, for a given solution, (y, x), we express both z_i and η_i in terms of the decision variables. That is

$$z_i = \sum_{j \in J} x_{ij} \quad \text{and} \quad \eta_i = \sum_{j \in J} \xi_j x_{ij}$$

In addition, define generically $\mathbb{P}_i(s)$ as the conditional probability of serving a demand customer assigned to facility $i \in I$ given that the total number of demand customers assigned to facility i is s (i.e. $\eta_i = s$). For $i \in I$, we have:

$$\mathbb{P}_i(s) = \frac{\min\{K_i, s\}}{s} = \begin{cases} 1 & \text{if } s \le K_i \\ \frac{K_i}{s} & \text{otherwise} \end{cases}$$

For a given *a priori* solution we can, thus, evaluate its recourse function. Therefore, we can write:

$$\begin{aligned}
&\mathbb{E}_\xi(\text{Service cost} + \text{Penalty cost}) \\
&= \sum_{i \in I} \sum_{s=0}^{z_i} \mathbb{P}[\eta_i = s] \times \mathbb{E}[\text{Service cost} + \text{penalty cost} | \eta_i = s] \\
&= \sum_{i \in I} \sum_{s=0}^{K_i} \mathbb{P}[\eta_i = s] \times \mathbb{E}[\text{Service cost} | \eta_i = s] \\
&\quad + \sum_{i \in I} \sum_{s=K_i+1}^{z_i} \mathbb{P}[\eta_i = s] \times \mathbb{E}[\text{Service cost} + \text{penalty cost} | \eta_i = s]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in I}\sum_{s=0}^{K_i}\left[\mathbb{P}[\eta_i = s]\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]c_{ij}x_{ij}\right] \\
&\quad + \sum_{i\in I}\sum_{s=K_i+1}^{z_i}\left[\mathbb{P}[\eta_i = s]\left(\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]\mathbb{P}_i(s)c_{ij}x_{ij} + g(s-K_i)\right)\right] \\
&= \sum_{i\in I}\sum_{s=0}^{z_i}\left[\mathbb{P}[\eta_i = s]\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]\mathbb{P}_i(s)c_{ij}x_{ij}\right] \\
&\quad + \sum_{i\in I}\left[g\sum_{s=K_i+1}^{z_i}\mathbb{P}[\eta_i = s](s-K_i)\right] \\
&= \sum_{i\in I}\sum_{s=0}^{z_i}\left[\mathbb{P}[\eta_i = s]\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]\frac{\min\{K_i, s\}}{s}c_{ij}x_{ij}\right] \\
&\quad + g\sum_{i\in I}\sum_{s=K_i+1}^{z_i}\mathbb{P}[\eta_i = s](s-K_i)
\end{aligned}
$$

Thus, the objective function (1) becomes

$$
\begin{aligned}
&\sum_{i\in I} f_i y_i + \sum_{i\in I}\sum_{s=0}^{z_i}\left[\mathbb{P}[\eta_i = s]\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]\frac{\min\{K_i, s\}}{s}c_{ij}x_{ij}\right] \\
&+ g\sum_{i\in I}\sum_{s=K_i+1}^{z_i}\mathbb{P}[\eta_i = s](s-K_i)
\end{aligned}
$$

In the general case, when the probabilities p_j are not necessarily equal for all customers, the probability distribution of the η_i variables is quite involved. For a given solution (y, x) recall that $J_i = \{j \in J : x_{ij} = 1\}$. Then,

$$
\begin{aligned}
\mathbb{P}[\eta_i = s] &= \sum_{S\subset J_i:|S|=s}\prod_{j\in S}p_j\prod_{j'\notin S}(1-p_{j'}) \qquad (7)\\
&= \frac{1}{s}\sum_{j\in J_i}p_j\mathbb{P}[\eta_i - \xi_j = s-1] \\
&= \cdots \\
&= \frac{1}{s!}\sum_{j_1\in J_i}\left(p_{j_1}\sum_{j_2\in J_i\setminus\{j_1\}}\left[p_{j_2}\cdots\sum_{j_s\in J_i\setminus\{j_1,\ldots,j_{s-1}\}}\left(p_{j_s}\prod_{j\neq j_1,j_2,\ldots,j_s}(1-p_j)\right)\right]\right).
\end{aligned}
$$

However, when $p_j = p\ \forall j \in J$, η_i is a Binomial random variable with parameters z_i and p. Thus, $\mathbb{P}[\eta_i = s] = \binom{z_i}{s}p^s(1-p)^{z_i-s}$, $s = 0,\ldots,z_i$. In this case, taking into account that $(\eta_i - 1)|_{(\xi_j=1)} \sim Bin(z_i - 1, p)$, we have:

$$
\begin{aligned}
\mathbb{P}[\xi_j = 1|_{\eta_i=s}] &= \frac{\mathbb{P}[\eta_i = s|_{\xi_j\,=\,1}]\mathbb{P}[\xi_j = 1]}{\mathbb{P}[\eta_i = s]} = \frac{\mathbb{P}[\eta_i - 1 = s-1||_{\xi_j\,=\,1}]\mathbb{P}[\xi_j = 1]}{\mathbb{P}[\eta_i = s]} \\
&= \frac{\binom{z_i-1}{s-1}p^s(1-p)^{z_i-s}p}{\binom{z_i}{s}p^s(1-p)^{z_i-s}} = \frac{s}{z_i}.
\end{aligned}
$$

and we obtain the following closed expression for the objective function:

$$\sum_{i\in I} f_i y_i + \sum_{i\in I}\sum_{s=0}^{z_i}\left[\binom{z_i}{s} p^s(1-p)^{z_i-s}\frac{\min\{K_i,s\}}{z_i}\sum_{j\in J} c_{ij}x_{ij}\right]$$
$$+\sum_{i\in I}\left[g\sum_{s=K_i+1}^{z_i}\left(\binom{z_i}{s} p^s(1-p)^{z_i-s}(s-K_i)\right)\right]. \tag{8}$$

2.1 Normal approximation for the probability distribution of total demand for open plants.

The evaluation of the expression (7) becomes unaffordable as z_i increases. However, for values of z_i large enough, the central limit theorem allows us to approximate the distribution of η_i by a normal law.

Indeed, since the expected value of η_i is $\mu_i = \sum_{j\in J_i} p_j$, and its variance, $\sigma_i^2 = \sum_{j\in J_i} p_j(1-p_j)$, we have

$$\frac{\eta_i - \mu_i}{\sigma_i} \approx N(0,1).$$

In addition, for a fixed $j \in J_i$, let $\widehat{\eta}_i^j = \sum_{k\in J_i\setminus j} \xi_k$. Then

$$\mathbb{E}[\widehat{\eta}_i^j] = \sum_{k\in J_i\setminus j} p_k = \widehat{\mu}_i^j;$$

and

$$var[\widehat{\eta}_i^j] = \sum_{k\in J_i\setminus j} p_k(1-p_k) = \widehat{\sigma}_i^j.$$

Thus, we also have

$$\frac{\widehat{\eta}_i^j - \widehat{\mu}_i^j}{\widehat{\sigma}_i^j} \approx N(0,1).$$

For a given solution to problem P, (y,x), we denote by $\widehat{I}$ the index set of opened facilities.

Proposition 1 *Let (y,x) be a solution to problem P. Then, for all $i \in \widehat{I}$ and $j \in J_i$,*

$$\mathbb{P}[\xi_j = 1|\eta_i = s] \sim \frac{\widehat{\alpha}_i^j(s)}{\alpha_i(s)} p_j,$$

where

$$\alpha_i(s) = \Phi\left(\frac{s+0.5-\mu_i}{\sigma_i}\right) - \Phi\left(\frac{s-0.5-\mu_i}{\sigma_i}\right),$$

and

$$\widehat{\alpha}_i^j(s) = \Phi\left(\frac{s-1+0.5-\widehat{\mu}_i^j}{\widehat{\sigma}_i^j}\right) - \Phi\left(\frac{s-1-0.5-\widehat{\mu}_i^j}{\widehat{\sigma}_i^j}\right).$$

Proof:

$$\mathbb{P}[\xi_j = 1|\eta_i = s] = \frac{\mathbb{P}[\eta_i = s|_{\xi_j = 1}]\mathbb{P}[\xi_j = 1]}{\mathbb{P}[\eta_i = s]}.$$

We have,

$$
\begin{aligned}
\mathbb{P}[\eta_i = s] &= \mathbb{P}\left[\frac{\eta_i - \mu_i}{\sigma_i} = \frac{s - \mu_i}{\sigma_i}\right] \\
&\approx \Phi\left(\frac{s + 0.5 - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{s - 0.5 - \mu_i}{\sigma_i}\right) \\
&= \alpha_i(s).
\end{aligned}
$$

Also,

$$
\begin{aligned}
\mathbb{P}[\eta_i = s|_{\xi_j} = 1] &= \mathbb{P}[\widehat{\eta}_i^j = s - 1] \\
&= \mathbb{P}\left[\frac{\widehat{\eta}_i^j - \widehat{\mu}_i^j}{\widehat{\sigma}_i^j} = \frac{s - 1 - \widehat{\mu}_i^j}{\widehat{\sigma}_i^j}\right] \\
&\approx \Phi\left(\frac{s - 1 + 0.5 - \widehat{\mu}_i^j}{\widehat{\sigma}_i^j}\right) - \Phi\left(\frac{s - 1 - 0.5 - \widehat{\mu}_i^j}{\widehat{\sigma}_i^j}\right) \\
&= \widehat{\alpha}_i^j(s).
\end{aligned}
$$

Therefore,

$$
\mathbb{P}[\xi_j = 1|\eta_i = s] = \frac{\widehat{\alpha}_i^j(s)}{\alpha_i(s)} p_j, \text{ and the result follows.}
$$

□

Proposition 2 *For any solution (y, x) to formulation P for which z_i is large enough, for all $i \in \widehat{I}$, the objective function can be approximated by*

$$
\sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} \left(c_{ij} x_{ij} p_j \sum_{s=0}^{z_i} \frac{\min\{K_i, s\}}{s} \widehat{\alpha}_i^j(s) \right) + g \sum_{i \in I} \sum_{s=K_i+1}^{z_i} \alpha_i(s)(s - K_i).
$$

Proof: The result holds by Proposition 1 since

$$
\begin{aligned}
&\sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{s=0}^{z_i} \left[\mathbb{P}[\eta_i = s] \sum_{j \in J} \mathbb{P}[\xi_j = 1|\eta_i = s] \frac{\min\{K_i, s\}}{s} c_{ij} x_{ij} \right] \\
&+ g \sum_{i \in I} \sum_{s=K_i+1}^{z_i} [\mathbb{P}[\eta_i = s](s - K_i)] \\
= &\sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{s=0}^{z_i} \left[\alpha_i(s) \sum_{j \in J} \frac{\widehat{\alpha}_i^j(s)}{\alpha_i(s)} p_j \frac{\min\{K_i, s\}}{s} c_{ij} x_{ij} \right] \\
&+ g \sum_{i \in I} \sum_{s=K_i+1}^{z_i} \alpha_i(s)(s - K_i)
\end{aligned}
$$

□

3 The Stochastic Assignment Subproblem

In this section we address the assignment problem that results when the set of operating facilities is known. Let $\widehat{I} \subset I$ denote the set of facilities that are opened and assume that $\sum_{i \in \widehat{I}} \ell_i \leq |J|$, to guarantee that there exists a feasible assignment on set $\widehat{J}$. We want to find an assignment of customers within the set of open facilities of minimum expected cost. This problem can be expressed as:

$$
\begin{aligned}
(AP(\widehat{I}))\quad \min \quad & \sum_{i\in\widehat{I}}\sum_{s=0}^{z_i}\left[\mathbb{P}[\eta_i = s]\sum_{j\in J}\mathbb{P}[\xi_j = 1|\eta_i = s]\frac{\min\{K_i, s\}}{s}c_{ij}x_{ij}\right] \\
& +g\sum_{i\in\widehat{I}}\sum_{s=K_i+1}^{z_i}\mathbb{P}[\eta_i = s](s-K_i) && (9)\\
\text{s. t.}\quad & \sum_{i\in\widehat{I}} x_{ij} = 1 \qquad \forall\, j \in J && (10)\\
& \ell_i \le \sum_{j\in J} x_{ij} \qquad \forall\, i \in \widehat{I} && (11)\\
& x_{ij} \in \{0,1\} \qquad \forall\, i \in \widehat{I}, j \in J && (12)
\end{aligned}
$$

For a given set of operating facilities $\widehat{I}$ it is not a simple task to solve $AP(\widehat{I})$ optimally. For this reason we propose a method for obtaining good quality solutions to the problem by solving a deterministic minimum cost flow problem in an auxiliary network $N(\widehat{I}) = (V, A)$ defined as follows:

- The set V has a vertex associated with each customer $j \in J$, two vertices associated with each open facility $i \in \widehat{I}$, and one additional vertex, that we denote v_0. Let us denote $V = \{v_r : r \in R\}$, where $R = \{J\} \cup \widehat{I} \cup \widehat{I}' \cup \{0\}$, being $\widehat{I}'$ a copy of the set of indices of the facilities in $\widehat{I}$.

- The set of arcs contains arcs (v_j, v_i), $\forall j \in J, \forall i \in \widehat{I}$; arcs $(v_i, v_{i'})$, where $i \in \widehat{I}$, $i' \in \widehat{I}'$ and $v_{i'}$ is the copy of vertex v_i; and arcs $(v_{i'}, v_0)$, $\forall i' \in \widehat{I}'$.

- The vector of productions/demands $b = (b_r)_{r\in R}$ is defined as follows: all vertices v_j, $j \in J$ are production ones with $b_j = 1$. Vertices v_i, $i \in \widehat{I}$, are transhipment ones, i.e., $b_i = 0$, $i \in \widehat{I}$. The remaining vertices are demand ones: $b_{i'} = \ell_i$, $\forall i \in \widehat{I}$ and $b_0 = n - \sum_{i\in\widehat{I}}\ell_i$ (recall that $n = |J|$). Thus, network N is balanced.

- Given the above productions and demands, the flows through arcs (v_j, v_i) will only take values 0 or 1, $\forall j \in J, i \in \widehat{I}$. They will take value 1 to represent that customer j is assigned to plant i. Thus, the flow through arcs $(v_i, v_{i'})$ will represent the number of customers allocated to facility i.

- Arcs $(v_i, v_{i'})$, $i \in \widehat{I}$, have a maximum capacity $u_i = \max\left\{K_i, \frac{n}{\overline{p}}\frac{K_i}{\sum_{t\in\widehat{I}}K_t}\right\}$, where $\overline{p}$ denotes the average of all p_j, $j \in J$. The remaining arcs have no upper limit on their maximum flow. The rationale behind the previous expression is to choose auxiliary capacities large enough to allow to assign all the customers to the set $\widehat{I}$, and proportional to the service capacities, K_i.

- We define a cost function $\widehat{c}$ on the arcs such that all arcs have cost zero apart from the arcs (v_j, v_i), whose cost is $\widehat{c}_{ji} = c_{ij}\frac{p_j}{K_i}$.

Figure 1 depicts a general network built as described above. Note that the solution to the minimum cost flow problem in the above network gives the optimal solution to the assignment problem when the capacities of the facilities, K_i ($i \in I$) are so high that the capacity constraints on the plants are never binding. In the general case, the cost function $\widehat{c}$ weights the original assignment costs penalizing assignments to facilities with small capacities, and giving more weight to customers with higher probability of requesting the service.

4 The FLPBD under perfect information

In this section we study the FLPBD under perfect information. That is, we study the problem that results from a particular realization of the random vector, which is the same as to consider that we

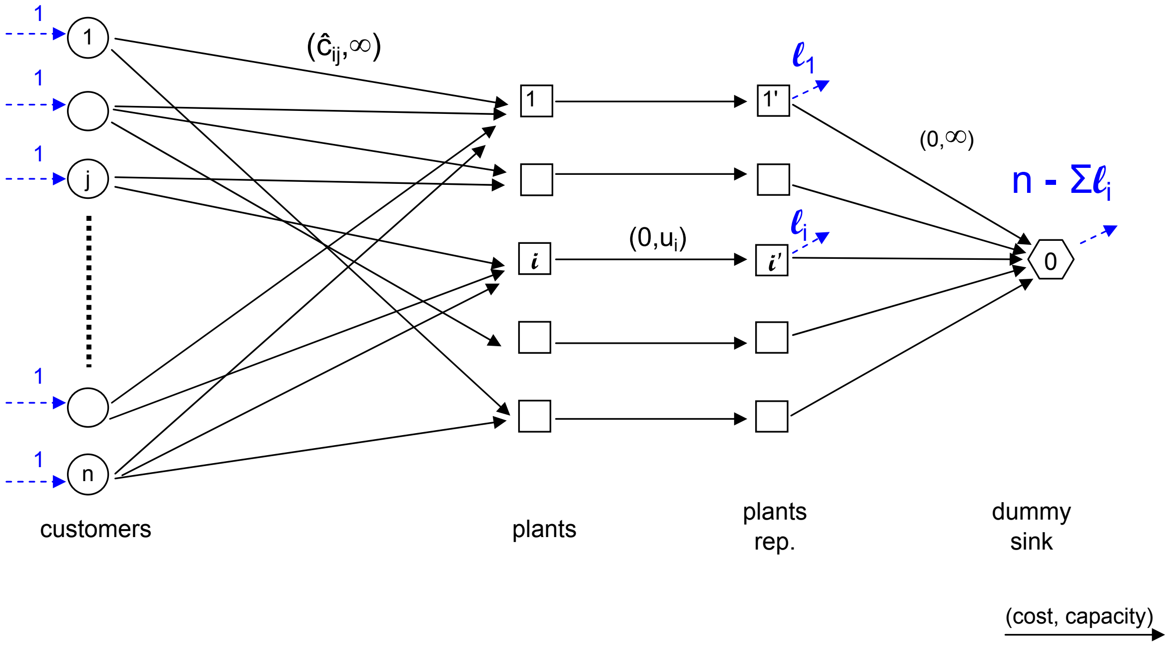

Figure 1: Auxiliary network to obtain an approximate solution to $AP(\widehat{I})$.

know in advance which customers will in fact request the service. We next study how to find an optimal solution of the problem for each realization of the random vector. By solving such a problem for all possible realizations of the random vector we obtain a population of optimal values, each of which with some probability. Finding this population and its corresponding probability distribution is known in the literature as the *distribution problem* (see for instance, Birge and Louveaux [8]). In particular, the mean value of the above mentioned distribution gives the so-called *wait-and-see value*. In practice, in general, it is very difficult to find the exact distribution of the above mentioned population as well as the wait-and-see value. This is also true for the particular case of the FLPBD. However, the possibility of solving the problem under perfect information allows us to estimate this latter value, for instance, by simulation. By estimating the wait-and-see value we get a way for measuring the degree of stochasticity of a particular instance although, as pointed out in Birge and Louveaux [8] no such measure is absolute for evaluating how stochastic an instance is.

In the FLPBD, a realization of the customers demands is not enough to define perfect information. Given that the number of customers with demand assigned to an open facility i may exceed its capacity K_i, and given that the number of customers that can be served from facility i cannot exceed K_i, for having perfect information we also need to know the criterion according to which we decide whether of not a customer with demand is going to be served. In this paper we assume that we use a first come first serve policy. Hence, perfect information is given by both the set of customers that have demand and also the order in which requests of service "arrive" at the system.

For a given realization ξ, denote by $J_\xi = \{j \in J : \xi_j = 1\}$ the set of indices of demand customers, $n_\xi = |J_\xi|$, and by $(\pi(1), \dots, \pi(n_\xi))$ the order by which their requests of service arrive, so that $\pi(q)$ is the index of the customer corresponding to the q-th request for service, $q = 1, \dots, n_\xi$. For formulating the FLPBD under perfect information, in addition to the y variables, that represent the location decisions, and the x variables, that represent *a priori* assignments of customers, we define a new set of variables, denoted w, that represent the served customers in the *a posteriori* solution, i.e. the demand customers that are actually served in the solution to the problem.

$$w_{ij} = \begin{cases} 1 & \text{if customer } j \text{ is served from facility } i \\ 0 & \text{otherwise} \end{cases} \qquad (i \in I,\ j \in J_\xi)$$

The problem under perfect information can then be formulated as follows:

$$(PI(\xi,\pi)) \quad \min \quad \sum_{i\in I} f_i y_i + \sum_{i\in I}\sum_{j\in J_\xi} c_{ij} w_{ij} + g \sum_{j\in J_\xi}\left(1 - \sum_{i\in I} w_{ij}\right) \tag{13}$$

$$\text{s. to:} \quad \sum_{i\in I} x_{ij} = 1 \qquad j \in J \tag{14}$$

$$x_{ij} \leqslant y_j \qquad i \in I, j \in J \tag{15}$$

$$\sum_{j\in J_\xi} w_{ij} \leqslant K_i \qquad i \in I \tag{16}$$

$$w_{ij} \leqslant x_{ij} \qquad i \in I, j \in J_\xi \tag{17}$$

$$\ell_i y_i \leqslant \sum_{j\in J} x_{ij} \qquad i \in I \tag{18}$$

$$w_{i\pi(s)} - w_{i\pi(r)} \leqslant 2 - x_{i\pi(r)} - x_{i\pi(s)} \qquad i \in I, r < s \in \{1,\ldots,n_\xi\} \tag{19}$$

$$y_i, w_{ij} \in \{0,1\} \qquad i \in I, j \in J_\xi \tag{20}$$

$$x_{ij} \in \{0,1\} \qquad i \in I, j \in J \tag{21}$$

In this formulation, constraints (14) and (15) together force each customer to be assigned to an operating facility. Constraints (16) state the service capacity of each plant, and constraints (18) avoid opening a plant that does not have enough customers assigned to it. Constraints (17) ensure that services are provided according to the assignment of customers to plants. Finally, constraints (19) assure that in each plant, a first call first serve policy is used to decide which customers to serve.

As mentioned before, having formulated the problem under perfect information it is possible to estimate the wait-and-see value, which is the expected value of $PI(\xi,\pi)$. To do so one possibility is to simulate the uncertain data, namely the customers that in fact have demand as well as the order by which they call to be served and to solve model $PI(\xi,\pi)$ with some general purpose software. By doing so for several simulations and by evaluating the average of the values obtained, we get an estimate of the wait-and-see value. The difference between the value of the best solution known for P and this estimate gives us an estimate of the expected value of perfect information.

5 A heuristic procedure

As mentioned in the previous section, one advantage of formulating the problem under perfect information (13)-(21) is to have the possibility of using simulation to obtain the optimal solutions to different realizations of the random vector. If we do so, the $\{x,y\}$ part of an optimal solution to each $PI(\xi,\pi)$ problem is, in fact, an *a priori* solution for the stochastic problem. Therefore, we can evaluate the objective function (8) using this solution. If the number of simulations is large enough, the best solution found (in terms of the objective function (8)) can be taken as an approximation of the optimal solution of the original stochastic problem. This methodology is a possibility for obtaining feasible solutions to the FLPBD. Nevertheless, it is a pure simulation procedure in the sense that feasible solutions are obtained by a simulation procedure and not using any well-structured methodology that takes advantage of the problem. Moreover, many simulations may be required before a good feasible solution is obtained.

In what follows, we propose a heuristic procedure for obtaining feasible solutions to the FLPBD. In the computational experiments section the results obtained using this heuristic will be compared with the ones obtained using the simulation procedure described above.

The heuristic we propose consists of three phases. The first one is a constructive phase that builds a feasible solution in a greedy fashion and does a rudimentary initial assignment. The second phase obtains the *a priori* assignment by solving the minimum cost flow problem associated with the assignment problem $AP(\widehat{I})$, where $\widehat{I}$ the set of indices of plants opened in the first phase. Finally, the third phase is an improvement one that performs Local Search on the assignment obtained in the second phase. The second phase was already described in Section 3. Accordingly, we detail now the first and third phases.

5.1 Constructive Phase

In the constructive phase we temporarily relax the constraints requiring each open plant to have a minimum number of assigned customers. Initially, one single facility is chosen to be opened and all customers are assigned to it. In a general step k a new plant is opened and some customers are reassigned to it, giving raise to a feasible solution $(\tilde{y}^k, \tilde{x}^k)$. To choose the plant to open we estimate, for each closed plant, the incremental cost that would be incurred by opening it. The chosen plant is the one leading to the lowest negative incremental cost, if any. The procedure stops when all incremental costs are non-negative.

In each step k, denote by I^k the set of facilities that are open, and by $i(j)$ the plant to which customer j is currently assigned ($j \in J$). Let also s_i^k denote the number of customers assigned to a facility $i \in I^k$ (i.e. $s_i^k = |\{i : i(j) = i\}|$). Then, for each closed plant $i \in I \setminus I^k$, we compute the incremental cost for opening it, say Δ_i^k, as the sum of the set-up cost of facility i plus an estimation of the variation in the assignment costs when reassigning customers from its current assignment, taking into account the expected saving in the penalty for unserved customers. If one or several closed facilities have a negative incremental cost, we choose to open the one with the most negative incremental cost. The procedure continues until all the closed facilities have a non-negative incremental cost. In the initialization step, for each plant i, the maximum number of customers we allow to be assigned to facility i is $u_i = \max\left\{K_i, \frac{n}{\overline{p}} \frac{K_i}{\sum\limits_{t \in I} K_t}\right\}$ following the same rationale as in the definition of the network of Figure 1 in Section 3 ($\overline{p}$ denotes the average of the values p_j, $j \in J$). We recall that z_i denotes the number of customers assigned to facility i ($i \in I$). The procedure can be formalized as follows:

Constructive phase for obtaining a feasible solution to FLPBD

1. Initialization:
   - Compute the initial auxiliary capacities u_i, $i \in I$, as follows
   $$u_i = min\left\{n, max\left\{K_i, \frac{nK_i}{\overline{p}\sum_{t \in I} K_t}\right\}\right\}$$
   - For $i \in I$ let $\Delta_i^0 = \frac{f_i + \overline{c}_i}{u_i}$, where $\overline{c}_i$ denotes the average assignment cost to facility i.
   - Let $i^0 \in I$ be such that $\Delta_{i^0}^0 = \min_{i \in I} \Delta_i^0$.
   - $I^0 = \{i^0\}$; $i(j) = i^0$, $\forall j \in J$; $z_{i^0} = n$, $z_i = 0$, $i \in I \setminus \{i^0\}$; $k = 1$.
2. While ($k = 1$ or $\Delta_{i^{k-1}}^{k-1} < 0$) do
   - For each $i \in I$, $j \in J$, compute the value $\delta_{ij} = c_{ij} - c_{i(j)j} - p_j \times g\frac{(z_{i(j)} - K_{i(j)})^+}{z_{i(j)}}$.
   - Sort the customers by non-decreasing order of the values δ_{ij}.
     Denote by $j_{[q]}$ ($q = 1, \dots, n$) the q-th customer after sorting.
   - For $i \in I$, $\Delta_i^k = f_i + \sum\limits_{t=1}^{r} \delta_{ij_{[t]}}$, with $r = \min\{\max\{\ell_i, \max\{q : \delta_{ij_{[q]}} < 0\}\}, u_i\}$
   - Let $i^k \in I \setminus I^{k-1}$ be such that $\Delta_{i^k}^k = \min_{i \in I \setminus I^{k-1}} \Delta_i^k$.
   - If $\Delta_{i^k}^k < 0$
     $I^k := I^{k-1} \cup \{i^k\}$; $z_{i(j)} = z_{i(j)} - 1$, $j = j_{[1]}, \dots, j_{[r]}$; $i(j) = i^k$, $j = j_{[1]}, \dots, j_{[r]}$;
     $z_{i^k} = z_{i^k} + r$; $k := k + 1$

As mentioned before, at termination of the constructive phase, we apply a second phase where an *a priori* assignment is obtained by solving the minimum cost flow problem of Section 3 as an approximation of the assignment problem $AP(\widehat{I^k})$.

5.2 Local Search

In the local search phase we consider two types of moves: reassignments and interchanges. In both cases we estimate the variation in the total cost, instead of computing it exactly. The exact value of a solution is computed after the constructive phase, and after the local search phase.

The two neighborhoods that we use in the local search are defined by reassignments of customers within the set of open plants, and by interchanges of assignments of pairs of customers, respectively. Below, we propose an estimate of the variation of the objective function value in each case:

Reassignments

An estimate for the variation in cost when reassigning customer j from plant $i_1 = i(j)$ to i_2 is:

$$\delta_{j,i_2} = c_{i_2 j} - c_{i_1 j} + g\left(\frac{p_j}{z_{i_2}+1}\max\{z_{i_2}+1-K_{i_2},0\} - \frac{p_j}{z_{i_1}}\max\{z_{i_1}-K_{i_1},0\}\right)$$

In this expression, the term $\frac{p_j}{z_{i_2}+1}\max\{z_{i_2}+1-K_{i_2},0\}$ is an estimate of the increase of the expected penalty in plant i_2, that will now have one more customer, whereas the term $\frac{p_j}{z_{i_1}}\max\{z_{i_1}-K_{i_1},0\}$ is an estimate of the reduction of the expected penalty in plant i_1.

Interchanges

When interchanging the assignment of two customers j_1, j_2, if we take $i_1 = i(j_1), i_2 = i(j_2)$ and we assume (without loss of generality) that $p_{j_1} > p_{j_2}$, the estimate for the cost variation is:

$$\delta_{j_1 j_2} = c_{i_1 j_2} + c_{i_2 j_1} - c_{i_1 j_1} - c_{i_2 j_2} + (p_{j_1} - p_{j_2})g\left(\frac{1}{z_{i_1}}\max\{z_{i_1}-K_{i_1},0\} - \frac{1}{z_{i_2}}\max\{z_{i_2}-K_{i_2},0\}\right)$$

Note that for the situation in which $p_j = p\,\forall j$, the penalty term is 0.

6 Computational Experiments

In this section we report the computational experience performed to evaluate the methodologies presented above applied to the homogeneous case ($p_j = p$, $j \in J$).

Due to the fact that no benchmark instances exist for the FLPBD, we considered two sets of instances existing in the literature for the capacitated single sourcing facility location problem namely, those presented by Holmberg *et al.* [16] and Díaz and Fernández [14] (publicly available at http://www-eio.upc.es/∼elena/sscplp/).

For each instance in each set, the demands of the customers were ignored, since now they follow a Bernoulli probability distribution, and new capacities K_i have been generated, coherent with the Bernoulli demands, as follows. Each instance produced four instances for the FLPBD according to:

1. $p_j = 0.25\ \forall j \in J$, $\ell_i = 0\ \forall i \in I$, and K_i are generated as described below $\forall i \in I$.
2. $p_j = 0.25\ \forall j \in J$. For each $i \in I$, K_i is generated as described below. Then, ℓ_i is set equal to the minimum between the nearest integers to $K_i/2$ and $n/4$.
3. $p_j = 0.75\ \forall j \in J$, $\ell_i = 0\ \forall i \in I$, and K_i are generated as described below $\forall i \in I$.
4. $p_j = 0.75\ \forall j \in J$. For each $i \in I$, K_i is generated as described below. Then, ℓ_i is set equal to the minimum between the nearest integers to $K_i/2$ and $n/4$.

For each instance, the capacities of the plants are generated as follows:

1. Generate a random number ρ in the set $\{1,2,3,4,5\}$, which represents a lower bound on the value that we accept for each K_i, $i \in I$.

2 For each plant $i \in I$

2.1 Define $\gamma_i = \frac{f_i}{\bar{c}_i}$, where $\bar{c}_i$ denotes the average assignment cost to facility i.

2.2 Define $\Gamma = \sum_{i \in I} \gamma_i$.

2.3 Define $\lambda_i = \frac{3}{2}\vartheta_i \times \frac{n \times p}{\Gamma}$, where ϑ_i is a random number in $[0.9\gamma_i; 1.1\gamma_i]$.

2.4 If $\lambda_i < \rho$ set $K_i = \rho$; else, if $\lambda_i > n$ set $K_i = n$; else, set $K_i = \lambda_i$.

In general terms, the capacities of the facilities are generated in such a way that the expected overall capacity is equal or close to $1.5np$ (recall that we are considering $p_j = p$, $j \in J$). Note that if $K_i = \lambda_i$ $\forall i \in I$, then the expected overall capacity is exactly $1.5np$, which is not the case if some $\lambda_i < \rho$ or $\lambda > n$. However, we expect the final value to be close to the target.

Note also that for each instance we first set the values p_j ($j \in J$) and then we determine K_i for all $i \in I$ according to the steps above. Instances of types 1 and 2 share the same vector of K_i. The same happens with instances of types 3 and 4. This way, in both cases, for the same vector of K_i two possible vectors of ℓ_i are considered.

Finally, for each instance, the penalty cost g was set equal to $\max_{i \in I, j \in J} c_{ij}$.

All the instances have a label with three components that allow an easy identification:

- A letter indicating the original set upon which the instance was built. 'D' denotes an instance from the set considered in Díaz and Fernández [14] and 'H' denotes one instance coming from the set presented in Holmberg *et al.* [16].
- The label of the original instance.
- The type of instance for FLPBD (1, 2, 3, or 4) as described above.

For example, 'D_p2_3' refers to an instance of type 3 built from instance p2 taken from Díaz and Fernández [14].

The computational results we present are based on instances p1-p33 from Díaz and Fernández [14] and instances p1-p24 from Holmberg *et al.* [16]. Taking into account the four types of instances built from each of these, in total we have 228 instances. The number of locations and the number of customers for these instances are given in the following table:

	Instance	$\lvert I \rvert$	$\lvert J \rvert$
Instances from Díaz and Fernández [14]	p1–p6	10	20
	p7–p17	15	30
	p18–p25	20	40
	p26–p33	20	50
Instances from Holmberg *et al.* [16]	p1–p12	10	50
	p13–p24	20	50

As it was mentioned in Section 4, in this paper we estimate the wait-and-see value by simulation. In particular, we consider the average of the optimal values of $PI(\xi, \pi)$ solved for the different realizations randomly generated for the random vectors. These optimal values are obtained using ILOG CPLEX 11.0.

It is a well-known result from stochastic programming (see, for instance, Kall and Wallace [20]) that in a minimization problem, the wait-and-see value can never be greater than the optimal value of the stochastic problem. However, this is not necessarily true (and in general it is not) for all realizations of the random vector. Accordingly, it may happen that the above mentioned average (that estimates the wait-and-see value) is greater than the value of some feasible solution found. When this is the case we take for the wait-and-see estimate the value of such feasible solution.

We defined as a stopping criterion for the simulation procedure an average variance below 10% of the average itself because this is a good indication that a very 'stable' value was achieved for the average. However, to overcome situations in which the first iterations lead to very similar values, we set a minimum number of 30 simulations. On the other hand, to make sure that the process would stop within a reasonable number of simulations we set a maximum number of 7500 simulations.

All computational tests were performed on a Pentium(R) 4, 3.2 GHz, 1.0 GB of Ram. All procedures were integrated in a in C++ code. For solving the network flow problem of Section 3 we used the callable library and the Network Optimizer of the general solver ILOG CPLEX 11.0 [18]. For solving the problem under perfect information ($PI(\xi, \pi)$) proposed in Section 4 (and, consequently, to obtain an estimation of the wait-and-see value) the same solver was combined with ILOG Concert 2.5 Technology [17].

For the two basic sets of instances, Figures 2 and 3 depict the gaps (%) with respect to the wait-and-see estimated value of the best solution found (during the simulation procedure) as well as of the solution found by the heuristic. In particular, Figure 2 refers to the instances 'H' while Figure 3 refers to the instances 'D'. Observing these two figures one realizes that the gaps (with respect to the wait-and-see estimated value) are much larger in instances 'H' than in instances 'D'. This is an indication that the first set of instances may be more 'stochastic' than the second one, which is not surprising when we look closely at the instances. In instances 'H', the ranges for the different set-up costs are small (instances H_p1 - H_p13) or even zero (instances H_p14 - H_p24). Moreover, the set-up costs have a magnitude similar to the demand satisfaction costs (often we find the latter greater than the former). Accordingly, in the deterministic version of the problem, there can be "many" quasi-optimal solutions (more or less similar in terms of their value). However, in the stochastic version, each realization might have a different optimal solution depending on where the demand is located. In this situation the optimal solution to the deterministic version does not coincide often with the optimal solution associated with a realization in the stochastic version. On the contrary, in instances 'D' the ranges for the different set-up costs are large, which, in turn, are considerably larger than the demand satisfaction costs. Therefore, often, some combinations of open facilities might be "good" for many different realizations of the random vector. In such cases, the optimal solution to the deterministic version might be also optimal for a realization. In fact, in this problem, more variability in the data (and in the magnitude of the set up costs versus demand satisfaction costs) lead to 'less stochasticity'.

Another conclusion that can be drawn by observing Figures 2 and 3 is that a change in the values ℓ_i does not have a great impact on the quality solutions obtained (either by the simulation procedure or by the heuristic). However, this is not the case when p increases. In Figures 2 and 3 note the differences in the scales used in upper part and in the lower part of the figures. In fact, the gaps with respect to the wait-and-see value of the solutions obtained are smaller for larger values of p_j.

In addition to Figures 2 and 3, detailed information on the computational tests performed can be found in Tables 1-4. In these tables, the first column presents the name of the instances. Columns 2-4 are associated with the wait-and-see value. Column 2 presents the estimate of this value, Column 3 presents the cpu time (in seconds) required for the simulations and Column 4 presents the number of simulations. Columns 5-7 refer to the best solution found during the simulations. In Column 5 we can observe the value of the best solution found whereas in Column 6 we can find the number of simulations performed until this solution was found. Column 7 presents the gap between the value of the best solution found and the wait-and-see estimated value. Columns 8 and 9 present the value obtained by the heuristic proposed in Section 5 and the gap between this value and the wait-and-see estimated value. In the last column one can see an estimate of the Expected Value of Perfect Information (EVPI). Note that Columns 7 and 9 detail the data depicted in Figures 2 and 3.

Observing Tables 1-4 we see that the number of simulations performed to estimate the wait-and-see value was much below the limit of 7500 that had been imposed. This is an indication that the average used to estimate the wait-and-see value becomes stable in a number of runs much lower that the imposed limit. Moreover, a clear difference can be found between 'H' instances and 'D' instances. The number of runs performed for estimating the wait-and-see value was much lower in the latter case as well as the time required to do it. Therefore, it is not surprising that the same type of conclusions can be drawn for the number of runs performed before the best solution was found (Column 6). Note, also, that the number of runs required until the best solution was found is much lower than the total number of runs. This means that a very good solution is obtained early in the simulation procedure.

By observing the last column in Tables 1-4 (which gives an estimate of the EVPI) we confirm the conclusion that was drawn above about the difference in the 'degree of stochasticity' between instances 'H' and 'D'. In fact, the EVPI is a measure (even if not always very robust) of the 'degree of stochasticity' of an instance (see, for instance, Birge and Louveaux [8]). As we can see, the EVPI is quite small in instances 'D' and often is equal to 0.

It is very important to observe the performance of the heuristic in the two sets of instances. In comparison with the best solution found during the simulation procedure, the heuristic performs much better in instances 'H' than in instances 'D'. This is very important because it gives evidence that the heuristic performs better in instances with a greater 'degree of stochasticity', which is a very important feature. In fact, the heuristic was designed for the stochastic case and a good behavior in instances with 'more stochasticity' is highly desirable. It should be pointed out that the cpu time required by the heuristic was always below 1 second (and this is the reason why the this cpu time is not presented in Tables 1-4).

As mentioned before, the possibility of working with the FLPBD in the non-homogeneous situation, relies on the approximation of the objective function stated in Proposition 2. Only for instances with a very small $|J|$ would we be able to evaluate the quality of the approximation. Accordingly, no results are presented for this situation because for instances with a size similar to the ones presented in Tables 1-4, it is already unaffordable to compute exactly the recourse function, which means that the quality of the approximation could not be measured. Therefore, any computational experience we presented for such case would be meaningless in the sense that the quality of the results could only be measured for instances where the conditions for the approximation do not hold.

7 Conclusions and directions for further research

In this paper we studied a facility location problem with Bernoulli demands. We proposed a two-stage stochastic programming model for the problem. We showed that the recourse function has a closed form. For the non-homogeneous case the exact evaluation of the recourse function becomes unaffordable when a large number of customers is assigned *a priori* to the operating facilities. In this case we used the central limit theorem to show that the recourse function can be approximated. We studied the stochastic assignment problem arising when a decision has been made on the operating facilities. Again, we propose approximating the optimal solution to this problem, which can be accomplished by solving a minimum cost network flow problem on an auxiliary network. We study the problem under perfect information. Finally, we propose a heuristic for finding an *a priori* solution to the problem. We present results from computational experience for the homogeneous case. The results indicate that it is possible to estimate the wait-and-see value as well as the expected value of perfect information with an affordable computational effort. Moreover, a good *a priori* solution can be found by using the simulation procedure that estimates the wait-and-see value. For instances with a large degree of 'stochasticity' the heuristic proposed can obtain even better results with a very small computational effort.

In the future, it would be interesting to study the possibility of extending the results and methodologies presented in this paper to more general situations namely, by considering other probability distributions for the demands and a multi-period setting for locating the facilities. Another opened direction for further research regards the algorithmic use of the normal approximation of the recourse function for the non-homogeneous case.

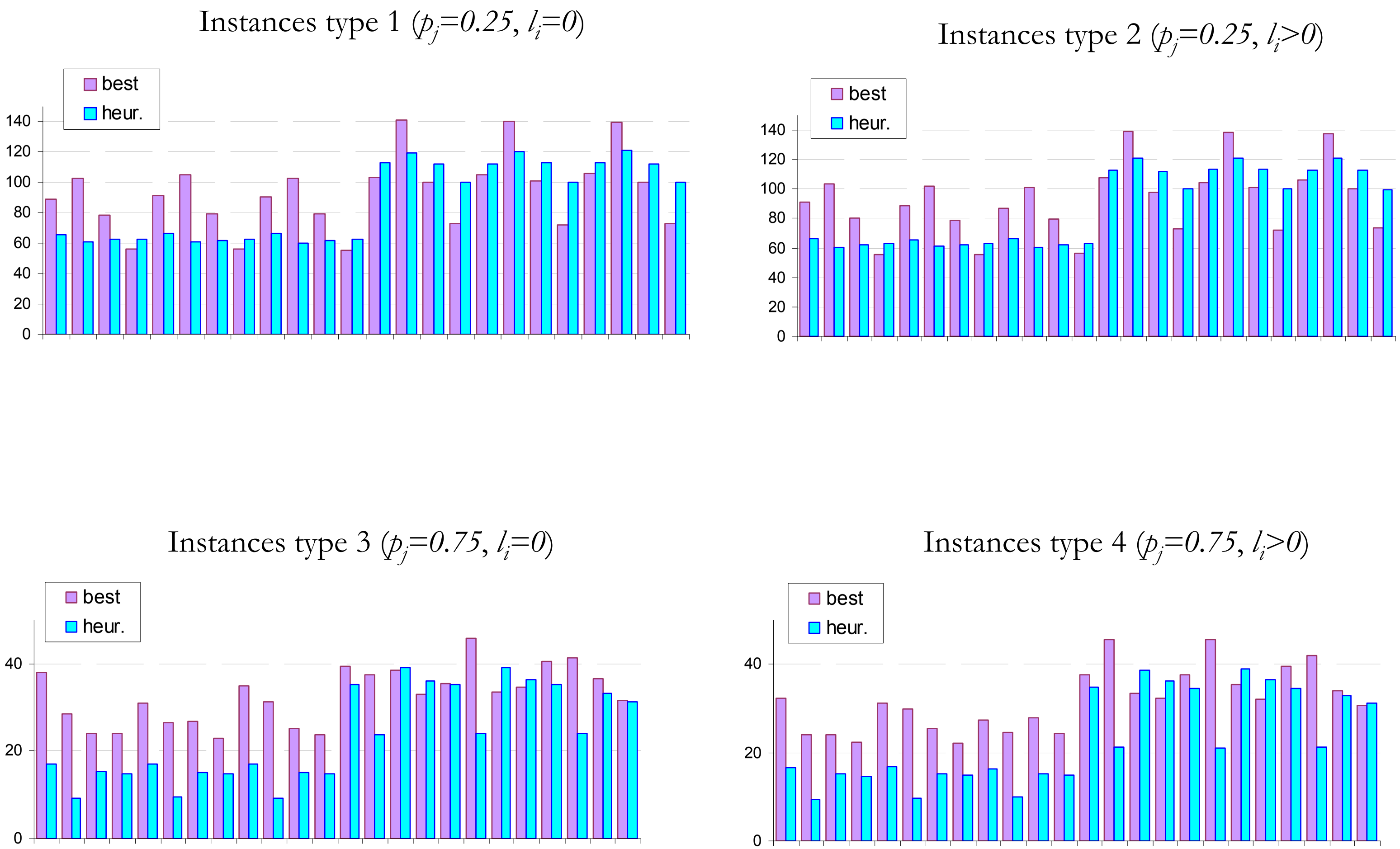

Figure 2: Gaps (%) with respect to the wait-and-see estimate - Instances from Holmberg *et al.* [16].

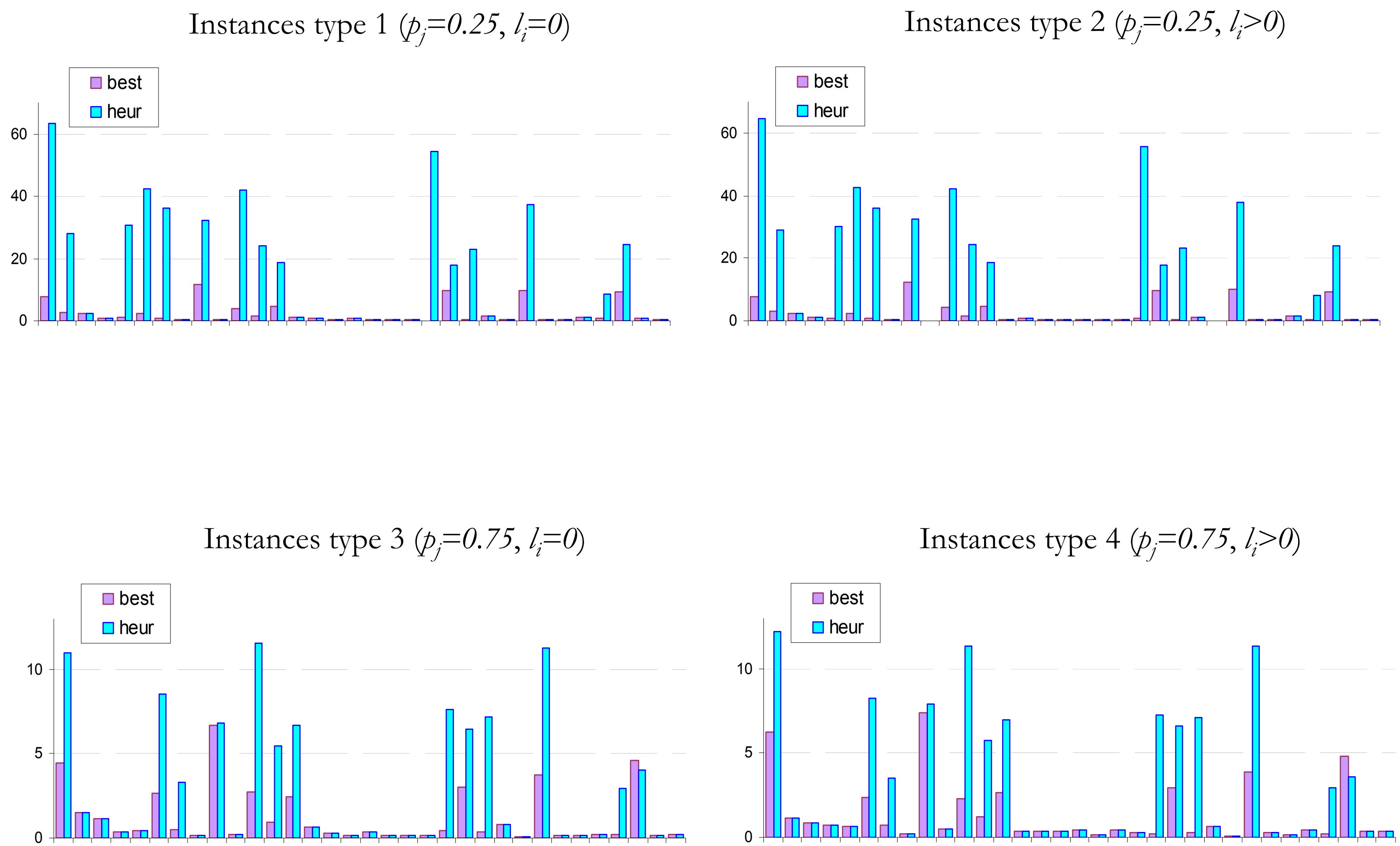

Figure 3: Gaps (%) with respect to the wait-and-see estimate - Instances from Díaz and Fernández [14].

	Wait-and-see value			Best solution - Simulations			Heuristic		EVPI
Instance	Estimate	Time	Runs	Value	Runs	Gap-WS (%)	Value	Gap-WS (%)	Value
D_p1_1	509,7	43,8	629	549,2	4	7,8	833,2	63,5	39,5
D_p2_1	682,5	36,1	566	699,3	1	2,5	873,8	28,0	16,8
D_p3_1	734,8	25,7	441	761,1	1	3,6	761,1	3,6	26,4
D_p4_1	743,0	24,3	547	743,0	1	0,0	743,0	0,0	0,0
D_p5_1	593,0	22,4	561	603,5	1	1,8	779,1	31,4	10,5
D_p6_1	504,8	61,9	700	531,2	3	5,3	739,6	46,5	26,5
D_p7_1	801,3	115,8	725	804,6	2	0,4	1084,1	35,3	3,3
D_p8_1	1208,0	46,5	475	1208,4	1	0,0	1208,4	0,0	0,4
D_p9_1	665,5	135,4	677	739,0	22	11,0	874,2	31,4	73,5
D_p10_1	1365,4	22,3	370	1373,2	1	0,6	1373,2	0,6	7,8
D_p11_1	738,0	188,1	744	765,8	1	3,8	1045,8	41,7	27,8
D_p12_1	750,1	144,9	740	758,7	1	1,2	929,4	23,9	8,7
D_p13_1	713,0	157,5	763	740,7	4	3,9	840,4	17,9	27,8
D_p14_1	937,4	70,8	615	944,6	1	0,8	944,6	0,8	7,2
D_p15_1	1003,5	66,7	530	1023,4	1	2,0	1023,4	2,0	19,9
D_p16_1	1214,5	45,8	454	1228,4	1	1,1	1228,4	1,1	13,9
D_p17_1	1378,3	35,0	388	1390,6	1	0,9	1390,6	0,9	12,3
D_p18_1	1231,7	148,0	616	1231,7	1	0,0	1231,7	0,0	0,0
D_p19_1	1591,0	99,8	499	1591,0	1	0,0	1591,0	0,0	0,0
D_p20_1	1599,9	113,4	497	1599,9	1	0,0	1599,9	0,0	0,0
D_p21_1	1006,4	353,2	810	1006,4	1	0,0	1552,9	54,3	0,0
D_p22_1	866,2	374,7	780	954,6	389	10,2	1026,1	18,5	88,4
D_p23_1	1075,1	307,1	690	1075,6	1	0,1	1318,9	22,7	0,6
D_p24_1	1273,1	169,6	570	1287,1	2	1,1	1287,1	1,1	13,9
D_p25_1	1419,5	105,4	516	1419,5	1	0,0	1419,5	0,0	0,0
D_p26_1	1130,3	654,6	686	1238,2	279	9,6	1550,4	37,2	108,0
D_p27_1	1574,7	213,4	625	1586,1	1	0,7	1586,1	0,7	11,4
D_p28_1	1498,2	229,5	661	1502,8	1	0,3	1502,8	0,3	4,6
D_p29_1	1595,3	233,4	518	1609,8	1	0,9	1609,8	0,9	14,5
D_p30_1	1356,1	562,7	655	1360,9	2	0,4	1467,9	8,3	4,9
D_p31_1	1115,9	624,8	768	1233,7	61	10,6	1408,0	26,2	117,8
D_p32_1	1649,1	205,0	559	1663,9	1	0,9	1663,9	0,9	14,8
D_p33_1	1870,2	174,6	494	1891,9	1	1,2	1891,9	1,2	21,7
H_p1_1	3906,6	632,5	2292	7293,2	781	86,7	6496,7	66,3	2590,1
H_p2_1	3305,2	328,4	1838	6624,7	1825	100,4	5342,5	61,6	2037,4
H_p3_1	4526,0	986,3	2672	7995,8	1072	76,7	7342,5	62,2	2816,6
H_p4_1	5686,4	1992,3	3364	8963,6	1700	57,6	9342,5	64,3	3277,2
H_p5_1	3926,0	628,8	2314	7332,2	1911	86,8	6496,7	65,5	2570,7
H_p6_1	3335,1	381,0	2093	6776,5	1490	103,2	5342,5	60,2	2007,5
H_p7_1	4540,0	932,7	2485	8144,7	1243	79,4	7342,5	61,7	2802,5
H_p8_1	5724,8	1933,0	3343	8992,5	877	57,1	9342,5	63,2	3267,8
H_p9_1	3925,1	626,9	2289	7396,0	1532	88,4	6496,7	65,5	2571,6
H_p10_1	3316,1	347,1	1905	6837,2	114	106,2	5342,5	61,1	2026,4
H_p11_1	4550,4	973,7	2567	8106,4	1571	78,2	7342,5	61,4	2792,1
H_p12_1	5723,5	1874,1	3275	8932,8	1202	56,1	9342,5	63,2	3209,3
H_p13_1	3833,7	1344,5	2061	7937,3	536	107,0	8231,5	114,7	4103,6
H_p14_1	3118,6	653,9	1531	7738,7	979	148,2	6888,9	120,9	3770,3
H_p15_1	4407,9	2233,0	2201	8986,8	272	103,9	9431,1	114,0	4578,9
H_p16_1	5726,7	6360,1	2927	9851,2	1173	72,0	11470,1	100,3	4124,5
H_p17_1	3864,5	1375,1	1996	7967,3	603	106,2	8231,5	113,0	4102,8
H_p18_1	3117,2	593,9	1478	7572,7	1416	142,9	6888,9	121,0	3771,7
H_p19_1	4450,3	2261,6	2092	8738,9	702	96,4	9431,1	111,9	4288,6
H_p20_1	5730,7	6412,1	2960	9807,5	658	71,1	11470,1	100,2	4076,8
H_p21_1	3852,2	1379,1	2063	7846,6	1805	103,7	8231,5	113,7	3994,4
H_p22_1	3107,5	553,4	1405	7715,7	12	148,3	6888,9	121,7	3781,4
H_p23_1	4448,5	2299,2	2164	8882,9	654	99,7	9431,1	112,0	4434,5
H_p24_1	5715,7	6130,9	2920	9823,5	1888	71,9	11470,1	100,7	4107,8

Table 1: Computational experience for type 1 instances ($p_j = 0.25$, $j \in J$ and $\ell_i = 0$, $i \in I$).

	Wait-and-see value			Best solution - Simulations			Heuristic		EVPI
Instance	Estimate	Time	Runs	Value	Runs	Gap-WS (%)	Value	Gap-WS (%)	Value
D_p1_2	497,0	48,6	700	549,2	4	10,5	839,0	68,8	52,2
D_p2_2	668,5	27,1	483	699,3	3	4,6	873,8	30,7	30,9
D_p3_2	732,3	22,5	457	761,1	6	3,9	761,1	3,9	28,8
D_p4_2	721,2	22,5	501	743,0	1	3,0	743,0	3,0	21,7
D_p5_2	595,2	25,6	643	603,5	1	1,4	779,1	30,9	8,3
D_p6_2	520,4	56,6	687	531,2	1	2,1	739,6	42,1	10,9
D_p7_2	794,5	107,8	625	804,6	1	1,3	1084,1	36,4	10,1
D_p8_2	1206,6	48,5	456	1208,4	1	0,1	1208,4	0,1	1,7
D_p9_2	669,3	131,4	611	739,0	154	10,4	871,3	30,2	69,7
D_p10_2	1373,2	28,0	426	1373,2	1	0,0	1373,2	0,0	0,0
D_p11_2	729,1	202,0	709	765,8	12	5,0	1045,8	43,4	36,6
D_p12_2	746,9	164,1	746	758,7	2	1,6	929,4	24,4	11,8
D_p13_2	702,9	173,8	760	740,7	2	5,4	840,4	19,6	37,8
D_p14_2	940,1	71,0	582	944,6	1	0,5	944,6	0,5	4,5
D_p15_2	1023,4	72,4	539	1023,4	1	0,0	1023,4	0,0	0,0
D_p16_2	1228,4	46,5	430	1228,4	1	0,0	1228,4	0,0	0,0
D_p17_2	1388,3	36,5	395	1390,6	1	0,2	1390,6	0,2	2,3
D_p18_2	1219,6	148,4	604	1231,7	1	1,0	1231,7	1,0	12,1
D_p19_2	1574,3	90,7	452	1591,0	1	1,1	1591,0	1,1	16,7
D_p20_2	1599,9	108,9	453	1599,9	1	0,0	1599,9	0,0	0,0
D_p21_2	1006,4	360,5	753	1006,4	1	0,0	1552,9	54,3	0,0
D_p22_2	867,2	380,3	759	962,3	690	11,0	1025,7	18,3	95,1
D_p23_2	1075,6	349,6	717	1075,6	1	0,0	1318,9	22,6	0,0
D_p24_2	1272,7	186,2	586	1287,1	2	1,1	1287,1	1,1	14,3
D_p25_2	1419,5	116,6	530	1419,5	1	0,0	1419,5	0,0	0,0
D_p26_2	1103,0	694,8	705	1238,2	656	12,3	1550,4	40,6	135,3
D_p27_2	1580,7	236,3	640	1586,1	1	0,3	1586,1	0,3	5,4
D_p28_2	1492,6	222,8	602	1502,8	1	0,7	1502,8	0,7	10,2
D_p29_2	1575,9	246,0	532	1609,8	1	2,2	1609,8	2,1	33,9
D_p30_2	1360,9	689,3	752	1360,9	1	0,0	1467,9	7,9	0,0
D_p31_2	1146,4	785,8	742	1256,8	645	9,6	1408,0	22,8	110,4
D_p32_2	1663,9	212,2	557	1663,9	1	0,0	1663,9	0,0	0,0
D_p33_2	1887,2	176,1	452	1891,9	1	0,3	1891,9	0,2	4,7
H_p1_2	3890,6	576,2	2174	7541,5	2014	93,8	6496,7	67,0	2606,2
H_p2_2	3332,8	354,6	1963	6689,2	1099	100,7	5342,5	60,3	2009,8
H_p3_2	4525,1	894,6	2614	8130,4	201	79,7	7342,5	62,3	2817,4
H_p4_2	5736,9	1983,1	3434	9015,0	2804	57,1	9342,5	62,8	3278,1
H_p5_2	3901,9	618,3	2271	7431,3	1925	90,5	6496,7	66,5	2594,8
H_p6_2	3319,7	350,6	1954	6794,2	313	104,7	5342,5	60,9	2022,8
H_p7_2	4545,1	881,1	2460	7996,9	921	75,9	7342,5	61,5	2797,4
H_p8_2	5738,4	1881,3	3333	8981,4	1007	56,5	9342,5	62,8	3243,0
H_p9_2	3889,4	650,3	2456	7282,1	676	87,2	6496,7	67,0	2607,3
H_p10_2	3341,3	360,3	1994	6778,6	335	102,9	5342,5	59,9	2001,2
H_p11_2	4522,2	850,3	2463	8089,4	2397	78,9	7342,5	62,4	2820,4
H_p12_2	5724,2	1824,6	3276	8999,9	864	57,2	9342,5	63,2	3275,7
H_p13_2	3848,1	1539,1	1951	7865,9	1560	104,4	8231,5	113,9	4017,8
H_p14_2	3116,3	703,1	1559	7500,5	1054	140,7	6888,9	121,1	3772,6
H_p15_2	4472,5	2929,9	2163	8744,3	115	95,5	9431,1	110,9	4271,8
H_p16_2	5745,6	10167,7	2892	9962,5	416	73,4	11470,1	99,6	4216,9
H_p17_2	3849,2	1531,9	1996	7925,9	1338	105,9	8231,5	113,8	4076,7
H_p18_2	3138,6	704,3	1510	7355,9	166	134,4	6888,9	119,5	3750,2
H_p19_2	4430,5	2877,1	2198	8897,7	1673	100,8	9431,1	112,9	4467,2
H_p20_2	5730,0	10947,1	3042	9918,3	2795	73,1	11470,1	100,2	4188,3
H_p21_2	3891,9	1615,8	2066	8150,7	1971	109,4	8231,5	111,5	4258,8
H_p22_2	3123,6	696,6	1491	7423,1	352	137,7	6888,9	120,5	3765,3
H_p23_2	4449,8	2910,1	2222	8966,5	698	101,5	9431,1	111,9	4516,7
H_p24_2	5726,5	9131,1	2831	9938,0	705	73,5	11470,1	100,3	4211,4

Table 2: Computational experience for type 2 instances ($p_j = 0.25$, $j \in J$. ℓ_i equal to the minimum between the integers closest to $K_i/2$ and $n/4$, $(i \in I)$).

	Wait-and-see value			Best solution - Simulations			Heuristic		EVPI
Instance	Estimate	Time	Runs	Value	Runs	Gap-WS (%)	Value	Gap-WS (%)	Value
D_p1_3	1481,7	196,4	234	1525,6	18	3,0	1620,0	9,3	44,0
D_p2_3	1651,7	140,7	244	1688,4	1	2,2	1688,4	2,2	36,7
D_p3_3	1730,8	82,2	233	1749,9	1	1,1	1749,9	1,1	19,1
D_p4_3	1721,0	42,3	237	1731,7	1	0,6	1731,7	0,6	10,7
D_p5_3	1589,5	60,8	246	1592,1	1	0,2	1592,1	0,2	2,6
D_p6_3	1473,2	297,6	275	1519,9	5	3,2	1607,0	9,1	46,7
D_p7_3	2268,0	571,7	318	2289,5	1	1,0	2352,6	3,7	21,5
D_p8_3	2693,2	162,6	178	2693,2	1	0,0	2693,2	0,0	0,0
D_p9_3	2064,5	612,1	227	2195,7	91	6,4	2189,8	6,1	125,3
D_p10_3	2858,1	94,3	202	2858,1	1	0,0	2858,1	0,0	0,0
D_p11_3	2184,9	864,3	290	2242,9	84	2,7	2435,4	11,5	58,0
D_p12_3	2222,6	636,8	264	2243,6	2	1,0	2343,8	5,5	21,0
D_p13_3	2154,7	689,8	255	2217,6	197	2,9	2307,8	7,1	62,9
D_p14_3	2419,1	243,0	209	2429,5	1	0,4	2429,5	0,4	10,4
D_p15_3	2508,2	228,0	206	2508,2	1	0,0	2508,2	0,0	0,0
D_p16_3	2713,2	164,8	193	2713,2	1	0,0	2713,2	0,0	0,0
D_p17_3	2865,4	142,8	204	2875,5	2	0,4	2875,5	0,4	10,1
D_p18_3	3210,5	706,5	208	3211,7	1	0,0	3211,7	0,0	1,2
D_p19_3	3571,0	353,8	194	3571,0	1	0,0	3571,0	0,0	0,0
D_p20_3	3565,8	449,8	195	3579,9	1	0,4	3579,9	0,4	14,1
D_p21_3	2964,9	1231,7	220	2986,4	1	0,7	3200,4	7,9	21,5
D_p22_3	2822,3	2168,7	252	2886,4	57	2,3	2983,8	5,7	64,1
D_p23_3	3046,1	1193,1	215	3055,6	1	0,3	3264,3	7,2	9,5
D_p24_3	3256,1	1013,3	226	3267,1	2	0,3	3267,1	0,3	10,9
D_p25_3	3399,5	422,6	230	3399,5	1	0,0	3399,5	0,0	0,0
D_p26_3	3529,1	5948,0	322	3655,0	304	3,6	3937,5	11,6	125,9
D_p27_3	4049,9	817,1	220	4061,1	1	0,3	4061,1	0,3	11,2
D_p28_3	3977,8	1076,6	223	3977,8	1	0,0	3977,8	0,0	0,0
D_p29_3	4084,8	1407,7	206	4084,8	2	0,0	4084,8	0,0	0,0
D_p30_3	3815,3	8194,3	300	3835,9	1	0,5	3941,4	3,3	20,6
D_p31_3	3587,3	5415,8	253	3721,6	149	3,7	3712,8	3,5	125,5
D_p32_3	4134,8	1055,0	226	4138,9	1	0,1	4138,9	0,1	4,1
D_p33_3	4343,5	622,6	211	4366,9	1	0,5	4366,9	0,5	23,5
H_p1_3	7401,5	315,7	350	9836,0	114	32,9	8648,4	16,8	1246,9
H_p2_3	6344,2	122,3	294	7605,4	264	19,9	6883,7	8,5	539,5
H_p3_3	7707,0	490,2	389	9406,6	216	22,1	8883,7	15,3	1176,7
H_p4_3	8998,6	1000,1	421	11901,6	145	32,3	10393,4	15,5	1394,8
H_p5_3	7433,1	366,6	414	9139,6	62	23,0	8648,4	16,4	1215,3
H_p6_3	6301,7	125,5	309	7530,0	128	19,5	6883,7	9,2	582,0
H_p7_3	7701,0	482,6	345	10000,8	113	29,9	8883,7	15,4	1182,7
H_p8_3	9073,5	1249,8	489	11712,3	80	29,1	10393,4	14,5	1319,9
H_p9_3	7369,2	358,7	403	10057,0	268	36,5	8648,4	17,4	1279,2
H_p10_3	6315,5	121,6	299	8094,7	212	28,2	6883,7	9,0	568,2
H_p11_3	7751,6	499,9	369	9490,1	318	22,4	8883,7	14,6	1132,1
H_p12_3	9055,7	993,9	386	11829,2	131	30,6	10393,4	14,8	1337,8
H_p13_3	8678,0	1617,8	439	12543,1	310	44,5	11742,7	35,3	3064,8
H_p14_3	7022,0	1105,1	412	9579,9	353	36,4	8677,5	23,6	1655,5
H_p15_3	9529,2	2429,5	626	13410,4	338	40,7	13332,5	39,9	3803,4
H_p16_3	12156,6	5434,9	810	16119,8	137	32,6	16517,5	35,9	3963,2
H_p17_3	8636,5	1897,5	523	11816,4	359	36,8	11742,7	36,0	3106,2
H_p18_3	7023,6	872,8	333	10638,6	331	51,5	8677,5	23,5	1653,9
H_p19_3	9578,5	2244,0	558	12628,4	141	31,8	13332,5	39,2	3049,9
H_p20_3	12076,5	4967,6	741	16456,1	85	36,3	16517,5	36,8	4379,6
H_p21_3	8687,2	2148,6	575	12115,6	136	39,5	11742,7	35,2	3055,5
H_p22_3	6976,8	995,6	387	9849,3	196	41,2	8677,5	24,4	1700,8
H_p23_3	9937,0	2575,8	568	13149,7	120	32,3	13204,4	32,9	3212,7
H_p24_3	12489,3	3828,8	760	16158,6	327	29,4	16404,4	31,3	3669,4

Table 3: Computational experience for type 3 instances ($p_j = 0.75$, $j \in J$ and $\ell_i = 0$, $i \in I$).

	Wait-and-see value			Best solution - Simulations			Heuristic		EVPI
Instance	Estimate	Time	Runs	Value	Runs	Gap-WS (%)	Value	Gap-WS (%)	Value
D_p1_4	1449,0	144,6	272	1529,6	115	5,6	1620,0	11,8	80,6
D_p2_4	1683,6	65,2	182	1688,4	2	0,3	1688,4	0,3	4,8
D_p3_4	1727,8	50,8	228	1749,9	3	1,3	1749,9	1,3	22,1
D_p4_4	1716,8	42,2	228	1731,7	2	0,9	1731,7	0,9	14,9
D_p5_4	1578,2	62,2	256	1592,1	1	0,9	1592,1	0,9	13,9
D_p6_4	1483,9	178,9	230	1519,9	3	2,4	1607,0	8,3	36,0
D_p7_4	2268,2	542,9	276	2289,5	1	0,9	2352,6	3,7	21,3
D_p8_4	2679,8	229,1	265	2693,2	1	0,5	2693,2	0,5	13,4
D_p9_4	2023,5	806,3	265	2183,7	264	7,9	2189,8	8,2	160,2
D_p10_4	2858,1	83,0	172	2858,1	1	0,0	2858,1	0,0	0,0
D_p11_4	2171,9	687,1	229	2236,4	214	3,0	2435,4	12,1	64,5
D_p12_4	2208,1	593,1	228	2243,6	2	1,6	2343,8	6,1	35,5
D_p13_4	2155,9	589,5	219	2215,0	142	2,7	2307,8	7,0	59,1
D_p14_4	2399,5	306,2	255	2429,5	1	1,3	2429,5	1,3	30,0
D_p15_4	2487,0	274,3	243	2508,2	1	0,9	2508,2	0,9	21,2
D_p16_4	2713,2	213,4	230	2713,2	1	0,0	2713,2	0,0	0,0
D_p17_4	2875,5	147,1	192	2875,5	1	0,0	2875,5	0,0	0,0
D_p18_4	3211,7	627,1	188	3211,7	1	0,0	3211,7	0,0	0,0
D_p19_4	3555,7	406,8	220	3571,0	1	0,4	3571,0	0,4	15,3
D_p20_4	3569,4	500,4	225	3579,9	1	0,3	3579,9	0,3	10,5
D_p21_4	2959,4	1314,5	209	2986,4	2	0,9	3197,3	8,0	27,0
D_p22_4	2787,5	2517,1	277	2882,4	274	3,4	2983,8	7,0	94,8
D_p23_4	3049,8	1717,9	282	3055,6	1	0,2	3264,3	7,0	5,8
D_p24_4	3233,0	1147,2	266	3267,1	1	1,1	3267,1	1,1	34,1
D_p25_4	3399,5	428,7	219	3399,5	1	0,0	3399,5	0,0	0,0
D_p26_4	3513,6	6118,8	288	3667,4	42	4,4	3937,5	12,1	153,8
D_p27_4	4061,1	848,1	218	4061,1	1	0,0	4061,1	0,0	0,0
D_p28_4	3977,8	1396,2	249	3977,8	1	0,0	3977,8	0,0	0,0
D_p29_4	4065,4	1627,9	202	4084,8	1	0,5	4084,8	0,5	19,4
D_p30_4	3815,3	8104,2	254	3835,9	1	0,5	3941,4	3,3	20,6
D_p31_4	3586,6	7797,8	276	3754,4	247	4,7	3712,8	3,5	126,2
D_p32_4	4119,8	1096,8	215	4138,9	1	0,5	4138,9	0,5	19,1
D_p33_4	4327,7	526,0	188	4366,9	1	0,9	4366,9	0,9	39,3
H_p1_4	7403,5	369,1	386	9250,8	338	25,0	8606,3	16,2	1202,8
H_p2_4	6288,6	136,1	315	6967,7	41	10,8	6883,7	9,5	595,1
H_p3_4	7736,9	513,6	393	9219,0	205	19,2	8883,7	14,8	1146,8
H_p4_4	9082,3	1080,9	416	11086,0	75	22,1	10393,4	14,4	1311,1
H_p5_4	7357,1	369,7	379	9817,8	214	33,5	8606,3	17,0	1249,2
H_p6_4	6295,2	128,4	301	7840,0	287	24,5	6883,7	9,3	588,5
H_p7_4	7694,4	566,0	404	9325,8	41	21,2	8883,7	15,5	1189,3
H_p8_4	9035,1	1195,0	481	10856,1	390	20,2	10393,4	15,0	1358,3
H_p9_4	7386,1	343,1	369	9402,9	20	27,3	8606,3	16,5	1220,2
H_p10_4	6255,6	125,4	294	7430,2	144	18,8	6883,7	10,0	628,1
H_p11_4	7691,0	550,8	416	10346,8	65	34,5	8883,7	15,5	1192,7
H_p12_4	9067,6	1240,6	475	10874,2	270	19,9	10393,4	14,6	1325,9
H_p13_4	8705,6	2094,6	518	11838,6	431	36,0	11701,5	34,4	2995,9
H_p14_4	7008,3	990,2	368	9931,0	86	41,7	8480,5	21,0	1472,2
H_p15_4	9597,1	2143,2	514	12666,6	377	32,0	13322,8	38,8	3069,5
H_p16_4	12148,0	5087,8	765	15713,6	626	29,4	16507,8	35,9	3565,6
H_p17_4	8714,5	2007,6	501	11250,8	441	29,1	11701,5	34,3	2536,2
H_p18_4	7013,9	977,2	353	9820,1	332	40,0	8480,5	20,9	1466,6
H_p19_4	9587,3	2276,5	545	12481,8	101	30,2	13322,8	39,0	2894,4
H_p20_4	12109,9	5047,8	784	16077,1	258	32,8	16507,8	36,3	3967,3
H_p21_4	8683,9	1793,0	443	12791,2	253	47,3	11701,5	34,7	3017,6
H_p22_4	6989,2	940,2	348	10815,6	124	54,8	8480,5	21,3	1491,3
H_p23_4	9921,3	2859,5	608	13328,2	57	34,3	13205,8	33,1	3284,5
H_p24_4	12527,0	4170,0	751	16449,2	113	31,3	16405,8	31,0	3878,9

Table 4: Computational experience for type 4 instances ($p_j = 0.75$, $j \in J$. ℓ_i equal to the minimum between the integers closest to $K_i/2$ and $n/4$, $i \in I$).

Acknowledgement

This research has been partially supported by the *Spanish Ministry of Science and Education*, Project MTM 2006-14961-C05-01, and by the *Portuguese Science Foundation*, Project POCTI - ISFL - 1 - 152.

References

[1] M. Albareda-Sambola, E. Fernández, and G. Laporte. Heuristic and lower bounds for a stochastic location routing problem. *European Journal of Operational Research*, 179:940–955, 2007.

[2] M. Albareda-Sambola, M.H. van der Vlerk, and E. Fernández. Exact solutions to a class of stochastic generalized assignment problems. *European Journal of Operational Research*, 173:465–487, 2006.

[3] I. Averbakh and O. Berman. Minimax regret p-center location on a network with demand uncertainty. *Location Science*, 5:247–254, 1997.

[4] I. Averbakh and O. Berman. Minmax regret median location on a network under uncertainty. *INFORMS Journal on Computing*, 12:104–110, 2000.

[5] I. Averbakh and O. Berman. An improved algorithm for the minmax regret median problem on a tree. *Networks*, 41:97–103, 2003.

[6] O. Berman and D. Simchi-Levi. Finding the optimal a priori tour and location of a traveling salesman with non homogeneous customers. *Transportation Science*, 22:148–154, 1988.

[7] L. Bianchi and A. M. Campbell. Extension of the 2-p-opt and 1-shift algorithms to the heterogeneous probabilistic traveling salesman problem. *European Journal of Operational Research*, 176:131–144, 2007.

[8] J.F. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer Series in Operations Research. Springer, New York, 1997.

[9] M.L. Brandeau and S.S. Chiu. An overview of representative problems in location research. *Management Science*, 35:645–674, 1989.

[10] E. Carrizosa, E. Conde, and M. Muñoz. Admission policies in loss queueing models with heterogeneus arrivals. *Management Science*, 44:311–320, 1998.

[11] E. Carrizosa and S. Nickel. Robust facility location. *Mathematical Methods in Operations Research*, 58:331–349, 2003.

[12] E. Conde. A branch and bound algorithm for the minimax regret spanning arborescence. *Journal of Global Optimization*, 37:467–480, 2007.

[13] E. Conde. Minmax regret location-allocation problem on a network under uncertainty. *European Journal of Operational Research*, 179:1025–1039, 2007.

[14] J.A. Díaz and E. Fernández. A branch-and-price algorithm for the single source capacitated plant location problem. *Journal of the Operational Research Society*, 53:728–740, 2002.

[15] E. Fernández, Y. Hinojosa, and J. Puerto. Filtering policies in loss queuing-location problems. *Annals of Operations Research*, 136:259–283, 2005.

[16] K. Holmberg, M. Rönnqvist, and D. Yuan. An exact algorithm for the capacitated facility location problem with single sourcing. *European Journal of Operational Research*, 113:544–559, 1999.

[17] ILOG Concert Technology 2.5 User's Manual. ILOG, Inc., Incline Village, Nevada, 2007. http://www.ilog.com/products/optimization/tech/concert.cfm.

[18] ILOG CPLEX 11.0 User's Manual. ILOG, Inc., Incline Village, Nevada, 2007. http://www.ilog.fr/products/cplex/.

[19] P. Jaillet. A priori solution of a traveling salesman problem in which a random subset of the customers are visited. *Operations Research*, 36:929–936, 1988.

[20] P. Kall and S. Wallace. *Stochastic Programming*. John Wiley & Sons, New York, 1994.

[21] G. Laporte, F.V. Louveaux, and H. Mercure. An exact solution for the a priori optimization of the probabilistic traveling salesman problem. *Operations Research*, 42:543–549, 1994.

[22] G. Laporte, F.V. Louveaux, and L. Van Hamme. Exact solution to a location problem with stochastic demands. *Transportation Science*, 28:95–103, 1994.

[23] F. Louveaux. Stochastic location analysis. *Location Science*, 1:127–154, 1993.

[24] F.V. Louveaux and D. Peeters. A dual-based procedure for stochastic facility location. *Operations Research*, 40:564–573, 1992.

[25] V. Marianov and C. ReVelle. The queueing maximal availability location problem: a model for the siting of emergency vehicles. *European Journal of Operational Research*, 93:110–120, 1996.

[26] C. ReVelle. Review, extension and prediction in emergency service siting. *European Journal of Operational Research*, 40:58–69, 1989.

[27] M. Riis and K.A. Andersen. Capacitated network design with uncertain demand. *INFORMS Journal on Computing*, pages 247–260, 2002.

[28] L.V. Snyder. Facility location under uncertainty: A review. *IIE Transactions*, 38:537–554, 2006.

[29] C. Toregas, R. Swain, and C. ReVelle. The location of emergency service facilities. *Operations Research*, 19:1363–1373, 1971.