REGULAR ARTICLE

On the impact of spatial pattern, aggregation, and model parameters in planar Huff-type competitive location and design problems

**Boglárka Tóth · Frank Plastria · José Fernández ·
Blas Pelegrín**

Abstract   A chain wants to set up a single new facility in a planar market where similar facilities of competitors, and possibly of its own chain, are already present. Fixed demand points split their demand probabilistically over all facilities in the market proportionally with their attraction to each facility, determined by the different perceived qualities of the facilities and the distances to them, through a gravitational or logit type model. Both the location and the quality (design) of the new facility are to be found so as to maximize the profit obtained for the chain. Several types of constraints and costs are considered. Applying an interval analysis based global optimization method on several spatial patterns in a quasi-real-world environment,

B. Tóth (✉)
Department of Differential Equations, Budapest University of Technology and Economics,
Budapest, Hungary
e-mail: bog@math.bme.hu

F. Plastria
MOSI, Department of Mathematics, Operational Research,
Statistics and Informations Systems for Management,
Vrije Universiteit Brussel, Brussels, Belgium
e-mail: Frank.Plastria@vub.ac.be

J. Fernández · B. Pelegrín
Department of Statistics and Operations Research, University of Murcia, Murcia, Spain
e-mail: josefdez@um.es

B. Pelegrín
e-mail: pelegrin@um.es

the behaviour of optimal solutions is investigated when changes are made in the basic model parameters. The study yields valuable insight for modellers into the impact of spatial pattern and various model parameters of the model on the resulting location and design decision. Spatial patterns differ in distribution of demand, of own and/or competing facilities, and of facility qualities. Studied model parameters include push force effects, investment restrictions and aggregation of demand.

1 Introduction

When a retail chain considers entering or extending its presence in a market, the two major questions are 'how' and 'where'. 'How' indicates the facility design part of the decision, required when several options exist in the type of outlet and/or the goods and services that will be available there. 'Where' clearly states that a location has also to be found. In order to be able to evaluate the market share resulting from the entry of the new facility, one needs to consider the way consumers behave in their choice between facilities offering similar goods/services. Many quite different proposals exist in the literature, as extensively explained by Drezner-Eiselt (2001). The goal pursued by the chain is to increase its profit, through an increase of its expected market share and without a too high investment in the facilities' design. This leads to optimization questions of various difficulty levels, depending on the modelling assumptions, see, e.g. the surveys of Eiselt et al. (1993), Eiselt-Laporte (1996) and Plastria (2001). Observe that in most of the literature only the question of location is investigated, whereas we would like to determine design and location simultaneously, since these two features cannot be separated. A remarkable exception is the work of Aboolian et al. (2007), who consider a related problem, although in discrete space. In a non-competitive setting, Dasci-Laporte (2005), Verter (2002), Verter-Dasci (2002) have studied the problem of simultaneous optimization of the plant location and technology selection decisions in a multi-product environment. There is another major question to be considered when locating a new outlet, the 'when'. However, this dynamic component is out of scope of this research.

The general situation we investigate in this paper is static, without looking ahead in time towards possible later newcomers (as in Drezner-Drezner 1998; Sáiz et al. 2008) or changes in the market (as in Drezner-Drezner 2002b; Averbakh-Berman 2003). We consider essential goods only, in other words, demand is assumed to be known and fixed (see Berman-Krass (2002) for a discrete competitive location model with elastic (variable) demand, and McGarvey-Cavalier (2005) for a planar location space). According to the terminology introduced by Hakimi (1983), we study a $(1|X_m)$-medianoid problem: given a set of m pre-existing facilities that are located at a set of points X_m, we would like to determine where a competitor should locate a single new facility so as to maximize his profit. Notice, however, that in this paper some of the pre-existing facilities may belong to the locating chain, and that the objective is maximization of expected profit, instead of market share.

Within these restrictions and considering a single new facility only, we investigate one of the most complicated situations for designing and locating a facility: Huff-type customer behaviour in a continuous region of the plane. This model assumes, following Reilly (1931, 1953), that customers will patronize all available facilities offering the adequate service in a probabilistic way, with probability proportional to their attraction to this facility, as compared to the other existing facilities. We assume that attraction depends on the customer's view of the quality offered by the facility and its distance to it, thereby generalizing the proposals of Huff (1964), Nakanishi-Cooper (1974) and Jain and Mahajan (1979). The chosen site for the new facility determines the distances from the potential customers, while the chosen design influences their quality assessments. Taking both effects together yields the probabilities with which the different customers will make use of the new facility. This then allows to evaluate the expected total market share of the new facility. No pricing policy is assumed in this paper. The influence of the prices in the patronizing behaviour of the customers is assumed to be included in the perception of the quality of the facilities.

The resulting model is a highly nonlinear optimization problem, involving several modelling decisions and many corresponding parameters to be estimated, the influence of which remains virtually unexplored. It is also intuitively clear that the number, location and relative size or design, both of the existing chain outlets and of its competitors, should have a significant impact on the location decision for a new facility of the chain. Existing literature says almost nothing about this impact, and we may say that the model's behaviour is not really well understood. It is the aim of this study to make a contribution to this understanding, by investigating the changes in optimal design/location when the model environment and parameters change. To this end one must first of all be able to solve the problem optimally in a reliable way.

But even for a fixed design, the objective function has a complicated multimodal shape as can be seen for example in the graph shown by Drezner (1994). The situation is complicated further when considering, as we do here, the presence of locational constraints, unavoidable in any realistic description of a real-world problem. For such functions and/or constraints, traditional analytic solution techniques or more model-specific iterative methods (see Drezner 1994) are unable to find guaranteed optimal solutions: what is obtained is a proposal of a site apparently close to a local optimal solution, without clear indication of the actual degree of (non)optimality, both in terms of objective value as in terms of spatial position. More precise information can then only be gathered through a more or less blind repetition of the method from varying starting solutions. It is not uncommon that very different solutions are then produced and an actual global optimum is not necessarily found. Even if it is found, one does not know so.

Additionally, the model typically contains many parameters, which are quite difficult to estimate precisely. On the one hand, it is therefore important to know how stable the solution is with respect to these parameters, in order to be able to appreciate whether the uncertainty in certain parameters may be ignored or not. This calls for a sensitivity analysis of the optimal solution in terms of changes in the parameters. But how can one correctly appreciate the results of such an analysis without a clear

distinction between jumps in the solution due to parameter modification rather than due to algorithm inaccuracy? Traditional local optimization approaches therefore are not acceptable for an adequate sensitivity analysis. One way of tackling this difficulty was described by Fernández-Pelegrín (2000): they incorporate the uncertainty on the parameter values as interval-valued data, specifying both an upper and a lower bound on each uncertain parameter, and apply interval analysis-based optimization tools, which construct a close approximation of the set of all possible optimal solutions for any setting of each parameter within its bounds.

On the other hand, recognizing that the model captures only approximately the real-world situation, Plastria (1992) argues that in practice a good idea of the full set of almost-optimal solutions is more useful than knowledge of an optimal solution, and proceeds to obtain such information by extending the global optimization method BSSS devised by Hansen et al. (1985).

Both methods, as well as a more recent variant based on triangle subdivisions (Drezner-Drezner 2004; Drezner-Suzuki 2004), were described for a pure location problem only. In a recent paper (Fernández et al. 2003), a solution strategy was developed for the full location-design problem using 'rigorous' global optimization based on interval analysis (see, e.g. Hansen 1992; Kearfott 1996; Ratschek-Rokne 1988), and including several novel acceleration techniques. This approach yields a complete view of the set of close-to-optimal solutions to any desired degree of accuracy, and therefore lends itself better to the proposed sensitivity analysis. To the best of our knowledge, no other available local or global optimization method allows to obtain the same amount of reliable information. However, a number of new interpretation difficulties arise, as will be explained in the sequel.

This tool has allowed us to investigate the behaviour of the optimal solution of Huff-type competitive location and design problem under various circumstances. The experimental design was made in such a way as to maximize the information gathered while keeping the computational burden within reasonable limits. We have done all experiments on data sets based on a single set of real-world demand data from the region of Murcia in southeastern Spain and recorded the variation of the optimal objective value and of the optimal solution set, both in design and in location.

Broadly speaking, we have studied the effects of

- the market structure, comparing chain superiority with strong competition in Sect. 4,
- the aggregation of demand in Sect. 5,
- homogeneous versus widely different facility qualities in Sect. 7,
- the location-cost relative to the quality cost in Sect. 8,
- and a restricted budget for initial investments in Sect. 9.

A precise description of these experiments and their results is given in separate sections, after the model has been fully formalized in the next section. The paper concludes with an overview of the experience gathered through these results and a number of recommendations for analysts and practitioners. Note: in the extended version of this paper Tóth et al. (2008) the effect of some more technical modelling parameters are also investigated.

2 The Huff-type competition model

2.1 The formal model

A chain wants to set up a single new facility in a given area of the plane $S \subset \mathbb{R}^2$. There already exist m facilities in the vicinity offering the same good or product. The first k of these belong to the chain, and the next $m - k$ facilities belong to competing chains. Since it does not matter in this model whether the competing facilities are of the same competing chain or of different competing chains, we will assume the former to be the case. Note that $k = 0$ means that the entering chain, which we will call 'the chain', is new to the market, whereas $k = m$ indicates simple expansion of the chain without competition. Existing facility $j = 1, \ldots, m$ is located at point f_j and has a known design, expressed by a quality $\alpha_j > 0$. The unknown location of the new facility will be denoted by x, and its quality by α.

Demand is concentrated at n demand points with known locations p_i and buying power w_i ($i = 1, \ldots, n$).

The distance between the demand point i and the new (resp. existing) facility (j) is denoted by d_{ix} (resp. d_{ij}), and are considered to be Euclidean throughout. Each demand point i has its own subjective perception of distance, obtained by application of some positive non-decreasing function g_i to actual distance. The inverse function $1/g_i$ is called the *distance decay function*. Several typical perceived distance functions g_i are found in literature, such as $g_i(d_{ix}) = e^{\lambda_i d_{ix}}$ (see Hodgson 1981) or $g_i(d_{ix}) = (d_{ix})^{\lambda_i}$ (see Drezner 1994), with $\lambda_i > 0$ a given parameter.

Attraction is modelled multiplicatively as quality divided by perceived distance, i.e. the attraction that demand point p_i feels for facility f_j is given by $\alpha_j/g_i(d_{ij})$, and for the new facility by $\alpha/g_i(d_{ix})$. This generalizes the 'law of retail gravitation' of Reilly (1931, 1953), who considered the perceived distance to be the squared distance. In this paper, we made the same assumption. Quality was first estimated as store surface by Huff (1964), and later several other store characteristics were incorporated by Jain and Mahajan (1979) and Nakanishi-Cooper (1974) [for details see Drezner-Eiselt (2001)].

In the spirit of Huff (1964), we consider that the patronizing behaviour of customers is probabilistic, that is, demand points split their buying power among the facilities proportionally to the attraction they feel for them. By these assumptions the total market share captured by the chain is given by

$$M(x, \alpha) \equiv \sum_{i=1}^{n} w_i \frac{\frac{\alpha}{g_i(d_{ix})} + \sum_{j=1}^{k} \frac{\alpha_j}{g_i(d_{ij})}}{\frac{\alpha}{g_i(d_{ix})} + \sum_{j=1}^{m} \frac{\alpha_j}{g_i(d_{ij})}}.$$

The chain's market share gives rise to expected sales through a strictly increasing function $F(\cdot)$. Operating costs for the new facility are given by a function $G(x, \alpha)$ of both its location x and its quality α.

The profit of the chain is obtained as sales minus costs. Since operating costs of the existing facilities of the own chain are assumed to be constant, we take into account costs of the new facility only.

The problem to be solved may therefore be written as follows

$$\begin{cases} \max \ \Pi(x, \alpha) \equiv F(M(x, \alpha)) - G(x, \alpha) \\ \text{s.t.} \ \ d_{ix} \geq d_i^{\min} \ \forall i \\ \qquad \alpha \in [\alpha_{\min}, \alpha_{\max}] \\ \qquad x \in S \subset \mathbb{R}^2 \end{cases} \qquad (1)$$

where the set S denotes the feasible space where the new facility is to be located. The parameters $d_i^{\min} > 0$ and $\alpha_{\min} > 0$ are given thresholds, guaranteeing on the one hand that the new facility is not located on top of demand point i and on the other hand that it has a minimum level of quality. The parameter $\alpha_{\max}$ is the maximum quality that a facility may have in practice. Note that in case some legal protection rules exist to guarantee a minimal market to any installed facility, one could also incorporate a minimum allowed distance to competitors, or another kind of constraint on the location.

In this study F was chosen to be linear, $F(M(x, \alpha)) = c \cdot M(x, \alpha)$, meaning each unit of goods sold yields a fixed income c. The cost-function G is assumed to have the separable form $G(x, \alpha) = G_1(x) + G_2(\alpha)$, with $G_1(x) = \sum_{i=1}^{n} \frac{w_i}{(d_{ix})^{\varphi_{i0}} + \varphi_{i1}}$ and $G_2(\alpha) = e^{\frac{\alpha}{\beta_0} + \beta_1} - e^{\beta_1}$. Notice how $G_1(x)$ increases as x approaches any demand point, expressing that in centres of high population concentrations the operational cost of the facility will be higher due to the value of land and premises, thus raising the cost of buying or renting the site. Also the minimal distance to demand point constraints $d_{ix} \geq d_i^{\min} > 0$ ensure the existence and differentiability of $G_1(x)$ at all *feasible* x. On the other hand, the cost of quality $G_2(\alpha)$ is an increasing convex function with $G_2(0) = 0$: the higher the quality of the facility, the higher are the costs, at an increasing rate.

As discussed in Fernández et al. (2003) many other income and cost functions may be considered and handled by the general algorithm proposed there.

2.2 Representation of the results

Results of an optimization problem may be understood in several ways. On the one hand there is the optimal value found, and on the other hand one may look at the optimal solution. In our location-design problem, we are clearly more interested in the latter, since it is the choice of the site that matters most, although the actual profit figure is also important for judging the profitability of the project. Therefore, we would like to represent both types of results, so as to be able to judge the sensitivity of these results when the parameter settings change.

Theoretically, this seems relatively simple, particularly for the optimal values: single values, which are easily compared. But this does not take into account that the outcome of a traditional local optimization algorithm usually only gives an approximation for the optimal value, very often without any information about the precision of the approximation or, even worse, without guarantee of having found (an approximation of) a globally optimal value. However, with the global optimization method we have used, we do have quite precise information about the global optimal value: during the optimization stage both an upper and lower bound on the global optimal value

are available, and gradually improved until their difference is below a prespecified precision ε. The resulting upper bound is therefore guaranteed to be an overestimation within ε-accuracy of the actual global optimal value. In our computations, we used $\varepsilon = 0.05$. It follows that when we compare the 'optimal' objective values obtained for two different parameter settings, the conclusions can be trusted, as long as they differ by significantly more than 0.05 units.

The situation is more complicated in variable space, when we look at the optimal sites and corresponding qualities found. Traditional methods simply give an approximate solution, which might be close to a local optimal solution (but not necessarily a globally optimal one), without information about how far it might be from such a solution (except possibly some very fuzzy indication one might derive from a Lipschitz constant, which is often awfully overestimated), nor about the possible existence of other (near)-optimal solutions. The second stage of our global optimization algorithm is aimed at producing an outer approximation (up to a precision η) of the set R_δ of δ-optimal solutions, i.e. with objective value at least a fraction $1 - \delta$ of the optimal value. To be more precise, it yields a list of solution boxes, and for each box an upper and lower bound on the objective values at any solution within the box. The final result is a list of all boxes lying fully within this δ-optimal region and those on its 'boundary', i.e. probably containing solutions that are not δ-optimal, but not too far, as indicated by the second precision parameter η. These solutions are all guaranteed to lie within $R_{\delta+\eta(1-\delta)}$. This is illustrated graphically in Fig. 1 for a two-dimensional problem. The first (inner) set IR_δ contains all boxes for which the lower bound is at least a fraction $1 - \delta$ of the (approximate) global optimal value. The second set (surrounding) $OR_\delta \setminus IR_\delta$ consists of boxes with upper bound at least reaching this fraction $1 - \delta$ and the lower bound below it, but not by more than $\eta\%$ [for details, see Plastria (1992)]. This implies that in case the objective is quite flat around a global optimum, these boxes may cover a non-negligible area. The advantage of this approach is that in case alternative near-global optimal solutions exist anywhere else in the area, this will immediately be apparent from the results, since the lists of boxes then define several disconnected regions. In our computational tests we used $\delta = 0.01$ and 0.002.

A more precise reporting of the near-optimal region calls for a graphical representation of the resulting lists of boxes in variable space, which leads to new difficulties, due
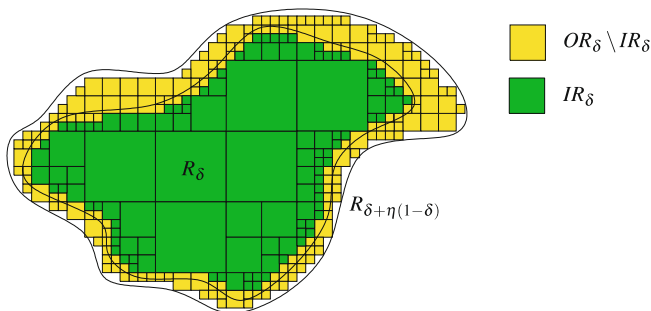

Fig. 1 Inner and outer approximation of the near optimal region

to their three-dimensional shape. We resolve this by reporting the information in three two-dimensional maps, each showing the interactions between some pair of variables. This is obtained by adding two windowpanes on the right and bottom of the basic spatial map, allowing us to view all three two-dimensional projections of the three-dimensional solution set; see, e.g. Fig. 4. The map itself shows the two-dimensional spatial part, without the quality (x_1 increases from left (0) to right (10) and x_2 increases from top (0) to bottom (10)). The right-hand pane shows the quality and vertical space part (quality increases from left (0.5) to right (5)). The bottom pane shows the quality and horizontal space part (quality increase from top (0.5) to bottom (5)).

Note that often the spatial solution set is quite small or falls almost exactly on the boundary of some forbidden region, and is therefore not well visible. One may however know where it is, by combining the regions shown in the two side panes.

3 Base data

The real-world data consists of an area around the city of Murcia in southeastern Spain. The city of Murcia is the capital of the autonomous region of Murcia (A.R.M.), a province with $11314 \, km^2$ and over 1 million inhabitants. More than half of the inhabitants of A.R.M. live in the city of Murcia or in the villages close to it.

A working radius of 25 km around the main city was considered to be suitable for use in a Huff-type model. This choice was motivated by two concerns. On the one hand, customers are assumed to patronize *all* existing facilities, so none of the facilities should be too far away from a population centre (e.g. within a 15 min drive), because otherwise none of its inhabitants would go there. On the other hand, in the region at hand, increasing the radius of the circle to 30 km increases the number of covered inhabitants only slightly, whereas reducing the radius to 20 km decreases it considerably.

Within the circle of radius 25 km, centred at Murcia, live 557,746 inhabitants of A.R.M. and another 74,812 inhabitants from the neighboring province of Alicante, on the east of A.R.M. (see Fig. 2). These 632,558 inhabitants form our primary set of customers. They are distributed over 71 population centres, with population varying between 1,138 and 178,013 inhabitants. In this study we have considered each population centre as a demand point, with buying power proportional to its total population (one unit of buying power per 17.800 inhabitants). Their position and population can be seen in Fig. 3a: each demand point is shown as a black dot, and also a grey circle, the radius of which is proportional to the buying power. In Fig. 3b the aggregated case is depicted, as discussed in Sect. 5. Note that here the grey circles also show the forbidden regions, but contrary to (a), do not represent the demand proportionally.

Our basic data deals with the location of supermarkets. Currently there are five supermarkets present in the area: two from a first chain, the 'Small Chain', and three from another chain, the 'Large Chain'. Figure 3 shows the location of each supermarket as viewed from some chain's point of view: firms belonging to the own chain are marked by a ×, and firms from the other chain are shown by a ⊡ marker on the map.

The feasible set S was taken exactly as depicted in Fig. 3, i.e. the smallest rectangle containing all demand points. This is approximately a square centred at Murcia and
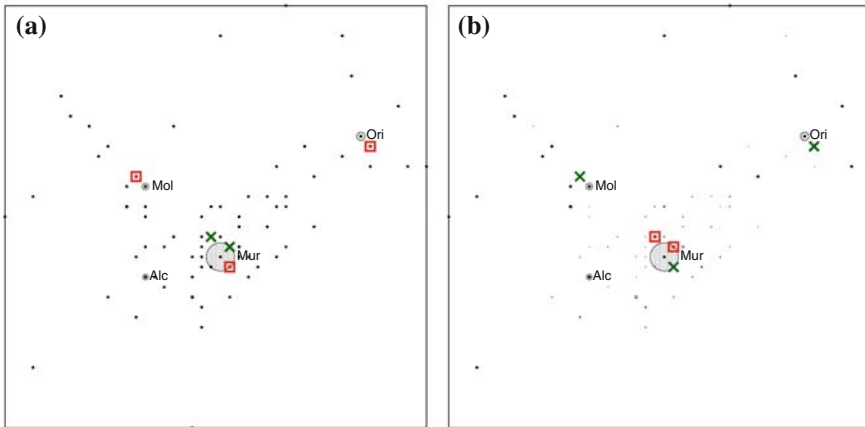
Fig. 2 Population centres in the neighbourhood of the city of Murcia

of sides close to 45 km. It is reduced further by some forbidden regions around each population centre as defined by $d_i^{\min}$ (see below).

The coordinates of the population centres and the supermarkets were obtained with the geographical information system called VisualMap (see Visual G.I.S. Engineering 2003), and were rescaled from coordinates $[200, 245] \times [243, 285]$ to an approximate standard square $[0, 10] \times [0, 9.3]$. Thus, the units correspond approximately to 4.5 km. Note that, as usual in computer graphics, the origin is at the top left, and the vertical axis runs downwards. The population data were obtained from the *Regional Centre of Statistics of Murcia* (see Anuario Murcia 2003) and correspond to the year 2002, rescaled to $]0, 10]$. The minimum distance $d_i^{\min}$ at which the new facility must be from the population centre i was chosen to be $\frac{w_i}{\xi}$, where ξ is set to 30. This means that for a demand point with 1,500 inhabitants ($w_i = 0.067$) $d_i^{\min} = 0.00223$ ($\sim$10 m), and for the largest demand point (Murcia, $w_i = 10$) $d_i^{\min} = 0.\dot{3}$ ($\sim$1.5 km). All final data, together with all the used parameter values discussed below, may be found in the Appendix of the extended version of this paper Tóth et al. (2008).

Fig. 3 Existing facilities, demand points (*black*) and forbidden regions (*grey*), **a** full problem for the Small Chain. **b** aggregated problem for the Large Chain

Approximate values for the quality parameters α_j for the five existing supermarkets were obtained through a small survey among ten persons who had visited all five supermarkets. Each respondent was asked to rank the supermarkets in increasing order of their perceived quality and to indicate the difference in quality between any two consecutive supermarkets in their ranking by a score between 1 (small) and 4 (big). That information yielded individual marks for all supermarkets, by starting from a lowest mark of 1 and adding each difference score to obtain the mark of the next supermarket in their ranking. Finally, these marks were rescaled to the interval [3, 4], because, according to all the respondents, in a proposed scale from 0.5 to 5 all considered supermarkets have a quality above the mean (2.75). But they all still have a large margin for improvement. The qualities α_j we finally considered are the average rescaled marks over the ten respondents. The optimization of quality was done in the interval $[\alpha_{\min}, \alpha_{\max}] = [0.5, 5]$. In that scale, 1 is considered to be the quality of small-size low-quality supermarkets, and 5 the quality of very big-size big-quality supermarkets. The range 3 to 4 for the existing facilities shows that all these supermarkets are of similar type (size and quality above average) with small differences.

Due to the lack of real data from the chains (they consider those data sensitive for them and are not willing to make them public), the other parameters have been validated in an ad hoc way to obtain 'reasonable' results. In particular, we used the following initial settings: the income per unit sold $c = 12$, in the location cost function G_1 we chose the parameter $\varphi_{i0} = 2$ for all i, while value of φ_{i1} decreases as the population increases. Finally, the parameters of the cost of quality function G_2 were initially set to $\beta_0 = 7$ and $\beta_1 = 3.75$.

4 Impact of market structure

The basic data described above have been used to define three quite different competitive market structures:

Newcomer: a new chain enters the market with one new facility, which has to compete with all five existing facilities,

Small Chain: the smaller chain wants to strengthen its two existing facilities by adding a new one to compete on a more equal basis against the three facilities of the Large Chain,

Large Chain: the larger chain wants to extend its dominance in the market by moving from the current three against two to a four against two facility situation.

The results obtained in these three problem instances are described in Table 1 and depicted in Figs. 4 and 5 at the left.

Table 1 Results for the basic problems

|	Newcomer	Small Chain	Large Chain
Optimum (best found)	44.93	210.39	242.96
δ-opt interval	[44.39, 45.02]	[207.87, 210.81]	[240.05, 243.44]
Market share	$0 \rightarrow 8.47$	$15.77 \rightarrow 18.70$	$19.74 \rightarrow 21.14$
Best point found	(4.82, 6.11, 5.00)	(8.40, 3.18, 1.43)	(3.29, 6.48, 0.52)
Hull of results			
x_1	[4.78, 4.88]	[8.32, 8.57]	[3.07, 3.57]
x_2	[5.99, 6.19]	[2.98, 3.22]	[6.19, 6.70]
α	[4.56, 5.00]	[0.69, 2.47]	[0.50, 1.70]
Second region			
x_1		[3.25, 3.31]	[4.77, 5.45]
x_2		[4.26, 4.36]	[5.61, 6.26]
α		[1.34, 2.08]	[2.01, 4.23]
Volume of the δ-opt	1.7e−04	2.1e−02	1.8e−01
Time in seconds	137.07	165.55	340.32
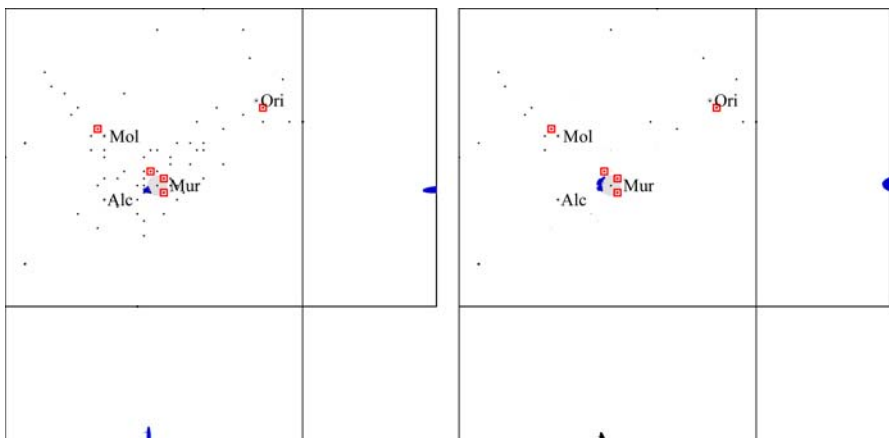
Fig. 4 Near optimal regions for the Newcomer with unaggregated (*left*) and aggregated (*right*) demand
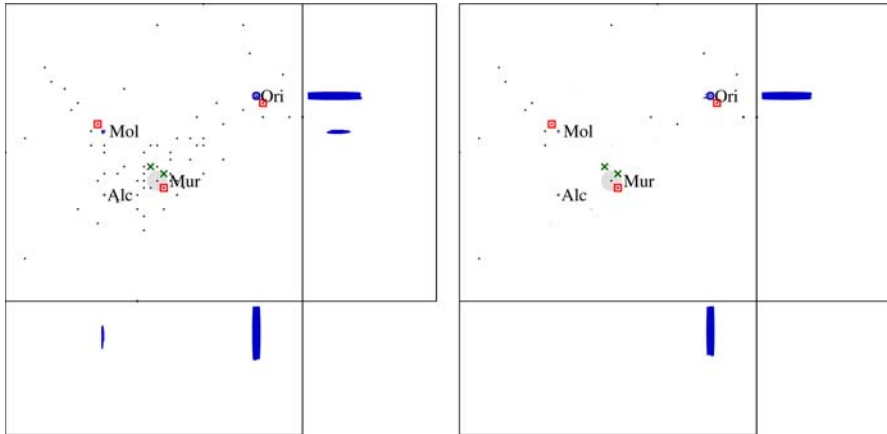
Fig. 5 Near optimal regions for the Large Chain with unaggregated (*left*) and aggregated (*right*) demand

All the computational results in this and later sections have been obtained under Linux on a Pentium IV with 2.66 GHz CPU and 2GB memory. The algorithm has been implemented in C++ using the interval arithmetic in the PROFIL/BIAS library (Knüppel 1993), and the automatic differentiation of the C++ Toolbox library (Hammer et al. 1995).

In the table, we first see the maximal value found for the profit of the locating chain, which is guaranteed to be within $\varepsilon = 0.05$ units of the global optimal value. Next, we indicate the interval of objective values reached within the $\delta = 1\%$-optimal region, up to a precision of $\eta = 0.2\%$.

Bear in mind that the objective values concern the total income from the chain minus the costs of the new facility. Due to the quite different sizes of the locating chain, this value is not directly comparable between the three cases.

A much better point of comparison is obtained when looking at the change in market shares after entry of the new facility, which is indicated on the third line of the table. The original market shares were estimated by applying the model parameter settings to the location of the existing facilities only, using our estimation of their quality. The Newcomer's single facility would be able to obtain a market share of 8.5 out of a total of approximately 35.5, i.e. 24%, a quite amazing success. The Small Chain would see its market share expand from its (estimated) current value of 44–53%, thus doing slightly better than its competitor with the same number of facilities. This shows, incidentally, that the Large Chain was not optimally located. Expansion of the Large Chain increases its current market share of 56% to only 59%, which is rather low considering that this results from twice as many facilities as the competition.

This may be explained when looking at the corresponding best solutions found on the fourth row of the table, indicated as the triplet (horizontal position x_1, vertical position x_2, and quality α), or in the graphical representation in Figs. 4 and 5 at the left.

The Newcomer's best strategy turns out to be to go for the largest demand concentration, although the competition there is quite fierce. That is why its quality should

be taken as high as possible: the best quality is at the maximum value of 5. The facility should be sited next to the central and main city of Murcia, at the southwestern boundary of its forbidden region, which is as far as possible from both competitors already present around the city. Since the competitors do not have quality better than 4, it obtains almost one-third of Murcia's market and a quite large part of the demand points on the southwest. This comes with an important cost of around 57 though, as can be evaluated by the difference between sales ($12 \times 8.5 = 102$) and actual profit (45).

The Small Chain, however, is already well and exclusively present around Murcia, and can therefore better aim at the next largest and cheaper city, Orihuela at the northeast of Murcia, in direct competition with the existing facility of the Large Chain next to it. At a low quality around 1.4, it quite easily grabs a good part of its competitor's market. This explains the relatively better gain/cost ratio as compared to the Newcomer's strategy.

For the Large Chain, the most interesting place is at Alcantarilla, only slightly to the west of Murcia. This region is currently badly served, and here minimum quality is sufficient to take over the southeastern market.

The next rows in the table give much more precise information about the near-optimal location and qualities: for each of the three basic variables of the problem it indicates the 'hull', i.e. the range within which it varies within the near-optimal region. One must, however, be quite careful with this hull: the near-optimal region may be disconnected, which happens if there exist more than one local optimum, yielding almost the same global optimal value. In such a case, the 'naive' hull would enclose all of these connected components and would be uninformative, since it would yield wide overestimations of the range. But the disconnectedness would not be apparent. Therefore, the hull is then given for each connected component separately.

One might argue that this should not happen very easily. Though many local optima may exist, there usually is a clear winner and only rarely more than one near-global optimum. This is perhaps true for pure location problems and is then restricted to some exceptional circumstances such as almost symmetrical situations. Consideration of the quality issue as part of the decision, as is done here, changes the setting. This independent choice variable offers the additional freedom of allowing to more easily obtain similar close to optimal profits at different sites. And this is what is observed here. This phenomenon happens in two of the three instances. Both the Small Chain and the Large Chain have two almost equivalent possibilities, which appear as the 'second region' in the table, and are also clearly visible in the left part of Figs. 5 and 6. No more than two connected components were observed in our experiments, but with other data such may certainly appear.

The penultimate row in each table gives the total volume of the three-dimensional boxes defining the near-optimal region. The ratio between the third root of this value and the size of the near optimal objective interval gives a more precise impression of the flatness of the objective function around the optimum value, although boundary effects on quality somewhat cloud the issue. Still, it is quite clear that the Large Chain has the flattest objective, the Small Chain case comes next, and in the Newcomer's case the objective seems quite steep.
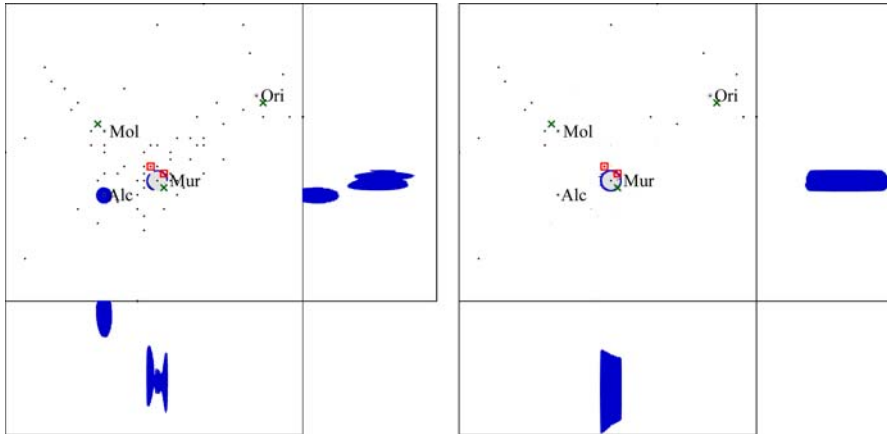
Fig. 6 Near optimal regions for the Small Chain with unaggregated (*left*) and aggregated (*right*) demand

Finally, in the last row we give the required computational time to find the near-optimal regions, which varied from a few seconds to several minutes.

As already observed, and clearly visible in the figures, the two last cases are quite interesting since they show that there are two near-optimal regions, which are quite far apart. The Small Chain has two almost equivalent options, first to locate at the northeast of Murcia, more precisely at Orihuela, or to locate at the northwest, at Molina. In both cases, it will be in direct competition with the facilities of the Large Chain in these same areas.

Similarly, the Large Chain also has two main options, which are almost equivalent, but of quite different type: either locating an additional facility around the main city Murcia and increasing its presence against the two-facility strong competition there, or grabbing the badly served market around the region's fourth city Alcantarilla, west of Murcia. The second and third largest markets (Orihuela and Molina) are not of interest to this chain since it already well covers them with its existing facilities. Observe also that several positions around Murcia's border are available, needing slightly different but rather upper average quality, whereas Alcantarilla is already conquered at almost minimum quality.

These last two examples demonstrate the power of global optimization: *all* (nearly) global optima are found. This is extremely useful when alternative location criteria are to be considered. A usual local optimization method might have found any of these global optima or perhaps a worse local optimum, but will probably not have found out that there are also almost equivalent solutions totally elsewhere.

5 The impact of aggregation

We have also constructed a reduced data set by aggregation of the demand. The aim is to study to what extent the aggregation may influence the solution to the problem. This is a very relevant question in practice, where the data may be available at different

levels of aggregation at very different costs. One should then have some idea if it is worth the effort/money to obtain the more detailed unaggregated data, and if this can affect the results considerably. Such errors incurred by aggregation have been studied intensively for several types of location models, see, e.g. Francis et al. (2002). But for competitive Huff-like location models, no theoretical or empirical information seems to be available as yet, even in the most recent survey Francis et al. (2008). The question has never been considered before with quality as part of the decision.

5.1 Aggregation design

An aggregated data set was obtained as follows. Several population centres do not have a city hall, so for public administrative tasks they depend on another town's city hall. The 21 towns with city hall form our reduced set of demand points, with population obtained by aggregating all population centres at the town on which they administratively depend. The resulting reduced demand distribution is shown in Fig. 3b. This way of aggregating is quite typical. Our two sources of data, Anuario Murcia (2003) and Visual G.I.S. Engineering (2003) give the data grouped by city hall, as in our aggregated case, while for the larger cities they are subdivided by population centres, as in our non-aggregated case.

In the aggregated problems, exactly the same forbidden regions were kept as in the corresponding unaggregated instances, including the forbidden regions around the population centres that are not demand points anymore after aggregation. Therefore, the results are directly comparable with the unaggregated instances, allowing to see the effect of the concentration of the demand by the aggregation.

5.2 Results

Three aggregated versions were considered, considering the same market structures as in the previous section, but now with the aggregated demand. The results obtained for these instances are described in Table 2 and are depicted in Figs. 4, 5 and 6 at the right.

For the Newcomer, the figure seems to show almost exactly the same results after aggregation: the best site is at Murcia at the western boundary of the forbidden region and using top quality almost exclusively. A closer look, in particular at the right pane, shows that the vertical position is less precisely determined than in the unaggregated case. The optimal profit, as found in the table, has increased by 18%, though. Both phenomena are explained by the demand at Murcia, which has now more than doubled, whereas it was spread around this city without the aggregation. On the one hand, this yields a less precise positioning of the optimum and a higher income evaluation. The latter is also clearly shown by the increase in market share. One may also observe that the objective has become steeper around the optimum.

For both the Small Chain and the Large Chain one observes that only one of the two alternative near-optimal solution regions have remained when the demand is aggregated. In the first instance, it is the former global optimum at Orihuela that remains as the only good choice. The former second option around Molina was very small,

Table 2 Results for the aggregated problems

	Aggr. Newcomer	Aggr. Small Chain	Aggr. Large Chain
Optimum (best found)	53.11	210.19	246.90
δ-opt interval	[53.00, 53.75]	[207.67, 210.61]	[243.95, 247.40]
Market share	$0 \rightarrow 9.00$	$15.52 \rightarrow 18.62$	$19.70 \rightarrow 23.73$
Best point found	(4.78, 5.92, 5.00)	(8.53, 3.05, 1.36)	(4.79, 6.02, 3.54)
Hull of results			
x_1	[4.78, 4.90]	[8.33, 8.56]	[4.76, 5.45]
x_2	[5.70, 6.12]	[3.00, 3.21]	[5.61, 6.30]
α	[4.75, 5.00]	[0.69, 2.33]	[2.15, 4.93]
Volume of the δ-opt	8.4e−05	1.6e−02	5.3e−02
Time in seconds	48	57	190

and now differs too much in value from the global optimum to remain visible at the $\delta = 1\%$ level. This indicates that the differences between the local optima have been sharpened in value by the aggregation. The slightly smaller volume of the near-optimal region, together with a virtually unchanged objective value also indicates that the objective has become somewhat sharper at the optimum.

For the Large Chain, however, it is the second option in the central Murcia that has taken the clear lead. This may be explained by the fact that Murcia has grown the most, and by far, through the aggregation, and thus has gained more in competitive interest than the smaller market around Alcantarilla. The estimation of the gain in market share has indeed been strongly influenced by the aggregation in this case.

This last result is quite interesting, because it shows that the decision may be strongly influenced by the fact that aggregated data were used.

As a final observation, note that the computational times have been more or less halved. This may be seen as basically due to the reduction of 71–21 demand points, which has proportionally reduced each objective function evaluation.

6 Comparing lists of boxes

The remainder of the paper is devoted to several types of sensitivity analysis, every time carried out for each of the three unaggregated instances and their three aggregated counterparts, yielding six base instances, varying in installed base of the locating firm and in the dispersion of demand.

In each analysis a particular model parameter is modified sequentially, and the intention is to measure the corresponding change in optimal value and in optimal solution. However, as explained in Sect. 2.2, these results are not simple values or points, but lists of boxes. In order to compare the sets of optimal solutions for two or more instances of our model, we must be able to judge the 'likeness' of the outcomes we obtain in each case. Therefore, we need a dissimilarity measure between lists of (pairwise non-overlapping) boxes.

According to the standard notation for interval analysis, lower case is reserved for scalar quantities or vectors, while intervals are indicated by boldface, where vector intervals are boxes (set theoretic products of intervals). A list of boxes $\mathcal{L}$ in $\mathbb{R}^d$ is then a finite set of non-overlapping d-dimensional boxes. Clearly any box $x = (x_i)_{i=1,\ldots,d}$ has a volume vol(x), and the volume of a list is the sum of volumes of its boxes: vol($\mathcal{L}$) $= \sum_{x \in \mathcal{L}}$ vol(x).

Let us introduce first the following asymmetric connector distance from a box x to the box y:

$$\Delta(x, y) = \max_{x \in x} \min_{y \in y} \|x - y\|$$

where $\|.\|$ is the standard Euclidean norm. We then define the following dissimilarity measure for two lists $\mathcal{L}$ and $\mathcal{L}'$:

$$\text{diss}(\mathcal{L}, \mathcal{L}') = \sum_{x \in \mathcal{L}} \frac{\text{vol}(x)}{\text{vol}(\mathcal{L})} \min_{y \in \mathcal{L}'} \Delta(x, y) + \sum_{y \in \mathcal{L}'} \frac{\text{vol}(y)}{\text{vol}(\mathcal{L}')} \min_{x \in \mathcal{L}} \Delta(y, x).$$

Applied to two single boxes (singleton lists of boxes), this formula yields the symmetrization of Δ as follows

$$\text{diss}(x, y) = \Delta(x, y) + \Delta(y, x).$$

For instance, in dimension 1, i.e. for two (disjoint) intervals on a line, this corresponds to the sum of the lengths of the two arrows in Fig. 7. It is similar to the traditional Hausdorff distance between the intervals, but uses a sum instead of a maximum. This means that the one-dimensional dissimilarity roughly doubles the distances between the corresponding bounds.

Notice that diss($\mathcal{L}, \mathcal{L}'$) is not a metric, but still has some nice properties. For instance, diss($\mathcal{L}, \mathcal{L}'$) $= 0$ iff $\cup \mathcal{L} = \cup \mathcal{L}'$. A full study falls outside the scope of this paper, but it should be clear that lists describing close shapes will be much less dissimilar than lists of very different shape or lying far apart. Therefore, this dissimilarity is quite appropriate to our purpose to be able to compare in a synthetic way the outcome from optimizing two (apparently slightly different) objectives, as obtained by changing some parameter setting.

This dissimilarity measure allows to compare the results in variable space. However, for objective values one should not only be able to see the amount of change, but also the direction of the change, either improvement or worsening. Since the results come as intervals of values known to bracket the corresponding optimal value, we have to compare two such intervals. However, since the bracket has an almost fixed
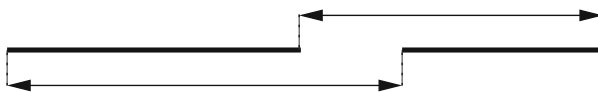

Fig. 7 The dissimilarity measure for two intervals

relative size, both extremes will either have increased, indicated by a +, or worsened, indicated by −. In the few cases when the extremes change with different signs, which only happen when the change is very small, no sign notations are used.

7 Impact of facility quality range

The first sensitivity analysis concerns the influence of the quality of each existing facility, as given by the parameters α_j. These were varied in several ways as described by the following scenarios:

Original In the base data, as given in the Appendix of the extended version of this paper Tóth et al. (2008), all existing qualities lie in the range [3, 4]. Since the new facility's quality is allowed to vary between 0.5 and 5, this means that it may be chosen as having a quality quite different from the qualities already present in the market.

Rescaled In this scenario, we rescaled all the α_j to the interval [1, 4.5], to make the existing facilities' range of qualities closer to the new facility's possibilities.

S>L Here the Small Chain has much higher quality facilities than the Large Chain. This is obtained by setting $\alpha_j = 4$ for all the existing facilities of the Small Chain and $\alpha_j = 2$ for all the existing facilities of the Large Chain.

S<L This is the opposite situation, where the largest chain offers the highest quality: Large Chain's existing facilities are now of quality 4 and those of the Small Chain of quality 2.

Closer better Now all existing facilities close to Murcia (two of the Large Chain and one of the Small Chain) have high quality 4, and those far from Murcia have quality of only 2.

Closer worse This final scenario inverses the previous scheme and gives the lowest quality 2 to all existing facilities in the surroundings of the city of Murcia, but high quality 4 to all those far from the main city.

Table 3 gives an overview of the comparative results for each of the six considered cases: Newcomer, Small Chain, Large Chain, each both in unaggregated and in aggregated form. This table is organised as follows.

In all cases, unaggregated and aggregated, the top row recalls the obtained range of objective values of the original problem reached within the near-optimal region given by the final list of solution boxes. All other rows show dissimilarities between the newly found list of boxes enclosing the optimal solution set with another list of boxes, and this in four ways: '3D' considers both quality and location variables together, while 'Loc' (resp. 'Qual') considers only the locational aspect (resp. only the quality aspect) by using the projections of all boxes on the two-dimensional location space (resp. one-dimensional quality space), and finally 'Obj' shows the dissimilarity in objective value ranges, using the sign convention explained in Sect. 6. In the rows 'rescaled', 'S>L', 'S<L', 'closer better' and 'closer worse' the comparison is made with the solution of the original problem.

We first observe that in most cases nothing much happens with respect to location. None of the quite important variations in the quality of the existing facilities have significant influence on the optimal spatial solution. In particular, the Small Chain

Table 3 Sensitivity to the quality of the existing facilities

	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.
	Newcomer [44.39, 45.02]				Small Chain [207.87, 210.81]				Large Chain [240.05, 243.44]			
rescaled	0.001	0.006	0.007	+8.2	0.042	0.002	0.118	+42.2	0.079	0.066	0.202	−60.3
S>L	0.002	0.008	0.010	+27.9	0.005	0.001	0.019	+82.0	1.274	0.814	2.283	−47.7
S<L	0.000	0.000	0.001	+12.8	0.005	0.017	0.079	−111.3	0.081	0.177	0.104	+77.4
closer better	0.000	0.002	0.002	+2.5	0.003	0.000	0.011	+9.9	0.251	0.844	0.264	−2.3
closer worse	0.078	0.002	0.064	+44.1	0.010	0.023	0.132	−37.7	1.282	0.564	1.983	+34.9
	Aggr. Newcom. [53.0, 53.75]				Aggr. Small Ch. [207.7, 210.6]				Aggr. Large Ch. [243.9, 247.4]			
rescaled	0.001	0.000	0.005	+4.6	3.089	0.001	3.125	+45.9	0.002	0.034	0.081	−55.5
S>L	0.001	0.000	0.002	+30.6	0.916	0.001	0.923	+88.2	0.003	0.146	0.251	−42.9
S<L	0.000	0.000	0.001	+19.2	2.058	1.160	2.263	−121.4	0.056	0.080	0.157	+78.8
closer better	0.000	0.000	0.002	−0.2	0.586	0.000	0.590	+2.5	0.000	0.001	0.004	−0.7
closer worse	0.002	0.000	0.004	+59.6	0.003	0.002	0.008	−30.5	0.001	0.039	0.071	+38.9

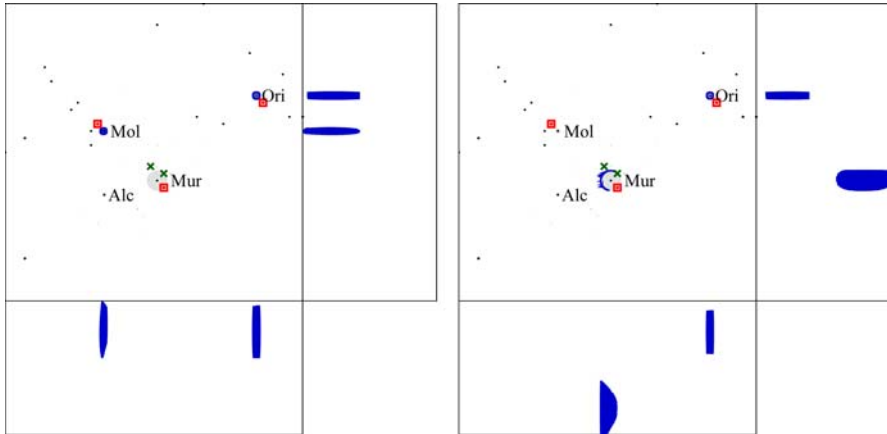**Fig. 8** Near optimal regions for the Small Chain with aggregated demand: *left* 'rescaled', *right* 'S<L'

with unaggregated demand systematically keeps its two almost equivalent options, a quite remarkable fact. The best solution for the Newcomer is also remarkably stable throughout.

In the Small Chain case with aggregated demand, however, we see appreciable dissimilarities appear with the corresponding base case. Two of such results are shown in Fig. 8. The first remark is that the original optimal solution around Orihuela shown at the right of Fig. 6 always still exists.

'rescaled' When rescaling the qualities one observes the reappearance of the second near optimal region around Molina, which had been lost by the aggregation of demand (compare with Fig. 6). This site is in direct competition with the existing Large Chain's facility in the same area, which by rescaling of its original quality of 3.1 has now a quality around 1. In other words, the range of interesting qualities for the entering facility is still spread around its direct competitor's quality.

'S<L' When the competing Large Chain has better quality facilities, we see a totally different second possibility appearing for the Small Chain: locate a third facility around Murcia, now of competitively high quality, to grab a larger share of this most important demand point from the competitor. Apparently the existing two facilities are of too low comparative quality to do this job correctly.

For the Large Chain under unaggregated demand, some smaller changes arise in the optimal solution, in particular in the scenarios 'S>L', 'closer better' and 'closer worse' (see Fig. 9).

'S>L' When the competing (small) chain has a better quality facilities the Large Chain's best option is to go into direct competition for the largest demand in Murcia, helping out its comparatively weak facility there and doing so at high quality. The alternative option appearing in the base case is not present anymore.

'closer better' When there are already three high quality facilities near Murcia, part of its demand can be gained only with the highest quality. Since that is too expensive, in this situation it is much better to go for Alcantarilla.
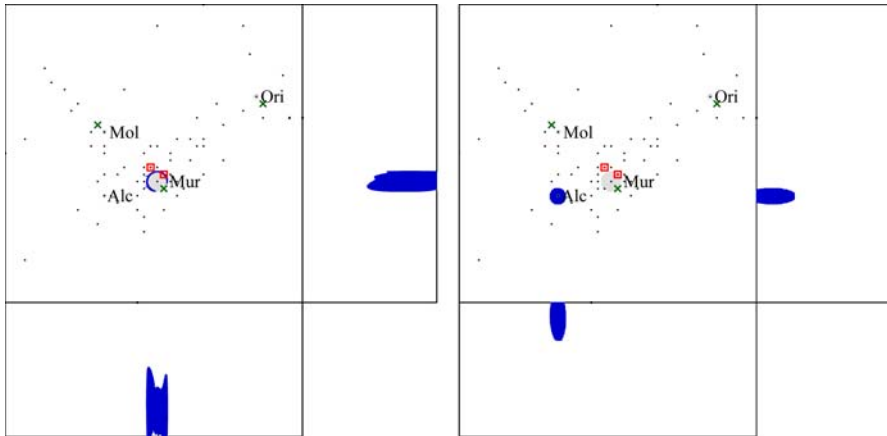
Fig. 9 Near optimal regions for the Large Chain with unaggregated demand: *left* 'S>L', *right* 'closer better'

'closer worse' Finally, when the own facility and both competing ones close to Murcia have a lower quality, it would be in the interests of the Large Chain to add a second facility around Murcia with a rather high quality, because this sufficiently grabs its opponent's market share.

We conclude that for a very small player the quality of existing facilities does not strongly influence its decision: it systematically goes towards the large demand areas where it is not yet present, taking over part of the competitor's market. A medium size player should be somewhat more careful in its appreciation of the facility qualities present in the market. Particularly when demand is quite concentrated it may otherwise miss some chances. But for chains having many and dispersed facilities (as the Large Chain has) almost any new facility may lead to a lot of cannibalization (loss of market share of the existing facilities of the chain), and therefore the qualities of own and competing facilities should be as accurate as possible if a correct decision is to be made. This seems to be particularly the case in a dispersed market.

With respect to optimal objective values, however, one observes quite strong variations in general, even when the optimal solution itself did not change much. Keep in mind that what is given in the table is the *change* in objective value, as compared to the original near-optimal interval, and expressed by the dissimilarity. In most cases, we see clear increases or decreases by relatively large amounts (although our dissimilarity roughly doubles the distances). Profit may increase by almost 50% compared to the original case (see the Newcomer in scenario 'closer worst') or reduced by over 25% (see the Small Chain in case 'S<L'). This implies that the evaluation of the expected profit is extremely sensitive to the existing facilities' quality assessment. Since any small chain has to attain a given minimum level of increase in profit in order to survive, qualities of the existing facilities will have to be estimated very precisely, if one wants to have confidence in the estimation of future profit.

From the cases 'S>L' and 'S<L', we can see that the chain that increases the quality of its facilities also has an increase in profit, and vice versa, when the competitor

increases the quality, the profit drops. These increases are almost perfectly proportional to the number of existing facilities with higher quality, but the drops are not: the Small Chain loses relatively much more profit, when the Large Chain increases its three facilities' quality, than the Large Chain in the opposite situation.

One may also observe that when existing facilities close to the global optimum increase (decrease) their quality, the optimal profit of the new facility decreases (increases) accordingly.

To understand the case of rescaled qualities, we should recall the original qualities of the existing facilities. In the case of the Small Chain, the qualities are 3.25 and 4, whereas for the Large Chain they are 3.1, 3.5 and 3.15. When we rescaled [3,4] to [1,4.5] the lower bound 3 goes down to 1, but the upper bound 4 goes up to only 4.5. Thus, the rescaled facilities of the Large Chain have lost more quality, when compared to the facilities of the Small Chain. This explains why the Newcomer and the Small Chain increase their profit, whereas the Large Chain sees a decrease.

8 Location cost versus quality cost

Next, we studied the influence of the relative importance between the two cost function elements, the location cost given by G_1 and the cost of quality, as described by G_2. This was done using the modified cost function $G(x, \alpha) = \mu_x G_1(x) + \mu_\alpha G_2(\alpha)$, with $\mu_\alpha = 2 - \mu_x$. The original case corresponds to $\mu_x = \mu_\alpha = 1$. The results are given in Table 4, following a structure similar to the previous tables.

One may expect the following phenomenon. When μ_x is very small, quality is quite expensive, so rather low quality settings will be chosen. When μ_x increases, and thus

Table 4 Cost sensitivity results changing parameter μ_x ($\mu_\alpha = 2 - \mu_x$)

| μ_x	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.
1.0	Newcomer [44.39, 45.02]				Small Chain [207.87, 210.81]				Large Chain [240.05, 243.44]			
0.1	0.008	3.363	3.400	−34.2	0.022	0.180	0.507	−4.7	0.265	1.461	0.557	+8.1
0.5	0.002	1.682	1.717	−24.1	0.005	0.076	0.185	−3.5	0.264	1.300	0.446	+4.4
0.9	0.000	0.043	0.037	−6.2	0.000	0.004	0.010	−1.0	0.051	0.182	0.079	+0.9
1.1	0.000	0.020	0.018	+6.4	0.000	0.004	0.011	+1.1	0.007	0.033	0.043	0.0
1.5	0.001	0.104	0.090	+31.9	0.010	0.436	0.773	+8.6	1.279	1.929	3.612	+20.2
1.9	0.002	0.126	0.114	+57.4	0.033	4.419	4.519	+31.7	1.278	2.921	4.498	+45.7
1.0	Aggr. Newcom. [53.0, 53.75]				Aggr. Small Ch. [207.7, 210.6]				Aggr. Large Ch. [243.9, 247.4]			
0.1	0.002	2.779	2.798	−46.0	2.017	0.176	2.293	−5.0	3.092	4.365	5.452	−7.1
0.5	0.001	1.045	1.073	−30.6	0.730	0.073	0.851	−3.6	0.618	0.708	1.542	−9.4
0.9	0.000	0.007	0.010	−6.8	0.000	0.004	0.007	−0.9	0.000	0.011	0.021	−3.2
1.1	0.000	0.002	0.004	+6.8	0.000	0.004	0.009	+1.1	0.000	0.006	0.021	+3.9
1.5	0.001	0.021	0.027	+34.2	0.004	0.299	0.462	+7.8	0.000	0.571	0.780	+28.0
1.9	0.003	0.031	0.044	+61.5	0.677	3.518	3.965	+26.6	0.000	1.079	1.329	+55.4
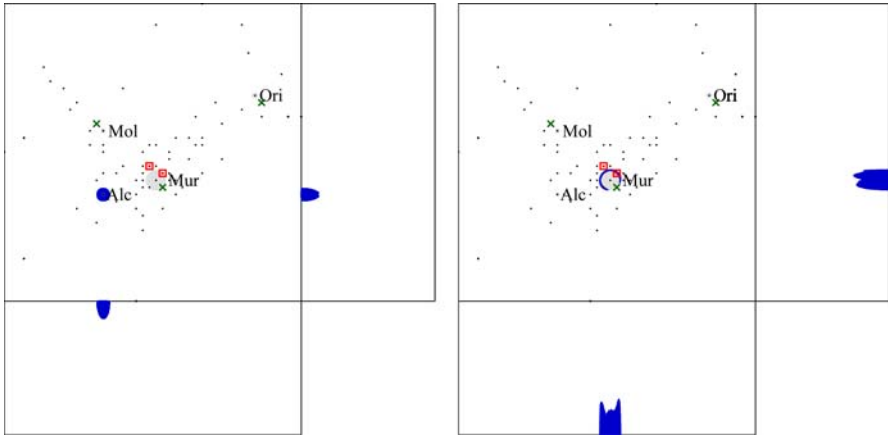
Fig. 10 The results for Large Chain for $\mu_x = 0.5$ (*left*) and $\mu_x = 1.5$ (*right*)

μ_α decreases, it becomes cheaper to increase quality, if possible, which will increase sales as long as the site does not change, thus increasing profits. This is exactly what happens in both Newcomer instances and in the unaggregated Small Chain instance, where the best location remains stable throughout.

For the Newcomer the maximum quality is almost reached in the base case (see Fig. 4). When μ_x increases above 1 both site and quality remain fixed, resulting in a corresponding linear increase in profit. When μ_x decreases, however, we see a steady decrease of optimal quality and a less than linear decrease in profit.

For the Small Chain the location remains stable too, but now quality starts at around minimum value and very slowly increases towards the still low 1.4 in the basic case. This slow increase goes on but accelerates towards the end, not quite reaching the maximum of 5, even when quality is cheapest. As the quality was quite similar at the two δ-optimal sites, Molina and Orihuela, both places remain optimal with the only change in quality. The same holds in the aggregated case, although at the extremes values of μ_x Molina also becomes part of the δ-optimal region, as in the unaggregated case.

For the Large Chain, in the base case, the optimal solution is at Alcantarilla with low quality, and also at Murcia with relatively high quality. When μ_x decreases Murcia is no longer a suitable place to locate: the high quality needed there becomes too expensive, see Fig. 10 left for $\mu_x = 0.5$. When μ_x increases, however, Alcantarilla is no longer a suitable place, because a location in Murcia at the higher quality that now becomes affordable obtains more and more profit, see Fig. 10 right for $\mu_x = 1.5$. Note also that this shift in optimal site may lead to increasing profits in both directions.

With the concentrated demand, this last effect is more pronounced around the same shift in the best site, but it occurs already at the much lower value $\mu_x = 0.5$.

9 The impact of budget restrictions

In our final study, we investigated the influence of a limited budget on the optimal solution. This was done as follows. The optimal solutions (x^*, α^*) of the base cases

Table 5 Results on budget restriction

|	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.	Loc.	Qual.	3D	Obj.
	Newcomer [44.39, 45.02]				Small Chain [207.87, 210.81]				Large Chain [240.05, 243.44]			
95%	0.000	0.081	0.098	−0.4	0.127	0.192	0.443	−0.1	0.490	1.711	0.905	−4.5
85%	0.001	0.876	0.904	−2.4	0.128	0.311	0.584	−0.5	4.914	1.445	4.732	−14.3
75%	0.002	1.906	1.932	−6.4	0.129	0.482	0.761	−1.6	5.155	1.327	5.381	−16.0
65%	0.005	3.023	3.050	−12.9	0.130	0.713	1.000	−3.8	5.525	1.439	5.826	−16.9
55%	0.012	4.247	4.273	−22.8	0.132	1.049	1.339	−7.9	6.281	1.526	6.610	−16.9
	Aggr. Newcom. [53.0, 53.75]				Aggr. Small Ch. [207.7, 210.6]				Aggr. Large Ch. [243.9, 247.4]			
95%	0.000	0.186	0.224	−1.1	0.001	0.203	0.325	−0.1	0.000	0.331	0.299	−0.2
85%	0.000	1.082	1.102	−4.7	0.002	0.344	0.478	−0.6	0.000	0.617	0.566	−1.0
75%	0.000	2.062	2.075	−10.3	0.003	0.494	0.645	−2.0	0.001	0.970	0.982	−3.0
65%	0.001	3.137	3.158	−18.5	0.005	0.752	0.923	−4.6	0.001	1.643	1.709	−6.3
55%	0.002	4.295	4.318	−30.9	0.046	1.052	1.214	−9.5	0.914	2.625	3.531	−11.4

yielded a corresponding maximally needed budget $B^* = G(x^*, \alpha^*)$. Then, we added a budget constraint of the form $G(x, \alpha) \leq B$, and solved these problems by reducing the available budget B to 95, 85, ..., 55% of B^* for the corresponding base case. The results are summarized in Table 5. All values are shown as dissimilarities as compared to the corresponding base case near-optimal region.

Note that when B increases, we obtain a relaxation of the problem at hand, so maximal profit cannot decrease. Therefore each 'Obj' column in the table should show decreasing negative values.

When the budget decreases slightly, the only effect observed is that quality is lowered in order to stay within the budget, resulting in a corresponding loss in sales, thus in profit. But at some point, the budget becomes so small that lowering quality does not suffice anymore and one also has to move to cheaper places. This typically means moving away from the population centres. In the Newcomer and Small Chain instances, for example, the optimal location remains fixed, and we observe only a reduction in the quality of the new facility as B decreases. Additionally, the Small Chain loses one of its two near-optimal locations in the move from 100 to 95%.

This spatial outward effect is fully due to the locational cost part in our model, and may become so large that the facility is pushed completely towards the boundaries of the feasible region. This happens for example for the Large Chain case at very low budgets, as illustrated in Fig. 11. The inner region is now unfeasible due to the added constraint, and one obtains a wide boundary region at low quality. Observe that the sheer size of this near-optimal region indicates a very flat objective function. This flatness is also the reason why the optimal profit is almost insensitive for a whole range of low budgets.

Fig. 11 Results for Large Chain when the budget is restricted by 55%
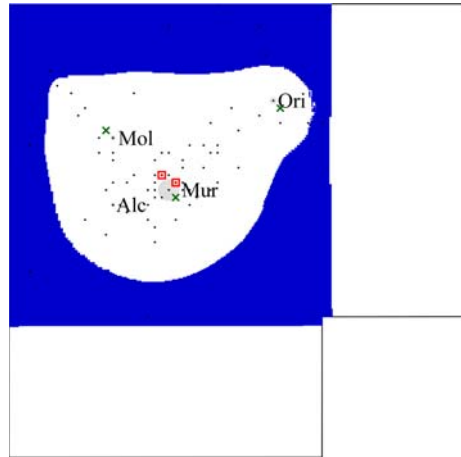
Overall the profit losses due to budget restrictions may be considered as reasonable in absolute terms in all cases. However, the relative losses due to budget constraints for the Newcomer move up to 25%. The Small Chain loses the least. This is because it starts out in the unconstrained base case by choosing a not too expensive location (Orihuela) at rather low quality. Thus, when the budget decreases it can cope by a small decrease of its quality and still capture most of its original market share. In contrast, the Newcomer has to decrease its quality by much more, so it loses more market share, while the Large Chain has to move out to much cheaper locations because reducing quality cannot compensate for the reduction in budget.

10 Conclusions

Not all results collected have led to significant conclusions, but we felt it important to give the reader access to the full study results. It should be considered as a first broad exploratory step towards a better understanding of Huff-type competitive models. In view of the relatively modest size of each particular sensitivity analysis, the conclusions remain, of course, tentative. But they have allowed to prove that several parameters of the model play a significant role, and should therefore be estimated with care in real applications in order to have confidence in the decisions suggested by the model.

Let us conclude by summarizing the main conclusions we consider to have obtained.

– First of all, it is very important to decide the location and the quality simultaneously, because each one may influence the other considerably (see also Tóth et al. 2009).
– Our experiments have shown the importance of using global optimization methods, in particular methods that guarantee that the found optimum is global and that they construct a region of near-optimality. This region is a great help to the decision maker to choose the best location and quality, in particular by indicating possible alternatives.

– All studied parameters have a direct impact on profit evaluations, and most of them turned out to be able to sometimes strongly modify the best choice. Therefore, estimation of data and parameters should always be done carefully, particularly when the profit forecasts are requested to be accurate.
– The technical parameters influencing the tradeoff between costs and income are very important.
– When demand is aggregated there is an almost systematic tendency to underestimate optimal profits. A more important effect is that the objective function usually becomes sharper with aggregation, and therefore may lead to the disappearance of good alternate optimal solutions.
– A newcomer in the market works in a considerably more hilly environment, and usually only a single best option appears, remaining quite insensitive to data changes.
– For cases when the owner's chain is already present, wrongly approximated data can cause wrong estimation, not only in profit but in location and design as well. Thus, apart from accurate approximation of the data and parameters, the decision maker has to look not only at the globally optimal location and design, but also at the near-optimal regions. We observed that the Large Chain has the flattest function, and thus larger near-optimal regions, often showing more than one almost equivalent options. The best choice among these may vary with relatively small parameter changes. For these larger chains, one may also recommend computing a larger near-optimal region and select from the possible new insights according to the chain's requirements and facilities.

References

Aboolian R, Berman O, Krass D (2007) Competitive facility location and design problem. Eur J Oper Res 182:40–62
Averbakh I, Berman O (2003) An improved algorithm for the minmax regret median problem on a tree. Networks 41:97–103
Berman O, Krass D (2002) Locating multiple competitive facilities: spatial interaction models with variable expenditures. Ann Oper Res 111:197–225
Centro Regional de Estadística de Murcia. Anuario estadístico de la Región de Murcia (2003) Dirección General de Economía, Planificación y Estadística, Centro Regional de Estadística de Murcia, 2003
Dasci A, Laporte G (2005) An analytical approach to the facility location and capacity acquisition problem under demand uncertainty. J Oper Res Soc 56:397–405
Drezner T (1994) Optimal continuous location of a retail facility, facility attractiveness, and market share: an interactive model. J Retail 70:49–64
Drezner T, Drezner Z (1998) Facility location in anticipation of future competition. Locat Sci 6:155–173
Drezner T, Drezner Z (2002) Retail facility location under changing market conditions. IMA J Manag Math 13:283–302
Drezner T, Drezner Z (2004) Finding the optimal solution to the Huff based competitive location model. Comput Manag Sci 1:193–208
Drezner T, Eiselt HA (2001) Consumers in competitive location models. In: Drezner Z, Hamacher HW (eds) Facility location: applications and theory. Springer, Heidelberg, pp 151–178
Drezner Z, Suzuki A (2004) The big triangle small triangle method for the solution of non-convex facility location problems. Oper Res 52:128–135
Eiselt HA, Laporte G (1996) Sequential location problems. Eur J Oper Res 96:217–231
Eiselt HA, Laporte G (1998) Demand allocation functions. Locat Sci 6:175–187

Eiselt HA, Laporte G, Thisse J-F (1993) Competitive location models: a framework and bibliography. Trans Sci 27:44–54

Fernández J, Pelegrín B (2000) Sensitivity analysis in continuous location models via interval analysis. Stud Locat Anal 14:121–136

Fernández J, Pelegrín B, Plastria F, Tóth B (2007) Solving a Huff-like competitive location and design model for profit maximisation in the plane. Eur J Oper Res 179:1274–1287

Francis RL, Lowe TJ, Tamir A (2002) Demand point aggregation of location models. In: Drezner Z, Hamacher H (eds) Facility location: applications and theory. Springer, Berlin

Francis RL, Lowe TJ, Rayco B, Tamir A (2008) Aggregation error for location models: survey and analysis. Ann Oper Res (to appear). doi:10.1007/s10479-008-0344-z

Hakimi SL (1983) On locating new facilities in a competitive environment. Eur J Oper Res 12:29–35

Hammer R, Hocks M, Kulisch U, Ratz D (1995) C++ toolbox for verified computing I: basic numerical problems: theory, algorithms, and programs. Springer, Berlin

Hansen E (1992) Global optimization using interval analysis. Marcel Dekker, New York

Hansen P, Peeters D, Richard D, Thisse JF (1985) The minisum and minimax location problems revisited. Oper Res 33:1251–1265

Hodgson MJ (1981) A location-allocation model maximizing consumers' welfare. Reg Stud 15:493–506

Huff DL (1964) Defining and estimating a trading area. J Mark 28:34–38

Jain AK, Mahajan V (1979) Evaluating the competitive environment in retailing using multiplicative competitive interactive models. In: Sheth J (ed) Research in marketing. JAI, Greenwich, pp 217–235

Kearfott RB (1996) Rigorous global search: continuous problems. Kluwer, Dordrecht

Knüppel O (1993) PROFIL/BIAS—a fast interval library. Computing 53:277–287

McGarvey RG, Cavalier TM (2005) Constrained location of competitive facilities in the plane. Comput Oper Res 32:359–378

Nakanishi M, Cooper LG (1974) Parameter estimate for multiplicative interactive choice model: least square approach. J Mark Res 11:303–311

Plastria F (1992) GBSSS: the generalized big square small square method for planar single-facility location. Eur J Oper Res 62:163–174

Plastria F (2001) Static competitive facility location: an overview of optimisation approaches. Eur J Oper Res 129:461–470

Ratschek H, Rokne J (1988) New computer methods for global optimization. Ellis Horwood, Chichester

Reilly WJ (1931) The law of retail gravitation, 2nd edn. Knickerbocker Press, New York

Reilly WJ (1953) The law of retail gravitation. Pilsbury Publishers, New York

Sáiz ME, Hendrix EMT, Fernández J, Pelegrín B (2008) On a branch-and-bound approach for a Huff-like Stackelberg location problem. OR Spectrum, (to appear) (doi:10.1007/s00291-008-0133-8)

Tóth B, Fernández J, Pelegrín B, Plastria F (2009) Sequential versus simultaneous approach in the location and design of two new facilities using planar Huff-like models. Comput Oper Res 36:1393–1405

Tóth B, Plastria F, Fernández J, Pelegrín B (2008) Sensitivity analysis of the optimal solutions to Huff-type competitive location and design problems. Optimization Online. http://www.optimization-online.org/DB_HTML/2008/11/2155.html

Verter V (2002) An integrated model for facility location and technology acquisition. Comput Oper Res 29:583–592

Verter V, Dasci A (2002) The plant location and flexible technology acquisition problem. Eur J Oper Res 136:366–382

Visual G.I.S. Engineering (2003) VisualMap. Basauri 17 (La Florida) 28023 Madrid