Annals of Operations Research 110, 167–181, 2002

Conditional Median: A Parametric Solution Concept for Location Problems

WŁODZIMIERZ OGRYCZAK ogryczak@ia.pw.edu.pl
Warsaw University of Technology, Institute of Control and Computation Engineering, 00-665 Warsaw, Poland

MARIUSZ ZAWADZKI zawadzki@mimuw.edu.pl
Warsaw University, Institute of Informatics, 02-097 Warsaw, Poland

Abstract. Classical approaches to location problems are based on the minimization of the average distance (the median concept) or the minimization of the maximum distance (the center concept) to the service facilities. The median solution concept is primarily concerned with the spatial efficiency while the center concept is focused on the spatial equity. The k-centrum model unifies both the concepts by minimization of the sum of the k largest distances. In this paper we investigate a solution concept of the conditional median which is a generalization of the k-centrum concept taking into account the portion of demand related to the largest distances. Namely, for a specified portion (quantile) of demand we take into account the entire group of the corresponding largest distances and we minimize their average. It is shown that such an objective, similar to the standard minimax, may be modeled with a number of simple linear inequalities. Equitable properties of the solution concept are examined.

Keywords: location, multiple criteria, efficiency, equity, median, center, cent-dian, k-centrum, conditional median

1. Introduction

A host of operational models has been developed to deal with facility location optimization (cf. [7,9,13]). Most classical location studies focus on the minimization of the mean (or total) distance (the median concept) or the minimization of the maximum distance (the center concept) to the service facilities [19]. The median solution concept is primarily concerned with the spatial efficiency. As based on averaging, it often provides solutions where remote and low-population density areas are discriminated in terms of accessibility to facilities, as compared with centrally situated and high-population density areas. For this reason, while locating public services the center solution concept is usually applied to minimize the maximum distance (travel time) between any consumer and the closest facility. As the minimax objective primarily addresses the geographical equity issues, this approach is of particular importance in spatial organization of emergency service systems.

The center approach is consistent with the Rawlsian [25] theory of justice, especially when additionally specified as the lexicographic center [20]. On the other hand, locating a facility at the center may cause a large increase in the total distance thus generating a substantial loss in spatial efficiency. This has led to a search for some compromise solution concept. Halpern [10] introduced the λ-cent-dian as a parametric solution concept based on the convex combination of the two objectives representing the minisum and the minimax approaches. Unfortunately, due to the lack of convexity, the solution concept of λ-cent-dian may fail to provide a compromise location in the case of discrete problems [21].

Another compromise solution concept was introduced by Slater [29] as the so-called k-centrum where the sum of the k largest distances is minimized. If $k = 1$ the model reduces to the standard center concept while with $k = m$ it is equivalent to the classical median model. Early works on the solution concept [1,2,29] were focused on the case of the discrete single facility location on tree graphs. Peeters [24] studied the single facility problem on a graph and he introduced the full classification of the related criteria and solution concepts. Consistently with typical distribution characteristics, four optimization criteria on outcomes (distances) were introduced: upper (lower) k-median where the sum of the k largest (smallest) outcomes was minimized, and upper (lower) k-center where the k largest (smallest) outcome itself was minimized. According to this classification, the k-centrum should be rather called the upper k-median. Recently, Tamir [30] has presented polynomial time algorithms for solving the multiple facility k-centrum (upper k-median) on path and tree graphs, while Ogryczak and Tamir [23] have shown that the upper k-median criterion can be modeled with an auxiliary set of simple linear inequalities thus simplifying several k-centrum models.

The k-centrum concept is restricted to unweighted problems. Although some weights are used to scale the specific distances [30] (which may be considered as a definition of distance dependent outcomes), the demand weights as defining the distribution of clients are not considered. In this paper we introduce a parametric generalization of the k-centrum concept applied to weighted problems by taking into account the portion of demand related to the largest outcomes (distances). Namely, for a specified portion β of demand we take into account the entire β portion (quantile) of the largest outcomes and we consider their average as the *(worst) conditional β-mean* outcome. Following the classification by Peeters [24], we call an *(upper) conditional median* every location pattern which minimizes the corresponding conditional mean outcome. According to this definition the concept of conditional median is based on averaging restricted to the portion of the worst outcomes. When parameter β approaches 0, the conditional β-mean tends to the largest outcome and the conditional median becomes the center. On the other hand, for $\beta = 1$ the corresponding conditional median becomes the standard median. For the unweighted location problems and $\beta = k/m$, the conditional β-mean represents the average of the k largest outcomes thus modeling the k-centrum solution concept.

One of the disadvantages of the minimax approach to location problems is that it is too crude and many quite different feasible solutions may be optimal with respect to the minimax criterion. While using standard algorithmic tools to identify the minimax solution, one of many solutions is selected randomly. It causes that the centers are highly unstable [26]. Furthermore, it often turns out that the distribution of spatial units in relation to the location of facilities may make the minimax criterion partially passive. It arises, for instance, when an isolated spatial unit is located at a considerable distance from all the locations of facilities. Minimization of the maximum distance is then reduced to the minimization of the distance of that single isolated spatial unit [15] leaving other location decisions unoptimized. The concept of conditional median, due to averaging within the group of the worst outcomes, reduces this flaw of the center approach.

While locating public facilities, the issue of equity is becoming important. Equity is, essentially, an abstract socio-political concept that implies fairness and justice [27]. Nevertheless, equity can be quantified [3] and several equity maximizing approaches to location problems have been developed and analyzed [6,8,14,16,17]. Moreover, the concept of equitable multiple criteria optimization [12] is well suited for the locational analysis [11,22]. The center model is an equitable approach [20]. It turns out that the concept of conditional median preserves this property allowing simultaneously for wider modeling of equitable preferences with the parameter.

The paper is organized as follows. In the next section we give a formal definition of the conditional median solution concept. We show there that, similarly to the standard center and the k-centrum [23], the conditional median may be found by solving an optimization problem with a linear objective and a number of auxiliary linear inequalities. In section 3 we compare the conditional median to the classical parametric solution concept of the λ-cent-dian [10]. It turns out that the former much better allows us to model a compromise between the center and the median approaches. Finally, in section 4 we discuss equitable properties of the conditional median. The conditional mean outcomes turn out to be closely related to the absolute Lorenz curve which implies the equitable properties of the corresponding solution concept.

2. The solution concept

The generic location problem that we consider may be stated as follows. There is given a set $I = \{1, 2, \ldots, m\}$ of m clients (service recipients). Each client is represented by a specific point in the geographical space. There is also given a set Q of location patterns (location decisions). For each client i ($i \in I$) a real valued function $f_i(\mathbf{x})$ of the location pattern $\mathbf{x}$ has been defined. This function, called the individual objective function, measures the outcome (effect) $y_i = f_i(\mathbf{x})$ of the location pattern for client i [17]. In the simplest problems an outcome usually expresses the distance. However, we emphasize to the reader that we do not restrict our considerations to the case of outcomes measured as distances. They can be measured (modeled) as travel time, travel costs as well as in a more subjective way as relative travel costs (e.g., travel costs by clients incomes) or ultimately as the levels of clients dissatisfaction (individual disutility) of locations. In typical formulations of location problems related to desirable facilities a smaller value of the outcome (distance) means a better effect (higher service quality or client satisfaction). This remains valid for location of obnoxious facilities if the distances are replaced with their complements to some large number. Therefore, without loss of generality, we can assume that each individual outcome y_i is to be minimized, and the location problem may be stated as the following multiple criteria minimization problem:

$$\min\left\{\mathbf{f}(\mathbf{x}):\ \mathbf{x} \in Q\right\} \tag{1}$$

where $\mathbf{f} = (f_1, \ldots, f_m)$ is a vector of the individual objective functions which measure the outcome (effect) $y_i = f_i(\mathbf{x})$ of the location pattern $\mathbf{x}$ for client i.

Typically, some additional weights $w_i > 0$ are included into the location model to represent the service demand. Integer weights can be directly interpreted as numbers of unweighted clients located at exactly the same place (with distances 0 among them). Theoretically, one may consider that the weighted problem is transformed (disaggregated) to the unweighted one (with all the demand weights equal to 1). Such a disaggregation is possible for integer as well as rational weights, but it usually dramatically increases the problem size. Therefore, we consider solution concepts which can be applied directly to the weighted problem. Since the demand weights describe the distribution of outcomes (distances), we will use the normalized weights

$$\bar{w}_i = w_i \Big/ \sum_{i=1}^{m} w_i \quad \text{for } i = 1, \ldots, m, \tag{2}$$

rather than the original quantities w_i. Note that, in the case of unweighted problem (all $w_i = 1$), all the normalized weights are given as $\bar{w}_i = 1/m$.

A wide gamut of location problems can be considered within the framework of model (1). We do not assume any special form of the feasible set while analyzing properties of the solution concepts. Similarly, we do not assume any special form of the individual objective functions nor their special properties (like convexity). Therefore, the results of our analysis apply to various location problems covering discrete as well as continuous location decisions. The only assumption specifies the finite set of clients.

The classical approaches to location problems use either the median or the center solution concept. Both the median and the center solution concepts are well defined for location models using weights $w_i > 0$ to represent service demand (several clients at the same geographical point). Exactly, for the weighted location problem, we consider, the *median* solution concept is defined by the minimization of the objective function expressing the *mean* (average) outcome

$$\mu(\mathbf{y}) = \sum_{i=1}^{m} \bar{w}_i y_i$$

but it is also equivalent to the minimization of the total outcome $\sum_{i=1}^{m} w_i y_i$. The *center* solution concept is defined by the minimization of the objective function representing the *maximum* (worst) outcome

$$M(\mathbf{y}) = \max_{i=1,\ldots,m} y_i$$

and it is not affected by the demand weights at all.

A natural generalization of the maximum outcome $M(\mathbf{y})$ is the (worst) conditional mean outcome defined as the mean of the specified size (quantile) of the worst (largest) outcomes. For the simplest case of the unweighted location problem one may distinguish the k largest outcomes (the k worst-off clients) and define the conditional mean outcome as the mean of the k distinguished outcomes. This can be mathematically formalized as follows. First, we introduce the ordering map $\Theta : R^m \to R^m$ such that $\Theta(\mathbf{y}) = (\theta_1(\mathbf{y}), \theta_2(\mathbf{y}), \ldots, \theta_m(\mathbf{y}))$, where $\theta_1(\mathbf{y}) \geqslant \theta_2(\mathbf{y}) \geqslant \cdots \geqslant \theta_m(\mathbf{y})$ and there exists a permutation τ of set I such that $\theta_i(\mathbf{y}) = y_{\tau(i)}$ for $i = 1, \ldots, m$. The *(worst) conditional k/m-mean outcome $M_{k/m}(\mathbf{y})$* is given then as

$$M_{k/m}(\mathbf{y}) = \frac{1}{k} \sum_{i=1}^{k} \theta_i(\mathbf{y}) \quad \text{for } k = 1, \ldots, m. \tag{3}$$

The minimization of the criterion (3) defines the upper k-median model [24] (or k-centrum).

The quantity $\theta_1(\mathbf{y})$, representing the largest distance, can easily be computed with auxiliary linear inequalities:

$$\begin{aligned} \theta_1(\mathbf{y}) = \min\ & t \\ \text{s.t.}\quad & y_i \leqslant t \quad \text{for } i = 1, \ldots, m. \end{aligned}$$

A similar formula can be given for any $\theta_k(\mathbf{y})$ although requiring the use of integer (binary) variables. Namely, for any $k = 1, 2, \ldots, m$, the following formula is valid:

$$\begin{aligned} \theta_k(\mathbf{y}) = \min\ & t \\ \text{s.t.}\quad & y_i \leqslant t + S z_i,\ z_i \in \{0, 1\} \quad \text{for } i = 1, \ldots, m, \\ & \sum_{i=1}^{m} z_i \leqslant k - 1, \end{aligned}$$

where S is a sufficiently large constant (larger than any possible difference between various individual outcomes y_i). This allows us to define, for any $0 < \beta \leqslant 1$, the *conditional β-maximum* outcome $C_\beta(\mathbf{y})$ as:

$$\begin{aligned} C_\beta(\mathbf{y}) = \min\ & t \\ \text{s.t.}\quad & y_i \leqslant t + S z_i,\ z_i \in \{0, 1\} \quad \text{for } i = 1, \ldots, m, \\ & \sum_{i=1}^{m} \bar{w}_i z_i < \beta. \end{aligned}$$

The minimization of $C_\beta(\mathbf{y})$ leads to the *conditional β-center* as a generalization of the k-center solution concept [24].

A similar approach can be used to generalize the upper k-median model. For the special cases of $k = 1$ and $k = m$ we get $M_{1/m}(\mathbf{y}) = \theta_1(\mathbf{y}) = M(\mathbf{y})$ and

$M_{m/m}(\mathbf{y}) = (1/m)\sum_{i=1}^{m} \theta_i(\mathbf{y}) = (1/m)\sum_{i=1}^{m} y_i = \mu(\mathbf{y})$, respectively, thus representing the classical criteria. For any $k = 1, 2, \ldots, m$, the following formula is valid:

$$M_{k/m}(\mathbf{y}) = \min\left(t + \frac{1}{k}\sum_{i=1}^{m} d_i\right)$$
$$\text{s.t.} \quad y_i \leqslant t + d_i,\ d_i \geqslant 0 \quad \text{for } i = 1, \ldots, m,$$
$$d_i \leqslant Sz_i,\ z_i \in \{0, 1\} \quad \text{for } i = 1, \ldots, m,$$
$$\sum_{i=1}^{m} z_i \leqslant k - 1$$

with an arbitrary large constant S. Again, this allows us to define, for any $0 < \beta \leqslant 1$, the *conditional β-mean* outcome $M_\beta(\mathbf{y})$ as:

$$M_\beta(\mathbf{y}) = \min\left(t + \frac{1}{\beta}\sum_{i=1}^{m} \bar{w}_i d_i\right)$$
$$\text{s.t.} \quad y_i \leqslant t + d_i,\ d_i \geqslant 0 \quad \text{for } i = 1, \ldots, m,$$
$$d_i \leqslant Sz_i,\ z_i \in \{0, 1\} \quad \text{for } i = 1, \ldots, m, \qquad (4)$$
$$\sum_{i=1}^{m} \bar{w}_i z_i < \beta.$$

The minimization of $M_\beta(\mathbf{y})$ defines the *conditional β-median* solution concept.

Note that, due to the finite distribution of outcomes y_i $(i = 1, \ldots, m)$, the optimization (4) is well defined. A linear programming model for $M_{k/m}(\mathbf{y})$ computation has been recently given in [23]. The following theorem generalizes this result showing that for any β the integer variables (and the corresponding constraints) in (4) are redundant and the conditional β-mean can be found by a simple linear programming minimization.

Theorem 1. For any outcome vector $\mathbf{y} \in R^m$ with the corresponding demand weights w_i, and for any real value $0 < \beta \leqslant 1$, the conditional β-mean outcome is given by the following linear program:

$$M_\beta(\mathbf{y}) = \min\left\{t + \frac{1}{\beta}\sum_{i=1}^{m} \bar{w}_i d_i\colon\ y_i \leqslant t + d_i,\ d_i \geqslant 0,\ \text{for } i = 1, \ldots, m\right\}. \qquad (5)$$

Proof. Consider an optimal solution $(t, d_1, \ldots, d_m)$ of problem (5), where the number of positive variables d_i is minimal. Let $I_+ = \{i\colon d_i > 0\}$. Define $z_i = 1$ for $i \in I_+$ and $z_i = 0$ for $i \notin I_+$. If $\sum_{i=1}^{m} \bar{w}_i z_i < \beta$, then one gets a solution to (4). Otherwise, by introducing $\tilde{t} = t + \Delta$, $\tilde{d}_i = d_i - \Delta$ for $i \in I_+$, $\tilde{d}_i = d_i$ for $i \notin I_+$, and $\Delta = \min_{i \in I_+} d_i$, one gets $\beta\tilde{t} + \sum_{i=1}^{m} \bar{w}_i \tilde{d}_i \leqslant \beta t + \sum_{i=1}^{m} \bar{w}_i d_i$. Hence, we obtain an optimal solution to (5), where the number of positive variables d_i is strictly smaller than $|I_+|$. This completes the proof. □

It follows from theorem 1 that the conditional β-median for the weighted location problem can be found as the optimal solution to the following problem:

$$\min\left\{t + \frac{1}{\beta}\sum_{i=1}^{m} \bar{w}_i d_i \colon \mathbf{x} \in Q;\ f_i(\mathbf{x}) \leqslant t + d_i,\ d_i \geqslant 0,\ \text{for } i = 1, \ldots, m\right\}, \tag{6}$$

or in a more compact form:

$$\min\left\{t + \frac{1}{\beta}\sum_{i=1}^{m} \bar{w}_i \big(f_i(\mathbf{x}) - t\big)^+ \colon \mathbf{x} \in Q\right\},$$

where $(\cdot)^+$ denotes the nonnegative part of a number.

For the special case of an unweighted location problem, one gets the conditional k/m-median model:

$$\min\left\{t + \frac{1}{k}\sum_{i=1}^{m} d_i \colon \mathbf{x} \in Q;\ f_i(\mathbf{x}) \leqslant t + d_i,\ d_i \geqslant 0,\ \text{for } i = 1, \ldots, m\right\} \tag{7}$$

which is the same as the computational formulation of the k-centrum model introduced in [23]. Hence, theorem 1 and formulation (6) generalize the k-centrum formulation of [23] allowing to consider demand weights but preserving the simple structure and dimension of the optimization problem.

3. Conditional median versus cent-dian

The conditional median concept provides a compromise between the center and the median models. Table 1 shows the quality of this compromise in terms of model (1). It provides the average percentage distribution of outcomes (i.e., $y_i = f_i(\mathbf{x})$ for $i = 1, \ldots, m$)

Table 1
Average distribution of outcomes.

Objective	Percentage distribution of outcomes ($m = 50,\ p = 3$)										
	0–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51+
M	8.0	7.6	10.6	13.2	15.8	18.6	13.4	10.2	2.6	0.0	0.0
$M_{0.1}$	9.0	5.8	10.0	16.6	15.8	19.6	14.2	7.4	1.6	0.0	0.0
$M_{0.2}$	8.6	6.2	11.0	16.2	18.4	21.0	10.8	5.4	2.2	0.2	0.0
$M_{0.3}$	8.6	7.8	10.4	15.6	19.4	20.6	9.6	5.2	2.6	0.2	0.0
$M_{0.4}$	9.0	8.0	11.6	16.8	20.6	17.6	8.0	5.6	1.6	0.8	0.4
$M_{0.5}$	8.8	8.6	14.0	20.0	18.0	14.6	6.4	5.2	2.0	1.4	1.0
$M_{0.7}$	9.0	9.2	16.0	19.6	16.0	13.8	5.4	5.2	2.6	2.0	1.2
$M_1 = \mu$	10.2	11.2	16.0	16.6	13.8	12.6	8.2	5.6	3.4	1.6	0.8

for compromise locations obtained by varying parameter β in the objective M_β to define respective conditional medians. There are also included the results for the center and the median as the limiting cases. The outcomes are evaluated for 10 discrete location problems consisting of $m = 50$ client locations with randomly generated distance matrices and three facilities ($p = 3$) to be placed among the client locations. We have generated m random (uniformly distributed) integer points with coordinates ranging from 0 to 100. The distances among the points have been defined by the Euclidean metric and then rounded to integers. Possible outcomes are partitioned into clusters of range five. Each row represents the average distribution for a particular β. Exactly, each field gives the percentage of outcomes within a given range in 10 optimal locations.

It is clear that percentage of low outcomes increases with β (first four columns). For small β, percentage of large outcomes is forced to zero. With β increasing, large outcomes occur incidentally, however, their overall percentage is smaller (last five columns). Compromise solutions close to the median maximize the number of shortest distances allowing some outcomes to exceed the minimax optimal value.

Halpern [10] has introduced the λ-cent-dian as a parametric solution concept based on minimization of the convex combination of two objectives:

$$H_\lambda(\mathbf{y}) = \lambda M(\mathbf{y}) + (1 - \lambda)\mu(\mathbf{y}) \quad \text{for } 0 \leqslant \lambda \leqslant 1. \tag{8}$$

The location pattern $\mathbf{x} \in Q$ which minimizes $H_\lambda(\mathbf{f}(\mathbf{x}))$ is called *λ-cent-dian*. The λ-cent-dian covers as special cases the center ($\lambda = 1$) and the median ($\lambda = 0$) solution concepts.

Both $H_\lambda(\mathbf{y})$ for $0 \leqslant \lambda \leqslant 1$ and $M_\beta(\mathbf{y})$ for $0 < \beta \leqslant 1$ provide us with a tool for modeling various compromise solutions between the center and the median. However, opposite to the λ-cent-dian, the conditional median is not based on the direct weighting of criteria. Therefore, the latter is also well applicable to discrete location problems where the former may fail to generate a compromise location, due to the lack of convexity. We illustrate this with the following example [21].

Example 2. Let us consider a single facility location problem defined on the network presented in figure 1. We consider all the clients located at vertices v_i ($i = 1, 2, \ldots, 8$) with the service demand defined by weights: $w_1 = w_2 = 47$ and $w_3 = w_4 = w_5 = w_6 = w_7 = w_8 = 1$. Note that these data define the entire set of median solutions on the edge $[v_1, v_2]$. Among them, point a (in the middle of edge $[v_1, v_2]$) is the median with the best value of the maximum distance. The network has the unique center at point c (in the middle of edge $[v_5, v_6]$). Point b, in the middle of edge $[v_3, v_4]$, seems to be a very interesting compromise location. We consider three points a, c and b as possible location of a single service center.

Outcome vectors for locations c, a and b are given as $\mathbf{y}^c = (5, 5, 3, 3, 1, 1, 5, 5)$, $\mathbf{y}^a = (1, 1, 3, 3, 5, 5, 9, 9)$ and $\mathbf{y}^b = (3.1, 3.1, 1.1, 1.1, 3.1, 3.1, 7.1, 7.1)$, respectively. Note that $M(\mathbf{y}^a) = 9$, $M(\mathbf{y}^c) = 5$ and $M(\mathbf{y}^b) = 7.1$ whereas (taking into account the weights w_i) $\mu(\mathbf{y}^a) = 1.28$, $\mu(\mathbf{y}^c) = 4.88$ and $\mu(\mathbf{y}^b) = 3.14$. One may easily verify that,

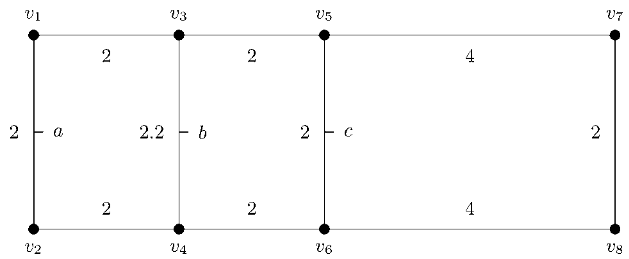

Figure 1. Sample network for example 2.

for any $0 \leqslant \lambda \leqslant 1$, $H_\lambda(\mathbf{y}^a) < H_\lambda(\mathbf{y}^b)$ or $H_\lambda(\mathbf{y}^c) < H_\lambda(\mathbf{y}^b)$. Thus, point b cannot be found as a λ-cent-dian for any $0 \leqslant \lambda \leqslant 1$.

Now, let us analyze the conditional medians. For β close to 0 one gets $M_\beta(\mathbf{y}) = M(\mathbf{y})$. For instance, with $\beta = 0.01$ we get: $M_{0.01}(\mathbf{y}^a) = 9$, $M_{0.01}(\mathbf{y}^c) = 5$, $M_{0.01}(\mathbf{y}^b) = 7.1$ and c is the corresponding conditional median. For large β, say $\beta = 0.5$, we get: $M_{0.5}(\mathbf{y}^a) = 1.56$, $M_{0.5}(\mathbf{y}^c) = 5$, $M_{0.5}(\mathbf{y}^b) = 3.26$ which means that a is the corresponding conditional median. There exist, however, values of β generating point b as the corresponding conditional median. In particular, for $\beta = 0.05$ we get: $M_{0.05}(\mathbf{y}^a) = 6.2$, $M_{0.05}(\mathbf{y}^c) = 5$, $M_{0.05}(\mathbf{y}^b) = 4.7$ which justifies location b as the corresponding conditional median.

For a further comparison between the conditional median and the cent-dian models, we have run computational experiments on discrete location problems. Given a distance matrix and number p, we search for compromise locations that are optimal with respect to the various conditional median and cent-dian models defined by varying parameter β in the objective M_β and parameter λ in the objective H_λ, respectively. As a comparison basis, we have used several distance matrices and $p = 1, 2, 3$. Indeed, we have analyzed 25 randomly generated distance matrices (10 of size $m = 25$, 10 of size $m = 50$ and 5 of size $m = 100$) as well as 5 distance matrices derived from OR-Lib [4] (pmed1, pmed2, pmed3, pmed4 and pmed5, all of size $m = 100$). For a given distance matrix and given value of p, 12 conditional medians (β varying from 0.01 to 1.00) and 12 cent-dians ($\lambda = 1 - \beta$) were found. Among them 2 represent the center and the median, respectively, while the remaining 10 (for each solution concept) represent compromise solutions (however, not necessarily distinct). The results of comparison between the models are shown in table 2 presenting the numbers of generated distinct solutions. Each table field is an average number of distinct locations (including 2 representing the center and the median solutions) found for 10 instances of distance matrix.

It turns out that varying parameter λ does not lead to many location patterns. Note that, in the case of locating a single facility ($p = 1$), 10 different intermediate values of λ do not generate any compromise solution for all 10 instances of size $m = 25$ and $m = 50$, respectively. Only in 5 percent cases varying the cent-dian parameter λ has resulted in more than one compromise solution. On the other hand, solving the conditional median instances with varying β yields much more (distinct) compromise locations. Thus, we

Table 2
Average number of distinct location patterns (center and median included).

Problem size	Distinct solutions (by varying β in M_β)			Distinct solutions (by varying λ in H_λ)		
	$p = 1$	$p = 2$	$p = 3$	$p = 1$	$p = 2$	$p = 3$
$m = 25$	2.4	3.7	4.5	2.0	2.4	2.6
$m = 50$	2.2	4.8	5.7	2.0	2.4	2.9
$m = 100$	3.6	5.1	6.2	2.1	2.8	3.1

may conclude that the conditional median concept is a more flexible tool for modeling compromise location patterns than the cent-dian approach.

4. Relation to equitable optimization

The issue of equity is important in many location decisions. Equity is usually quantified with the so-called inequality measures to be minimized. Inequality measures were primarily studied in economics [27]. However, Marsh and Schilling [17] describe twenty different measures proposed in the literature to gauge the level of equity in facility location alternatives. The simplest inequality measures are based on the absolute measurement of the spread of outcomes. In economics one usually considers relative inequality measures normalized by mean outcome. Among many inequality measures perhaps the most commonly accepted by economists is the Gini index (Lorenz measure), which has been also analyzed in the location context [8,14,16]. One can easily notice that a direct minimization of typical inequality measures (especially relative ones) contradicts the minimization of individual outcomes. As noticed by Erkut [8], it is rather a common flaw of all the relative inequality measures that while moving away from the clients to be serviced one gets better values of the measure as the relative distances become closer to one-another. As an extreme, one may consider an unconstrained continuous (single-facility) location problem and find that the facility located at (or near) infinity will provide (almost) perfectly equal service (in fact, rather lack of service) to all the clients.

The notion of equitable multiple criteria optimization [12] introduces the preference structure that complies with both the efficiency (Pareto-optimality) and with the inequality measurement rules (in particular the Pigou–Dalton approach) [27]. It is well suited for the locational analysis [11,22]. The equitable optimization can be mathematically formalized as follows. First, we introduce the left-continuous right tail cumulative distribution function:

$$F_{\mathbf{y}}(d) = \sum_{i=1}^{m} \bar{w}_i \delta_i(d) \quad \text{where } \delta_i(d) = \begin{cases} 1 & \text{if } y_i \geqslant d, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

which for any real (distance) value d provides the measure of outcomes greater or equal to d. Next, we introduce the quantile function $F_{\mathbf{y}}^{(-1)}$ as the right-continuous inverse of the cumulative distribution function $F_{\mathbf{y}}$:

$$F_{\mathbf{y}}^{(-1)}(v) = \sup\{\eta\colon\ F_{\mathbf{y}}(\eta) \geqslant v\} \quad \text{for } 0 < v \leqslant 1.$$

By integrating $F_{\mathbf{y}}^{(-1)}$ one gets:

$$F_{\mathbf{y}}^{(-2)}(0) = 0 \quad \text{and} \quad F_{\mathbf{y}}^{(-2)}(v) = \int_0^v F_{\mathbf{y}}^{(-1)}(\alpha)\,\mathrm{d}\alpha \quad \text{for } 0 < v \leqslant 1, \tag{10}$$

where $F_{\mathbf{y}}^{(-2)}(1) = \mu(\mathbf{y})$. Graphs of functions $F_{\mathbf{y}}^{(-2)}(v)$ (with respect to v) take the form of concave curves (figure 2), the *(upper) absolute Lorenz curves*. In the special case of uniform demand weights $\bar{w}_i = 1/m$, the absolute Lorenz curve is completely defined by the values $F_{\mathbf{y}}^{(-2)}(k/m) = (1/m)\sum_{i=1}^{k} \theta_i(\mathbf{y})$ for $k = 1, \ldots, m$.

The absolute Lorenz curves define the relation (partial order) of the equitable dominance. The equitable dominance is originally defined by axioms of efficiency, impartiality and the Pigou–Dalton principle of transfers [12,22]. Nevertheless, due to the results of the majorization theory [18] and its generalizations [28], it can be expressed with inequalities on the absolute Lorenz curves. Exactly, outcome vector $\mathbf{y}' \in Y$ equitably dominates $\mathbf{y}'' \in Y$, if and only if $F_{\mathbf{y}'}^{(-2)}(v) \leqslant F_{\mathbf{y}''}^{(-2)}(v)$ for all $v \in (0; 1]$ where at least one strict inequality holds. We say that a location pattern $\mathbf{x} \in Q$ is *equitably efficient* (is an equitably efficient solution of the problem (1)), if and only if there does not exist any $\mathbf{x}' \in Q$ such that $\mathbf{f}(\mathbf{x}')$ equitably dominates $\mathbf{f}(\mathbf{x})$.

In income economics the Lorenz curve is a cumulative population versus income curve. A perfectly equal distribution of income has the diagonal line as the Lorenz curve and no outcome vector can be better. The absolute Lorenz curves, used in the equitable optimization, are unnormalized taking into account also values of outcomes.

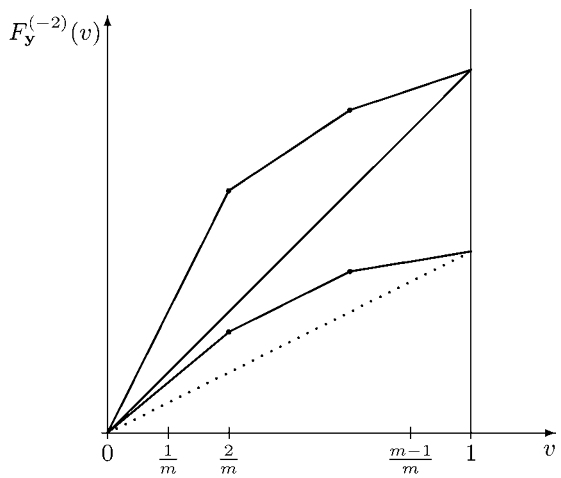

Figure 2. The (upper) absolute Lorenz curves.

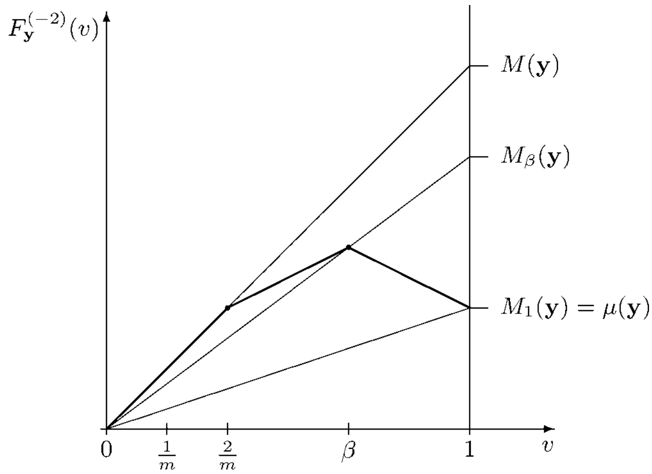

Figure 3. Absolute Lorenz curve and the conditional mean outcomes.

Vectors of equal outcomes are distinguished according to the value of outcomes. They are graphically represented with various ascent lines in figure 2. Hence, with the relation of equitable dominance an outcome vector of small unequal outcomes may be preferred to an outcome vector with large equal outcomes. This allows to overcome the common flaws of the approaches based on a strict inequality minimization [8].

Recall that the worst conditional β-mean outcome is defined as the mean of the β-quantile of the worst outcomes. Hence, $M_\beta(\mathbf{y}) = (1/\beta)\int_0^\beta F_{\mathbf{y}}^{(-1)}(\alpha)\,\mathrm{d}\alpha$ while $\int_0^\beta F_{\mathbf{y}}^{(-1)}(\alpha)\,\mathrm{d}\alpha = F_{\mathbf{y}}^{(-2)}(\beta)$ is the value of the absolute Lorenz curve at $v = \beta$. Therefore, as shown in figure 3, the conditional mean outcome represents the projection of the point of the Lorenz curve onto the vertical line at $v = 1$. Since the conditional mean outcome $M_\beta(\mathbf{y})$ is proportional to the value of the absolute Lorenz curve at a specific point β, an outcome vector $\mathbf{y}' \in Y$ equitably dominates $\mathbf{y}'' \in Y$, if and only if $M_\beta(\mathbf{y}') \leqslant M_\beta(\mathbf{y}'')$ for all $\beta \in (0; 1]$, where at least one strict inequality holds. Hence, the following assertion is valid.

Theorem 3. For any $0 < \beta < 1$, except for location patterns with identical conditional mean outcome $M_\beta(\mathbf{y})$, every location pattern $\mathbf{x} \in Q$ that is minimal for $M_\beta(\mathbf{f}(\mathbf{x}))$ is an equitably efficient solution of the location problem (1).

The perfectly equal outcome vector generates its absolute Lorenz curve as the ascent line connecting points $(0, 0)$ and $(1, \mu(\mathbf{y}))$. Hence, the space between the curve and its ascent line (the chord) represents the dispersion (and thereby the inequality) of $\mathbf{y}$ in comparison to the perfectly equal outcome vector of $\mu(\mathbf{y})$. It is called the *dispersion space* [22]. Various size parameters of the dispersion space are considered as summary characteristics of inequality – the so-called inequality measures. They allow to build the corresponding bicriteria mean/equity models [5,16,22]. Note that vertical diameter of the dispersion space at point β is given as $\beta(M_\beta(\mathbf{y}) - \mu(\mathbf{y}))$. The following assertion justifies the related mean/equity approach as an equitable optimization technique.

Theorem 4. For any $0 < \beta < 1$, except for location patterns with identical mean $\mu(\mathbf{y})$ and conditional mean outcome $M_\beta(\mathbf{y})$, every efficient solution to the bicriteria problem

$$\min\{(\mu(\mathbf{f}(\mathbf{x})), M_\beta(\mathbf{f}(\mathbf{x}))): \mathbf{x} \in Q\} \tag{11}$$

is an equitably efficient solution of the location problem (1).

Proof. Let $\mathbf{x}^0 \in Q$ be an efficient solution of problem (11). Suppose there exists a location pattern $\mathbf{x}' \in Q$ such that $\mathbf{y}' = \mathbf{f}(\mathbf{x}')$ equitably dominates $\mathbf{y}^0 = \mathbf{f}(\mathbf{x}^0)$. Then $F^{(-2)}_{\mathbf{y}'}(v) \leqslant F^{(-2)}_{\mathbf{y}^0}(v)$ for all $v \in (0; 1]$. Hence, in particular, $F^{(-2)}_{\mathbf{y}'}(\beta) \leqslant F^{(-2)}_{\mathbf{y}^0}(\beta)$ and $F^{(-2)}_{\mathbf{y}'}(1) \leqslant F^{(-2)}_{\mathbf{y}^0}(1)$. Thus, $M_\beta(\mathbf{f}(\mathbf{x}')) \leqslant M_\beta(\mathbf{f}(\mathbf{x}^0))$ and $\mu(\mathbf{f}(\mathbf{x}')) \leqslant \mu(\mathbf{f}(\mathbf{x}^0))$. This, together with the fact that $\mathbf{x}^0$ is efficient to the bicriteria optimization (11), implies $\mu(\mathbf{f}(\mathbf{x}')) = \mu(\mathbf{f}(\mathbf{x}^0))$ and $M_\beta(\mathbf{f}(\mathbf{x}')) = M_\beta(\mathbf{f}(\mathbf{x}^0))$ which completes the proof. □

For more detailed modeling of the equitable preferences one may use multiple criteria model with several conditional mean criteria for various values of the parameter β or a combination of such criteria.

5. Concluding remarks

In this paper we have introduced a solution concept of the conditional median which is a parametric generalization of the k-centrum concept (upper k-median according to the classification by Peeters [24]) taking into account the portion of demand related to the largest distances. The conditional median is shown to be much more effective in modeling various compromise location preferences than the classical cent-dian approach [10] (especially, in the case of discrete location problems). Moreover, the conditional mean outcome, used to define the solution concept of the conditional median, is closely related to the absolute Lorenz curve which implies the equitable properties of the solution concept.

Minimization of the conditional mean, similar to the standard minimax approach, may be modeled with a number of simple linear inequalities. Our limited experiments with the use of a simple general purpose MIP code show that the conditional median usually needs a computational effort larger than that for the median but smaller than that for the center. Certainly, large-scale real-life location problems will require some specialized algorithms. Therefore, research on efficient specialized algorithms for conditional medians of various specific types of location problems should be continued. Recall that for the specific case of unweighted problems (k-centrum), polynomial algorithms were shown for location on path and tree graphs [30] as well as for rectilinear problems [23].

This paper has focused on location problems. However, the location decisions are analyzed from the perspective of their effects for individual clients. Therefore, the general concept of the proposed conditional mean outcome can be used for optimization of various systems which serve many users. In particular, it offers a new promising approach to the equitable resource allocation problems.

Acknowledgment

The authors are indebted to Prof. Arie Tamir for his helpful comments.

References

[1] G. Andreatta and F.M. Mason, k-eccentricity and absolute k-centrum of a tree, European Journal of Operational Research 19 (1985) 114–117.
[2] G. Andreatta and F.M. Mason, Properties of the k-centrum in a network, Networks 15 (1985) 21–25.
[3] A.B. Atkinson, On the measurement of inequality, Journal of Economic Theory 2 (1970) 244–263.
[4] J.E. Beasley, A note on solving large p-median problems, European Journal of Operational Research 71 (1994) 270–273.
[5] O. Berman, Mean-variance location problems, Transportation Science 24 (1990) 287–293.
[6] O. Berman and E.H. Kaplan, Equity maximizing facility location schemes, Transportation Science 24 (1990) 137–144.
[7] J. Current, H. Min and D. Schilling, Multiobjective analysis of facility location decisions, European Journal of Operational Research 49 (1990) 295–307.
[8] E. Erkut, Inequality measures for location problems, Location Science 1 (1993) 199–217.
[9] R.L. Francis, L.F. McGinnis and J.A. White, *Facility Layout and Location: An Analytical Approach* (Prentice-Hall, Englewood Cliffs, NJ, 1992).
[10] J. Halpern, Finding minimal center-median convex combination (cent-dian) of a graph, Management Science 24 (1978) 534–544.
[11] M.M. Kostreva and W. Ogryczak, Equitable approaches to location problems, in: *Spatial Multicriteria Decision Making and Analysis: A Geographic Information Sciences Approach*, ed. J.-C. Thill (Ashgate, Brookfield, 1999) pp. 103–126.
[12] M.M. Kostreva and W. Ogryczak, Linear optimization with multiple equitable criteria, RAIRO Recherche Opérationnelle 33 (1999) 275–297.
[13] R.F. Love, J.G. Morris and G.O. Wesolowsky, *Facilities Location: Models and Methods* (North-Holland, New York, 1988).
[14] O. Maimon, An algorithm for the Lorenz measure in locational decisions on trees, Journal of Algorithms 9 (1988) 583–596.
[15] J. Malczewski and W. Ogryczak, A multiobjective approach to the reorganization of health-service areas: A case study, Environment and Planning A 20 (1988) 1461–1470.
[16] M.B. Mandell, Modelling effectiveness-equity trade-offs in public service delivery systems, Management Science (1991) 467–482.
[17] M.T. Marsh and D.A. Schilling, Equity measurement in facility location analysis: a review and framework, European Journal of Operational Research 74 (1994) 1–17.
[18] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, 1979).
[19] R.L. Morrill and J. Symons, Efficiency and equity aspects of optimum location, Geographical Analysis 9 (1977) 215–225.
[20] W. Ogryczak, On the lexicographic minimax approach to location problems, European Journal of Operational Research 100 (1997) 566–585.
[21] W. Ogryczak, On cent-dians of general networks, Location Science 5 (1997) 15–28.

[22] W. Ogryczak, Inequality measures and equitable approaches to location problems, European Journal of Operational Research 122 (2000) 374–391.

[23] W. Ogryczak and A. Tamir, Minimizing the sum of the k largest functions in linear time, Information Processing Letters, forthcoming.

[24] P.H. Peeters, Some new algorithms for location problems on networks, European Journal of Operational Research 104 (1998) 299–309.

[25] J. Rawls, *The Theory of Justice* (Harvard University Press, Cambridge, 1971).

[26] D. Richard, H. Beguin and D. Peeters, The location of fire stations in a rural environment: A case study, Environment and Planning A 22 (1990) 39–52.

[27] A. Sen, *On Economic Inequality* (Clarendon Press, Oxford, 1973).

[28] A.F. Shorrocks, Ranking income distributions, Economica 50 (1983) 3–17.

[29] P.J. Slater, Centers to centroids in a graph, Journal of Graph Theory 2 (1978) 209–222.

[30] A. Tamir, The k-centrum multi-facility location problem, Discrete Applied Mathematics 109 (2001) 293–307.