ISSN 2007-9737

A Generalization of the Averaged Hausdorff Distance

Andrés Vargas, Johan Bogoya

Pontificia Universidad Javeriana, Bogotá,
Colombia

$\{$a.vargasd, jbogoya$\}$@javeriana.edu.co

Abstract. The averaged Hausdorff distance Δ_p is an inframetric which has been recently used in evolutionary multiobjective optimization (EMO). In this paper we introduce a new two-parameter performance indicator $\Delta_{p,q}$ which generalizes Δ_p as well as the standard Hausdorff distance. For $p, q \geqslant 1$ the indicator $\Delta_{p,q}$ (that we call the (p, q)-*averaged distance*) turns out to be a proper metric and preserves some of the Δ_p advantages. We proof several properties of $\Delta_{p,q}$, and provide a comparison with Δ_p and the standard Hausdorff distance. For simplicity we restrict ourselves to finite sets, which is the most common case, but our results can be extended to the continuous case.

Keywords. Averaged Hausdorff distance, generational distance, inverted generational distance, multiobjective optimization, performance indicator, power means.

1 Introduction

In most cases, the solution of a multiobjective optimization problem (MOP), is a subset of $\mathbb{R}^n$ called *Pareto set* (P). Sometimes P can be computed analytically but, typically, the use of a numerical algorithm is necessary to find a reasonable finite size approximation.

To establish the accuracy of an EMO-algorithm trying to approximate this Pareto set or its image, the *Pareto front* (PF), we can measure the distance between the algorithm outcome set A and the respective Pareto set or front. Since, in general, a specific distance to PF can be attained for different sets A, this method will not produce unique solutions.

For any metric space X, the standard Hausdorff distance d_H (see [3, 6]) is a metric for $\mathcal{P}_{\mathsf{c}}(X)$ (the family of all possible compact subsets of X). Intuitively, if $d_H(X, Y)$, is small it means that every $x \in X$ is close to some $y \in Y$ and vice versa. The metric d_H is used in Brownian motion [15], matrix theory [1], dynamical systems [2], or fractal geometry [7], among other research areas. In the theory of evolutionary multiobjective optimization (EMO), the closeness of a set A to certain PF determines the approximation (called *convergence* in the EMO literature) of the outcome, and the closeness of PF to A determines the spread (maximal gap).

The metric d_H is rarely used by the EMO community because its values allow for undesirable ambiguities. An illustrative example is that a large value of $d_H(PF, A)$, can indicate both: a "bad" approximation or a "good" one with at least one outlier (see Figure 1). Instead of d_H, several alternative indicators have been introduced, e.g. the hypervolume indicator [17], the R indicators [10], or the averaged Hausdorff distance Δ_p introduced in [13]. The adequacy of the use of different indicators has been studied in [14]. Among them, the performance indicator Δ_p has the advantage of not punishing heavily the outliers and to produce solutions evenly spread around PF (which is highly desirable [16]), but the disadvantage of not satisfying the triangle inequality. In other words, Δ_p is not a metric but a semimetric with relaxed triangle inequality, which we will refer to as an inframetric. This terminology, explained with more detail at the end of the Section 2, does not conflict with related (but different) notions and it has also been used in computer science (see [8]).

In this paper we introduce a new indicator $\Delta_{p,q}$, that we call the (p, q)-averaged Hausdorff distance, or simply (p, q)-averaged distance for brevity.

When $p, q \geqslant 1$ this distance turns out to be a proper metric and preserves the principal advantages of Δ_p. When $|p|, |q| \geqslant 1$ (but not $p, q \geqslant 1$), we show that $\Delta_{p,q}$ satisfies a relaxed triangle inequality, turning it into an inframetric. Moreover $\Delta_{p,q}$ is related to the p-averaged Hausdorff distance Δ_p and the standard Hausdorff distance d_H in the following way:

$$\Delta_{p,-\infty}(\cdot\,,\cdot) := \lim_{q\to-\infty} \Delta_{p,q}(\cdot\,,\cdot) = \Delta_p(\cdot\,,\cdot), \quad (1.1)$$

and

$$: \Delta_{\infty,-\infty}(\cdot\,,\cdot) := \lim_{\substack{p\to\infty \\ q\to-\infty}} \Delta_{p,q}(\cdot\,,\cdot) = d_H(\cdot\,,\cdot). \quad (1.2)$$

The remainder of this paper is organized as follows: Section 2, presents some basic preliminaries, including well-known properties of q-power means, the Generational Distance GD, the Inverted Generational Distance IGD, and their p-averaged versions GD_p and IGD_p, respectively. This section concludes with a review of the inframetric properties of the p-averaged Hausdorff distance Δ_p. Section 3, presents $\mathrm{GD}_{p,q}$ and $\mathrm{IGD}_{p,q}$ which are modifications of GD_p and IGD_p, respectively. We introduce here the (p,q)-averaged distance $\Delta_{p,q}$ and prove several properties, including a result related to Pareto-compliance. Section 4, presents some numerical results showing the behavior of $\Delta_{p,q}$ as an indicator. Finally, in Sections 5 and 6, we present our conclusions and future work proposals, respectively.

2 Preliminaries

2.1 Multiobjective Optimization

For a vector valued function $f\colon X \subset \mathbb{R}^n \to \mathbb{R}^\ell$, the multiobjective optimization problem (MOP), under consideration requires the simultaneous minimization of its ℓ component functions $f_1, \dots, f_\ell$. A solution is *optimal* when the elements of the image $Y = f(X)$, are *nondominated* in the sense of Pareto [11], which derives from a partial order in $\mathbb{R}^\ell$, whose definition we recall below for the convenience of the reader.

Let $(Y, \preceq)$ be a subset $Y \subset \mathbb{R}^\ell$ equipped with the partial order $\preceq$ defined for $y, y' \in Y$ by:

$$y \preceq y' \;\text{ if and only if }\; y_i \leqslant y'_i \;\text{ for all }\; i = 1, \dots, \ell.$$

An element $y \in Y$ is said to be *dominated by* $y' \in Y$ and denoted $y' \prec y$, if $y' \preceq y$ and $y \neq y'$. Moreover, $y \in Y$ is *dominated by* $Y' \subset Y$, written $Y' \preceq y$, if there exists some $y' \in Y'$ such that $y' \preceq y$, otherwise it is said to be *nondominated* by it, $Y' \npreceq y$. A subset $Y' \subset Y$ is *dominated* by a subset $Y'' \subset Y$, and written $Y'' \preceq Y'$, if for every $y' \in Y'$ there exist some $y'' \in Y''$ such that $y'' \preceq y'$. If this is not the case Y' is said to be *nondominated* by Y'' and denoted $Y' \npreceq Y''$.

If $Y = f(X)$ is the objective space of some MOP with decision space $X \subset \mathbb{R}^n$ and objective function $f\colon X \to \mathbb{R}^\ell$, its *Pareto front* is defined as the set $Y^* := \{y \in Y \mid \nexists y' \in Y \colon y' \prec y\}$ of nondominated elements. An element $x \in X$ is called *Pareto-optimal* if its image is nondominated, i.e., $f(x) \in Y^*$, and the set X^* of all Pareto-optimal points is called *Pareto set*.

Finally, if $Y \subset \mathbb{R}^\ell$, we say that a performance indicator given by a function $\mathcal{I}\colon \mathcal{P}(Y) \to \mathbb{R}$ is *Pareto-compliant* if for subsets $A, B \subset Y$ the strict dominance condition $A \preceq B$ and $B \npreceq A$ implies the relation $\mathcal{I}(A) \leqslant \mathcal{I}(B)$ (or in a stronger version $\mathcal{I}(A) < \mathcal{I}(B)$). We refer the reader to [18] for details. In section 3.2 we provide a brief discussion of the compliance of the indicator associated to $\Delta_{p,q}$ with Pareto-related optimality criteria.

2.2 Power Means

For a finite set $X = \{x_i\}_{i=1}^N \subset [0, \infty)$ and a non-zero real q, the q-average or the q-power mean of X is given by:

$$\mathcal{M}^q(X) := \left(\frac{1}{N} \sum_{i=1}^N x_i^q \right)^{\frac{1}{q}}.$$

Remark. When the set of values taken by an indexing quantity has been explicitly specified, e.g. $i \in I := \{1, \dots, N\}$, for convenience, the following abbreviated notation will be used:

$$\mathop{\mathcal{M}^q}_{i\in I}(x_i) := \mathcal{M}^q(\{x_i\}_{i\in I}) = \mathcal{M}^q(X).$$

ISSN 2007-9737

A comprehensive reference on the theory and properties of means is e.g. [5], where proofs of the statements presented in this section can be found.

It is well-known that limit cases of power means recover familiar quantities, for example:

$$\lim_{q\to 0} \mathcal{M}^q(X) = \left(\prod_{i=1}^{N} x_i\right)^{\frac{1}{N}},$$

is the standard geometric mean of X, additionally:

$$\lim_{q\to\infty} \mathcal{M}^q(X) = \max\{x_1, \dots, x_N\}, \quad \text{and}$$
$$\lim_{q\to -\infty} \mathcal{M}^q(X) = \min\{x_1, \dots, x_N\}.$$

The special case $q = -1$ corresponds to the *harmonic mean* and it is also of our interest:

$$\operatorname{harm}(X) \coloneqq \mathcal{M}^{-1}(X).$$

Now, we can define the q-average of a finite set for any q in the extended real line $\overline{\mathbb{R}} \coloneqq [-\infty, \infty]$. Let $Y \coloneqq \{y_i\}_{i\in I}$ be a finite subset of $[0, \infty)$. The following properties hold for power means:

1. If $x_i \leqslant y_i$ for all indices $i \in I$, then for any $q \in \overline{\mathbb{R}}$:

$$\mathcal{M}^q(X) \leqslant \mathcal{M}^q(Y). \tag{2.1}$$

2. For $p, q \in \overline{\mathbb{R}}$, if $p \leqslant q$, then:

$$\mathcal{M}^p(X) \leqslant \mathcal{M}^q(X). \tag{2.2}$$

3. If $A = (a_{ij})$ denotes an array of indexed positive elements with $i \in I$ and $j \in J$, then their p-average satisfies:

$$\mathcal{M}^p(A) = \mathop{\mathcal{M}^p}_{i\in I}\big(\mathop{\mathcal{M}^p}_{j\in J}(a_{ij})\big) = \mathop{\mathcal{M}^p}_{j\in J}\big(\mathop{\mathcal{M}^p}_{i\in I}(a_{ij})\big). \tag{2.3}$$

4. For $p \geqslant 1$ it follows from Minkowski inequality that:

$$\mathop{\mathcal{M}^p}_{i\in I}(x_i + y_i) \leqslant \mathop{\mathcal{M}^p}_{i\in I}(x_i) + \mathop{\mathcal{M}^p}_{i\in I}(y_i). \tag{2.4}$$

5. The harmonic mean admits the bound:

$$\operatorname{harm}(X) \leqslant N \min(X). \tag{2.5}$$

2.3 Averaged Hausdorff Distance

Suppose that A and B belong to the family of finite subsets of $\mathbb{R}^n$, denoted by $\mathcal{P}_0(\mathbb{R}^n)$, and that $p \geqslant 1$. Recall that the "modified" generational distance (GD_p) and the inverted generational distance (IGD_p) are defined by power means as follows (see [13]):

$$\begin{aligned}\mathrm{GD}_p(A, B) &\coloneqq \left(\frac{1}{N_A}\sum_{i=1}^{N_A} d(a_i, B)^p\right)^{\frac{1}{p}},\\ &= \left(\frac{1}{N_A}\sum_{i=1}^{N_A} \min_{j=1..N_B}\left\{d(a_i, b_j)^p\right\}\right)^{\frac{1}{p}},\end{aligned}$$

and:

$$\mathrm{IGD}_p(A, B) \coloneqq \mathrm{GD}_p(B, A),$$

where $d(\cdot,\cdot)$ stands for the standard Euclidean metric. Let us denote by $\Delta_p \colon \mathcal{P}_0(\mathbb{R}^n) \times \mathcal{P}_0(\mathbb{R}^n) \to \mathbb{R}$ the so-called averaged Hausdorff distance, i.e.:

$$\Delta_p(A, B) \coloneqq \max\{\mathrm{GD}_p(A, B), \mathrm{IGD}_p(A, B)\}.$$

It has been established in [13] that Δ_p does not satisfy the triangle inequality but a weaker version given by:

$$\Delta_p(A, C) \leqslant N^{\frac{1}{p}}\big(\Delta_p(A, C) + \Delta_p(B, C)\big),$$

where $N \geqslant 1$ is a constant with $|A|, |B|, |C| \leqslant N$. Because of this, Δ_p is not a proper metric but a semimetric with relaxed triangle inequality. This notion has appeared with several conflicting names in the literature, among which *inframetric* is probably the one with less friction with pre-existing terminology.

There are two related conditions that relax the triangle inequality for a function $d \colon X \times X \to [0, \infty)$. Namely, the existence of a constant $C > 0$ such that for any points $a, b, c \in X$ one of the following properties hold:

1. The C-relaxed triangle inequality:

$$d(a, b) \leqslant C\big(d(a, c) + d(c, b)\big).$$

2. The C-inframetric inequality:

$$d(a, b) \leqslant C \max\big\{d(a, c), d(c, b)\big\}.$$

Computación y Sistemas, Vol. 22, No. 2, 2018, pp. 331–345
doi: 10.13053/CyS-22-2-2950

Since condition (2) implies (1), with the same constant C and, reciprocally, the C-relaxed triangle inequality implies the $2C$-inframetric one, it is clear that both conditions are equivalent for appropriate choice of constants. Hence, a semimetric satisfying anyone of these conditions will be called an inframetric.

3 The (p, q)-Averaged Distance

In order to simplify the forthcoming calculations we use the following abbreviation:

$$\sum_{i=1}^{N} x_i := \frac{1}{N}\sum_{i=1}^{N} x_i = \mathcal{M}^1(\{x_1, \dots, x_N\}),$$

to denote the average of $x_1, \dots, x_N \in [0, \infty]$.

Definition 3.1. For $p, q \in \mathbb{R}\backslash\{0\}$, the generational (p, q)-distance $\mathrm{GD}_{p,q}(A, B)$ between two finite sets $A = \{a_i\}_{i=1}^{N_A}$ and $B = \{b_j\}_{j=1}^{N_B}$ in $\mathbb{R}^n$ is given by:

$$\mathrm{GD}_{p,q}(A, B) := \left(\sum_{i=1}^{N_A}\left(\sum_{j=1}^{N_B} d(a_i, b_j)^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}.$$

When $p < 0$ or $q < 0$ it will always be assumed for consistency that $A \cap B = \varnothing$. The indicator $\mathrm{GD}_{p,q}(A, B)$ can be extended for values of $p = 0$ and/or $q = 0$ by taking the limits when $p \to 0$ and/or $q \to 0$, respectively. In such cases, the properties mentioned in the previous section suggest the following appropriate definitions:

$$\mathrm{GD}_{p,0}(A, B) := \left(\sum_{i=1}^{N_A}\left(\prod_{j=1}^{N_B} d(a_i, b_j)\right)^{\frac{p}{N_B}}\right)^{\frac{1}{p}},$$

for $p \neq 0$,

$$\mathrm{GD}_{0,q}(A, B) := \left(\prod_{i=1}^{N_A}\left(\sum_{j=1}^{N_B} d(a_i, b_j)^q\right)^{\frac{1}{q}}\right)^{\frac{1}{N_A}},$$

for $q \neq 0$, and

$$\mathrm{GD}_{0,0}(A, B) := \left(\prod_{i=1}^{N_A}\left(\prod_{j=1}^{N_B} d(a_i, b_j)\right)^{\frac{1}{N_B}}\right)^{\frac{1}{N_A}}.$$

We can also calculate $\mathrm{GD}_{p,q}$ when $p \to \pm\infty$ and/or $q \to \pm\infty$, by changing the respective sum for a minimum or a maximum. In particular, if $A \cap B = \varnothing$, we have the nice relation:

$$\lim_{q\to-\infty} \mathrm{GD}_{p,q}(A, B) = \mathrm{GD}_p(A, B). \tag{3.1}$$

Note that the definition of $\mathrm{GD}_{p,q}$ has two drawbacks, namely $\mathrm{GD}_{p,q}(A, B)$ does not necessarily vanish if $A = B$, and in general $\mathrm{GD}_{p,q}(A, B) \neq \mathrm{GD}_{p,q}(B, A)$, thus this indicator does not define a metric. In order to get a proper metric we introduce the following notion.

Definition 3.2. The (p, q)-*averaged distance* is the map $\Delta_{p,q} \colon \mathcal{P}_0(\mathbb{R}^n) \times \mathcal{P}_0(\mathbb{R}^n) \to \mathbb{R}$ given by:

$$\Delta_{p,q}(A, B) := \max\{\mathrm{GD}_{p,q}(A, B\backslash A), \mathrm{GD}_{p,q}(B, A\backslash B)\}.$$

If $A \cap B = \varnothing$ then $\mathrm{GD}_p(A, B) = \mathrm{GD}_p(A, B\backslash A)$, therefore using (3.1) and Definition 3.2 we easily obtain:

$$\lim_{q\to-\infty} \Delta_{p,q}(A, B) = \Delta_p(A, B). \tag{3.2}$$

More generally, $\Delta_{p,-\infty}(A, B) \geqslant \Delta_p(A, B)$ always holds. We point out that similarly to the relation:

$$\mathrm{GD}_p(A, B) = N_A^{-\frac{1}{p}} \|D_{AB}\|_p,$$

between the generational distance $\mathrm{GD}_p(A, B)$ and the matrix ℓ_p-norm of the distance matrix $(D_{AB})_{ij} := d(a_i, b_j)$, we also have the following relation between the (p, q)-generational distance $\mathrm{GD}_{p,q}(A, B)$ and the matrix $\ell_{p,q}$-norm $\|D_{AB}\|_{p,q}$ (defined precisely as $\mathrm{GD}_{p,q}$ but with standard sums $\sum$ instead of the normalized ones $\sum$, see e.g. [9]):

$$\begin{aligned}\mathrm{GD}_{p,q}(A, B) &= \mathcal{M}^p_{i\in I}\big(\mathcal{M}^q_{j\in J}(d(a_i, b_j))\big),\\ &= N_A^{-\frac{1}{p}} N_B^{-\frac{1}{q}} \|D_{AB}\|_{p,q},\end{aligned} \tag{3.3}$$

where $I := \{1, \dots, N_A\}$ and $J := \{1, \dots, N_B\}$.

3.1 Metric Properties

Now we study the behavior of the distances $\mathrm{GD}_{p,q}$ and $\Delta_{p,q}$ from a metric perspective. Since the elements of the distance matrix do not satisfy a relation of the kind $D_{AC} = D_{AB} + D_{BC}$ (and

ISSN 2007-9737

in general these matrices are of different sizes), their properties are not immediate from (3.3) and the triangle inequality for $\ell_{p,q}$-norms. Here, we provide self-contained proofs for future reference and completeness.

Theorem 3.3. *For parameters $p, q \in [1, \infty)$ the generational (p, q)-distance $\mathrm{GD}_{p,q}$ satisfies the triangle inequality, namely:*

$$\mathrm{GD}_{p,q}(A, C) \leqslant \mathrm{GD}_{p,q}(A, B) + \mathrm{GD}_{p,q}(B, C),$$

for any finite sets $A, B, C \subset \mathbb{R}^n$.

Proof. Let assume that $A = \{a_i\}_{i=1}^{N_A}$, $B = \{b_j\}_{j=1}^{N_B}$, and $C = \{c_k\}_{k=1}^{N_C}$. From the triangle inequality for the Euclidean metric $d(\cdot\,,\cdot)$ on $\mathbb{R}^n$ we know that for arbitrary values of the indices $i \in I \coloneqq \{1, \ldots, N_A\}$, $j \in J \coloneqq \{1, \ldots, N_B\}$, and $k \in K \coloneqq \{1, \ldots, N_C\}$ it holds that:

$$d(a_i, c_k) \leqslant d(a_i, b_j) + d(b_j, c_k).$$

Let us abbreviate these quantities by $\delta_{ik} \coloneqq d(a_i, c_k)$, $\delta_{ij} \coloneqq d(a_i, b_j)$, and $\delta_{jk} \coloneqq d(b_j, c_k)$, where the indices i, j, k will be understood to take values in the sets I, J, and K defined above, respectively. Then, for any of them:

$$\delta_{ik} \leqslant \delta_{ij} + \delta_{jk}.$$

By (2.1), we can take the q-average over all $k \in K$ at both sides of the previous inequality to obtain:

$$\begin{aligned}\underset{k\in K}{\mathcal{M}^q}(\delta_{ik}) &\leqslant \underset{k\in K}{\mathcal{M}^q}(\delta_{ij} + \delta_{jk}),\\ &\leqslant \underset{k\in K}{\mathcal{M}^q}(\delta_{ij}) + \underset{k\in K}{\mathcal{M}^q}(\delta_{jk}),\end{aligned}$$

where the last line follows by using Minkowski inequality (2.4) for $q \geqslant 1$. Since the averaged quantities in the first term are independent of k, we have for all $i \in I$ and $j \in J$ that:

$$\underset{k\in K}{\mathcal{M}^q}(\delta_{ik}) \leqslant \delta_{ij} + \underset{k\in K}{\mathcal{M}^q}(\delta_{jk}). \tag{3.4}$$

Here, we consider two cases for the parameters $p, q \in [1, \infty)$ independently.

Case $p \leqslant q$: Under this assumption, we take at both sides of (3.4) the p-average over all $i \in I$, and using (2.4) for $p \geqslant 1$ at the RHS (right-hand side), we get:

$$\underset{i\in I}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{ik})\big) \leqslant \underset{i\in I}{\mathcal{M}^p}(\delta_{ij}) + \underset{i\in I}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big).$$

Due to the independence on $i \in I$ of the p-averaged quantity in the last term at the RHS, this simplifies to:

$$\underset{i\in I}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{ik})\big) \leqslant \underset{i\in I}{\mathcal{M}^p}(\delta_{ij}) + \underset{k\in K}{\mathcal{M}^q}(\delta_{jk}),$$

the LHS (left-hand side) is precisely $\mathrm{GD}_{p,q}(A, C)$. Now, we take at both sides the p-average over all $j \in J$, use that the LHS is independent of j, and employ (2.4) for $p \geqslant 1$ at the RHS, to find:

$$\begin{aligned}\mathrm{GD}_{p,q}(A, C) &\leqslant \underset{j\in J}{\mathcal{M}^p}\big(\underset{i\in I}{\mathcal{M}^p}(\delta_{ij})\big) + \underset{j\in J}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\\ &\leqslant \underset{i\in I}{\mathcal{M}^p}\big(\underset{j\in J}{\mathcal{M}^p}(\delta_{ij})\big) + \underset{j\in J}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\end{aligned}$$

where the interchange at the last line is allowed by property (2.3). Now, property (2.2) for $p \leqslant q$ ensures that in the first term at the RHS we can replace the inner $\mathcal{M}^p$ by $\mathcal{M}^q$ to get an equal or larger quantity, therefore:

$$\begin{aligned}\mathrm{GD}_{p,q}(A, C) &\leqslant \underset{i\in I}{\mathcal{M}^p}\big(\underset{j\in J}{\mathcal{M}^q}(\delta_{ij})\big) + \underset{j\in J}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\\ &= \mathrm{GD}_{p,q}(A, B) + \mathrm{GD}_{p,q}(B, C),\end{aligned}$$

which proves the claim.

Case $q \leqslant p$: Let us take at both sides of (3.4) the q-average over all $j \in J$. Using that the LHS is independent of j, and (2.4) for $q \geqslant 1$ at the RHS, we obtain:

$$\begin{aligned}\underset{k\in K}{\mathcal{M}^q}(\delta_{ik}) &\leqslant \underset{j\in J}{\mathcal{M}^q}(\delta_{ij}) + \underset{j\in J}{\mathcal{M}^q}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\\ &\leqslant \underset{j\in J}{\mathcal{M}^q}(\delta_{ij}) + \underset{j\in J}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\end{aligned}$$

where the change of $\mathcal{M}^q$ by $\mathcal{M}^p$ in the last-term of the RHS is justified by property (2.2) for $q \leqslant p$.

Finally, we take at both sides the p-average over all $i \in I$, employ (2.4) for $p \geqslant 1$ at the RHS and use that the last term is independent of $i \in I$, to write:

$$\begin{aligned}\mathrm{GD}_{p,q}(A, C) &\leqslant \underset{i\in I}{\mathcal{M}^p}\big(\underset{j\in J}{\mathcal{M}^q}(\delta_{ij})\big) + \underset{j\in J}{\mathcal{M}^p}\big(\underset{k\in K}{\mathcal{M}^q}(\delta_{jk})\big),\\ &= \mathrm{GD}_{p,q}(A, B) + \mathrm{GD}_{p,q}(B, C).\end{aligned}$$

$\square$

The corollary below states that the indicator $\Delta_{p,q}$ is a semimetric that becomes a metric if $p, q \geqslant 1$. When this is not the case but still $|p|, |q| \geqslant 1$, the theorem that follows assures that $\Delta_{p,q}$ is at least inframetric and provides the associated constants.

Corollary 3.4. *For $p, q \in \overline{\mathbb{R}}$ the (p,q)-averaged distance $\Delta_{p,q}$ is a semimetric on disjoint members of the family $\mathcal{P}_0(\mathbb{R}^n)$ of finite subsets of $\mathbb{R}^n$. Moreover, for $p, q \in [1, \infty]$, the (p,q)-averaged distance is a proper metric on disjoint subsets of $\mathcal{P}_0(\mathbb{R}^n)$.*

Proof. From Definition 3.2, it is easy to see that $\Delta_{p,q}(A,B) \geqslant 0$ as well as:

$$\Delta_{p,q}(A,B) = \Delta_{p,q}(B,A),$$

for every $A, B \in \mathcal{P}_0(\mathbb{R}^n)$ and all $p, q \in \overline{\mathbb{R}}$. Moreover, it is also clear from Definition 3.1 that $\mathrm{GD}_{p,q}(A, B\backslash A) = 0$ if and only if $A = \varnothing$ or $B \subseteq A$, hence from Definition 3.2 we find, for $A, B \neq \varnothing$, that:

$$\Delta_{p,q}(A,B) = 0 \quad \text{if and only if} \quad A = B.$$

These properties show that $\Delta_{p,q}$ is a semimetric for any $p, q \in \overline{\mathbb{R}}$.

Since the maximum of two functions satisfying the triangle inequality also satisfies it, Theorem 3.3 shows that $\Delta_{p,q}$ satisfies the triangle inequality for all $p, q \in [1, \infty)$ and the cases p or q equal to ∞ follow by taking the appropriate limits. □

Theorem 3.5. *For any $p, q \in \mathbb{R}$ with $|p|, |q| > 1$ the generational (p,q)-distance $\mathrm{GD}_{p,q}$ satisfies a relaxed triangle inequality. Explicitly:*

$$\mathrm{GD}_{p,q}(A,C) \leqslant N^{\frac{1}{r}} \big(\mathrm{GD}_{p,q}(A,B) + \mathrm{GD}_{p,q}(B,C)\big),$$

for all $A, B, C \in \mathcal{P}_0(\mathbb{R}^n)$, any constant $N \geqslant 1$ such that $|A|, |B|, |C| \leqslant N$ and where r is given by:

$$\frac{1}{r} := \frac{1}{|p|} + \frac{1}{|q|}.$$

Proof. For arbitrary $p \neq 0$, let us assume that $q < 0$, so that $|q| = -q$. We can write:

$$\begin{aligned}
\mathrm{GD}_{p,|q|}(A,B) &= \left(\sum_{i=1}^{N_A}\left(\left[\sum_{j=1}^{N_B}\big[d(a_i,b_j)^q\big]^{-1}\right]^{-1}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}},\\
&= \left(\sum_{i=1}^{N_A}\Big(N_B^{-2}\operatorname*{harm}_{j=1..N_B}\big\{d(a_i,b_j)^q\big\}\Big)^{\frac{p}{q}}\right)^{\frac{1}{p}},
\end{aligned}$$

which combined with the property (2.5) gives us:

$$\begin{aligned}
\mathrm{GD}_{p,|q|}(A,B) &\leqslant \left(\sum_{i=1}^{N_A}\Big(N_B^{-1}\min_{j=1..N_B}\big\{d(a_i,b_j)^q\big\}\Big)^{\frac{p}{q}}\right)^{\frac{1}{p}},\\
&= N_B^{\frac{1}{-q}}\left(\sum_{i=1}^{N_A}\Big(\min_{j=1..N_B}\big\{d(a_i,b_j)^q\big\}\Big)^{\frac{p}{q}}\right)^{\frac{1}{p}},\\
&\leqslant N_B^{\frac{1}{|q|}}\left(\sum_{i=1}^{N_A}\left(\sum_{j=1}^{N_B} d(a_i,b_j)^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}},\\
&= N_B^{\frac{1}{|q|}}\,\mathrm{GD}_{p,q}(A,B).
\end{aligned}$$

An analogous inequality holds for arbitrary $q \neq 0$ if we assume that $p < 0$. Therefore, we can write in general:

$$\mathrm{GD}_{|p|,|q|}(A,B) \leqslant N^{\frac{1}{r}}\,\mathrm{GD}_{p,q}(A,B),$$

where $N_A, N_B \leqslant N$ and:

$$r := \begin{cases} |\min\{p,q\}| & \text{if } pq < 0,\\ \frac{1}{2}\operatorname{harm}\{|p|,|q|\} & \text{if } p < 0,\ q < 0.\end{cases}$$

Notice that in both cases r can be chosen to take the second value, i.e., $\frac{1}{r} := \frac{1}{|p|} + \frac{1}{|q|}$, when the sharpness of the constants does not matter. Finally, when $|p|, |q| \geqslant 1$ we employ the triangle inequality for $\mathrm{GD}_{|p|,|q|}$ to conclude that:

$$\begin{aligned}
\mathrm{GD}_{p,q}(A,C) &\leqslant \mathrm{GD}_{|p|,|q|}(A,C),\\
&\leqslant \mathrm{GD}_{|p|,|q|}(A,B) + \mathrm{GD}_{|p|,|q|}(B,C),\\
&\leqslant N^{\frac{1}{r}}\big(\mathrm{GD}_{p,q}(A,B) + \mathrm{GD}_{p,q}(B,C)\big),
\end{aligned}$$

which finishes the proof. □

From Corollary 3.4 we know that $\Delta_{p,q}$ is a metric for $p, q \geqslant 1$. More generally, the following corollary states that $\Delta_{p,q}$ is an inframetric when $|p|, |q| \geqslant 1$.

Corollary 3.6. *For any $p, q \in \mathbb{R}$ with $|p|, |q| \geqslant 1$ the (p,q)-averaged distance $\Delta_{p,q}$ satisfies the following relaxed triangle inequality on disjoint subsets $A, B, C \in \mathcal{P}_0(\mathbb{R}^n)$:*

$$\Delta_{p,q}(A,C) \leqslant N^{\frac{1}{r}}\big(\Delta_{p,q}(A,B) + \Delta_{p,q}(B,C)\big),$$

for any constant $N \geqslant 1$ such that $|A|, |B|, |C| \leqslant N$ and where $\frac{1}{r} := \frac{1}{|p|} + \frac{1}{|q|}$.

Proof. The corollary follows immediately from Theorem 3.5 and Definition 3.2. □

Theorem 3.7. *Let* $A, B \in \mathcal{P}_0(\mathbb{R}^n)$ *and suppose that* $p \leqslant p'$, $q \leqslant q'$, *then:*

$$\Delta_{p,q}(A,B) \leqslant \Delta_{p',q}(A,B), \quad \textit{and}$$
$$\Delta_{p,q}(A,B) \leqslant \Delta_{p,q'}(A,B).$$

Proof. If follows easily from Definition 3.2 and two applications of property (2.2). □

Corollary 3.8. *Let* $A, B \in \mathcal{P}_0(\mathbb{R}^n)$ *and suppose that* $p \leqslant p'$, *then:*

$$\Delta_p(A,B) \leqslant \Delta_{p'}(A,B).$$

Proof. We obtain the corollary by using (3.2) and taking the limit $q \to -\infty$ in Theorem 3.7. For the convenience of the reader we present here an alternative, self-contained proof useful in the continuous case. Let X be a measure space with finite μ-measure. For any function $f\colon X \to \mathbb{C}$ in the Lebesgue space $L^r(X)$ with $r \geqslant 1$, a simple modification of the Hölder inequality tells us that:

$$\left(\int_X |f|\, d\mu\right)^r \leqslant \mu(X)^{r-1} \int_X |f|^r d\mu. \tag{3.5}$$

For any $p, p' \in \mathbb{R}$ with $1 \leqslant p \leqslant p'$ we have $\dfrac{p'}{p} \geqslant 1$, so with $r = \dfrac{p'}{p}$ and $f = g^p$ in (3.5) we obtain:

$$\left(\int_X |g^p|\, d\mu\right)^{\frac{1}{p}} \leqslant \mu(X)^{\frac{1}{p}-\frac{1}{p'}} \left(\int_X |g^{p'}|\, d\mu\right)^{\frac{1}{p'}}. \tag{3.6}$$

Taking as X the set $A = \{a_i\}_{i=1}^N$, as $g\colon \mathbb{R}^n \to \mathbb{R}$ the function given by $g(x) := d(x, B)$, and as μ the discrete measure μ_d on A:

$$\mu_d(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A; \end{cases}$$

the inequality (3.6) becomes:

$$\left(\sum_{i=1}^N d(a_i, B)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^N d(a_i, B)^p\right)^{\frac{1}{p'}}.$$

That is $\mathrm{GD}_p(A,B) \leqslant \mathrm{GD}_{p'}(A,B)$, which easily implies $\Delta_p(A,B) \leqslant \Delta_{p'}(A,B)$. □

3.2 Pareto-Compliance

A discussion of the Pareto-compliance for the averaged GD_p, IGD_p and Δ_p-indicators appeared in [13, Sec. III]. Similar observations can be made for the corresponding (p,q)-indicators introduced in this work. Here we will concentrate only in providing a complete proof of a result (analogous to Proposition 3 there) that describes the behavior of the $\mathrm{GD}_{p,q}$-indicator. The assumptions required are stronger than the compliance notion defined in Section 2.1 but they are useful to understand what will be needed in a very general situation.

If a MOP problem has an associated objective function $f\colon X \subset \mathbb{R}^n \to \mathbb{R}^\ell$ whose objective space $Y = f(X)$ has a Pareto front Y^*, and $A \subset Y$ denotes an approximating subset, the explicit $\mathrm{GD}_{p,q}$ and $\Delta_{p,q}$-performance indicators assigned to A are given, respectively, by:

$$\mathcal{I}^{\mathrm{GD}}_{p,q}(A) := \mathrm{GD}_{p,q}(A, Y^*), \ \text{ and}$$
$$\mathcal{I}^{\Delta}_{p,q}(A) := \Delta_{p,q}(A, Y^*).$$

For convenience, given $x, y \in \mathbb{R}^\ell$ and $q \in \overline{\mathbb{R}}$ we will abbreviate $\delta_q(x) := \mathcal{M}^q{}_{y \in Y^*}\, d(x,y)$.

Theorem 3.9. *Let* $p, q \in \mathbb{R}$ *and suppose that a pair of distinct finite subsets* $A, B \subset Y$ *satisfy:*

1. $A \preceq B$, *i.e.,* $\forall b \in B$, $\exists a \in A$ *such that* $a \preceq b$.
2. $\forall a \in A$, $\exists b \in B$ *such that* $a \preceq b$ *and with* $\delta_q(b) \leqslant \mathcal{M}^p\{\delta_q(b') \mid a \npreceq b',\ b' \in B\}$.
3. $\exists a \in A \backslash B$, $\exists b \in B \backslash A$ *such that* $a \prec b$.
4. $\forall a \in A$, $\forall b \in B$ *the following property holds,*

$$a \prec b \implies \delta_q(a) < \delta_q(b).$$

then $\mathcal{I}^{\mathrm{GD}}_{p,q}(A) < \mathcal{I}^{\mathrm{GD}}_{p,q}(B)$.

Proof. Let us assume that $A = \{a_i\}_{i=1}^{N_A}$, where its elements are arranged in a nonincreasing order with respect to $\delta_q(a_i)$, this means that:

$$i < j \implies \delta_q(a_j) \leqslant \delta_q(a_i).$$

By conditions 1 and 2, we can decompose B into a partition $B = B_1 \cup \cdots \cup B_m$ of subsets defined recursively for $i = 1, \ldots, m$ (with $1 \leqslant m \leqslant N_A$), as:

$$B_i := \{b \in B \backslash (B_0 \cup \cdots \cup B_{i-1}) \mid a_i \preceq b\},$$

ISSN 2007-9737

with $B_0 \coloneqq \varnothing$. Let N_i denote the size of B_i satisfying $1 \leqslant N_i \leqslant N_B = \sum_{i=1}^m N_i$ and arrange the elements of each subset $B_i = \{b_k^{(i)}\}_{k=1}^{N_i}$ in a nondecreasing order with respect to $\delta_q(b_k^{(i)})$, i.e., for each fixed i:

$$k < k' \implies \delta_q(b_k^{(i)}) \leqslant \delta_q(b_{k'}^{(i)}).$$

Note that from the construction and condition 4, for $i = 1, \ldots, m$, we have that:

$$b \in B_i \implies \delta_q(a_i) \leqslant \delta_q(b).$$

In particular, for the first element $b_1^{(i)}$ of each B_i, which minimizes the set $\{\delta_q(b) \mid b \in B_i\}$ we have $\delta_q(a_i) \leqslant \delta_q(b_1^{(i)})$. Moreover, condition 2 necessarily implies:

$$\delta_q(b_1^{(i)})^p \leqslant \sum_{b \in B \setminus B_i} \delta_q(b)^p. \tag{3.7}$$

Due to the ordering of A and this observation,

$$\begin{aligned}\mathcal{I}_{p,q}^{\mathrm{GD}}(A)^p = \sum_{i=1}^{N_A} \delta_q(a_i)^p \leqslant \sum_{i=1}^{m} \delta_q(a_i)^p, \\ \leqslant \sum_{i=1}^{m} \delta_q(b_1^{(i)})^p. \end{aligned} \tag{3.8}$$

But the inequality still holds if, for any given value of $j = 1, \ldots, m$, we replace at the RHS of (3.8) the element $b_1^{(j)}$ in the j-th term by any other $b_k^{(j)} \in B_j$ (with $k = 2, \ldots, N_j$) while keeping all the remaining terms fixed. Therefore, there are N_j possible choices for this element, and in consequence N_j different inequalities for any given $j = 1, \ldots, m$:

$$\mathcal{I}_{p,q}^{\mathrm{GD}}(A)^p \leqslant \frac{1}{m}\left(\sum_{\substack{i=1\\i\neq j}}^{m} \delta_q(b_1^{(i)})^p + \delta_q(b_{k_j}^{(j)})^p\right),$$

where $k_j = 1, \ldots, N_j$. When varying $j = 1, \ldots, m$, this procedure yields a total of $\sum_{j=1}^m N_j = N_B$ inequalities with the same LHS, and using (2.1) we can take the average of all of them. Since the LHS remains the same, we obtain:

$$\begin{aligned}\mathcal{I}_{p,q}^{\mathrm{GD}}(A)^p &\leqslant \frac{1}{N_B}\sum_{j=1}^{m}\sum_{k_j=1}^{N_j}\frac{1}{m}\left(\sum_{\substack{i=1\\i\neq j}}^{m} \delta_q(b_1^{(i)})^p + \delta_q(b_{k_j}^{(j)})^p\right), \\ &= \sum_{j=1}^{m}\frac{1}{N_B}\left(\sum_{\substack{i=1\\i\neq j}}^{m} N_j\,\delta_q(b_1^{(i)})^p + \sum_{k_j=1}^{N_j}\delta_q(b_{k_j}^{(j)})^p\right).\end{aligned}$$

Notice that conditions 3 and 4 imply that the previous inequality has to be strict since the LHS contains all the elements of A and the RHS all the elements of B.

Rearranging and counting the terms, we get that:

$$\begin{aligned}&\mathcal{I}_{p,q}^{\mathrm{GD}}(A)^p, \\ &< \sum_{j=1}^{m}\frac{1}{N_B}\left(\sum_{\substack{i=1\\i\neq j}}^{m} N_j\,\delta_q(b_1^{(i)})^p + \sum_{k_j=1}^{N_j}\delta_q(b_{k_j}^{(j)})^p\right), \\ &= \frac{1}{m}\sum_{j=1}^{m}\sum_{\substack{i=1\\i\neq j}}^{m}\frac{N_j}{N_B}\,\delta_q(b_1^{(i)})^p + \frac{1}{N_B}\sum_{j=1}^{m}\sum_{b\in B_j}\delta_q(b)^p,\end{aligned}$$

which after a reordering in the first term becomes:

$$\begin{aligned}&= \frac{1}{m}\sum_{i=1}^{m}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{N_j}{N_B}\,\delta_q(b_1^{(i)})^p + \frac{1}{m}\sum_{b\in B}\delta_q(b)^p, \\ &= \frac{1}{m}\sum_{i=1}^{m}\left(\frac{N_B - N_i}{N_B}\right)\delta_q(b_1^{(i)})^p + \frac{1}{m}\sum_{b\in B}\delta_q(b)^p,\end{aligned}$$

but using (3.7) we obtain the inequality:

$$\begin{aligned}&\leqslant \frac{1}{m}\sum_{i=1}^{m}\left(\frac{N_B - N_i}{N_B}\right)\sum_{b\in B\setminus B_i}\delta_q(b)^p + \frac{1}{m}\sum_{b\in B}\delta_q(b)^p, \\ &= \frac{1}{mN_B}\sum_{i=1}^{m}\sum_{b\in B\setminus B_i}\delta_q(b)^p + \frac{1}{m}\sum_{b\in B}\delta_q(b)^p, \\ &= \left(\frac{m-1}{m}\right)\sum_{b\in B}\delta_q(b)^p + \frac{1}{m}\sum_{b\in B}\delta_q(b)^p, \\ &= \sum_{b\in B}\delta_q(b)^p.\end{aligned}$$

Thus, $\mathcal{I}_{p,q}^{\mathrm{GD}}(A) < \mathcal{I}_{p,q}^{\mathrm{GD}}(B)$ as expected. □

The behavior of $\mathcal{I}_{p,q}^{\Delta}$ is more delicate due to the definition of $\Delta_{p,q}$, but for disjoint subsets not intersecting Y^* similar arguments can be employed. We remark also that condition 3 is only needed to ensure an strict inequality in the conclusion. We would have obtained $\mathcal{I}_{p,q}^{\mathrm{GD}}(B)^p \leqslant \mathcal{I}_{p,q}^{\mathrm{GD}}(B)^p$ by removing condition 3 and weakening 4 to the condition:

4'. $\forall a \in A, \forall b \in B$: $a \preceq b \implies \delta_q(a) \leqslant \delta_q(b)$.

4 Numerical Experiments

4.1 Working with $\Delta_{p,q}$

In this section we consider a hypothetical Pareto front P given by the line segment from $(0,1)$ to $(1,0)$ in $\mathbb{R}^2$, i.e. the set of all points:

$$(t, 1-t) \in \mathbb{R}^2 \quad \text{for} \quad 0 \leqslant t \leqslant 1.$$

This is the same example considered in [13, p. 506], and enables us to make a comparison with values of Δ_p. In order to use the (p,q)-averaged distance, we discretize P', by taking 11 uniformly distributed points over P. We assume two archives: X_1 is obtained from P' by changing $(0,1)$ for $(0,10)$, including an outlier, and adding $\frac{1}{10}$ to the remaining ordinates. X_2 is obtained from P' by adding 5 to each ordinate. See Figure 1.

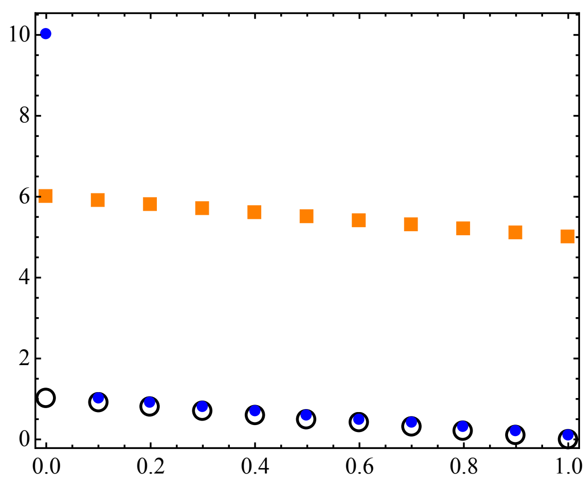

Fig. 1. A hypothetical Pareto front discretization P' (black circles) and two different archives: X_1 (blue dots) and X_2 (orange squares)

From [13], we know that:

$$\Delta_\infty(A,B) := \lim_{p\to\infty} \Delta_p(A,B)$$

coincides with the standard Hausdorff distance d_H.

In this case:

$$\Delta_1(P',X_1) = 0.9091, \qquad \Delta_1(P',X_2) = 4.5412,$$
$$d_H(P',X_1) = 9, \qquad d_H(P',X_2) = 5;$$

according to Corollary 3.8 and [13, p. 512], these values must increase as p increases.

Tables 1 and 2 show that we can find values of p and q such that the (p,q)-averaged distance does not punish heavily the outliers, for example $p=q=1$ or $p=1$ and $q=-1$. We remark that the values of $\Delta_{p,q}(P',X_1)$ do not present a significative change under variations of $q \leqslant 1$ for a fixed p.

Thus it is possible to work with $q=1$, in which case $\Delta_{p,q}$ is a metric according to Corollary 3.4, and still obtain values close to the ones given by the inframetric Δ_p, with the same $p \geqslant 1$.

For large values of p the behavior of $\Delta_{p,q}$ present the same disadvantages of Δ_p or of the standard Hausdorff distance. For example, in Table 1 it can be observed that all distances for $p \geqslant 5$ are useless because they imply that the distance from the discrete Pareto front P' to the archive X_1 is larger than its distance to the archive X_2.

Figure 1 suggests that this is an undesirable outcome.

Table 1. $\Delta_{p,q}(P',X_1)$ for different values of p and q

q \ p	1	2	5	10	20
$-\infty$	0.9091	2.7153	5.5714	7.0811	7.9831
-100	0.9272	2.7701	5.6839	7.2241	8.1443
-20	0.9537	2.8367	5.8202	7.3974	8.3396
-5	0.9895	2.8624	5.8705	7.4613	8.4117
-1	1.1131	2.8782	5.8848	7.4795	8.4322
1	1.3243	2.9112	5.8920	7.4886	8.4425
2	2.9277	2.9295	5.8956	7.4932	8.4476
5	5.8920	5.8956	5.9063	7.5068	8.4630
10	7.4886	7.4932	7.5068	7.5292	8.4882

Table 2. $\Delta_{p,q}(P',X_2)$ for different values of p and q

q \ p	1	2	5	10	20
$-\infty$	4.5412	4.5497	4.5751	4.6160	4.6867
-100	4.6442	4.6529	4.6790	4.7209	4.7933
-20	4.8425	4.8518	4.8795	4.9239	5.0003
-5	4.9624	4.9720	5.0007	5.0465	5.1250
-1	5.0008	5.0105	5.0394	5.0856	5.1646
1	5.0203	5.0301	5.0591	5.1055	5.1848
2	5.0301	5.0398	5.0690	5.1154	5.1949
5	5.0591	5.0690	5.0983	5.1450	5.2248
10	5.1055	5.1154	5.1450	5.1921	5.2725

ISSN 2007-9737

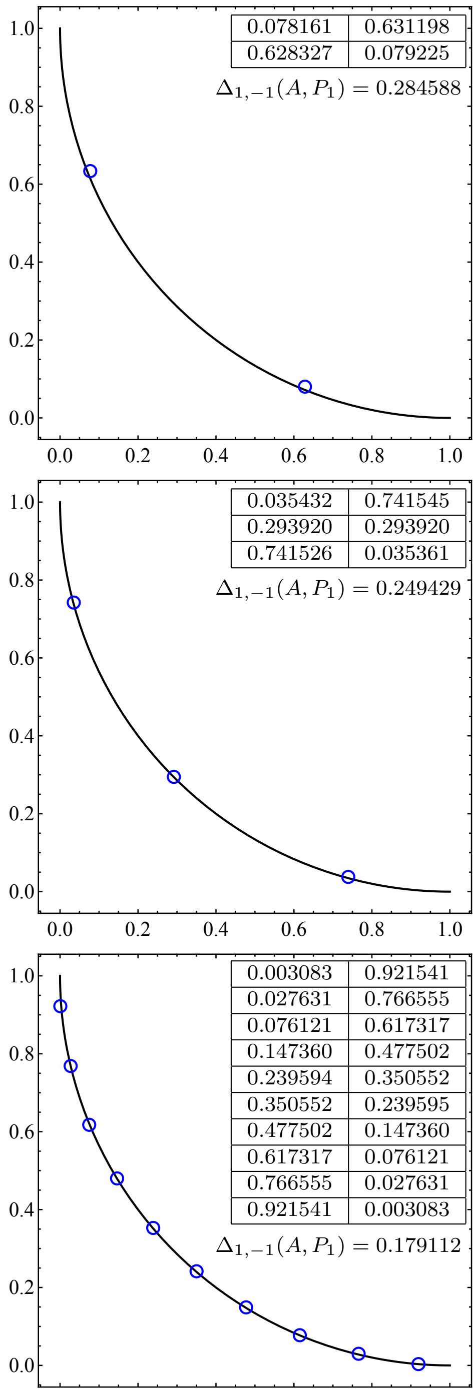

Fig. 2. Optimal $\Delta_{1,-1}$ archives A for the connected Pareto front P_1 given by (4.1) with 2, 3, and 10 elements (blue circles). Each figure includes the respective archive coordinates and the $\Delta_{1,-1}$ distance

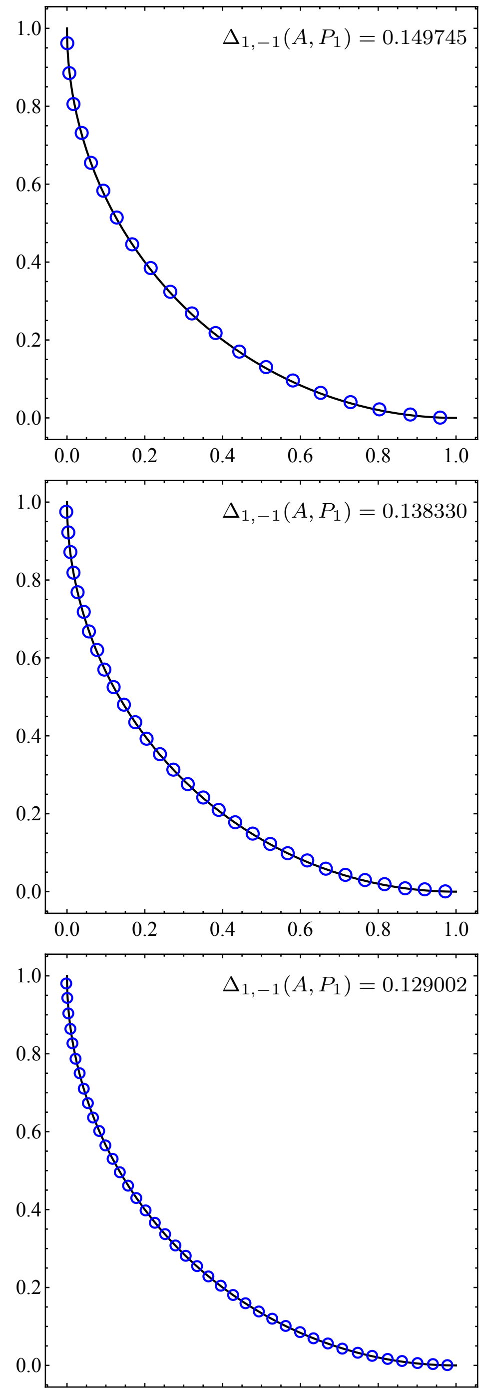

Fig. 3. Optimal $\Delta_{1,-1}$ archives A for the connected Pareto front P_1 given by (4.1) with 20, 30, and 40 elements (blue circles). Each figure includes the respective $\Delta_{1,-1}$ distance

ISSN 2007-9737

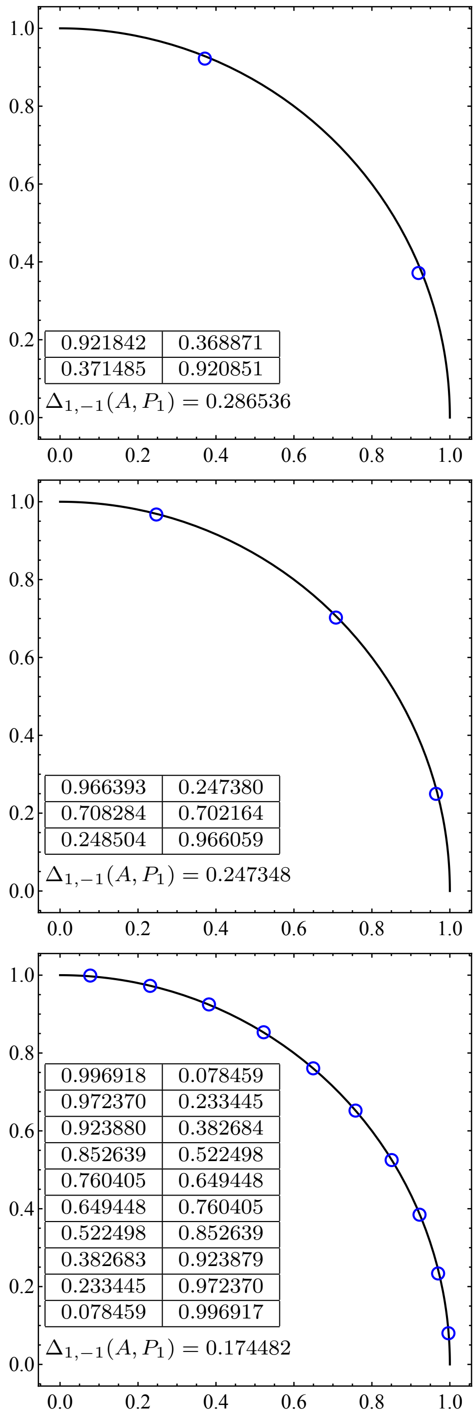

Fig. 4. Optimal $\Delta_{1,-1}$ archives A for the connected Pareto front P_2 given by (4.1) with 2, 3, and 10 elements (blue circles). Each figure includes the respective archive coordinates and the $\Delta_{1,-1}$ distance

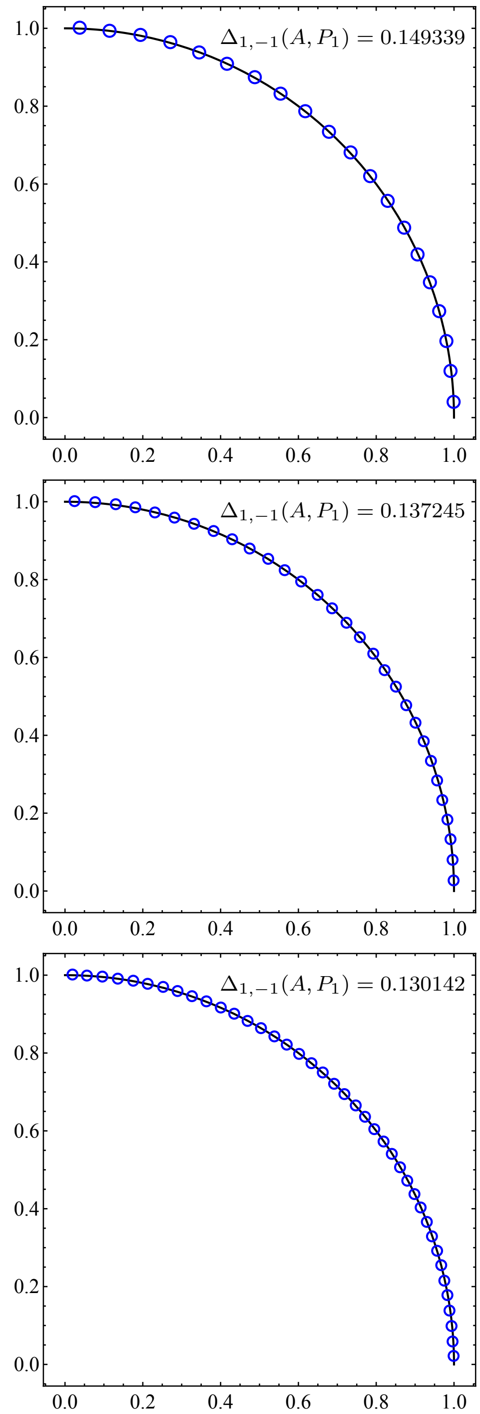

Fig. 5. Optimal $\Delta_{1,-1}$ archives A for the connected Pareto front P_2 given by (4.1) with 20, 30, and 40 elements (blue circles). Each figure includes the respective $\Delta_{1,-1}$ distance

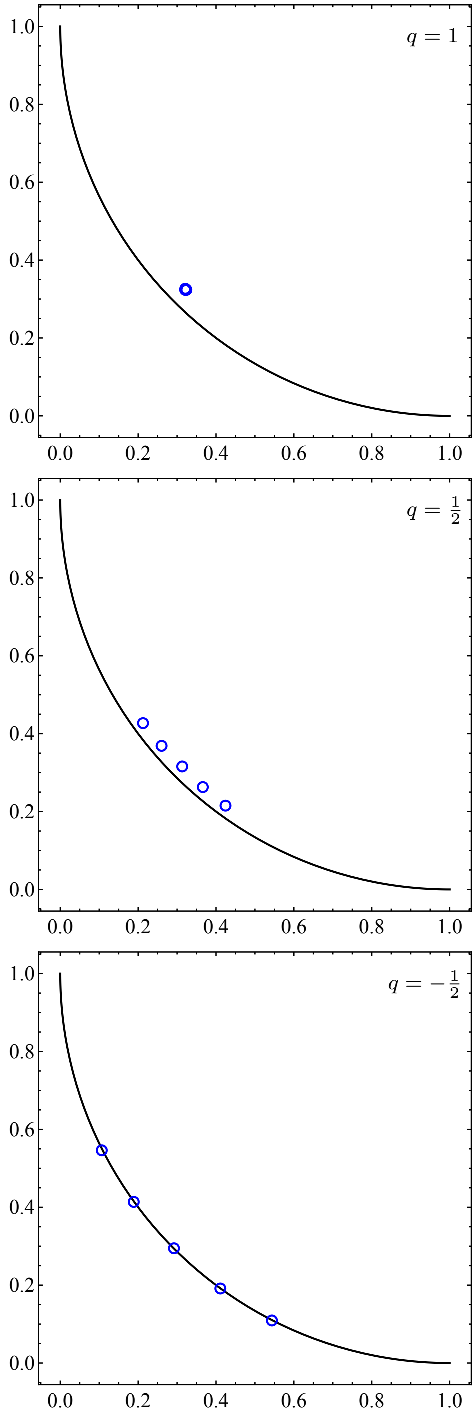

Fig. 6. Optimal $\Delta_{1,q}$ five point set archives A for the connected Pareto front P_1 given by (4.1) with $p = 1$ and different values of q

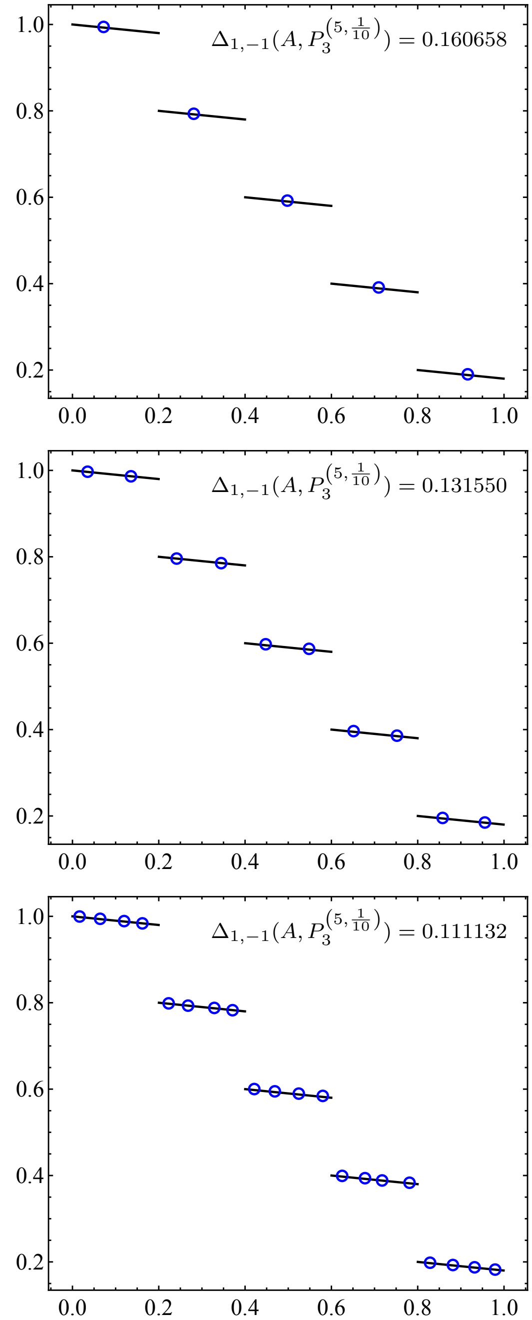

Fig. 7. Numerical optimal $\Delta_{1,-1}$ archive A for the disconnect step Pareto front $P_3^{(5)}$ given by (4.2) with 5, 10, and 20 elements. Each figure includes the respective $\Delta_{1,-1}$ distance

ISSN 2007-9737

Table 3. Percentage of the triangle inequality violations for different values of p and q. Here we randomly chose 80 sets, each one containing 2 points in $[0, 10]^2$, and verified the triangle inequality for all possible set permutations (that is 492960)

q \ p	1	2	5	10
-1	0.05396	0	0	0
-2	0.10265	0.00041	0	0
-5	0.28815	0.01217	0	0
-10	0.35622	0.05031	0.00041	0
-20	0.43046	0.08439	0.00446	0.00041

Table 4. The same as Table 3 but with sets containing 3 points in $[0, 10]^2$

q \ p	1	2	5	10
-1	0.00446	0	0	0
-2	0.02881	0	0	0
-5	0.15660	0.01379	0	0
-10	0.30388	0.05558	0.00609	0.00122
-20	0.40774	0.09250	0.01461	0.00609

Tables 3 and 4 show that $\Delta_{p,q}$ is close to a metric when $q \leqslant -1$ and $p \geqslant 1$. The percentage of the triangle inequality violations decreases as p increases or q decreases. Comparing both tables we can see, also, that this percentage decreases as the size of the sets increases.

4.2 Optimal Archives for Spherical Pareto Fronts

We now consider two standard Pareto sets: The convex and concave quarter-circle, see Figures 2, 3, 4, and 5.

$$P_1 = \left\{(\cos\theta + 1, \sin\theta + 1)\colon -\pi \leqslant \theta \leqslant -\frac{\pi}{2}\right\},$$
$$P_2 = \left\{(\cos\theta, \sin\theta)\colon 0 \leqslant \theta \leqslant \frac{\pi}{2}\right\}. \qquad (4.1)$$

To numerically find the optimal $\Delta_{p,q}$ archive of size M, we discretized the Pareto front with 1000 equidistant points (which is an acceptable discretization according to [12, p. 603]) and randomly choose an initial M-sized archive. Then we used a random-walk evolutionary algorithm moving one point at a time. Finally we refine the optimal archive with the "evenly spaced" construction suggested by [12, p. 607].

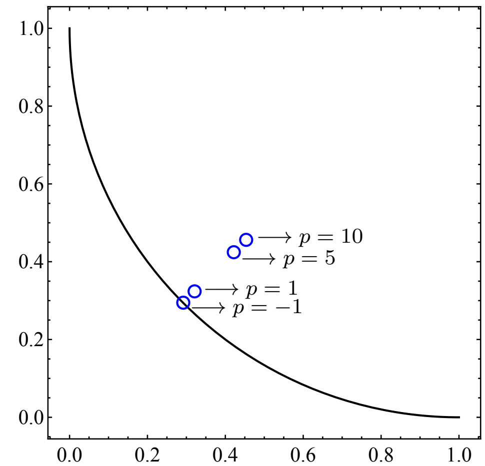

Fig. 8. Optimal $\Delta_{p,-1}$ one point archives A for the connected Pareto front P_1 given by (4.1) with $q = -1$ and different values of p. In all cases, the archives are located in the line $x = y$

When finding optimal $\Delta_{p,q}$ archives, our numerical experiments suggest a clear geometrical influence of the parameters p and q. When $p \geqslant -1$ increases the optimal archive moves away from the Pareto set (see Figure 8). For values of p in $(-\infty, -1)$ the optimal archive sets are basically the same. When $q \in [-1, 1]$ increases the optimal archive tends to lose dispersion, converging to one point. When $q \geqslant 1$ the optimal archive collapses to one point and when $q \in (-\infty, -1]$ the corresponding optimal archives are basically the same (see Figure 6).

4.3 Optimal Archives for Disconnected Pareto Sets

In this section we present the optimal $\Delta_{p,q}$ archives for a disconnected step Pareto front:

$$P_3^{(s,\gamma)} = \left\{\left(t, 1 - \gamma t + (\gamma - 1)\frac{\lfloor st \rfloor}{s}\right) : 0 \leqslant t \leqslant 1\right\}, \quad (4.2)$$

where s is the number of steps, $\gamma > 0$ is a small constant responsible for the step's twist, and $\lfloor \cdot \rfloor$ stands for the integer part function.

Figure 7 show numerical optimal $\Delta_{1,-1}$ archives of sizes 5, 10, and 20, respectively. In each case

doi: 10.13053/CyS-22-2-2950

(as in the previous section), the archive coordinates reveal that:

$$A \cap P_3^{\left(5, \frac{1}{10}\right)} = \varnothing,$$

i.e., the optimal archive points do not lie over the Pareto front but they are so close to it that this is hardly noticeable. It is also evident that the archives are evenly distributed along the Pareto front.

5 Conclusions

1. The indicator $\Delta_{p,q}$ generalizes the well-known averaged Hausdorff distance Δ_p (see (1.1) and (3.2)), it is still related with the standard Hausdorff distance (see (1.2)), and admits an expression in terms of the matrix $\ell_{p,q}$-norm $\|D_{AB}\|_{p,q}$ (see (3.3)).

2. For arbitrary values of $p, q \in \overline{\mathbb{R}}$, the indicator $\Delta_{p,q}$ is an inframetric on the space of finite subsets of $\mathbb{R}^n$, and when $p, q \in [1, \infty)$ it is a proper metric (see Corollary 3.4). With a proper metric the principle "the distance between two objects is the length of the shortest path joining them" is satisfied, thus, working with a metric has the advantage of avoiding unpleasant geometrical phenomena like the one shown in Figure 1.

3. For $p, q \in \mathbb{R}$ the $\mathrm{GD}_{p,q}$ and $\Delta_{p,q}$-performance indicators are compliant with an optimality associated to the dominance relations derived from the conditions in Theorem 3.9.

4. The parameters p and q play geometrical roles in the $\Delta_{p,q}$-optimal archive finding process, i.e., when p increases the optimal archive moves away from the Pareto set (see Figure 8) and when q increases, the optimal archive loses dispersion (see Figure 6). Thus the (p, q)-averaged distance can be calibrated to fulfill a large variety of optimization objectives.

5. Comparing our solutions with the optimal Δ_1 archives shown in [13], we conclude that they are very close and the procedure to calculate $\Delta_{p,q}$ is no harder than the one used for Δ_p, both analytically or numerically.

6 Future Work

1. Suppose that $p, q \in [1, \infty)$ are fixed. Given a Pareto front and an arbitrary archive, it would be useful to establish a procedure to find a shortest or best "path" of configurations joining the given archive with the optimal one. In principle, this is possible when we are working with a proper metric.

2. Section 3 shows that the metricity of $\Delta_{p,q}$ is a consequence of the properties of power means for appropriate values of p and q. This indicates viable ways in which one would be able to modify or generalize this indicator preserving its behavior.

3. A deeper study of the Pareto-type compliance of the $\mathrm{GD}_{p,q}$ and $\Delta_{p,q}$-indicators is desirable to better assess their characteristics and possible drawbacks. This is also very important for applications.

4. The details of the extension of $\mathrm{GD}_{p,q}$ and $\Delta_{p,q}$ to continuous sets and their properties are part of ongoing research and will appear in forthcoming publications (see [4]).

Acknowledgements

We would like to thank Prof. Oliver Schütze for introducing us to the subject of multiobjective optimization and for his warm friendship. We are also grateful with the referees, for the positive review and their useful comments and suggestions that helped to improve the presentation and results.

This work was partially supported by the project ID-PRY: 6736 of the Faculty of Sciences, Pontificia Universidad Javeriana, Bogotá, Colombia.

References

1. **Abyar, E. & Ghaemi, M. (2015).** Hausdorff measure of noncompactness of matrix operators on some sequence spaces of a double sequential band matrix. *J. Inequal. Appl.*, pp. 2015:406, 19.

2. **Aulbach, B., Rasmussen, M., & Siegmund, S. (2005).** Approximation of attractors of nonautonomous dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B*, Vol. 5, No. 2, pp. 215–238.
3. **Barnsley, M. F. (1993).** *Fractals everywhere*. Academic Press Professional, Boston, MA, second edition. Revised with the assistance of and with a foreword by Hawley Rising, III.
4. **Bogoya, J. M., Vargas, A., Cuate, O., & Schütze, O. (2018).** A (p, q)-averaged Hausdorff distance for arbitrary measurable sets. *Submitted*.
5. **Bullen, P. S. (2003).** *Handbook of means and their inequalities*, volume 560 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht. Revised from the 1988 original [P. S. Bullen, D. S. Mitrinović and P. M. Vasić, ıt Means and their inequalities, Reidel, Dordrecht; MR0947142].
6. **Burago, D., Burago, Y., & Ivanov, S. (2001).** *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.
7. **Falconer, K. (2003).** *Fractal geometry*. John Wiley & Sons, Inc., Hoboken, NJ, second edition. Mathematical foundations and applications.
8. **Fraigniaud, P., Lebhar, E., & L., V. (2008).** The inframetric model for the internet. *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, pp. 1085–1093.
9. **Goldberg, M. (1987).** Equivalence constants for l_p norms of matrices. *Linear and Multilinear Algebra*, Vol. 21, No. 2, pp. 173–179.
10. **Hansen, M. & Jaszkiewicz, A. (1998).** *Evaluating the quality of approximations to the non-dominated set*. IMM, Department of Mathematical Modelling, Technical University of Denmark.
11. **Pareto, V. (1971).** *Manual of Political Economy*. The Macmillan Press, London.
12. **Rudolph, G., Schütze, O., Grimme, C., Domínguez-Medina, C., & Trautmann, H. (2016).** Optimal averaged Hausdorff archives for bi-objective problems: theoretical and numerical results. *Comput. Optim. Appl.*, Vol. 64, pp. 589–618.
13. **Schütze, O., Esquivel, X., Lara, A., & Coello, C. (2012).** Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. *IEEE Trans. Evol. Comput.*, Vol. 16, No. 4, pp. 504–522.
14. **Siwel, J., Yew-Soon, O., Jie, Z., & Liang, F. (2014).** Consistencies and contradictions of performance metrics in multiobjective optimization. *IEEE Trans. Evol. Comput.*, Vol. 44, No. 12, pp. 2329–2404.
15. **Taylor, S. (1964).** The exact Hausdorff measure of the sample path for planar Brownian motion. *Proc. Cambridge Philos. Soc.*, Vol. 60, pp. 253–258.
16. **Zhang, Q. & Li, H. (2007).** MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.*, Vol. 11, No. 6, pp. 712–731.
17. **Zitzler, E. & Thiele, L. (1999).** Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.*, Vol. 3, No. 4, pp. 257–271.
18. **Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., & Da Fonseca, V. G. (2003).** Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Trans. Evol. Comput.*, Vol. 7, No. 2, pp. 117–132.

Article received on 01/02/2017; accepted on 28/08/2017.
Corresponding author is Andrés Vargas.