

To appear in *Statistics: A Journal of Theoretical and Applied Statistics*
Vol. 00, No. 00, Month 20XX, 1–24

A spatial-type interval-valued median for random intervals

Beatriz Sinova^{a,c*} and Stefan Van Aelst^{b,c}

^a*Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, 33071 Oviedo, Spain;*

^b*Department of Mathematics, KU Leuven, 3001 Leuven, Belgium;* ^c*Department of Applied Mathematics, Computer Science and Statistics, Ghent University, 9000 Gent, Belgium*

(Received November 2016)

To estimate the central tendency or location of a sample of interval-valued data, a standard statistic is the interval-valued sample mean. Its strong sensitivity to outliers or data changes motivates the search for more robust alternatives. In this respect, a more robust location statistic is studied in this paper. This measure is inspired by the concept of spatial median and makes use of the versatile generalized Bertoluzza's metric between intervals, the so called d_θ distance. The problem of minimizing the mean d_θ distance to the values the random interval takes, which defines the spatial-type d_θ -median, is analyzed. Existence and uniqueness of the sample version are shown. Furthermore, the robustness of this proposal is investigated by deriving its finite sample breakdown point. Finally, a real-life example from the Economics field illustrates the robustness of the sample d_θ -median, and simulation studies show some comparisons with respect to the mean and several recently introduced robust location measures for interval-valued data.

Keywords: interval-valued data; random interval; d_θ -median; algorithm; statistical robustness

AMS Subject Classification: 62G35; 62–07

1. Introduction and motivation

Nowadays, data is an essential source of information and knowledge. Statistical data analysis methodology is permanently evolving due to the emergence of new types of data in real-life studies and the development of tools to analyze more complex types of data. The first statistical techniques were simply designed to manage either quantitative or qualitative data. At present, we can find numerous statistical procedures to handle complex types of data such as functional data [1], fuzzy-valued data [2], incomplete/missing data [3], and several other types of data.

Among the types of real-life data that cannot be strictly classified as quantitative or qualitative, one often finds cases involving interval-valued data. Such data may have different origins, such as

- The representation of quality by perceptions, opinions or judgements (e.g., tastes, smells, beauty, etc.);
- Data which are exact observations of existing intrinsically interval-valued random attributes, like, for instance, attributes quantifying ranges, fluctuations, etc. along a

*Corresponding author. Email: sinovabeatriz@uniovi.es

period of time, a given parcel, etc. (e.g., the daily fluctuation of the systolic blood pressure, the monthly salary range of a worker over a year, and so on);

- Intrinsically real-valued data from which, to guarantee e.g. a certain level of confidentiality, only interval-valued information is recorded (for instance, in many surveys respondents are asked to indicate the interval containing their age, salary, etc.);
- Data which are obtained by aggregating a typically large dataset into one of a reduced size (this happens, for example, with interval-valued symbolic data [4, 5] which include interval variation and structure).

Our estimation method can be applied to any type of interval-valued data, irrespective of the source.

Any interval-valued observation can be identified with a two-dimensional vector for which the first component is its real-valued mid-point/centre and the second component is its nonnegative spread/radius. Hence, there is a one-to-one mapping between the space of nonempty compact intervals and the cone $\mathbb{R} \times [0, \infty)$. However, for statistical developments it is often inconvenient to identify the random mechanism generating interval values with a two-dimensional random vector with nonnegative second component. Common assumptions for multivariate techniques such as normality may not make sense or become too restrictive (e.g. elliptical symmetry) in the case of random intervals. A flexible approach to model interval-valued data is obtained by assuming that the data are outcomes of a random interval (that is, of an interval-valued random set), then possible interactions or dependencies between the two components can be better formalized and treated.

Data analysis involving random interval-valued data has already received some attention in the literature. Some of the problems that have been considered are regression analysis for interval-valued data [6–14], testing hypotheses with interval-valued data [15, 16] (or [11] in a more general situation), clustering interval-valued data [17–22], principal component analysis with interval-valued data [5, 23], modelling distributions for interval-valued data [24, 25], among others.

As for standard data, a useful summary of the information contained in a set of interval-valued observations (or in the distribution of a random interval) is its central tendency of location. One of the most commonly used location measures is the Aumann mean [26, 27]. The Aumann mean is supported by many valuable properties, including laws of Large Numbers, and is coherent with interval arithmetic. However, despite the very good probabilistic and statistical properties of this location measure, it is not always a reliable and representative measure of the center of interval-valued observations or the distribution of a random interval. This is due to its lack of robustness. Similarly as the mean in the real/vectorial-valued case, the Aumann mean is strongly influenced by data changes or the occurrence of outliers. This drawback motivates the search for a more robust measure of the center of random intervals.

In the real-valued setting, the most popular robust alternative for the mean is the median. However, since there is no universally accepted total order criterion over the space of non-empty compact intervals, similarly as in the case of vector-valued data, the median of interval-valued observations cannot be (universally) defined as a ‘middle position’ value within a ranked set. On the other hand, distances between intervals can be stated with sound support and intuitive interpretation. Therefore, one can attempt to extend the median to random intervals by defining it as the interval minimizing the mean distance to all the values the random interval can take.

For multivariate data, the spatial median (also called the L^1 -median, as introduced by Weber [28]) is a popular robust alternative for estimating the center of the multivariate

data when the data possibly contains outlying observations. The spatial median is defined as the point in the multivariate space with minimal average Euclidean distance to the observations. For more details and applications, see for instance [29–35].

To adapt the spatial median to interval-valued data we need to consider a suitable L^2 metric on this space. The approach in this paper is based on the versatile generalized metric introduced by Bertoluzza *et al.* [36], which can be expressed in a convenient and easy-to-handle parameterized way [6] (see also [37], for a recent review in a more general setting). The generalized metric includes as a special case the well-known Vitale L^2 metric [38] for interval values. The resulting estimator was first mentioned in [39] (a preliminary empirical study of its behavior) and its strong consistency was proven in [40]. One of the aims of this paper is to provide theoretical results for the existence and uniqueness of the spatial-type median. Furthermore, it will be shown that the resulting estimator is robust with a high breakdown point and has good finite-sample properties, also in comparison to alternative M-estimators of location, whose robustness in the interval-valued settings has already been shown in [41] (see [42] for the more general case of fuzzy-valued data). Note that the spatial-type interval-valued median is also an M-estimator of location with the absolute value function as loss function. However, the results in [41, 42] do not extend to the spatial-type interval-valued median because its loss function does not satisfy the conditions required by the representer theorem (see Theorem 3.3 in [41] or Theorem III-A.1 in [42]). Therefore, theoretical results such as existence and uniqueness are not yet available for the spatial-type median. Moreover, a comparison of the spatial-type median and M-estimators defined through the representer theorem is especially of interest.

Some other approaches to extend the spatial median to more general spaces have already been proposed in the literature. For example, Debruyne *et al.* [43] introduced an extension to feature spaces for kernel methods. This extension still uses the Euclidean norm and seeks for the element corresponding to a solution of the first order equations, while we define our extension of the spatial median as the interval minimizing the average distance to the values of the random interval. Moreover, in [43] neither existence nor uniqueness of the solution is investigated.

The rest of this paper is organized as follows: in Section 2, the basic concepts related to the interval-valued space, interval arithmetic and metrics for intervals are recalled. Random intervals and the Aumann mean are also reviewed. In Section 3, the d_θ -median, an L^2 -type median for random intervals inspired by the spatial median, is proposed and the existence and uniqueness of its sample version are shown. The robustness properties of the sample d_θ -median are investigated in Section 4 and two algorithms for its computation are explained in detail in Section 5. In Section 6, we illustrate its application on a real data example and we empirically compare its behaviour to the Aumann mean, two L^1 -type medians introduced in [44] and an M-estimator of location based on Hampel loss function proposed in [41]. Finally, Section 7 ends with some concluding remarks.

2. Preliminary concepts

Let $\mathcal{K}_c(\mathbb{R})$ denote the class of nonempty compact intervals from \mathbb{R} . Any interval K in the space $\mathcal{K}_c(\mathbb{R})$ can be characterized in terms of either its infimum and supremum, $K = [\inf K, \sup K]$, or its mid-point and spread or radius, $K = [\text{mid } K - \text{spr } K, \text{mid } K + \text{spr } K]$, where

$$\text{mid } K = \frac{\inf K + \sup K}{2}, \quad \text{spr } K = \frac{\sup K - \inf K}{2} \geq 0.$$

The usual interval arithmetic provides both the addition and the product by a scalar, which are the two most relevant operations from a statistical point of view. These operations are defined as follows.

- The Minkowski *sum* of $K, K' \in \mathcal{K}_c(\mathbb{R})$ is defined as the interval

$$K + K' = [\inf K + \inf K', \sup K + \sup K'].$$

It admits an alternative expression in terms of the second characterization:

$$\text{mid}(K + K') = \text{mid} K + \text{mid} K', \text{spr}(K + K') = \text{spr} K + \text{spr} K'.$$

- The *product of an interval $K \in \mathcal{K}_c(\mathbb{R})$ by a scalar $\gamma \in \mathbb{R}$* is defined as the element of $\mathcal{K}_c(\mathbb{R})$ such that

$$\begin{aligned} \gamma \cdot K &= \begin{cases} [\gamma \cdot \inf K, \gamma \cdot \sup K] & \text{if } \gamma \geq 0 \\ [\gamma \cdot \sup K, \gamma \cdot \inf K] & \text{otherwise} \end{cases} \\ &= [\gamma \cdot \text{mid} K - |\gamma| \cdot \text{spr} K, \gamma \cdot \text{mid} K + |\gamma| \cdot \text{spr} K]. \end{aligned}$$

In contrast to the real-valued case, an important complexity is that the space $\mathcal{K}_c(\mathbb{R})$ is not linear with these two operations, but only semilinear with a conical structure, because of the lack of an opposite element for the Minkowski addition. Hence, there does not exist a generally well-defined notion of difference of intervals preserving key properties of the real-valued case. For this reason it is not possible to treat intervals directly as two-dimensional vectors. Therefore, distances play a crucial role in statistical developments.

To measure the distance between two interval-valued observations, the following metrics have been considered among others:

- The *Hausdorff metric* [45] between two intervals $K, K' \in \mathcal{K}_c(\mathbb{R})$ given by

$$d_H(K, K') = |\text{mid} K - \text{mid} K'| + |\text{spr} K - \text{spr} K'|.$$

- The *1-norm metric*, introduced by Vitale [38]. The 1-norm distance between any two intervals $K, K' \in \mathcal{K}_c(\mathbb{R})$ is given by

$$\rho_1(K, K') = \frac{1}{2} |\inf K - \inf K'| + \frac{1}{2} |\sup K - \sup K'|.$$

This metric is topologically equivalent to the well-known Hausdorff metric above.

- The *d_θ metric* by Gil *et al.* [6] is given by

$$d_\theta(K, K') = \sqrt{(\text{mid} K - \text{mid} K')^2 + \theta \cdot (\text{spr} K - \text{spr} K')^2},$$

where $K, K' \in \mathcal{K}_c(\mathbb{R})$ and $\theta \in (0, \infty)$. It is often imposed that $\theta \in (0, 1]$, in order to weigh the deviation in location no less than the deviation in imprecision, as well as to make d_θ coincide with Bertoluzza *et al.*'s metric [36]. This last coincidence is especially helpful in choosing values for the weight θ . In this way, the common choice $\theta = 1/3$ corresponds to $d_{1/3}(K, K') = \sqrt{\int_{[0,1]} (K^{[\lambda]} - K'^{[\lambda]})^2 d\lambda}$, where $K^{[\lambda]} = \lambda \cdot \sup K + (1 - \lambda) \cdot \inf K$. On the other hand, $\theta = 1$ allocates a uniform relevance to the extreme points in each interval; actually, the d_1 metric coincides with the Euclidean distance in the two-dimensional space $\{(\text{mid} K, \text{spr} K) : K \in \mathcal{K}_c(\mathbb{R})\}$, and also with the 2-norm metric between intervals by Vitale [38]. Nevertheless, the theoretical results will be presented for the whole class of d_θ metrics.

Following the general random set approach, a *random interval* is usually defined as a Borel measurable mapping $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, where (Ω, \mathcal{A}, P) is a probability space, with respect to \mathcal{A} . The Borel σ -field generated on $\mathcal{K}_c(\mathbb{R})$ by the topology induced by any of the previous metrics (which are all of them strongly equivalent) is considered. As a consequence from the Borel measurability, crucial concepts in probabilistic and inferential developments, such as the (induced) distribution of a random interval or the stochastic independence of random intervals, are well-defined. Alternatively, a random interval can be defined in terms of real-valued random variables: X is a random interval if, and only if, both functions $\inf X : \Omega \rightarrow \mathbb{R}$ and $\sup X : \Omega \rightarrow \mathbb{R}$ (or equivalently, $\text{mid } X : \Omega \rightarrow \mathbb{R}$ and $\text{spr } X : \Omega \rightarrow [0, \infty)$) are real-valued random variables.

The *Aumann mean value* [26, 27], in case it exists, is given by the interval $E[X] = [E(\inf X), E(\sup X)]$ or, equivalently, in terms of the mid/spr representation, $E[X] = [E(\text{mid } X) - E(\text{spr } X), E(\text{mid } X) + E(\text{spr } X)]$. It can also be obtained as the Fréchet expectation [46] with respect to the d_θ metric, i.e., the Aumann expected value is the unique interval which minimizes, over $K \in \mathcal{K}_c(\mathbb{R})$, the expected squared distance $E[(d_\theta(X, K))^2]$.

In the next section we use the d_θ metric as L^2 -type distance between intervals and propose a robust measure of the location of the distribution of a random interval by minimizing the mean d_θ distance to the values the random interval takes.

3. The d_θ -median of a random interval: existence and conditions for uniqueness

Inspired by the spatial median as extension of the median to higher dimensional Euclidean spaces and even Banach spaces [47], we now study a generalization of the spatial median to random intervals based on the L^2 metric d_θ as in Sinova *et al.* [39, 40]. We call this location measure the d_θ -median to stress its dependence on the parameter involved in the metric.

Definition 3.1 The d_θ -median(s) of a random interval $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ is (are) the interval(s) $M_\theta[X] \in \mathcal{K}_c(\mathbb{R})$ such that

$$E(d_\theta(X, M_\theta[X])) = \min_{K \in \mathcal{K}_c(\mathbb{R})} E(d_\theta(X, K)),$$

whenever the involved expectations exist.

By using the empirical distribution of the data we obtain the sample d_θ -median estimator, which is thus defined as follows:

Definition 3.2 Let (X_1, \dots, X_n) be a random sample of size n from a random interval X and with realizations $\mathbf{x}_n = (x_1, \dots, x_n)$. The *sample d_θ -median* (or medians) $\widehat{M}_\theta[X]_n$ is (are) the random interval that takes, for \mathbf{x}_n , the interval value(s) $\widehat{M}_\theta[\mathbf{x}_n]$ that is (are) the solution(s) of the following optimization problem:

$$\min_{K \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^n d_\theta(x_i, K) = \min_{(y,z) \in \mathbb{R} \times [0, \infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2 + \theta \cdot (\text{spr } x_i - z)^2}. \quad (1)$$

It is shown in [40] that for simple random samples, the sample d_θ -median is a strongly consistent estimator of the d_θ -median of a random interval under general conditions. However, theoretical results about the existence and uniqueness of the spatial-type interval-valued median are still lacking. In this section, it is shown that the use of the d_θ metric

to define the median of a random interval allows us to guarantee its existence and to state conditions for its uniqueness.

As a supporting result to prove such existence and uniqueness of the sample d_θ -median, we first show the convexity of the objective function to be minimized in (1).

PROPOSITION 3.3 *For any $\mathbf{x}_n = (x_1, \dots, x_n) \subset K_c(\mathbb{R})$, the function given by $f(y, z) = \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2 + \theta \cdot (\text{spr } x_i - z)^2}$ for $(y, z) \in \mathbb{R} \times [0, \infty)$ is strictly convex unless all the sample points $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$ are collinear (in this case f is still a convex function).*

Proof. $D = \mathbb{R} \times [0, \infty)$ is convex and given any two different elements $(y_1, z_1), (y_2, z_2) \in D$ and $t \in (0, 1)$ we have that

$$\begin{aligned} & f(t(y_1, z_1) + (1-t)(y_2, z_2)) = f(ty_1 + (1-t)y_2, tz_1 + (1-t)z_2) \\ &= \frac{1}{n} \sum_{i=1}^n \left[t^2((\text{mid } x_i - y_1)^2 + \theta(\text{spr } x_i - z_1)^2) + (1-t)^2((\text{mid } x_i - y_2)^2 + \theta(\text{spr } x_i - z_2)^2) \right. \\ & \quad \left. + 2t(1-t)((\text{mid } x_i - y_1)(\text{mid } x_i - y_2) + \theta(\text{spr } x_i - z_1)(\text{spr } x_i - z_2)) \right]^{1/2}, \end{aligned}$$

and

$$\begin{aligned} t \cdot f(y_1, z_1) + (1-t) \cdot f(y_2, z_2) &= \frac{1}{n} \sum_{i=1}^n \sqrt{t^2((\text{mid } x_i - y_1)^2 + \theta(\text{spr } x_i - z_1)^2)} \\ & \quad + \frac{1}{n} \sum_{i=1}^n \sqrt{(1-t)^2((\text{mid } x_i - y_2)^2 + \theta(\text{spr } x_i - z_2)^2)}. \end{aligned}$$

Since not all the sample points are collinear, it holds for all i that

$$\left[(\text{mid } x_i - y_1)(\text{spr } x_i - z_2) - (\text{spr } x_i - z_1)(\text{mid } x_i - y_2) \right]^2 \geq 0,$$

this inequality being strict for at least one subindex $i_0 \in \{1, \dots, n\}$. Then,

$$\begin{aligned} & 2\theta(\text{mid } x_i - y_1)(\text{mid } x_i - y_2)(\text{spr } x_i - z_1)(\text{spr } x_i - z_2) \\ & \leq \theta(\text{mid } x_i - y_1)^2(\text{spr } x_i - z_2)^2 + \theta(\text{spr } x_i - z_1)^2(\text{mid } x_i - y_2)^2 \end{aligned}$$

with a strict inequality for the subindex i_0 . By adding the expression $(\text{mid } x_i - y_1)^2(\text{mid } x_i - y_2)^2 + \theta^2(\text{spr } x_i - z_1)^2(\text{spr } x_i - z_2)^2$ to both sides of the inequality and taking square roots, we obtain that

$$\begin{aligned} & (\text{mid } x_i - y_1)(\text{mid } x_i - y_2) + \theta(\text{spr } x_i - z_1)(\text{spr } x_i - z_2) \\ & \leq \left[(\text{mid } x_i - y_1)^2(\text{mid } x_i - y_2)^2 + \theta^2(\text{spr } x_i - z_1)^2(\text{spr } x_i - z_2)^2 \right. \\ & \quad \left. + \theta(\text{mid } x_i - y_1)^2(\text{spr } x_i - z_2)^2 + \theta(\text{spr } x_i - z_1)^2(\text{mid } x_i - y_2)^2 \right]^{1/2}, \end{aligned}$$

whence for $t \in (0, 1)$

$$t(1-t)((\text{mid } x_i - y_1)(\text{mid } x_i - y_2) + \theta(\text{spr } x_i - z_1)(\text{spr } x_i - z_2))$$

$$\leq [t^2((\text{mid } x_i - y_1)^2 + \theta(\text{spr } x_i - z_1)^2) + (1-t)^2((\text{mid } x_i - y_2)^2 + \theta(\text{spr } x_i - z_2)^2)]^{1/2}.$$

After adding now $t^2[(\text{mid } x_i - y_1)^2 + \theta(\text{spr } x_i - z_1)^2] + (1-t)^2[(\text{mid } x_i - y_2)^2 + \theta(\text{spr } x_i - z_2)^2]$ to both sides of the last inequality and by taking square roots, one can conclude that

$$f(t(y_1, z_1) + (1-t)(y_2, z_2)) < t \cdot f(y_1, z_1) + (1-t) \cdot f(y_2, z_2)$$

since the strict inequality holds for at least one subindex i_0 . Of course, if all the sample points are collinear, all the previous strict inequalities reduce simply to equalities. \square

By using Proposition 3.3 we now show the existence and conditions for the uniqueness of the sample d_θ -median.

THEOREM 3.4 *Given a random sample (X_1, \dots, X_n) from a random interval $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, the corresponding sample d_θ -median always exists. Moreover, the sample d_θ -median is unique for any sample realization $\mathbf{x}_n = (x_1, \dots, x_n)$ for which the two-dimensional sample points $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$ are not all collinear.*

Proof. Two cases will be distinguished, namely, when all the sample interval-valued data are just real numbers and when they are not.

- Consider the particular case in which all the sample interval-valued data x_i are just real numbers ($\text{spr } x_i = 0$ for all $i = 1, \dots, n$). Should this be the case, the aim would be to find the solution of the minimization problem

$$\min_{(y,z) \in \mathbb{R} \times [0,\infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2 + \theta z^2}.$$

By taking into account that

$$\begin{aligned} & \min_{(y,z) \in \mathbb{R} \times [0,\infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2 + \theta z^2} \\ & \geq \min_{y \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2} = \min_{y \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n |\text{mid } x_i - y|, \end{aligned}$$

the sample d_θ -median is the interval with mid-point $y = \text{Me}(\{\text{mid } x_i\}_{i=1}^n)$ and spread $z = 0$. As a consequence, the d_θ -median coincides with the median for the real-valued case.

- Now the case in which at least one sample interval-valued datum has not degenerated to a real number (i.e., there exists i_0 such that $\text{spr } x_{i_0} \neq 0$) will be studied.

Note that the possibility of considering a degenerated d_θ -median with null spread (i.e., a pair of the form $(y, 0)$ minimizing the objective function f) is excluded, because whatever the value of y may be, the objective function f is smaller for the pair $(y, 2 \cdot \min_{j: \text{spr } x_j \neq 0} \{\text{spr } x_j\})$:

$$\begin{aligned} f(y, 2 \cdot \min_{j: \text{spr } x_j \neq 0} \{\text{spr } x_j\}) &= \frac{1}{n} \sum_{i=1}^n \left[(\text{mid } x_i - y)^2 + \theta \left(\text{spr } x_i - 2 \cdot \min_{j: \text{spr } x_j \neq 0} \{\text{spr } x_j\} \right)^2 \right]^{1/2} \\ &= \frac{1}{n} \sum_{i=1}^n \left[(\text{mid } x_i - y)^2 + 4\theta \min_{j: \text{spr } x_j \neq 0} \{\text{spr } x_j\} \left(\min_{j: \text{spr } x_j \neq 0} \{\text{spr } x_j\} - \text{spr } x_i \right) + \theta(\text{spr } x_i)^2 \right]^{1/2} \\ &< \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y)^2 + \theta(\text{spr } x_i)^2} = f(y, 0). \end{aligned}$$

Therefore, D reduces to $\mathbb{R} \times (0, \infty)$. We now show that the minimum, if it exists, belongs to a closed rectangle. Convexity will then be used to prove the existence.

We use the following result from standard calculus. If a real-valued function of several variables f has a local maximum or minimum in an interior point (y_0, z_0) of its domain $D = \mathbb{R} \times (0, \infty)$ and the partial derivatives f_y and f_z exist at (y_0, z_0) , then it holds that $f_y(y_0, z_0) = f_z(y_0, z_0) = 0$. Therefore, if a local minimum is achieved at (y_0, z_0) , then one of the following situations must hold:

- The point associated with the local minimum belongs to the interior of D , and f_y and f_z exist at this point. In this case, $f_y(y_0, z_0) = f_z(y_0, z_0) = 0$, i.e.,

$$\sum_{i=1}^n \frac{\text{mid } x_i - y_0}{\sqrt{(\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2}} = \sum_{i=1}^n \frac{\text{spr } x_i - z_0}{\sqrt{(\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2}} = 0.$$

Moreover, we now check that in case such an extremum is achieved in the interior of D , $f(y_0, z_0)$ is a local minimum and not a local maximum. Using the second derivative criterion for local minimum, we prove that

- the partial derivatives

$$\begin{aligned} f_y(y, z) &= \frac{1}{n} \sum_{i=1}^n \frac{y - \text{mid } x_i}{\sqrt{(\text{mid } x_i - y)^2 + \theta(\text{spr } x_i - z)^2}} \\ f_z(y, z) &= \frac{1}{n} \sum_{i=1}^n \frac{\theta(z - \text{spr } x_i)}{\sqrt{(\text{mid } x_i - y)^2 + \theta(\text{spr } x_i - z)^2}} \end{aligned} \quad (2)$$

are continuous and have continuous partial derivatives on the domain $D \setminus \{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n \subset \mathbb{R}^2$;

- $f_{yy}(y_0, z_0) = \frac{1}{n} \sum_{i=1}^n \frac{\theta(\text{spr } x_i - z_0)^2}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} > 0$;
- finally, by using the Cauchy-Swartz inequality (strict inequality whenever not all the sample points are collinear),

$$\begin{aligned} f_{yy}(y_0, z_0) \cdot f_{zz}(y_0, z_0) &= \frac{\theta^2}{n^2} \left(\sum_{i=1}^n \frac{(\text{spr } x_i - z_0)^2}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} \right) \\ &\quad \cdot \left(\sum_{i=1}^n \frac{(y_0 - \text{mid } x_i)^2}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} \right) \\ &> \frac{\theta^2}{n^2} \left(\sum_{i=1}^n \frac{\text{spr } x_i - z_0}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} \right. \\ &\quad \left. \cdot \frac{y_0 - \text{mid } x_i}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} \right)^2 \\ &= \left(\frac{1}{n} \sum_{i=1}^n \frac{\theta(y_0 - \text{mid } x_i)(\text{spr } x_i - z_0)}{\sqrt{((\text{mid } x_i - y_0)^2 + \theta(\text{spr } x_i - z_0)^2)^3}} \right)^2 = (f_{yz}(y_0, z_0))^2. \end{aligned}$$

So, all the sufficient conditions of the criterion are fulfilled and there would be a local minimum at (y_0, z_0) .

- The local minimum is achieved at a point belonging to the interior of D but f_y or f_z do not exist at it. This can only happen if any of the square roots is equal to zero, i.e., if the point (y_0, z_0) belongs to the sample $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$.

By analyzing the possible situations, the following cases cannot hold:

- ✓ if $y_0 < \min_{1 \leq i \leq n} \{\text{mid } x_i\}$, then $f_y(y_0, z_0) < 0$ and (y_0, z_0) does not belong to the sample;
- ✓ if $y_0 > \max_{1 \leq i \leq n} \{\text{mid } x_i\}$, then $f_y(y_0, z_0) > 0$ and (y_0, z_0) does not belong to the sample;
- ✓ if $z_0 < \min_{1 \leq i \leq n} \{\text{spr } x_i\}$, then $f_z(y_0, z_0) < 0$ and (y_0, z_0) does not belong to the sample;
- ✓ if $z_0 > \max_{1 \leq i \leq n} \{\text{spr } x_i\}$, then $f_z(y_0, z_0) > 0$ and (y_0, z_0) does not belong to the sample.

Consequently, if (y_0, z_0) exists, it should belong to the rectangle

$$\left[\min_{1 \leq i \leq n} \{\text{mid } x_i\}, \max_{1 \leq i \leq n} \{\text{mid } x_i\} \right] \times \left[\min_{1 \leq i \leq n} \{\text{spr } x_i\}, \max_{1 \leq i \leq n} \{\text{spr } x_i\} \right].$$

As the restriction of the objective function f to this closed and bounded subset of \mathbb{R}^2 is continuous, the Weierstrass Theorem guarantees that f has at least one minimum within this subset. Indeed, the function f is convex, so any local minimum is also a global minimum. Since the function is strictly convex whenever not all the sample data are collinear, the global minimum is unique in such a case. \square

4. Robustness of the sample d_θ -median

We investigate the robustness of the sample d_θ -median by means of its finite sample breakdown point. The *finite sample breakdown point* (fsbp) is defined as the minimum proportion of sample data that needs to be perturbed to make the bias on the estimator arbitrarily large [48]. The fsbp of the spatial median for multivariate data has been investigated by Lopuhaä and Rousseeuw [49]. To investigate the fsbp of the d_θ -median, we first extend the definition of fsbp to the case of interval-valued data. Formally, the fsbp of the sample d_θ -median of a sample \mathbf{x}_n of size n from a random interval X can be defined as

$$\text{fsbp}(\widehat{M_\theta[X]}_n, \mathbf{x}_n, d_\theta) = \frac{1}{n} \min \left\{ k \in \{1, \dots, n\} : \sup_{\mathbf{y}_{n,k}} d_\theta(\widehat{M_\theta[\mathbf{x}_n]}, \widehat{M_\theta[\mathbf{y}_{n,k}]}) = \infty \right\},$$

where $\widehat{M_\theta[\mathbf{y}_{n,k}]}$ is the sample median of a sample $\mathbf{y}_{n,k}$ obtained from \mathbf{x}_n by perturbing at most k observations. Note that the d_θ -metric satisfies the condition $\sup\{d_\theta(K, K') : K, K' \in \mathcal{K}_c(\mathbb{R})\} = \infty$ which is required in the definition of the fsbp. Using this definition, the following result can be obtained.

THEOREM 4.1 *The finite sample breakdown point of the sample d_θ -median from a random interval X equals*

$$\text{fsbp}(\widehat{M_\theta[X]}_n, \mathbf{x}_n, d_\theta) = \frac{1}{n} \cdot \left\lfloor \frac{n+1}{2} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

Proof. The proof follows the lines of the derivation for the multivariate spatial median in

[49]. However, the situation is different because the d_θ -median is not translation equivariant in the same sense considered in that paper (due to the semilinearity of $\mathcal{K}_c(\mathbb{R})$). We first prove that

$$\text{fsbp}(\widehat{M_\theta[X]}_n, \mathbf{x}_n, d_\theta) > \frac{1}{n} \cdot \left(\left\lfloor \frac{n+1}{2} \right\rfloor - 1 \right) = \frac{1}{n} \cdot \left\lfloor \frac{n-1}{2} \right\rfloor,$$

i.e., $\sup_{\mathbf{y}_{n,k}} d_\theta \left(\widehat{M_\theta[\mathbf{x}_n]}, \widehat{M_\theta[\mathbf{y}_{n,k}]} \right) < \infty$ for all $k \in \mathbb{N}, k \leq \lfloor \frac{n-1}{2} \rfloor$. Now, let k be any natural number such that $k \leq \lfloor \frac{n-1}{2} \rfloor$. By the triangular inequality,

$$\sup_{\mathbf{y}_{n,k}} d_\theta \left(\widehat{M_\theta[\mathbf{x}_n]}, \widehat{M_\theta[\mathbf{y}_{n,k}]} \right) \leq \sup_{\mathbf{y}_{n,k}} d_\theta \left(\widehat{M_\theta[\mathbf{x}_n]}, \{0\} \right) + \sup_{\mathbf{y}_{n,k}} d_\theta \left(\{0\}, \widehat{M_\theta[\mathbf{y}_{n,k}]} \right).$$

The first term, which does not depend on the perturbed sample, is finite, so it suffices to show that the second term is also finite.

Put $\eta = \max_{1 \leq i \leq n} d_\theta(\{0\}, x_i)$ and let $B(\mathbf{0}, 2\eta)$ denote the closed ball with center $\mathbf{0} = (0, 0)$ and radius 2η . Denote the distance between $\widehat{M_\theta[\mathbf{y}_{n,k}]}$ and $B(\mathbf{0}, 2\eta)$ by

$$d = \inf_{\substack{(\text{mid } V, \text{spr } V) \in B(\mathbf{0}, 2\eta) \\ V \in \mathcal{K}_c(\mathbb{R})}} d_\theta \left(\widehat{M_\theta[\mathbf{y}_{n,k}]}, V \right),$$

so that $d_\theta(\{0\}, \widehat{M_\theta[\mathbf{y}_{n,k}]}) \leq d + 2\eta$. Then, for each of the replaced y_j 's, it holds by the triangular inequality that

$$d_\theta \left(y_j, \widehat{M_\theta[\mathbf{y}_{n,k}]} \right) \geq d_\theta(y_j, \{0\}) - d_\theta \left(\{0\}, \widehat{M_\theta[\mathbf{y}_{n,k}]} \right) \geq d_\theta(y_j, \{0\}) - (d + 2\eta). \quad (3)$$

Suppose that the distance between $\widehat{M_\theta[\mathbf{y}_{n,k}]}$ and $B(\mathbf{0}, 2\eta)$ is large, i.e., $d > 2\eta k$; this assumption will now be shown to be wrong.

Since $x_i \in B(\mathbf{0}, \eta)$, for each of the $n - k$ original x_l 's in $\mathbf{y}_{n,k}$,

$$d_\theta(x_l, \widehat{M_\theta[\mathbf{y}_{n,k}]}) \geq \eta + d \geq d_\theta(\{0\}, x_l) + d. \quad (4)$$

From (3) and (4),

$$\begin{aligned} \sum_{i=1}^n d_\theta(y_i, \widehat{M_\theta[\mathbf{y}_{n,k}]}) &= \sum_{\substack{i=1 \\ i: y_i \text{ replaced}}}^n d_\theta(y_i, \widehat{M_\theta[\mathbf{y}_{n,k}]}) + \sum_{\substack{i=1 \\ i: y_i \text{ original}}}^n d_\theta(y_i, \widehat{M_\theta[\mathbf{y}_{n,k}]}) \\ &\geq \sum_{\substack{i=1 \\ i: y_i \text{ replaced}}}^n (d_\theta(y_i, \{0\}) - (d + 2\eta)) + \sum_{\substack{i=1 \\ i: y_i \text{ original}}}^n (d_\theta(\{0\}, x_i) + d) \\ &= \sum_{i=1}^n d_\theta(y_i, \{0\}) - k(d + 2\eta) + \sum_{\substack{i=1 \\ i: x_i \text{ original}}}^n d = \sum_{i=1}^n d_\theta(y_i, \{0\}) + (n - k)d - k(d + 2\eta) \\ &\geq \sum_{i=1}^n d_\theta(y_i, \{0\}) + d - 2\eta k > \sum_{i=1}^n d_\theta(y_i, \{0\}) \end{aligned}$$

by using the assumption $d > 2\eta k$. However, this inequality cannot hold because $\widehat{M}_\theta[\mathbf{y}_{n,k}]$ minimizes the mean d_θ distance to all the interval-valued data from the sample $\mathbf{y}_{n,k}$. Therefore, $d \leq 2\eta k \leq 2\eta \lfloor \frac{n-1}{2} \rfloor$ and the inequality that proves the first part holds:

$$\sup_{\mathbf{y}_{n, \lfloor \frac{n-1}{2} \rfloor}} d_\theta \left(\{0\}, M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n-1}{2} \rfloor}}] \right) \leq d + 2\eta \leq 2\eta \left\lfloor \frac{n-1}{2} \right\rfloor + 2\eta = 2\eta \left\lfloor \frac{n+1}{2} \right\rfloor < \infty.$$

We now derive the other inequality, $\text{fsbp}(\widehat{M}_\theta[\widehat{X}]_n, \mathbf{x}_n, d_\theta) \leq \frac{1}{n} \cdot \left\lfloor \frac{n+1}{2} \right\rfloor$. That is, we have to prove that $\sup_{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}} d_\theta \left(\widehat{M}_\theta[\mathbf{x}_n], M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}}] \right) = \infty$.

It suffices to find a corrupted collection $\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*$ (by replacing at most $\lfloor \frac{n+1}{2} \rfloor$ points of \mathbf{x}_n) such that $d_\theta \left(\widehat{M}_\theta[\mathbf{x}_n], M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*}] \right) = \infty$. We will then replace $\lfloor \frac{n+1}{2} \rfloor$ observations of \mathbf{x}_n by copies of the point $y^{(m)} \in \mathcal{K}_c(\mathbb{R})$ with

$$\inf_{\substack{(\text{mid } V, \text{spr } V) \in B(\mathbf{0}, \eta) \\ V \in \mathcal{K}_c(\mathbb{R})}} d_\theta(y^{(m)}, V) = m,$$

where m is an arbitrary number in \mathbb{N} . Thus the new sample $\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*$ contains $q = n - \lfloor \frac{n+1}{2} \rfloor$ of the original points. We show that $M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*}] = y^{(m)}$.

Since $\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^* = (x_1, \dots, x_q, \overbrace{y^{(m)}, \dots, y^{(m)}}^{(n-q \text{ times})})$, for any $z \in \mathcal{K}_c(\mathbb{R}) \setminus \{y^{(m)}\}$,

$$\begin{aligned} \sum_{i=1}^n d_\theta(y_i^*, z) &= \sum_{i=1}^q d_\theta(x_i, z) + \sum_{i=q+1}^n d_\theta(y^{(m)}, z) = \sum_{i=1}^q d_\theta(x_i, z) + (n-q) \cdot d_\theta(y^{(m)}, z) \\ &\geq \sum_{i=1}^q d_\theta(x_i, z) + q \cdot d_\theta(y^{(m)}, z) = \sum_{i=1}^q \left(d_\theta(x_i, z) + d_\theta(y^{(m)}, z) \right) \geq \sum_{i=1}^q d_\theta(x_i, y^{(m)}) \\ &= \sum_{i=1}^q d_\theta(x_i, y^{(m)}) + \sum_{i=q+1}^n d_\theta(y^{(m)}, y^{(m)}) = \sum_{i=1}^n d_\theta(y_i^*, y^{(m)}). \end{aligned}$$

Therefore, and due to the fact that $\widehat{M}_\theta[\mathbf{x}_n] \in B(\mathbf{0}, \eta)$,

$$\sup_{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}} d_\theta \left(\widehat{M}_\theta[\mathbf{x}_n], M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*}] \right) \geq d_\theta \left(\widehat{M}_\theta[\mathbf{x}_n], M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}^*}] \right) = d_\theta(\widehat{M}_\theta[\mathbf{x}_n], y^{(m)}) \geq m.$$

Since $m \in \mathbb{N}$ could be chosen to be arbitrarily large, it follows that

$$\sup_{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}} d_\theta \left(\widehat{M}_\theta[\mathbf{x}_n], M_\theta[\widehat{\mathbf{y}_{n, \lfloor \frac{n+1}{2} \rfloor}}] \right) = \infty. \quad \square$$

Theorem 4.1 shows that the d_θ -median is highly robust, while the Aumann mean is not robust with fsbp equal to $1/n$ (since it is defined in terms of the mean of real-valued random variables).

5. Algorithms to compute the sample d_θ -median

Although existence and uniqueness have been proven in Section 3, an explicit expression for the d_θ -median cannot be provided in general. In this section, two algorithms to compute the sample d_θ -median will be presented. The first alternative is an improved version of a natural and well-known algorithm for computing the spatial median, the iteratively re-weighted least squares algorithm. The second proposal is an algorithm based on non-linear minimization. In the real-valued framework, several algorithms for computing the spatial median have been compared in [34] and it has been concluded that the non-linear minimization procedure clearly outperforms other approaches. Although the conclusions of the comparison study in [34] may not hold directly for interval-valued data, we have chosen the algorithm based on non-linear minimization as the second alternative to compute the d_θ -median, since in many situations the classical settings are indeed a good inspiration for developing techniques for random intervals.

5.1. Vardi and Zhang's modification of Weiszfeld's algorithm

We now propose an efficient algorithm to compute the sample d_θ -median, which is based on the iterative procedure that Weiszfeld [50] – its English translation and some annotations can be found in [51] – first suggested for Euclidean spaces and distances. Although Kuhn [52] showed the monotone convergence of Weiszfeld's algorithm whenever the starting point does not belong to the domain of attraction of the data points, it was later proved [53] that such domain may contain a continuum set and that it is not even guaranteed that the set of bad starting points may not be dense. To avoid any problem, the simple modification proposed in Vardi and Zhang [54] will be used instead.

Given a sample of interval-valued observations, and a fixed value of the parameter θ (chosen by the user), the algorithm allows us to compute the corresponding d_θ -median. Note that the d_θ -median is unique unless all the two-dimensional sample points $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$ are collinear, as shown in Theorem 3.4.

The algorithm is based on the fact that the objective function in the minimization problem (1) is differentiable at any point of the domain $\mathbb{R} \times [0, \infty)$ except for the sample points $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$. Therefore, the minimum in (1) will be reached at either a sample point or the point for which both partial derivatives equal zero. From (1) it easily follows that this point (y_0, z_0) satisfies

$$y_0 = \sum_{i=1}^n w_i(y_0, z_0) \text{mid } x_i \quad \text{and} \quad z_0 = \sum_{i=1}^n w_i(y_0, z_0) \text{spr } x_i,$$

with

$$w_i(y_0, z_0) = \frac{((\text{mid } x_i - y_0)^2 + \theta \cdot (\text{spr } x_i - z_0)^2)^{-1/2}}{\sum_{j=1}^n ((\text{mid } x_j - y_0)^2 + \theta \cdot (\text{spr } x_j - z_0)^2)^{-1/2}} \quad (5)$$

for $i = 1, \dots, n$. It can thus immediately be seen that the mid-point and the spread of the sample d_θ -median are weighted means of the mid-points and the spreads of the intervals in the sample, respectively. The steps of the algorithm are now described in detail.

Step 0. If the data are specified in terms of their inf/sup characterization, then first

compute the mid-point and spread of the interval-valued data:

$$\text{mid } x_i = \frac{\inf x_i + \sup x_i}{2}, \quad \text{spr } x_i = \frac{\sup x_i - \inf x_i}{2}, \quad \text{for } i = 1, \dots, n.$$

Choose the value of the parameter $\theta > 0$ to fix the metric between intervals.

Step 1. Fix the tolerance threshold $\varepsilon > 0$ for convergence. Set $m = 1$ and select a starting solution $(y_m, z_m) \in \mathbb{R} \times [0, \infty)$. Calculate the corresponding error

$$\text{Error}_m = \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y_m)^2 + \theta \cdot (\text{spr } x_i - z_m)^2}. \tag{6}$$

Step 2. Compute the weights $w_i(y_m, z_m)$ given in (5) for all $i = 1, \dots, n$ such that $(\text{mid } x_i, \text{spr } x_i) \neq (y_m, z_m)$. Let \tilde{T} be the following function with components \tilde{T}_1 and \tilde{T}_2 :

$$\tilde{T}(y, z) = (\tilde{T}_1(y, z), \tilde{T}_2(y, z)) = \sum_{\substack{i=1 \\ (\text{mid } x_i, \text{spr } x_i) \neq (y, z)}}^n w_i(y, z) (\text{mid } x_i, \text{spr } x_i)$$

for all $(y, z) \in \mathbb{R} \times [0, \infty)$. The update following the original Weiszfeld algorithm would be either the value $\tilde{T}(y_m, z_m)$ if $(y_m, z_m) \notin \{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$ or the value $(\text{mid } x_j, \text{spr } x_j)$ if $(y_m, z_m) = (\text{mid } x_j, \text{spr } x_j)$ for some $j \in \{1, \dots, n\}$. However, by using the modification in [54], the estimate will be updated by $(y_{m+1}, z_{m+1}) = T(y_m, z_m)$ with

$$T(y, z) = \max \left\{ 1 - \frac{\xi(y, z)}{r(y, z)}, 0 \right\} \tilde{T}(y, z) + \min \left\{ 1, \frac{\xi(y, z)}{r(y, z)} \right\} (y, z)$$

using the convention $0/0 = 0$ for the division $\xi(y, z)/r(y, z)$, ξ representing the indicator function of the set of sample observations $\{(\text{mid } x_i, \text{spr } x_i)\}_{i=1}^n$ and r being defined by

$$r(y, z) = \left(\sum_{i: x_i \neq [y-z, y+z]} 1/d_\theta^i \right) \sqrt{(\tilde{T}_1(y, z) - y)^2 + \theta(\tilde{T}_2(y, z) - z)^2}$$

with the abbreviation $d_\theta^i = d_\theta(x_i, [y - z, y + z])$.

Step 3. Calculate the corresponding error Error_{m+1} as given by (6). If the difference $\text{Error}_m - \text{Error}_{m+1}$ exceeds the tolerance ε , then increase m by 1 and go to Step 2. Otherwise, return the solution (y_{m+1}, z_{m+1}) .

The convergence of this algorithm when computing the spatial median in a Euclidean space by using Euclidean distances was shown by Vardi and Zhang [54]. The same reasoning can be applied when using the d_θ metric and $\mathcal{K}_c(\mathbb{R})$, so the proof is not repeated here.

THEOREM 5.1 *Given a realization $\mathbf{x}_n = (x_1, \dots, x_n)$ of a simple random sample from a random interval X , and any $P_0 \in \mathbb{R} \times [0, \infty)$, define $P_r := T^r(P_0) = T(T(\dots T(P_0)))$ for $r \in \mathbb{N}$. Then $\lim_{r \rightarrow \infty} P_r = \widehat{M}_\theta[\mathbf{x}_n]$.*

5.2. Algorithm based on non-linear minimization

As explained in [34], the minimization of (1) can be done using general numerical algorithms designed for non-linear optimization purposes, such as a Newton-type algorithm

(see [55]) with the gradient information from equations (2). The Hessian matrix is always approximated using the *Line Search* method, since it is a faster approach than its analytical calculation. The steps of the algorithm are the following:

Step 0. If the data are specified in terms of their inf/sup characterization, then first compute the mid-point and spread of the interval-valued data:

$$\text{mid } x_i = \frac{\inf x_i + \sup x_i}{2}, \quad \text{spr } x_i = \frac{\sup x_i - \inf x_i}{2}, \quad \text{for } i = 1, \dots, n.$$

Choose the value of the parameter $\theta > 0$ to fix the metric between intervals.

Step 1. Fix the tolerance threshold $\varepsilon > 0$ for convergence. Set $m = 1$ and select a starting solution $(y_m, z_m) \in \mathbb{R} \times [0, \infty)$. Denote by f the objective function of (1) and approximate its Hessian matrix $\mathbf{H}f(y_m, z_m)$ by \mathbf{H}_m^* , an approximate initial value of $\mathbf{H}_m^* = \mathbf{I}(y_m, z_m)^T$ is often sufficient to achieve rapid convergence. Compute the search direction $\mathbf{p} = -(\mathbf{H}_m^*)^{-1}(\nabla f(y_m, z_m))^T$.

Step 2. Use the *Line Search* strategy to find a step size λ_m which yields an adequately corresponding decrease in the objective function (1) and move from the current position (y_m, z_m) to $(y_{m+1}, z_{m+1}) = (y_m, z_m) + \lambda_m \mathbf{p}$. The details of the *Line Search* procedure can be found in Algorithm 6.3.1 of [55].

Step 3. Calculate the corresponding error

$$\begin{aligned} \text{Error}_{m+1} = & \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y_{m+1})^2 + \theta \cdot (\text{spr } x_i - z_{m+1})^2} \\ & - \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } x_i - y_m)^2 + \theta \cdot (\text{spr } x_i - z_m)^2} \end{aligned} \quad (7)$$

and update the approximation of the Hessian matrix by Broyden-Fletcher-Goldfarb-Shanno algorithm:

$$\mathbf{H}_{m+1}^* = \mathbf{H}_m^* + \frac{\mathbf{q}_m^T \mathbf{q}_m}{\mathbf{q}_m \mathbf{s}_m^T} - \frac{\mathbf{H}_m^{*T} \mathbf{s}_m^T \mathbf{s}_m \mathbf{H}_m^*}{\mathbf{s}_m \mathbf{H}_m^* \mathbf{s}_m^T},$$

where $\mathbf{s}_m = (y_{m+1} - y_m, z_{m+1} - z_m)$ and $\mathbf{q}_m = \nabla f(y_{m+1}, z_{m+1}) - \nabla f(y_m, z_m)$. If Error_{m+1} exceeds the tolerance ε , then increase m by 1 and go to Step 2. Otherwise, return the solution (y_{m+1}, z_{m+1}) .

6. Empirical behavior of the d_θ -median

6.1. Real data application: the influence of Brexit on IBEX 35

The robustness of the sample d_θ -median is now illustrated on a real data example and compared to the Aumann mean and alternative robust location measures from the literature.

Example 6.1 The IBEX 35 is the benchmark stock market index of the Bolsa de Madrid, Spain's principal stock exchange. It is administered and calculated by Sociedad de Bolsas (a subsidiary of the company which runs Spain's securities markets including the Bolsa de Madrid) and is composed of the 35 securities listed on the Stock Exchange Interconnection System of the four Spanish Stock Exchanges which were most liquid during the control period pursuant to the terms of the regulation. For more details, visit the web page of the Bolsa de Madrid (<http://www.bolsamadrid.es/ing/asp/Portada/Portada.aspx>).

In this example, we consider the daily fluctuation of the IBEX 35 during the past months of the second semester of 2016 (from 1 June to 24 October 2016). To measure such fluctuation, the daily minimum and maximum values of the index along this period of time have been recorded. Figure 1 represents the daily fluctuation for the considered period of time.



Figure 1. The daily fluctuation of the IBEX 35 Index from 01/06/2016 to 24/10/2016

In this example, we have set the parameter θ equal to the common choice $\theta = 1/3$. With this value of θ , it can be shown that the d_θ metric takes into account all the points of the intervals and weighs them all equally (for details, see [37]). The corresponding sample d_θ -median has been computed using the iterative algorithm proposed in Section 5.1 with the 1-norm median for interval-valued data (see [44]) as starting solution. The (sample) 1-norm median is defined as

$$\widehat{\text{Me}}[X]_n = [\widehat{\text{Me}}(\inf X)_n, \widehat{\text{Me}}(\sup X)_n].$$

In Figure 2 the d_θ -median is compared to the Aumann mean, the 1-norm median, the Hausdorff-type median (see [44]) and the M-estimator of location based on the Hampel loss function (see [41]). The (sample) Hausdorff-type median is defined as

$$\widehat{\text{M}}[X]_n = [\widehat{\text{Me}}(\text{mid } X)_n - \widehat{\text{Me}}(\text{spr } X)_n, \widehat{\text{Me}}(\text{mid } X)_n + \widehat{\text{Me}}(\text{spr } X)_n].$$

In case any of the medians of real-valued random variables involved in the definition of the 1-norm and Hausdorff-type medians is non-unique, the common convention of taking the mid-point of the interval of medians is adopted.

Finally, the Hampel M-estimator of location is defined as

$$\widehat{\text{H}}^M[X]_n = \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^n \rho_{a,b,c}(d_\theta(X_i, K))$$

with $\rho_{a,b,c}$ the loss function given by

$$\rho_{a,b,c}(x) = \begin{cases} x^2/2 & \text{if } 0 \leq |x| < a \\ a(|x| - a/2) & \text{if } a \leq |x| < b \\ \frac{a(|x| - c)^2}{2(b-c)} + \frac{1}{2}a(b+c-a) & \text{if } b \leq |x| < c \\ a(b+c-a)/2 & \text{if } c \leq |x|. \end{cases}$$

For this example the tuning parameters have been set equal to $a = 2$, $b = 4$ and $c = 8$. Except for the Aumann mean, all these measures have been proven to reach the highest possible value for the finite sample breakdown point, but an application to a real-life example like this one can indicate differences in finite-sample behavior.

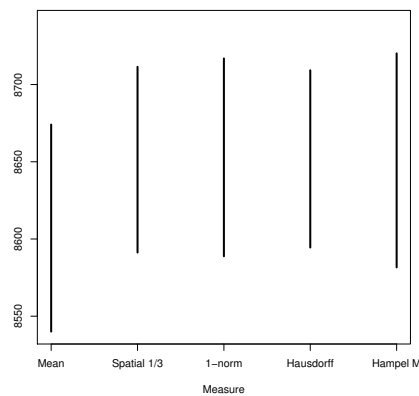


Figure 2. Interval-valued location estimates for the daily fluctuation of the IBEX 35 Index from 01/06/2016 to 24/10/2016

From Figure 2, we first notice that the mean interval is clearly lower than any other interval. This difference can be explained by looking at Figure 1. We can see that at the end of June and the beginning of July the daily interval-values of the IBEX 35 Index were much lower and, in some cases also wider, compared to the other periods. This period of instability was a consequence of the Brexit: the UK public voted to leave the European Union in a referendum on Thursday 23 June 2016. As Figure 1 clearly shows, the most outlying interval represents the fluctuation of the IBEX 35 on 24 June, the first day the IBEX 35 opened after the referendum results were known. We can notice that these interval measurements have a larger influence on the mean interval than on the other measures, as expected.

Table 1. Location estimates (rounded) for the IBEX 35 Index data

Measures	Mid-point	Spread
Mean	8607	67
$d_{1/3}$ -median	8651	60
1-norm	8653	64
Hausdorff	8651	57.5
Hampel M	8651	69

In relation to the comparison of the robust measures (the three interval-valued medians

and the M-estimator), Table 1 offers some additional information. From this table we can see that for this example the robust estimates all have (almost) the same midpoint, but show some differences in their spread. The Hampel M-estimate of location and the 1-norm median have a larger spread than the d_θ -median and the Hausdorff-type median. To investigate which estimate of the latter two medians has been affected more by the period of instability, we removed the interval observations of the IBEX 35 recorded on the first three opening days after the referendum, i.e., of 24, 27 and 28 June, and recalculated the robust location estimates for this reduced dataset:

- the d_θ -median is equal to $[8602, 8720]$;
- the Hausdorff-type median is equal to $[8607, 8721]$;

To further evaluate the effect of the outliers on the location estimators, we measured the distance between their value in the complete dataset and their value in the reduced dataset by using the d_θ -metric ($\theta = 1/3$). For the d_θ -median we obtain 10, while for the Hausdorff-type median we get 12.5. The Hausdorff-type median intervals with and without the most extreme observations are thus farther apart than the d_θ -median intervals.

6.2. Simulation study

The main purpose of this section now is to compare the newly introduced d_θ -median to the alternative location measures introduced above, i.e. the Aumann mean, the 1-norm and Hausdorff-type medians and the Hampel M-estimator of location, from an empirical point of view. For the d_θ -median we have used the common value of $\theta = 1/3$. For the M-estimator, the tuning parameters a , b and c in the Hampel loss function have been chosen according to the proposal in [56] to use the median, 75th and 85th percentiles of the distribution of sample distances (see [56] for details). Moreover, we also compare the performance of the solutions of the two algorithms for the d_θ -median developed in the previous section.

We compare the location measures in terms of their bias, variance and mean squared error (MSE). Based on N samples, these quantities can be estimated in the interval-valued setting by the formulas $\text{Bias} = d\left(\frac{1}{N} \sum_{i=1}^N \widehat{T}_i, T\right)$, $\text{Var} = \frac{1}{N} \sum_{i=1}^N \left(d^2\left(\widehat{T}_i, \frac{1}{N} \sum_{i=1}^N \widehat{T}_i\right)\right)$ and $\text{MSE} = \frac{1}{N} \sum_{i=1}^N \left(d^2(\widehat{T}_i, T)\right)$, respectively. Here, d denotes a metric between intervals, T represents the population value of a location measure and \widehat{T}_i denotes any of the estimates of T for the i th generated sample. We have computed the bias, variance and MSE in terms of the three different distances already used in this paper, namely, the ρ_1 -metric, d_H -metric and $d_{1/3}$ -metric. Therefore, each result in Tables 2 and 3 is given by a sequence $\rho_1\text{-metric} \star d_H\text{-metric} \star d_{1/3}\text{-metric}$ which represents the measurement of the quantity based on the three respective metrics.

The simulation setup can be described as follows.

Step 1. A random sample of 100 interval-valued observations is generated from a random interval X as follows:

Step 1.1 A random interval X is characterized as $X = [X_1 - X_2, X_1 + X_2]$ where $X_1 = \text{mid } X$ and $X_2 = \text{spr } X$ are two real-valued random variables. Hence, a random sample of X is obtained by randomly generating observations from X_1 and X_2 . The distributions considered for the real-valued random variables X_1 and X_2 are specified below.

Step 1.2 A proportion $\epsilon = 0, .1, .2$ or $.4$ of the observations in the sample generated in *Step 1.1* is then contaminated. The relative distance between the distributions of the regular and contaminated observations is controlled by a parameter C_D which can take the values $C_D = 0, 1, 5, 10$ or 100 in our simulations.

Step 2. Some of the population parameters T cannot be easily derived in an analytical way for random intervals. In that case, the population target is approximated by Monte Carlo simulation using $N = 10000$ replications of *Step 1* for the situation $(\epsilon, C_D) = (0, 0)$.

Step 3. For each combination (ϵ, C_D) , *Step 1* is repeated $N = 10000$ times and for each of these samples, the location estimates \hat{T}_i (i.e., the Aumann mean, the $d_{1/3}$ -median, the 1-norm median, the Hausdorff-type median and the M-estimator of location based on the Hampel loss function) are calculated.

Two different sets of distributions have been considered for the distributions of the regular and contaminated observations. In STUDY 1 the real-valued random variables have been generated according to

- $X_1 \rightsquigarrow \mathcal{N}(0, 1)$ and $X_2 \rightsquigarrow \chi_1^2$ for the non-contaminated observations,
- $X_1 \rightsquigarrow \mathcal{N}(0, 3) + C_D$ and $X_2 \rightsquigarrow \chi_4^2 / (C_D^2 + 1)$ for the contaminated observations,

whereas in STUDY 2 these variables are generated according to

- $X_1 \rightsquigarrow \beta(6, 1)$ and

$$X_2 \rightsquigarrow \begin{cases} \exp(100 + 4X_1) & \text{if } X_1 < .25, \\ \exp(200) & \text{if } .25 \leq X_1 \leq .75, \\ \exp(500 - 4X_1) & \text{if } X_1 > .75, \end{cases}$$

for the non-contaminated observations,

- $X_1 \rightsquigarrow \beta(1, 6)$ and

$$X_2 \rightsquigarrow \begin{cases} \exp(100 + 4X_1) / (C_D^2 + 1) & \text{if } X_1 < .25, \\ \exp(200) / (C_D^2 + 1) & \text{if } .25 \leq X_1 \leq .75, \\ \exp(500 - 4X_1) / (C_D^2 + 1) & \text{if } X_1 > .75. \end{cases}$$

for the contaminated observations.

Note that in both studies contaminated observations are generated to be outlying in both location and spread. These simulation results complement the empirical results presented in [41] where the focus was on M-estimators of location. The main conclusion in [41] was that there is no uniformly most appropriate location estimate, but the Hampel M-estimator of location often shows the best behavior with respect to either the bias, the variance or the mean squared error for the considered designs. The aim of our current empirical investigation is to show that the d_θ -median overall shows a stable behavior with performance that is consistently among the best estimators, and in realistic settings can outperform other interval-valued medians (STUDY 1) and even robust location estimators more generally (STUDY 2). Note that STUDY 1 was inspired by Study 1 - CASE1 in [41] while STUDY 2 is an adaptation of the simulation strategy proposed in [57] for the generation of trapezoidal fuzzy numbers whose aim was to mimic the human behavior when answering a fuzzy rating scale-based questionnaire.

6.2.1. STUDY 1

In STUDY 1 we compare the behavior of the d_θ -median with that of alternative interval-valued median estimators, namely the 1-norm and Hausdorff-type medians. We also included the Aumann mean, which is expected to yield the best behavior in absence of contamination, as a reference, and therefore allows to evaluate how much precision is lost when data are completely regular.

Table 2 shows the results of STUDY 1. The smallest value for each choice of ϵ and C_D (each row) and each metric has been highlighted in **bold**.

Table 2. Monte Carlo approximation (MC), and bias, variance (Var) and MSE of the mean, and the 1-norm, Hausdorff-type and $d_{1/3}$ -medians in STUDY 1. The values of the bias, variance and MSE have been multiplied by a factor 100.

ϵ	C_D		Mean	1-norm median	Hausdorff-type median	$d_{\theta=1/3}$ -Median
.0	0	MC	[-.99887, 1.00106]	[-.73865, .74095]	[-.46115, .46531]	[-.67129, .64727]
		Bias	.02 * .03 * .02	.19 * .30 * .15	.81 * .87 * .47	.29 * .35 * .18
		Var	2.54 * 4.83 * 1.67	2.31 * 4.60 * 1.91	2.24 * 4.38 * 1.92	2.03 * 4.01 * 1.69
		MSE	2.54 * 4.83 * 1.67	2.31 * 4.60 * 1.91	2.25 * 4.38 * 1.92	2.03 * 4.01 * 1.69
.1	0	MC	[-1.30527, 1.31453]	[-.84729, .85105]	[-.58653, .59071]	[-.79066, .76522]
		Bias	31.0 * 31.3 * 17.9	10.9 * 11.0 * 6.31	12.5 * 12.5 * 7.23	11.9 * 11.9 * 6.85
		Var	3.66 * 7.21 * 2.62	2.78 * 5.64 * 2.24	2.88 * 5.77 * 2.30	2.51 * 4.99 * 1.97
		MSE	12.6 * 20.7 * 5.82	3.82 * 7.49 * 2.63	4.28 * 8.09 * 2.83	3.76 * 7.08 * 2.44
.1	1	MC	[-.98969, 1.20299]	[-.74342, .8315]	[-.52057, .60258]	[-.69183, .74767]
		Bias	10.6 * 20.2 * 11.9	4.76 * 9.05 * 5.09	9.83 * 13.7 * 6.88	6.04 * 10.0 * 5.30
		Var	3.05 * 5.97 * 2.39	2.62 * 5.41 * 2.24	2.72 * 5.29 * 2.30	2.31 * 4.58 * 1.95
		MSE	4.72 * 9.38 * 3.81	2.99 * 6.07 * 2.50	3.65 * 7.09 * 2.78	2.77 * 5.41 * 2.23
.1	5	MC	[-.40688, 1.40955]	[-.55668, .89602]	[-.23694, .48870]	[-.50468, .75937]
		Bias	50.0 * 59.2 * 50.3	16.9 * 18.2 * 16.9	12.4 * 22.4 * 13.7	13.9 * 16.7 * 14.0
		Var	2.78 * 5.63 * 2.34	2.50 * 4.99 * 2.08	2.12 * 3.93 * 1.98	2.15 * 4.24 * 1.85
		MSE	26.9 * 42.2 * 27.6	4.98 * 9.41 * 4.93	4.02 * 8.52 * 3.85	3.89 * 7.38 * 3.82
.1	10	MC	[-.08976, 1.90527]	[-.55124, .93225]	[-.21933, .49225]	[-.51214, .79510]
		Bias	99.6 * 109 * 99.8	18.9 * 19.1 * 18.9	13.4 * 24.2 * 14.8	15.3 * 15.9 * 15.4
		Var	3.00 * 6.19 * 2.47	2.74 * 5.41 * 2.31	2.22 * 4.16 * 2.03	2.34 * 4.56 * 1.98
		MSE	101 * 131 * 102	5.86 * 10.9 * 5.89	4.42 * 9.47 * 4.22	4.39 * 8.18 * 4.34
.1	100	MC	[9.09925, 10.90346]	[-.56220, .92851]	[-.22206, .49260]	[-.52301, .79934]
		Bias	10 ³ * 1009 * 10 ³	18.2 * 18.8 * 18.2	13.3 * 23.9 * 14.7	15.0 * 15.2 * 15.0
		Var	2.89 * 5.71 * 2.34	2.48 * 5.02 * 2.10	2.04 * 3.97 * 1.91	2.09 * 4.23 * 1.78
		MSE	10 ⁴ * 10270 * 10 ⁴	5.34 * 10.2 * 5.41	4.27 * 9.01 * 4.06	4.05 * 7.70 * 4.04
.2	0	MC	[-1.60316, 1.59337]	[-.98052, .97776]	[-.76109, .74992]	[-.94213, .90299]
		Bias	59.8 * 60.4 * 34.5	23.9 * 24.2 * 13.8	29.2 * 29.3 * 16.9	26.3 * 27.1 * 15.2
		Var	4.49 * 9.08 * 3.50	3.37 * 6.71 * 2.74	3.95 * 7.92 * 3.08	3.07 * 6.19 * 2.46
		MSE	38.9 * 56.5 * 15.4	8.38 * 15.4 * 4.65	11.6 * 20.5 * 5.93	9.28 * 16.4 * 4.78
.2	1	MC	[-1.00005, 1.39407]	[-.75638, .9412]	[-.60366, .75294]	[-.73352, .86056]
		Bias	19.7 * 39.3 * 22.7	10.9 * 20.0 * 11.1	21.5 * 28.8 * 14.4	13.8 * 21.3 * 11.0
		Var	3.74 * 7.66 * 3.35	3.08 * 6.17 * 2.71	3.22 * 6.39 * 2.82	2.74 * 5.43 * 2.40
		MSE	9.61 * 21.1 * 8.48	4.70 * 9.50 * 3.94	7.44 * 14.8 * 4.89	4.71 * 9.59 * 3.61
.2	5	MC	[-.16977, 1.81859]	[-.36245, 1.10632]	[-.00716, .56785]	[-.31925, .91271]
		Bias	99.3 * 117 * 99.8	37.1 * 37.6 * 37.1	27.8 * 45.4 * 29.6	30.9 * 35.2 * 31.0
		Var	3.55 * 6.93 * 3.24	2.87 * 5.81 * 2.44	2.13 * 3.51 * 2.12	2.33 * 4.66 * 2.07
		MSE	101 * 143 * 103	15.8 * 24.7 * 16.2	10.4 * 23.1 * 10.9	11.4 * 18.3 * 11.7
.2	10	MC	[1.19173, 2.80876]	[-.33527, 1.19776]	[.06299, .56154]	[-.31330, .98972]
		Bias	1999 * 2190 * 2002	43.0 * 45.7 * 43.0	31.0 * 52.4 * 33.4	35.0 * 35.8 * 35.0
		Var	3.45 * 6.67 * 3.15	3.29 * 6.58 * 2.71	2.37 * 4.10 * 2.29	2.58 * 5.15 * 2.24
		MSE	402 * 487 * 404	20.6 * 32.4 * 21.2	12.7 * 30.3 * 13.4	14.1 * 22.1 * 14.5
.2	100	MC	[19.20247, 20.79971]	[-.33514, 1.19341]	[.06534, .56219]	[-.33052, 1.01026]
		Bias	1999 * 2020 * 2000	42.8 * 45.2 * 42.8	31.2 * 52.7 * 33.5	35.2 * 36.3 * 35.2
		Var	3.46 * 6.74 * 3.15	3.15 * 6.43 * 2.73	2.24 * 3.80 * 2.21	2.60 * 5.17 * 2.24
		MSE	4.10 ⁴ * 40840 * 4.10 ⁴	20.6 * 31.5 * 21.1	12.8 * 30.1 * 13.5	14.3 * 22.1 * 14.6
.4	0	MC	[-2.19779, 2.20479]	[-1.32187, 1.33136]	[-1.28725, 1.29099]	[-1.34905, 1.31808]
		Bias	120 * 120 * 69.4	58.7 * 59.0 * 33.9	82.6 * 82.6 * 47.7	67.4 * 67.8 * 38.9
		Var	7.23 * 14.8 * 5.82	5.48 * 11.0 * 4.51	6.37 * 12.7 * 4.65	4.75 * 9.52 * 3.65
		MSE	149 * 194 * 53.9	37.8 * 58.9 * 16.0	72.9 * 99.1 * 27.4	48.7 * 68.7 * 18.8
.4	1	MC	[-1.00221, 1.80061]	[-.78725, 1.22979]	[-.78765, 1.14069]	[-.82041, 1.16631]
		Bias	40.1 * 80.0 * 46.1	26.9 * 48.9 * 26.9	50.1 * 67.5 * 33.8	33.4 * 51.9 * 26.7
		Var	5.47 * 10.5 * 5.20	4.67 * 9.71 * 4.26	4.41 * 8.57 * 3.79	3.86 * 7.84 * 3.46
		MSE	27.6 * 70.7 * 26.4	13.5 * 31.3 * 11.5	28.0 * 53.7 * 15.20,	15.0 * 33.2 * 10.6
.4	5	MC	[1.35388, 2.66681]	[.27726, 1.76037]	[.57475, .97644]	[.29873, 1.39288]
		Bias	201 * 235 * 202	102 * 102 * 102	77.4 * 104 * 78.8	85.8 * 97.0 * 86.0
		Var	4.70 * 8.16 * 4.65	5.06 * 10.5 * 4.70	3.45 * 4.24 * 3.47	3.52 * 6.31 * 3.42
		MSE	408 * 559 * 412	107 * 136 * 108	63.3 * 111 * 65.6	76.7 * 101 * 77.4
.4	10	MC	[3.37595, 4.59969]	[.43903, 2.29608]	[.87188, 1.04121]	[.41975, 1.72101]
		Bias	399 * 437 * 399	137 * 156 * 137	95.4 * 133 * 97.9	108 * 109 * 108
		Var	4.57 * 7.89 * 4.48	7.71 * 17.2 * 6.60	3.51 * 3.99 * 3.52	3.85 * 7.53 * 3.61
		MSE	1593 * 1919 * 1598	192 * 267 * 194	94.6 * 181 * 99.4	120 * 144 * 121
.4	100	MC	[39.39936, 40.60287]	[.45825, 2.34548]	[.92426, 1.02835]	[.37239, 1.83676]
		Bias	4000 * 4039 * 4000	140 * 160 * 141	97.4 * 139 * 100	112 * 119 * 112
		Var	4.66 * 8.24 * 4.56	8.05 * 18.5 * 6.93	3.61 * 4.35 * 3.62	4.50 * 8.95 * 3.90
		MSE	1.6.10 ⁵ * 163200 * 1.6.10 ⁵	201 * 284 * 205	98.5 * 196 * 104	128 * 160 * 129

The results for the d_θ -median in this table have been obtained with Vardi and Zhang's modification of Weiszfeld's algorithm. For this setting, the results for the algorithm based on non-linear minimization were almost the same and hence lead to the same conclusions. From the results in Table 2 we can draw the following conclusions:

- Not surprisingly, the Aumann mean shows excellent behavior in absence of contamination, but the spatial median shows similar or even better behavior in terms of variance and MSE.
- As expected, the estimators of all three medians are much less influenced by contamination than the sample (Aumann) mean.
- Which median estimator is the best depends on different aspects, such as the performance criterion (and metric) and the contamination level. Clearly, there is no uniformly best estimator for the location of random intervals as is the case for real-valued data.
- For several cases, the bias, variance or MSE is smallest for the $d_{1/3}$ -median, which shows that the spatial-type median yields good estimates (in this setting, it achieves the lowest MSE values for moderate amounts of contamination, $\epsilon \leq .1$). More importantly, whenever the $d_{1/3}$ -median is not the best choice w.r.t. bias, variance or MSE, it is usually the second best choice with a performance that is close to the optimal estimator. This does not hold for either of the L^1 -type medians and makes the $d_{1/3}$ -median an overall safe choice with either the best performance or a performance close to the best possible choice.

6.2.2. STUDY 2

In STUDY 2 we also include the M-estimator of location for interval-valued data in the analysis to compare the behavior of the d_θ -median to other robust location estimators more generally. We use the M-estimator of location with Hampel's loss function because this loss function has shown good performance in [41], due to its three tuning parameters, which allow much flexibility to handle contamination. Table 3 shows the results of STUDY 2. Again, the smallest value for each choice of ϵ and C_D (each row) and each metric has been highlighted in **bold**. Note that the contamination level $\epsilon = .3$ has also been included in this setting to be able to better compare the behavior of the estimators in presence of higher amounts of contamination. The results for the d_θ -median in this table have been obtained with the algorithm based on non-linear minimization which clearly yielded better results in this setting.

From Table 3 we can see that there is again no uniformly best estimator for the location of random intervals. While the Aumann mean has a small variance, it is highly biased in presence of contamination, as expected. When comparing both bias and MSE, it is very interesting to see that while the Hampel M-estimator of location is the best option for moderate amounts of contamination ($\epsilon \leq .2$), the $d_{1/3}$ -median provides the best results for larger amounts of contamination ($\epsilon \geq .3$). This STUDY thus reveals that while the flexibility of the Hampel M-estimator allows it to yield the best performance for moderate contamination levels, it does pay a price at higher levels of contamination. Indeed the three medians, and especially the $d_{1/3}$ -median yield much better performance in that case.

7. Concluding remarks

In this paper, the d_θ -median, a concept of median for random intervals defined by means of an L^2 metric, has been studied. This notion was inspired by the well-known spatial

Table 3. Monte Carlo approximation (MC), and bias, variance (Var) and MSE of the mean, the 1-norm, Hausdorff-type and $d_{1/3}$ -medians and the M-estimator of location based on Hampel loss function in STUDY 2. The values of the bias, variance and MSE have been multiplied by a factor 1000.

ϵ	c_D		Mean	1-norm median	Hausdorff-type median	$d_{\theta=1/3}$ -Median	Hampel M-estimator
.0	0	MC	[.85441, .85951]	[.88813, .89216]	[.88854, .89178]	[.88815, .89203]	[.85650, .86157]
		Bias	.2 * .2 * .2	.7 * .7 * .7	.6 * .7 * .6	.8 * .8 * .8	.2 * .2 * .2
		Var	.2 * .2 * .2	.2 * .2 * .2	.2 * .2 * .2	.2 * .2 * .2	.1 * .2 * .1
		MSE	.2 * .2 * .2	.2 * .2 * .2	.2 * .2 * .2	.2 * .2 * .2	.1 * .2 * .1
.1	0	MC	[.78246, .78887]	[.87074, .87478]	[.87098, .87457]	[.87074, .87471]	[.85391, .85902]
		Bias	71.5 * 72.1 * 71.5	18.0 * 18.1 * 18.0	18.0 * 18.2 * 18.0	18.1 * 18.2 * 18.1	2.8 * 2.8 * 2.8
		Var	.2 * .2 * .2	.3 * .3 * .3	.3 * .3 * .3	.3 * .3 * .3	.2 * .2 * .2
		MSE	5.3 * 5.4 * 5.3	.6 * .7 * .6	.6 * .6 * .6	.6 * .6 * .6	.2 * .2 * .2
.1	1	MC	[.78312, .78861]	[.87100, .87502]	[.87131, .87473]	[.87102, .87494]	[.85404, .85911]
		Bias	71.3 * 71.5 * 71.3	17.8 * 17.8 * 17.8	17.8 * 17.9 * 17.8	17.9 * 17.9 * 17.9	2.7 * 2.7 * 2.7
		Var	.2 * .2 * .2	.3 * .3 * .3	.3 * .3 * .3	.3 * .3 * .3	.2 * .2 * .2
		MSE	5.3 * 5.3 * 5.3	.6 * .6 * .6	.6 * .6 * .6	.6 * .6 * .6	.2 * .2 * .2
.1	5	MC	[.78358, .78824]	[.87100, .87504]	[.87166, .87441]	[.87104, .87493]	[.85419, .85924]
		Bias	71.2 * 71.5 * 71.2	17.8 * 17.8 * 17.8	17.8 * 18.0 * 17.8	17.9 * 17.9 * 17.9	2.5 * 2.6 * 2.5
		Var	.2 * .2 * .2	.3 * .3 * .3	.3 * .3 * .3	.3 * .3 * .3	.2 * .2 * .2
		MSE	5.2 * 5.3 * 5.2	.6 * .6 * .6	.6 * .6 * .6	.6 * .6 * .6	.2 * .2 * .2
.1	10	MC	[.78344, .78803]	[.87076, .87479]	[.87142, .87415]	[.87079, .87466]	[.85395, .85898]
		Bias	71.4 * 71.7 * 71.4	18.0 * 18.0 * 18.0	18.0 * 18.2 * 18.0	18.1 * 18.1 * 18.1	2.8 * 2.8 * 2.8
		Var	.2 * .2 * .2	.3 * .3 * .3	.3 * .3 * .3	.3 * .3 * .3	.2 * .2 * .2
		MSE	5.3 * 5.3 * 5.3	.6 * .7 * .6	.6 * .6 * .6	.6 * .6 * .6	.2 * .2 * .2
.1	100	MC	[.78340, .78800]	[.87106, .87509]	[.87171, .87445]	[.87110, .87497]	[.85393, .85898]
		Bias	71.5 * 71.7 * 71.5	17.7 * 17.7 * 17.7	17.7 * 18.0 * 17.7	17.8 * 17.8 * 17.8	2.8 * 2.8 * 2.8
		Var	.2 * .2 * .2	.3 * .3 * .3	.3 * .3 * .3	.3 * .3 * .3	.2 * .2 * .2
		MSE	5.3 * 5.3 * 5.3	.6 * .6 * .6	.6 * .6 * .6	.6 * .6 * .6	.2 * .2 * .2
.2	0	MC	[.71025, .71796]	[.84612, .85014]	[.84614, .85015]	[.84610, .85022]	[.83567, .84100]
		Bias	143 * 144 * 143	42.7 * 42.7 * 42.7	42.7 * 43.1 * 42.7	42.7 * 42.8 * 42.7	20.9 * 21.0 * 20.9
		Var	.2 * .2 * .2	.4 * .4 * .4	.4 * .4 * .4	.4 * .4 * .4	.2 * .2 * .2
		MSE	20.6 * 21.0 * 20.6	2.2 * 2.3 * 2.2	2.2 * 2.3 * 2.2	2.2 * 2.3 * 2.2	.7 * .7 * .7
.2	1	MC	[.71113, .71703]	[.84639, .85041]	[.84658, .85023]	[.84639, .85041]	[.83525, .84038]
		Bias	143 * 144 * 143	42.4 * 42.4 * 42.4	42.4 * 42.6 * 42.4	42.5 * 42.5 * 42.5	21.4 * 21.4 * 21.4
		Var	.2 * .2 * .2	.4 * .4 * .4	.4 * .4 * .4	.4 * .4 * .4	.2 * .2 * .2
		MSE	20.6 * 20.8 * 20.6	2.2 * 2.3 * 2.2	2.2 * 2.2 * 2.2	2.2 * 2.2 * 2.2	.7 * .7 * .7
.2	5	MC	[.71202, .71624]	[.84642, .85045]	[.84731, .84958]	[.84645, .85036]	[.83485, .83977]
		Bias	143 * 144 * 143	42.4 * 42.4 * 42.4	42.4 * 42.8 * 42.4	42.4 * 42.5 * 42.4	21.9 * 22.0 * 21.9
		Var	.2 * .2 * .2	.4 * .4 * .4	.4 * .4 * .4	.4 * .4 * .4	.2 * .2 * .2
		MSE	20.6 * 20.7 * 20.6	2.2 * 2.3 * 2.2	2.2 * 2.2 * 2.2	2.2 * 2.2 * 2.2	.7 * .7 * .7
.2	10	MC	[.71259, .71669]	[.84716, .85117]	[.84809, .85028]	[.84720, .85110]	[.83544, .84032]
		Bias	143 * 143 * 143	41.6 * 41.6 * 41.6	41.6 * 42.1 * 41.6	41.7 * 41.7 * 41.7	21.4 * 21.5 * 21.4
		Var	.2 * .2 * .2	.4 * .4 * .4	.4 * .4 * .4	.4 * .4 * .4	.2 * .2 * .2
		MSE	20.5 * 20.6 * 20.5	2.1 * 2.2 * 2.1	2.1 * 2.2 * 2.1	2.1 * 2.2 * 2.1	.7 * .7 * .7
.2	100	MC	[.71227, .71635]	[.84650, .85056]	[.84747, .84965]	[.84656, .85048]	[.83492, .83982]
		Bias	143 * 143 * 143	42.3 * 42.3 * 42.3	42.2 * 42.8 * 42.2	42.3 * 42.3 * 42.3	21.9 * 22.0 * 21.9
		Var	.2 * .2 * .2	.4 * .4 * .4	.4 * .4 * .4	.4 * .4 * .4	.2 * .2 * .2
		MSE	20.6 * 20.7 * 20.6	2.2 * 2.3 * 2.2	2.2 * 2.2 * 2.2	2.3 * 2.4 * 2.3	.7 * .7 * .7
.3	0	MC	[.63827, .64727]	[.80854, .81266]	[.80834, .81287]	[.80858, .81309]	[.72857, .73579]
		Bias	214 * 216 * 214	80.2 * 80.3 * 80.2	80.2 * 80.9 * 80.2	80.0 * 80.3 * 80.0	127 * 128 * 127
		Var	.1 * .2 * .1	.6 * .7 * .6	.6 * .6 * .6	.6 * .6 * .6	1.5 * 1.5 * 1.5
		MSE	46.1 * 47.0 * 46.1	7.1 * 7.2 * 7.1	7.1 * 7.2 * 7.1	7.0 * 7.1 * 7.0	17.6 * 17.9 * 17.6
.3	1	MC	[.63957, .64586]	[.80885, .81296]	[.80897, .81284]	[.80895, .81322]	[.72293, .72868]
		Bias	214 * 215 * 214	79.9 * 80.0 * 79.9	79.9 * 80.2 * 79.9	79.8 * 80.0 * 79.8	133 * 134 * 133
		Var	.2 * .2 * .2	.6 * .7 * .6	.7 * .7 * .7	.6 * .7 * .6	1.4 * 1.4 * 1.4
		MSE	46.1 * 46.4 * 46.1	7.0 * 7.2 * 7.0	7.0 * 7.1 * 7.0	7.0 * 7.1 * 7.0	19.2 * 19.3 * 19.2
.3	5	MC	[.64086, .64463]	[.80887, .81302]	[.81005, .81185]	[.80913, .81319]	[.71858, .72277]
		Bias	214 * 215 * 214	79.9 * 79.9 * 79.9	79.8 * 80.6 * 79.9	79.7 * 79.8 * 79.7	139 * 139 * 139
		Var	.2 * .2 * .2	.6 * .7 * .6	.6 * .6 * .6	.6 * .6 * .6	1.3 * 1.3 * 1.3
		MSE	46.1 * 46.4 * 46.1	7.0 * 7.2 * 7.0	7.0 * 7.1 * 7.0	7.0 * 7.0 * 7.0	20.5 * 20.7 * 20.5
.3	10	MC	[.64088, .64450]	[.80865, .81277]	[.80992, .81149]	[.80889, .81294]	[.71827, .72236]
		Bias	215 * 215 * 215	80.1 * 80.2 * 80.1	80.1 * 80.9 * 80.1	79.9 * 80.0 * 79.9	139 * 139 * 139
		Var	.2 * .2 * .2	.6 * .7 * .6	.6 * .7 * .6	.6 * .6 * .6	1.3 * 1.3 * 1.3
		MSE	46.1 * 46.5 * 46.1	7.1 * 7.2 * 7.1	7.0 * 7.2 * 7.0	7.0 * 7.1 * 7.0	20.6 * 20.8 * 20.6
.3	100	MC	[.64112, .64469]	[.80900, .81314]	[.81030, .81186]	[.80923, .81328]	[.71867, .72272]
		Bias	214 * 215 * 214	79.7 * 79.8 * 79.7	79.7 * 80.5 * 79.7	79.6 * 79.7 * 79.6	139 * 139 * 139
		Var	.2 * .2 * .2	.6 * .7 * .6	.6 * .6 * .6	.6 * .6 * .6	1.3 * 1.3 * 1.3
		MSE	46.1 * 46.4 * 46.1	7.0 * 7.2 * 7.0	7.0 * 7.1 * 7.0	7.0 * 7.0 * 7.0	20.5 * 20.6 * 20.5
.4	0	MC	[.56605, .57635]	[.73780, .74523]	[.73887, .74402]	[.73904, .74557]	[.61304, .62239]
		Bias	286 * 289 * 286	149 * 151 * 149	149 * 150 * 149	149 * 150 * 149	242 * 244 * 242
		Var	.1 * .2 * .1	1.2 * 1.3 * 1.2	1.2 * 1.2 * 1.2	1.2 * 1.3 * 1.2	2.1 * 2.1 * 2.1
		MSE	81.9 * 83.4 * 81.9	23.5 * 24.2 * 23.5	23.5 * 23.8 * 23.5	23.3 * 23.7 * 23.3	60.4 * 61.5 * 60.4
.4	1	MC	[.56820, .57488]	[.73808, .74547]	[.73965, .74376]	[.73945, .74540]	[.61037, .61677]
		Bias	286 * 286 * 286	149 * 151 * 149	149 * 150 * 149	148 * 149 * 148	246 * 246 * 246
		Var	.2 * .2 * .2	1.2 * 1.4 * 1.2	1.2 * 1.2 * 1.2	1.2 * 1.3 * 1.2	1.7 * 1.7 * 1.7
		MSE	81.7 * 82.2 * 81.7	23.4 * 24.2 * 23.4	23.5 * 23.6 * 23.5	23.3 * 23.6 * 23.3	62.0 * 62.4 * 62.0
.4	5	MC	[.56960, .57294]	[.73804, .74544]	[.74098, .74237]	[.73983, .74520]	[.60795, .61148]
		Bias	286 * 287 * 286	149 * 151 * 149	149 * 150 * 149	148 * 149 * 148	250 * 250 * 250
		Var	.2 * .2 * .2	1.2 * 1.3 * 1.2	1.2 * 1.2 * 1.2	1.2 * 1.3 * 1.2	1.3 * 1.3 * 1.3
		MSE	81.9 * 82.4 * 81.9	23.4 * 24.1 * 23.4	23.4 * 23.7 * 23.4	23.2 * 23.5 * 23.2	63.6 * 64.0 * 63.6
.4	10	MC	[.57001, .57314]	[.73853, .74586]	[.74169, .74260]	[.74026, .74557]	[.60766, .61101]
		Bias	286 * 287 * 286	147 * 150 * 149	149 * 150 * 149	148 * 149 * 148	250 * 251 * 250
		Var	.2 * .2 * .2	1.2 * 1.3 * 1.2	1.2 * 1.2 * 1.2	1.2 * 1.2 * 1.2	1.3 * 1.3 * 1.3
		MSE	81.7 * 82.3 * 81.7	23.3 * 24.0 * 23.3	23.3 * 23.6 * 23.3	23.1 * 23.4 * 23.1	63.7 * 64.2 * 63.7
.4	100	MC	[.57008, .57313]	[.73822, .74558]	[.74140, .74226]	[.74004, .74536]	[.60753, .61082]
		Bias	286 * 287 * 286	149 * 151 * 149	149 * 150 * 149	148 * 149 * 148	250 * 251 * 250
		Var	.2 * .2 * .2	1.2 * 1.3 * 1.2	1.2 * 1.2 * 1.2	1.2 * 1.2 * 1.2	1.3 * 1.3 * 1.3
		MSE	81.7 * 82.3 * 81.7	23.4 * 24.1 * 23.4	23.4 * 23.8 * 23.4	23.2 * 23.5 * 23.2	63.8 * 64.3 * 63.8

median for multivariate data. The existence (and uniqueness in most cases) of the d_θ -median has been shown, as well as its robustness by means of its high breakdown point. With respect to its practical computation, there is no explicit solution for the d_θ -median, but two algorithms for its computation have been presented.

The usefulness of the d_θ -median for practical purposes is the same as the benefit of the median with respect to the mean value in the case of univariate data or of multivariate medians, e.g. based on depth, in the case of multivariate data. When there is no contamination, the best option is to work with the mean in order to take advantage of its numerous good statistical and probabilistic properties, especially when we want to use inferential techniques. However, when there are outliers or other deviations in the data, the mean is extremely sensitive to these anomalies and does not reflect well the center of the data distribution anymore. On the other hand, due to its robustness, the d_θ -median remains a reliable measure of the location of the random intervals and thus is much more convenient in this case. A comparison of the (d_θ -)mean and d_θ -median estimates is thus a good start to explore interval-valued data and assess whether there are any potential outliers among the interval data.

A first advantage of the d_θ -median is that the Aumann mean and the d_θ -median yield two estimates of the center of the interval-valued observations based on the same metric, which makes them directly comparable, in the same way as the mean and the median of univariate observations can be compared to each other. Moreover, our empirical results have shown that the d_θ -median can outperform other robust location measures proposed in the literature, such as the L^1 -type medians and the Hampel M-estimator of location in case of a high contamination fraction. More generally, our simulation experiments have revealed that the d_θ -median is a very reliable robust estimator of the location of interval-valued data. No estimator uniformly dominated its competitors, but even in cases where the d_θ -median is not the best estimator its performance is close to that of the optimal estimator.

Directions for further research include extending the d_θ -median to fuzzy number-valued observations, and performing a sensitivity analysis on the choice of the parameter θ . More simulation studies on the comparison of location measures for random intervals could be developed, in order to analyze in more detail the the performance of the different location estimators when varying aspects such as the sample size and the considered distributions. Finally, more advanced methods like median regression could be developed using the d_θ -median.

Acknowledgements

The authors are grateful to the reviewers and Associate Editor handling this paper, as well as to María Ángeles Gil, for their insightful comments and suggestions to improve this paper.

Disclosure

No potential conflict of interest was reported by the authors.

Funding

The research by Sinova was partially supported by/benefited from the Spanish Ministry of Economy and Competitiveness Grant MTM2013-44212-P, the Principality of Asturias/FEDER Grants GRUPIN14-101 and an Ayuda de Movilidad de Excelencia para Docentes e Investigadores de la Universidad de Oviedo. The research by Van Aelst was supported by the Internal Funds KU Leuven under Grants C16-15-068. Their financial support is gratefully acknowledged.

References

- [1] Kokoszka P, Young G. KPSS test for functional time series. *Statistics* 2016;50(5):957–973.
- [2] Ferraro MB, Giordani P. On possibilistic clustering with repulsion constraints for imprecise data. *Inform. Sci.* 2013;245:63–75.
- [3] Ling N, Liu Y, Vieu P. Conditional mode estimation for functional stationary ergodic data with responses missing at random. *Statistics* 2016;50(5):991–1013.
- [4] Bock HH, Diday E. *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information From Complex Data*. Berlin: Springer-Verlag; 2000.
- [5] Billard L, Diday E. From the Statistics of data to the Statistics of knowledge: Symbolic Data Analysis. *J. Am. Stat. Ass.* 2003;98:470–487.
- [6] Gil MA, Lubiano MA, Montenegro M, López-García MT. Least squares fitting of an affine function and strength of association for interval data. *Metrika* 2002;56:97–111.
- [7] Manski CF, Tamer E. Inference on regressions with interval data on a regressor or outcome. *Econometrica* 2002;70:519–546.
- [8] Gil MA, González-Rodríguez G, Colubi A, Montenegro M. Testing linear independence in linear models with interval-valued data. *Comput. Stat. Data Anal.* 2007;51(6):3002–3015.
- [9] Blanco-Fernández A, Corral N, González-Rodríguez G. Estimation of a flexible simple linear model for interval data based on set arithmetic. *Comput. Stat. Data Anal.* 2011;55(9):2568–2578.
- [10] Lima Neto EA, Cordeiro GM, de Carvalho FAT. Bivariate symbolic regression models for interval-valued variables. *J. Stat. Comput. Simul.* 2011;81(11):1727–1744.
- [11] González-Rodríguez G, Colubi A, Gil MA. Fuzzy data treated as functional data: A one-way ANOVA test approach. *Comput. Stat. Data Anal.* 2012;56:943–955.
- [12] Blanco-Fernández A, Colubi A, García-Bárcana M. A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables. *Inform. Sci.* 2013;247:109–122.
- [13] Fagundes RAA, De Souza RMCR, Cysneiros FJA. Robust regression with application to symbolic interval data. *Eng. Appl. Art. Intel.* 2013;26(1):564–573.
- [14] González-Rivera G, Lin W. Constrained regression for interval-valued data. *J. Bus. Econ. Stat.* 2013;31:473–490.
- [15] Montenegro M, Casals MR, Colubi A, Gil MA. Testing 'two-sided' hypothesis about the mean of an interval-valued random set. In: Dubois D, Lubiano MA, Prade H, Gil MA, Grzegorzewski P, Hryniewicz O, editors. *Soft Methods for Handling Variability and Imprecision*. Heidelberg: Springer; 2008. p. 133–139.
- [16] Nakama T, Colubi A, Lubiano MA. Two-way analysis of variance for interval-valued data. In: Borgelt C, González-Rodríguez G, Trutschnig W, Lubiano MA, Gil MA, Grzegorzewski P, Hryniewicz O, editors. *Combining Soft Computing and Statistical Methods in Data Analysis*. Heidelberg: Springer; 2010. p. 475–482.
- [17] De Carvalho FDAT, Brito P, Bock HH. Dynamic clustering for interval data based on L_2 distance. *Comput. Stat.* 2006;21(2):231–250.
- [18] D'Urso P, Giordani P. A robust fuzzy k-means clustering model for interval valued data. *Comput. Stat.* 2006;21(2):251–269.
- [19] Da Costa AFBF, Pimentel BA, De Souza RMCR. Clustering interval data through kernel-induced feature space. *J. Intell. Inf. Syst.* 2013;40(1):101–140.
- [20] D'Urso P, De Giovanni L, Massari R. Trimmed fuzzy clustering for interval-valued data. *Adv. Data Anal. Classif.* 2015;9:21–40.
- [21] D'Urso P, Leski L. Fuzzy C-ordered medoids clustering of interval-valued data. *Pattern Recognit.* 2016;58:9–67.
- [22] D'Urso P, De Giovanni L, Massari R, Cappelli C. Exponential distance-based fuzzy clustering for interval-valued data. *Fuzzy Optim. Decis. Making* 2017;16:51–70.
- [23] D'Urso P, Giordani P. A least squares approach to principal component analysis for interval valued data. *Chemometr. Intell. Lab.* 2004;70(2):179–192.
- [24] Brito P, Duarte Silva AP. Modelling interval data with Normal and Skew-Normal distributions. *J. Appl. Stat.* 2012;39(1):3–20.

- [25] Sun Y, Ralescu DA. A normal hierarchical model and minimum contrast estimation for random intervals. *Ann. Inst. Stat. Math.* 2015;67(2):313–333.
- [26] Aumann RJ. Integrals of set-valued functions. *J. Math. Anal. Appl.* 1965;12:1–12.
- [27] Molchanov I. Statistical problems for random sets. In: Goutsias J, editor. *Random sets: theory and applications*. Heidelberg: Springer; 1997. p. 27–45.
- [28] Weber A. *Über den Standort der Industrien*. Tübingen: Mohr; 1909.
- [29] Gower JC. Algorithm AS 78: The mediancentre. *Appl. Stat.* 1974;23:466–470.
- [30] Brown BM. Statistical uses of the spatial median. *J. Royal Stat. Soc. Ser. B* 1983;45(1):25–30.
- [31] Milasevic P, Ducharme GR. Uniqueness of the spatial median. *Ann. Stat.* 1987;15:1332–1333.
- [32] Roelant E, Van Aelst S. An L_1 -type estimator of multivariate location. *Stat. Meth. & Appl.* 2007;15:381–393.
- [33] Möttönen J, Nordhausen K, Oja H. Asymptotic theory of the spatial median. In: Antoch J, Hušková M, Sen PK, editors. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurečková*. Beachwood, Ohio, USA: Institute of Mathematical Statistics; 2010. p. 182–193.
- [34] Fritz H, Filzmoser P, Croux C. A comparison of algorithms for the multivariate L_1 -median. *Comput. Stat.* 2012;27(3):393–410.
- [35] Zuo Y. Multidimensional medians and uniqueness. *Comput. Stat. Data Anal.* 2013;66:82–88.
- [36] Bertoluzza C, Corral N, Salas A. On a new class of distances between fuzzy numbers. *Math. & Soft Comput.* 1995;2:71–84.
- [37] Trutschnig W, González-Rodríguez G, Colubi A, Gil MA. A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. *Inf. Sci.* 2009;179:3964–3972.
- [38] Vitale RA. L_p metrics for compact, convex sets. *J. Approx. Theory* 1985;45:280–287.
- [39] Sinova B, González-Rodríguez G, Van Aelst S. An alternative approach to the median of a random interval using an L^2 metric. In: Kruse R, Berthold MR, Moewes C, Gil MA, Grzegorzewski P, Hryniewicz O, editors. *Sinergies of Soft Computing and Statistics for Intelligent Data Analysis*. Heidelberg: Springer; 2013. p. 273–281.
- [40] Sinova B, Van Aelst S. On the consistency of a spatial-type interval-valued median for random intervals. *Stat. Probab. Lett.* 2015;100:130–136.
- [41] Sinova B. M-estimators of location for interval-valued random elements. *Chemometr. Intell. Lab.* 2016;156:115–127.
- [42] Sinova B, Gil MA, Van Aelst S. M-estimates of location for the robust central tendency of fuzzy data. *IEEE Trans. Fuzzy Syst.* 2016;24(4):945–956.
- [43] Debruyne M, Hubert M, Van Horebeek J. Detecting influential observations in Kernel PCA. *Comput. Stat. Data Anal.* 2010;54(12):3007–3019.
- [44] Sinova B, Van Aelst S. Comparing the medians of a random interval defined by means of two different L^1 metrics. In: Borgelt C, Gil MA, Sousa JMC, Verleysen M, editors. *Towards Advanced Data Analysis by Combining Soft Computing and Statistics*. Heidelberg: Springer; 2013. p. 75–86.
- [45] Hausdorff F. *Grundzüge der Mengenlehre*. Leipzig: Viet; 1914.
- [46] Fréchet M. Les éléments aléatoires de natures quelconque dans un espace distancié. *Ann. Inst. H. Poincaré.* 1948;10:215–310.
- [47] Cadre B. Convergent estimators for the L_1 -median of a Banach valued random variable. *Statistics* 2001;35:509–521.
- [48] Donoho DL, Huber PJ. The notion of breakdown point. In: Bickel PJ, Doksum K, Hodges Jr JL, editors. *A Festschrift for Erich L. Lehmann*. Belmont: Wadsworth; 1983.
- [49] Lopuhaä HP, Rousseeuw PJ. Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Stat.* 1991;19:229–248.
- [50] Weiszfeld E. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tōhoku Mathematics Journal* 1937;43:355–386.
- [51] Weiszfeld E, Plastria F. On the point for which the sum of the distances to n given points is minimum. *Ann. Oper. Res.* 2009;167:7–41.
- [52] Kuhn HW. A note on Fermat’s problem. *Math. Program.* 1973;4:98–107.
- [53] Chandrasekaran R, Tamir A. Open questions concerning Weiszfeld’s algorithm for the Fermat-Weber location problem. *Math. Program.* 1989;44:293–295.
- [54] Vardi Y, Zhang CH. A modified Weiszfeld algorithm for the Fermat-Weber location problem. *Math. Program.* 2001;90(3):559–566.
- [55] Dennis J, Schnabel R. *Numerical methods for unconstrained optimization and nonlinear equations*. New Jersey: Prentice Hall; 1983.
- [56] Kim JS, Scott CD. Robust kernel density estimation. *J. Mach. Learn. Res.* 2012;13:2529–2565.
- [57] De la Rosa de Saa S, Gil MA, González-Rodríguez G, López MT, Lubiano MA. Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE Trans. Fuzzy Syst.* 2015;23(1):111–126.