

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/303393328>

M-estimators of location for interval-valued random elements

Article in *Chemometrics and Intelligent Laboratory Systems* · May 2016

DOI: 10.1016/j.chemolab.2016.05.003

CITATIONS

2

READS

64

1 author:



Beatriz Sinova

University of Oviedo

39 PUBLICATIONS 353 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Statistical analyses of fuzzy rating scale-based data [View project](#)

M-estimators of location for interval-valued random elements

Beatriz Sinova^{a,*}

^a*Department of Statistics, O.R. and D.M., University of Oviedo, 33007 Oviedo, Spain*

Abstract

Among the new types of data emerging from real-life experiments, interval-valued ones are becoming very prevalent nowadays. In summarizing the location of interval-valued datasets, the Aumann mean is the most usual measure. This measure inherits almost all the nice properties of the mean value for real-valued datasets. Nevertheless, it also inherits a critical property, which is the one related to its high sensitivity to data changes or to the presence of outliers.

As an approach to measure the location of interval-valued datasets in a more robust way, the notion of M-estimators will be considered. Two applications on chemical data will be included to motivate and illustrate the problem. Finally, an empirical comparative study will be conducted to show the performance of the different types of M-estimators proposed in this work.

Keywords: interval-valued data, M-estimator, robustness, location measure, chemical data

1. Introduction and motivation

Interval-valued data often emerge from experiments with random elements associated with many real-life applications. For instance, they can arise as a consequence of either uncertain or incomplete information (e.g., when one can only
5 establish lower and upper bounds for the actual exact measurement/observation). They can also correspond to aggregate information either because of the magnitude of interest being the range or fluctuation of a real-valued attribute over a given time period or collection of individuals (as happens for many symbolic data) or because of confidentiality reasons. Alternatively, data can be essentially
10 interval-valued (like interval-censored ones). See [1], [2], [3], [4] and [5] as examples of different studies involving interval-valued data.

This work aims to analyze the location of interval-valued data robustly. The best known location measure in summarizing interval-valued data is the Aumann mean. This measure, which extends the mean value of a real-valued dataset,

*Corresponding author

Email address: `sinovabeatriz@uniovi.es` (Beatriz Sinova)

15 has been shown to preserve all the valuable properties for real-valued datasets, but it also inherits its lack of robustness.

In this paper, in order to establish robust location measures for interval-valued data, the M-estimation approach (which has shown a successful behavior when dealing with real-valued data) is to be considered. For this purpose, two
20 methodologies will be used:

- on one hand, interval-valued data can be identified through a certain vectorial-valued function with elements of a closed convex cone within a Hilbert space. Consequently, results and methods from Hilbert space-valued Data Analysis can be particularized provided that one guarantees
25 that the process does not move out of the cone;
- on the other hand, when either the preceding methodology fails or one can design exact techniques or supply a better approximation, *ad hoc* concepts and methods can be developed.

As a real-life example motivating this study, we can consider the following
30 one:

Example 1.1. This example has been chosen taking into account the importance of the chemical differentiation of wines in enology. The organic and inorganic compounds that are part of wine depend on many factors, e.g., ground, climate, variety of vine or enological practice. The considered dataset has been
35 published in [6] and consists of 33 Greek red and white commercial wines from the 1998 vintage. Some tables in [6] show numerous mineral, phenol and anthocyanin concentrations in such wines. Each concentration is given in terms of an interval whose mid-point (or center) is the achieved mean value and whose spread (or radius) is the associated standard deviation. Based on this information,
40 Kallithraka *et al.* [6] and D'Urso and Giordani [7] propose a Principal Component Analysis.

For this study, the caffeic acid concentrations have been chosen. Figure 1 represents the intervals corresponding to this phenol concentration for each of the 33 wines $(\omega_1, \dots, \omega_{33})$.

45 Figure 1 clearly shows that wines with codes 1, 6, 9, 10 and 16 can be viewed as 'outliers' because their caffeic acid mean concentration has been much higher (wines with codes 6, 9, 10 and 16) or the corresponding standard deviation has been much larger (wines with codes 1 and 16). The Aumann mean value for the whole dataset equals $\left[\frac{1}{33} \sum_{i=1}^{33} \inf X(\omega_i), \frac{1}{33} \sum_{i=1}^{33} \sup X(\omega_i) \right] = [64.5, 70.2]$.

50 By removing the 5 'outliers' identified above, the Aumann mean value equals $\left[\frac{1}{28} \sum_{j=1}^{28} \inf X(\omega'_j), \frac{1}{28} \sum_{j=1}^{28} \sup X(\omega'_j) \right] = [48.4, 52.7]$, which shows the unwanted effect outliers have on the estimate of the Aumann mean and the clear need for a more robust location measure for interval-valued data.

With the goal of developing M-estimators of location for interval-valued
55 datasets, this paper is structured as follows. In Section 2, the concept of interval-valued data will be defined and the main preliminary and supporting tools related to this notion will be recalled, the distinctive structural characteristics of

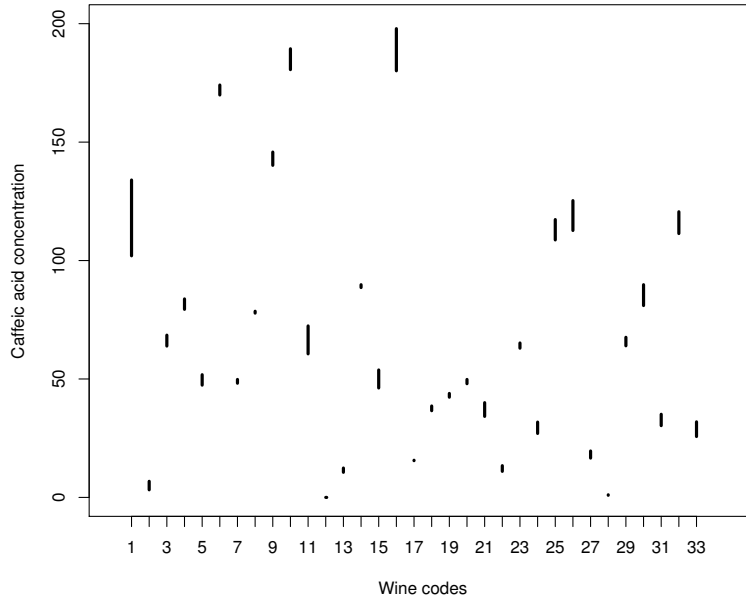


Figure 1: Caffeic acid concentration (mg/l; mean value \pm standard deviation) of Greek red and white wines

the space of study being highlighted. The extension of M-estimators of location for interval-valued data will be dealt with in Section 3. Two real-life examples
60 regarding chemical data will be detailed and analyzed in Section 4. In Section 5, an empirical analysis will be carried out to compare the robustness of the introduced location measures and to check how they improve the outputs with respect to the Aumann mean. Finally, some concluding remarks will be included in Section 6. The proofs of the theoretical results can be checked in
65 the Appendix.

2. Preliminary tools and supporting results

Let $\mathcal{K}_c(\mathbb{R})$ denote the space of nonempty compact intervals. The elements in $\mathcal{K}_c(\mathbb{R})$ will be referred to herein as *interval-valued data*.

Elements in $\mathcal{K}_c(\mathbb{R})$ can be uniquely characterized, among others, by the
70 vectorial *mid/spr representation* $\boldsymbol{\eta}$. It captures their trend and imprecision associating with each interval $K \in \mathcal{K}_c(\mathbb{R})$ the vector value $\boldsymbol{\eta}(K) = (\text{mid } K, \text{spr } K) \in \mathbb{R} \times [0, \infty)$, being $\text{mid } K = \text{mid-point/center of } K$ and $\text{spr } K = \text{spread/radius of } K$. Another commonly used characterization is the *inf/sup representation* $\boldsymbol{\iota}$ given by the vector value $\boldsymbol{\iota}(K) = (\text{inf } K, \text{sup } K) \in \{(y, z) \in \mathbb{R}^2 : y \leq z\}$ for each
75 $K \in \mathcal{K}_c(\mathbb{R})$.

Remark 2.1. Although interval-valued data can be characterized as vectorial data through η , ι or other representations, the purpose is not to directly apply multivariate but Hilbert space-valued robust statistical ideas and results whenever it is possible. This is due to the fact that multivariate robust developments concerning M-estimation are usually based on some assumptions on the joint distribution of the involved one-dimensional random components, these assumptions either not making sense or not being acceptable, realistic or general enough for the mid and spr of an interval-valued random element. In this way, the meaning of the random components of the η or ι characterizations of an interval-valued random element and the constraints between them would be mostly ignored in case they are treated as simple components of a random vector.

2.1. Arithmetic with interval-valued data

In developing statistics with interval-valued data, one of the key tools is given by the arithmetic. The elementary operations to be specified are the sum and the multiplication by scalars of interval values, and the most natural way to proceed consists of defining these operations as the image sets of the involved interval values through the corresponding operation (see [8]). Thus,

Definition 2.1. Let $K = [a, b], K' = [a', b'] \in \mathcal{K}_c(\mathbb{R})$. The *Minkowski sum* of K and K' is defined as the interval value

$$K + K' = [a + a', b + b'].$$

Definition 2.2. Let $K = [a, b] \in \mathcal{K}_c(\mathbb{R})$ and $\gamma \in \mathbb{R}$. The *product of K by the scalar γ* is defined as the interval value

$$\gamma \cdot K = \begin{cases} [\gamma a, \gamma b] & \text{if } \gamma \geq 0 \\ [\gamma b, \gamma a] & \text{otherwise.} \end{cases}$$

Remark 2.2. Although a difference operator could be defined by means of $K - K' = K + (-1) \cdot K'$, unlike what happens in the real-valued case, the interval $K - K$ would not coincide in general with the neutral element of the sum, $[0, 0]$. As a consequence, one can conclude that the space $(\mathcal{K}_c(\mathbb{R}), +)$ is a commutative semigroup, although not a group, and $(\mathcal{K}_c(\mathbb{R}), +, \cdot)$ is not a linear but a semilinear space (in fact, a closed convex cone). The semilinearity of the space will be a crucial feature when trying to develop statistical procedures.

2.2. Metrics between interval-valued data

In addition to $(\mathcal{K}_c(\mathbb{R}), +, \cdot)$ not being linear, one can easily verify that there is no ‘difference operation’ between interval values that is simultaneously well-defined and preserves the main properties of the difference between real values in connection with the sum. This drawback, which clearly differentiates real-
from interval-valued cases, can be overcome to some extent by incorporating suitable distances between values in $\mathcal{K}_c(\mathbb{R})$ as another basic tool.

Definition 2.3. [9] Let $K, K' \in \mathcal{K}_c(\mathbb{R})$. The **Hausdorff distance** between K and K' is given by

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

Definition 2.4. [10] Let $K, K' \in \mathcal{K}_c(\mathbb{R})$. The **1-norm distance** between K and K' is given by

$$\delta_1(K, K') = \frac{1}{2} |\inf K - \inf K'| + \frac{1}{2} |\sup K - \sup K'|.$$

Definition 2.5. [11, 12] Let $K, K' \in \mathcal{K}_c(\mathbb{R})$ and $\theta > 0$. The **d_θ -distance** between K and K' is given by

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

In particular, the d_1 metric equals the **δ_2 -distance**, which has been introduced by Vitale [10] and formerly given as

$$\delta_2(K, K') = \sqrt{\frac{(\inf K - \inf K')^2 + (\sup K - \sup K')^2}{2}}.$$

The value of θ weighs the relative importance assessed to deviations in imprecision in contrast to deviations in trends. Based on the equivalences in [12],
 110 the most common choices correspond to $\theta = 1/3$ (which allocates a uniform relevance to all the points in each interval) and $\theta = 1$ (which allocates a uniform relevance to the extreme points in each interval).

Following ideas similar to those in [13], we can identify interval- with some vectorial-valued data *via* the $\boldsymbol{\eta}$ representation and the d_θ metric. This supports
 115 what has been said in Remark 2.1 about the particularization of methods for the robust analysis of Hilbert space-valued random elements whenever the outputs remain within the cone $\mathbb{R} \times [0, \infty)$.

Remark 2.3. All these metrics are strongly equivalent. It should be emphasized that although an isometrical embedding can be stated through $\boldsymbol{\eta}$ (or $\boldsymbol{\iota}$)
 120 when $\mathcal{K}_c(\mathbb{R})$ is endowed with the d_H (or δ_1) metric, the final space the cone is included in is not a Hilbert but a Banach space.

2.3. Interval-valued random elements

Mathematical modeling is another essential tool in developing data analysis, and data are usually assumed to come from the repeated performance of a
 125 random mechanism. Fréchet [14] anticipated that future mathematics would have to incorporate new and unexpected sorts of objects quite beyond numbers and vectors, so he introduced random elements taking on values in metric spaces. In accordance with their current usage, a random element is defined to be a measurable function between a sample space and a metric space equipped with
 130 its Borel σ -algebra.

Compact random intervals (see [15]) determine a well-stated and supported model for the random mechanisms generating interval-valued data within the probabilistic setting. They integrate both randomness and imprecision, so that the first one affects the generation of experimental data, whereas the second
135 affects the nature of the experimental data which, for formal purposes, are assumed to be intrinsically interval-valued.

Definition 2.6. *Given a probability space (Ω, \mathcal{A}, P) , a mapping $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ is said to be a **compact random interval** associated with it if X is measurable with respect to \mathcal{A} and the Borel σ -algebra generated by the topology induced by
140 the Hausdorff metric on $\mathcal{K}_c(\mathbb{R})$ (or any of the topological equivalent metrics). Equivalently, a compact random interval X can be expressed as $X = [\inf X, \sup X]$, with $\inf X$ and $\sup X$ being random variables such that $\inf X \leq \sup X$.*

As a consequence from the Borel measurability, concepts such as the (induced) distribution of a (compact) random interval, some of its relevant parameters, as well as the stochastic independence of (compact) random intervals, can
145 be immediately obtained.

In this respect, if one attempts to summarize the distribution of a compact random interval, one can think about formalizing a location measure as follows:

Definition 2.7. [16] *Let (Ω, \mathcal{A}, P) be a probability space and $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ be an associated compact random interval such that $E(\inf X)$ and $E(\sup X)$ exist. The **Aumann mean** of X is the interval value $E[X] \in \mathcal{K}_c(\mathbb{R})$ such that $E[X] = [E(\inf X), E(\sup X)]$. Equivalently,*

$$\eta(E[X]) = E(\eta(X)) = (E(\text{mid } X), E(\text{spr } X)).$$

The Aumann mean satisfies several valuable properties similar to those in
150 the classical case (see for reviews [17], and also [13] in a more general setting), it is coherent with the usual interval arithmetic and it satisfies Strong Laws of Large Numbers (see, in the more general setting, [18]).

3. M-estimators of location for compact random intervals

M-estimators of location for real-valued data were introduced by Huber [19],
155 with the aim of limiting the influence of outliers in approaches like least squares. The crucial idea was the replacement of the squared errors by a (generally less rapidly increasing) loss function of the data and the parameter estimate. For this reason, M-estimators are considered intermediaries between the sample mean
160 and median, both of them being included as special cases. The extension of this idea to the interval setting can be stated as follows:

Definition 3.1. *Consider a metric space $(\mathcal{K}_c(\mathbb{R}), d)$ and let (Ω, \mathcal{A}, P) be a probability space, $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ be an associated compact random interval and ρ*

be a continuous loss function. The **M-location measure**, if it exists, is the interval $\mathbb{K}_P^M[\mathbb{X}] \in \mathcal{K}_c(\mathbb{R})$ defined as follows

$$\mathbb{K}_P^M[\mathbb{X}] = \arg \min_{\mathbb{K} \in \mathcal{K}_c(\mathbb{R})} J_P(\mathbb{K}) = \arg \min_{\mathbb{K} \in \mathcal{K}_c(\mathbb{R})} E[\rho(d(\mathbb{X}, \mathbb{K}))].$$

Remark 3.1. The function ρ in Definition 3.1 is a generalization of the square loss function in the least squares approach. Indeed, it is usually assumed to vanish at 0 and be non-decreasing. The continuity of the loss function ρ is commonly assumed (see, e.g., [20] or [21]).

Definition 3.2. Consider $(\mathcal{K}_c(\mathbb{R}), d)$ a metric space. Let (Ω, \mathcal{A}, P) be a Polish probability space, \mathbb{X} be an associated compact random interval $\mathbb{X} : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, $(\mathbb{X}_1, \dots, \mathbb{X}_n)$ be a simple random sample from \mathbb{X} and ρ be a continuous loss function. The **M-estimator of location** is the $\mathcal{K}_c(\mathbb{R})$ -valued statistic $\widehat{\mathbb{K}}^M[\mathbb{X}]_n$, if it exists, given by

$$\widehat{\mathbb{K}}^M[\mathbb{X}]_n = \arg \min_{\mathbb{K} \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^n \rho(d(\mathbb{X}_i, \mathbb{K})).$$

In particular, given a sample of observations from \mathbb{X} , $\mathbf{x}_n = (x_1, \dots, x_n)$, the (sample) **M-estimate of location** is the interval $\widehat{\mathbb{K}}^M[\mathbf{x}_n] \in \mathcal{K}_c(\mathbb{R})$ given by

$$\widehat{\mathbb{K}}^M[\mathbf{x}_n] = \arg \min_{\mathbb{K} \in \mathcal{K}_c(\mathbb{R})} J_n(\mathbb{K}) = \arg \min_{\mathbb{K} \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^n \rho(d(x_i, \mathbb{K})).$$

Proposition 3.1. The M-estimator of location is well-defined for the metric space $(\mathcal{K}_c(\mathbb{R}), d)$ with $d \in \{d_\theta, d_H, \delta_1\}$.

An important property of M-estimators of location for compact random intervals, the strong consistency, is now to be analyzed. The following result presents some suitable choices of ρ to guarantee the strong consistency, including some well-known loss functions like the ones proposed by Huber [22]

$$\rho_a(x) = \begin{cases} x^2/2 & \text{if } |x| \leq a \\ a(|x| - a/2) & \text{otherwise,} \end{cases}$$

with $a > 0$ a tuning parameter, and Hampel [23]

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

where the nonnegative parameters $a < b < c$ allow us to control the degree of suppression of the outliers. The smaller their values, the greater this degree.

170 **Theorem 3.2.** Consider the metric space $(\mathcal{K}_c(\mathbb{R}), d)$ with $d \in \{d_\theta, d_H, \delta_1\}$. Let X be a compact random interval associated with a Polish probability space (Ω, \mathcal{A}, P) . Under any of the following assumptions:

- ρ is non-decreasing, subadditive and unbounded,
- ρ has linear upper and lower bounds with the same slope (including Huber loss function),
- 175 • ρ is bounded (including Hampel loss function),

and whenever the M -location measure exists and is unique, the sample M -estimator of location is a strongly consistent estimator of the M -location measure, i.e.,

$$\lim_{n \rightarrow \infty} d(\widehat{\mathcal{K}^M[X]}_n, \mathcal{K}_P^M[X]) = 0 \quad \text{a.s. } [P].$$

The measurability and the consistency of the M -estimators of a compact random interval have been proven assuming their existence. In the next subsections, we will restrict the study to some loss functions for which the existence of M -estimators can be certainly guaranteed. Two different methodologies will be presented: the adaptation of already existing ideas and results in the literature for Hilbert space-valued data (Subsection 3.1) and some *ad hoc* developments for a natural choice of the loss function that does not fulfill the conditions required for the first methodology (Subsection 3.2).

185 3.1. M -estimators of location defined through the Representer Theorem

In this subsection, Kim and Scott's studies [24, 25] in the Hilbert space framework will be adapted, following the first methodology, to deal with interval-valued data. Their ideas have been developed for reproducing kernel Hilbert spaces, in the context of robust nonparametric density estimation, combining a traditional kernel density estimator with ideas from classical M -estimation through the interpretation of this estimator based on a radial, positive semi-definite kernel as a sample mean.

To lower the sensitivity of the sample mean to outliers, Kim and Scott suggest that we estimate it robustly *via* M -estimators yielding a robust kernel density estimator. Although Kim and Scott generalized their results to other Hilbert spaces, they have never moved out of the setting of kernel density estimation. However, these ideas and proofs can be easily adapted to general Hilbert spaces and, therefore, applicable to the interval-valued case through the isometrical embedding of $(\mathcal{K}_c(\mathbb{R}), d_\theta)$ into the cone $\mathbb{R} \times [0, \infty)$, with the corresponding metric induced by the isometry $\boldsymbol{\eta}$. In this section, such adaptation is to be shown and completed with several properties Kim and Scott have not examined.

The most crucial result in [25] for the particularization of M -estimators to the interval-valued case is the Representer Theorem, since it guarantees that the M -estimator of location associated with a compact random interval will remain in the corresponding parameter space, $\mathcal{K}_c(\mathbb{R})$. The Representer Theorem states that the M -estimate of location can be expressed as a convex linear combination of the sample components. Since the sample observations obviously belong to the cone and the linear combination is a closed operator, the implication follows.

210 **Theorem 3.3.** [Adapted from [25]] Consider the metric space $(\mathcal{K}_c(\mathbb{R}), d_\theta)$. Let (Ω, \mathcal{A}, P) be a probability space, $\mathbf{X} : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ be a compact random interval and $\mathbf{x}_n = (x_1, \dots, x_n)$ be a sample of independent observations from \mathbf{X} . Under the assumptions

R1. ρ is non-decreasing (increasing in a neighborhood of 0) and differentiable,
 215 $\rho(0) = 0$ and $\lim_{x \rightarrow 0} \rho(x)/x = 0$,

R2. Let $\phi(x) = \rho'(x)/x$, and assume the convention $\phi(0) = \lim_{x \rightarrow 0} \phi(x)$, where $\phi(0)$ is supposed to exist and be finite,

the $\mathcal{K}_c(\mathbb{R})$ -valued sample M-estimate of location exists and can be expressed as

$$\widehat{\mathcal{K}^M[\mathbf{x}_n]} = \sum_{i=1}^n w_i \cdot x_i,$$

where $w_i \geq 0$ and $\sum_{i=1}^n w_i = 1$. Furthermore, $w_i \propto \phi(d_\theta(x_i, \widehat{\mathcal{K}^M[\mathbf{x}_n]}))$. Indeed, under the additional assumption

220 R3. J_n is strictly convex (for which sufficient conditions are given by either

- ρ is strictly convex, or
- ρ is convex, strictly increasing, $n \geq 3$ and $\mathbf{A} = (\langle \boldsymbol{\eta}(x_i), \boldsymbol{\eta}(x_j) \rangle_\theta)_{i,j=1}^n$ is positive definite, with $\langle \boldsymbol{\eta}(x_i), \boldsymbol{\eta}(x_j) \rangle_\theta = \text{mid } x_i \cdot \text{mid } x_j + \theta \cdot \text{spr } x_i \cdot \text{spr } x_j$;

the necessary conditions

225 i) $\mathcal{K}^M[\mathbf{x}_n] = \sum_{i=1}^n w_i \cdot x_i$,

ii) $w_i \propto \phi(d_\theta(x_i, \widehat{\mathcal{K}^M[\mathbf{x}_n]}))$,

iii) $\sum_{i=1}^n w_i = 1$

are also sufficient for $\widehat{\mathcal{K}^M[\mathbf{x}_n]}$ to be the minimizer of $J_n(\mathbf{K}) = \frac{1}{n} \sum_{i=1}^n \rho(d_\theta(x_i, \mathbf{K}))$.

230 **Remark 3.2.** Notice that these conditions on ρ are not uncommon. Apart from ρ being non-decreasing and vanishing at 0, M-estimators of location with differentiable loss functions are usually chosen. On the other hand, $\rho'(0) = 0$ is a consequence of ρ' being odd (see [21]) and, in most interesting cases, $\phi(0)$ exists and is finite.

235 Notice that Huber's and Hampel's loss functions satisfy R1-R2 (and Huber may also fulfill R3).

The existence and uniqueness of the sample M-estimate of location for a sample of independent observations obtained from a compact random interval are guaranteed under assumptions R1 – R3 in the Representer Theorem. However, despite the fact that we can express the M-estimate as a weighted mean of
 240 the observations, such weights also depend on the M-estimate, so the problem does not have an explicit solution in general. To deal with these implicit formulas in practice, Kim and Scott [25] proposed an iterative algorithm that is an extension of the iteratively re-weighted least squares algorithm used in classical M-estimation (see [19]). Its notation will be now adapted to the interval setting.

245 *Step 1.* Take the initial weights $w_i^{(0)} \in \mathbb{R}$, for $i \in \{1, \dots, n\}$, such that $w_i^{(0)} \geq 0$ and $\sum_{i=1}^n w_i^{(0)} = 1$, and fix a tolerance ε .

Step 2. Generate a sequence $\{\mathbf{K}_{(k)}^M\}_{k \in \mathbb{N}}$ by iterating on the following procedure:

$$\mathbf{K}_{(k)}^M = \sum_{i=1}^n w_i^{(k-1)} \cdot \mathbf{x}_i, \quad w_i^{(k)} = \frac{\phi(d_\theta(\mathbf{x}_i, \mathbf{K}_{(k)}^M))}{\sum_{j=1}^n \phi(d_\theta(\mathbf{x}_j, \mathbf{K}_{(k)}^M))}.$$

Step 3. Terminate the algorithm when

$$\frac{|J_n(\mathbf{K}_{(k+1)}^M) - J_n(\mathbf{K}_{(k)}^M)|}{J_n(\mathbf{K}_{(k)}^M)} < \varepsilon.$$

The following property of M-estimators obtained under the Representer Theorem can be proven.

Proposition 3.4. Consider the metric space $(\mathcal{K}_c(\mathbb{R}), d_\theta)$. Let $\mathbf{x}_n = (x_1, \dots, x_n)$ denote a sample of independent observations from a compact random interval $\mathbf{X} : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ on a probability space (Ω, \mathcal{A}, P) . Moreover, let ρ be a continuous loss function fulfilling the assumptions R1-R3.

If $\mathbf{K} \in \mathcal{K}_c(\mathbb{R})$, then it holds that the M-estimator is translation equivariant, i.e.,

$$\mathbf{K}^M[\widehat{\mathbf{x}_n + \mathbf{K}}] = \widehat{\mathbf{K}^M[\mathbf{x}_n]} + \mathbf{K}.$$

Moreover, if ρ satisfies that $\phi(k \cdot x) \propto \phi(x)$ for any $k \in (0, \infty)$ (i.e., the coefficient of proportionality is independent of x), then for all $\gamma \in \mathbb{R}$ the M-estimator is also scale equivariant, i.e.,

$$\mathbf{K}^M[\widehat{\gamma \cdot \mathbf{x}_n}] = \gamma \cdot \widehat{\mathbf{K}^M[\mathbf{x}_n]}.$$

In general, M-estimators of location are not scale equivariant, since most loss functions do not satisfy the proportionality assumption in Proposition 3.4. Therefore, the measurement units will have to be chosen carefully.

In order to measure the robustness, the *finite sample breakdown point* has been considered. Although other tools like the influence function could be very informative, the lack of realistic and easy-to-use models for the distribution of compact random intervals makes such an approach non-viable at present. The intuitive idea behind the concept of finite sample breakdown point (denoted by fsbp) is that it represents the minimum proportion of observations that should be perturbed enough to make the corresponding estimate be arbitrarily far from the estimate obtained for the original non-contaminated sample. The mathematical formalization by Donoho and Huber [26] can be adapted to cover estimators taking values in general metric spaces as follows. The finite sample breakdown point of a general M-estimator of location $\widehat{\mathbf{K}^M[\mathbf{X}]_n}$ corresponding to a sample \mathbf{x}_n of size n from a compact random interval $\mathbf{X} : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ with $(\mathcal{K}_c(\mathbb{R}), d)$ a metric space fulfilling that $\sup_{K, K' \in \mathcal{K}_c(\mathbb{R})} d(K, K') = \infty$ is given by

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

where $\mathbf{y}_{n,k}$ denotes the sample obtained from \mathbf{x}_n by perturbing up to k of its elements. The following result shows an upper bound for the fsbp of M-estimators of location under the Representer Theorem and gives the exact value of their fsbp under some additional conditions on the loss function.

260 **Theorem 3.5.** *Consider the metric space $(\mathcal{K}_c(\mathbb{R}), d_\theta)$ and let X be a compact random interval associated with a Polish probability space (Ω, \mathcal{A}, P) . Moreover, let ρ be a continuous loss function satisfying the assumptions R1-R2 in Theorem 3.3, such that the corresponding sample M-estimator of location is unique. Then, the fsbp of the corresponding M-estimator of location is at most $\lfloor (n+1)/2 \rfloor / n$,
265 where $\lfloor \cdot \rfloor$ denotes the floor function. Furthermore, under any of the additional assumptions:*

- ρ admits linear upper and lower bounds with the same slope (e.g. Huber loss function)
- ρ has a finite upper bound C^* (e.g. Hampel loss function) and satisfies

$$\rho \left(\max_{1 \leq i, j \leq n} d_\theta(x_i, x_j) \right) < \frac{n - 2 \lfloor \frac{n-1}{2} \rfloor}{n - \lfloor \frac{n-1}{2} \rfloor - 1} \cdot C^*,$$

then the fsbp is exactly equal to $\lfloor (n+1)/2 \rfloor / n$.

270 The sufficient conditions over the loss function allowing us to guarantee the existence of sample M-estimates of location as well as their expression as convex linear combinations of the sample elements are fulfilled for different interesting choices of the loss function ρ , as it has just been verified. However, there are some other interesting choices of ρ for which such conditions fail and *ad hoc* developments should be considered.
275

In this respect, if one chooses $\rho(x) = |x|$, one cannot apply the results in this subsection. In Subsection 3.2, some *ad hoc* developments in the literature to get the (exact) M-estimates associated with this choice are to be recalled.

3.2. Ad-hoc M-estimators of location

280 In this subsection, a particular M-estimator of location on $\mathcal{K}_c(\mathbb{R})$ will be reviewed: the median. In this case $\rho(x) = |x|$, and we will consider the three distances introduced in Section 2.2: δ_1 and d_H as L^1 -type metrics and d_θ as L^2 -type.

285 For the two L^1 -type extensions of the concept of median to the interval-valued setting, the idea behind them is similar, the properties they fulfill mostly coincide and the way of operating with them is analogous. A detailed study of such concepts can be found in [27], [28], and also [29, 30] for the more general setting of random fuzzy numbers.

Definition 3.3. *Given a probability space (Ω, \mathcal{A}, P) and an associated compact random interval X , the δ_1 -median(s) (or 1-norm median) and the d_H -median(s) (or Hausdorff-type median) of X are the intervals*

$$\text{Me}[X] = \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} E(\delta_1(X, K)) \quad \text{and} \quad \text{M}[X] = \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} E(d_H(X, K)),$$

respectively, whenever the involved expectations exist.

290 The following result guarantees that at least one value for each of these medians always exists and it is rather easy to compute.

Theorem 3.6. *Given a probability space (Ω, \mathcal{A}, P) and an associated compact random interval X , the intervals $\text{Me}[X], M[X] \in \mathcal{K}_c(\mathbb{R})$ such that*

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

(which corresponds to the median suggested by De Souza and De Carvalho [31]) and

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

(which corresponds to the median suggested by De Carvalho et al. [32]) are a δ_1 -median and a d_H -median of X , respectively. In case any of the medians of the involved real-valued random variables is non-unique, the most usual convention of choosing the mid-point of the interval of medians is assumed to be considered.

From now on, these solutions will be the ones considered when mentioning the δ_1 - and d_H -medians to avoid an unnecessary cumbersome checking and to ease the study of their properties. Analogously, given a simple random sample (X_1, \dots, X_n) from a compact random interval X associated with a probability space (Ω, \mathcal{A}, P) , the **sample δ_1 -median** and the **sample d_H -median** will be respectively given by

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

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

following the same convention as in Theorem 3.6.

Remark 3.3. It should be pointed out that the use of the convention in Theorem 3.6 has a different mission when applied to the δ_1 -median and the d_H -median. In the first case, if we do not consider some valid conventions, the result can fail and not determine an interval. For example, consider the compact random interval X taking on the values $x_1 = [0, 2]$ and $x_2 = [3, 4]$, both with induced probabilities $P(X = x_1) = P(X = x_2) = 0.5$. Then, we have that $\text{Me}(\inf X)$ is any value in $[0, 3]$, whereas $\text{Me}(\sup X)$ is any value in $[2, 4]$, so a choice of the kind $\text{Me}(\inf X) = 2.5$ and $\text{Me}(\sup X) = 2$ would lead to an empty interval. On the other hand, any possible choice for the medians $\text{Me}(\text{mid } X)$ and $\text{Me}(\text{spr } X)$ would lead to a valid solution of d_H -median, so the convention in Theorem 3.6 serves to provide uniqueness.

Remark 3.4. In contrast to the median of random variables, both the δ_1 -median and the d_H -median do not necessarily correspond to one of the values of the compact random interval, even when the convention is not needed. Consider the compact random interval Y taking on values $y_1 = [0, 2]$, $y_2 = [2, 3]$ and $y_3 = [1, 5]$, all of them with probability $1/3$. Then, $\text{Me}[Y] = [1, 3]$ and $M[Y] = [1.5, 3.5]$.

In addition to the use of L^1 -type metrics, it should be noted that a well-known generalization of the median of real-valued random variables to the multivariate settings is the spatial median or mediancenter (see, for example, [33] or [34]), which is based on an L^2 -type metric.

Inspired by the spatial median as extension of the median to higher dimensional Euclidean spaces and even Banach spaces (see [35]), Sinova *et al.* [36] introduced the population and sample d_θ -medians on $\mathcal{K}_c(\mathbb{R})$.

Definition 3.4. *Given a probability space (Ω, \mathcal{A}, P) and an associated compact random interval $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, the **d_θ -median(s)** of X is (are) the interval value(s) $M_\theta[X] \in \mathcal{K}_c(\mathbb{R})$ such that*

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

whenever the involved expectation exists.

Definition 3.5. *Given a probability space (Ω, \mathcal{A}, P) , an associated compact random interval $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, and a simple random sample (X_1, \dots, X_n) from X , the **sample d_θ -median(s)** of X is (are) the interval-valued statistic(s)*

$$\begin{aligned} \widehat{M}_\theta[X]_n &= \arg \min_{K \in \mathcal{K}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^n d_\theta(X_i, K) \\ &= \arg \min_{(y,z) \in \mathbb{R} \times [0, \infty)} \frac{1}{n} \sum_{i=1}^n \sqrt{(\text{mid } X_i - y)^2 + \theta \cdot (\text{spr } X_i - z)^2}, \end{aligned}$$

where the target values of K , y and z depend in fact on (X_1, \dots, X_n) (although, for the sake of simplicity, this has been omitted from the notation).

The existence and uniqueness of the sample d_θ -median has been discussed in [37] and the strong consistency in [38].

Theorem 3.7. *Given a simple random sample (X_1, \dots, X_n) from a compact random interval $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$, the corresponding sample d_θ -median always exists and, moreover, it 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.*

In contrast to the L^1 -type medians for random intervals, the solution for the d_θ -median is implicit. In order to approximate it, a similar algorithm to the one proposed for M-estimators under the Representer Theorem (see Section 3.1) can be considered.

Regarding the inferential behavior of these medians, all of them are strongly consistent whenever the corresponding population medians $\text{Me}[X]$, $M[X]$ and $M_\theta[X]$ exist and are actually unique (i.e., they are unique without applying the convention in Theorem 3.6), and their finite sample breakdown point equals the value $\lfloor (n+1)/2 \rfloor / n$.

340 Indeed, the sample mean has the lowest possible breakdown point while the
sample median can withstand up to 50% of contamination. In consequence, the
definition of these three medians succeeds in inheriting the robustness properties
of the real-valued sample median.

4. Illustrative real-life examples

345 4.1. Greek wine data

In this subsection, we will come back to Example 1.1 introduced in Section 1.
It has already been shown that the Aumann mean was influenced by the atypical
values concerning the caffeic acid concentrations in wine. Therefore, the robust
location measures for random intervals proposed/recalled along this work will be
350 now computed to avoid the sensitivity to outliers when summarizing the central
tendency of such a phenol concentration. The (rounded) obtained estimates for
the complete dataset (1) are the following:

$$\begin{aligned} E^{(1)}[X] &= [64.47, 70.17], \\ (K_{\text{Huber}}^M)^{(1)}[X] &= [48.45, 52.73], \quad (K_{\text{Hampel}}^M)^{(1)}[X] = [47.25, 51.02], \\ \text{Me}^{(1)}[X] &= [48.2, 53.8], \quad M^{(1)}[X] = [47.8, 52.2], \quad M_{1/3}^{(1)}[X] = [48.07, 52.46]. \end{aligned}$$

The M-estimates of location have been computed by means of both Huber
and Hampel loss functions when the parameter θ involved in the d_θ metric
355 has been assumed to be equal to $1/3$. The choice of the values for the tuning
parameters in the Huber ($a = 1.345$) and Hampel ($a = 2, b = 4, c = 8$) loss
functions is common in the literature (see e.g. [39] and [40]).

By removing the most extreme values (corresponding to the wines with codes
1, 6, 9, 10 and 16), dataset (2), the Aumann mean is not so far from the
360 remaining location estimates:

$$\begin{aligned} E^{(2)}[X] &= [48.36, 52.66], \\ (K_{\text{Huber}}^M)^{(2)}[X] &= [46.77, 50.46], \quad (K_{\text{Hampel}}^M)^{(2)}[X] = [47.25, 51.02], \\ \text{Me}^{(2)}[X] &= [46.8, 49.8], \quad M^{(2)}[X] = [47.15, 50.75], \quad M_{1/3}^{(2)}[X] = [47.65, 50.10]. \end{aligned}$$

It can be checked that, although outliers make all the measures have a larger
mid-point, the Aumann mean value is the estimate which has been the most
perturbed by them:

$$\begin{aligned} d_{1/3}(E^{(1)}[X], E^{(2)}[X]) &= 16.80, \quad d_{1/3}\left((K_{\text{Huber}}^M)^{(1)}[X], (K_{\text{Huber}}^M)^{(2)}[X]\right) = 1.99, \\ d_{1/3}\left((K_{\text{Hampel}}^M)^{(1)}[X], (K_{\text{Hampel}}^M)^{(2)}[X]\right) &= 2.77 \cdot 10^{-4}, \\ d_{1/3}(\text{Me}^{(1)}[X], \text{Me}^{(2)}[X]) &= 2.80, \quad d_{1/3}(M^{(1)}[X], M^{(2)}[X]) = 1.08, \end{aligned}$$

$$d_{1/3}(M_{1/3}^{(1)}[X], M_{1/3}^{(2)}[X]) = 1.50.$$

The difference between the original estimate and the estimate after the removal of the outliers is clearly smaller than the one for the Aumann mean in all cases.

4.2. Water quality data

365 The second real-life example deals with water quality data. As Ongley [41] commented, detailed (real-valued) data in problems related with in-stream calculations or environmental quality issues that require temporal extrapolation (e.g. chemical loads) are of limited practical use, although they contain much information about the variability. Therefore, in practice many chemical pa-
 370 rameters are represented by the whole range of individual records (that is to say, interval-valued data) observed daily, such as ammonia, chloride, nitrate, phosphorus or suspended sediment concentrations, specific conductance and turbidity (see, for example, [42]). Among them, pH and dissolved oxygen are frequently measured when controlling the water quality, as in [43].

375 The considered dataset contains the interval-valued information (that is, the minima and maxima values achieved daily) related to the dissolved oxygen (mg/L) and the pH (standard units) measured in Delaware River at Chester (agency USGS, site 1477050, Pennsylvania, United States) from May 15, 2014 till June 15, 2014. A complete description, dataset and more detailed information can be found in the webpage of the United States Geological Survey
 380 (USGS) <http://waterdata.usgs.gov>. Collected data are shown in Table 1 and Figures 2 and 3.

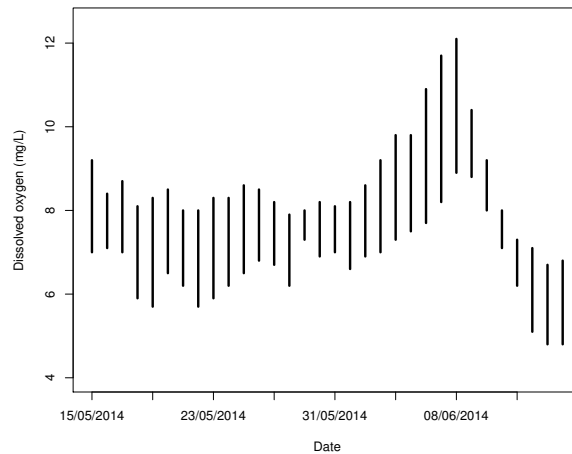


Figure 2: Dissolved oxygen (mg/L) measured in Delaware River at Chester

The mid-points/centers and spreads/radius of all the interval-valued location measures defined/recalled before have been computed and included in Table 2.
 385 The only difference with respect to the calculations in the previous example of

Table 1: Interval-valued information related to the dissolved oxygen (mg/L) and the pH (standard units) measured in Delaware River at Chester from May 15, 2014 till June 15, 2014

Date	O ₂	pH	Date	O ₂	pH
15/05/2014	[7.0, 9.2]	[6.9, 7.2]	31/05/2014	[7.0, 8.1]	[7.0, 7.1]
16/05/2014	[7.1, 8.4]	[7.0, 7.2]	01/06/2014	[6.6, 8.2]	[7.0, 7.1]
17/05/2014	[7.0, 8.7]	[6.9, 7.1]	02/06/2014	[6.9, 8.6]	[7.0, 7.2]
18/05/2014	[5.9, 8.1]	[6.9, 7.1]	03/06/2014	[7.0, 9.2]	[7.0, 7.3]
19/05/2014	[5.7, 8.3]	[6.9, 7.1]	04/06/2014	[7.3, 9.8]	[7.0, 7.4]
20/05/2014	[6.5, 8.5]	[7.0, 7.1]	05/06/2014	[7.5, 9.8]	[7.1, 7.3]
21/05/2014	[6.2, 8.0]	[6.9, 7.1]	06/06/2014	[7.7, 10.9]	[7.1, 7.6]
22/05/2014	[5.7, 8.0]	[6.9, 7.1]	07/06/2014	[8.2, 11.7]	[7.1, 7.9]
23/05/2014	[5.9, 8.3]	[6.8, 7.0]	08/06/2014	[8.9, 12.1]	[7.3, 8.2]
24/05/2014	[6.2, 8.3]	[6.8, 7.0]	09/06/2014	[8.8, 10.4]	[7.2, 7.6]
25/05/2014	[6.5, 8.6]	[6.8, 7.1]	10/06/2014	[8.0, 9.2]	[7.2, 7.3]
26/05/2014	[6.8, 8.5]	[6.8, 7.0]	11/06/2014	[7.1, 8.0]	[7.1, 7.2]
27/05/2014	[6.7, 8.2]	[6.8, 7.1]	12/06/2014	[6.2, 7.3]	[7.0, 7.2]
28/05/2014	[6.2, 7.9]	[6.8, 7.1]	13/06/2014	[5.1, 7.1]	[7.0, 7.1]
29/05/2014	[7.3, 8.0]	[7.0, 7.1]	14/06/2014	[4.8, 6.7]	[7.0, 7.1]
30/05/2014	[6.9, 8.2]	[7.0, 7.2]	15/06/2014	[4.8, 6.8]	[6.9, 7.1]

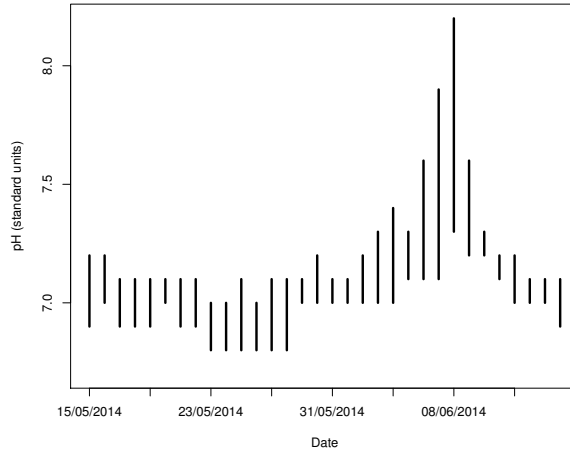


Figure 3: pH (standard units) measured in Delaware River at Chester

Greek wine data is that, due to the small distances between the observations, tuning parameters a , b and c have been fixed following Kim and Scott's suggestions (see [25] for more details). Notice that the values $a = 1.345$ (for Huber) and $a = 2$, $b = 4$ and $c = 8$ (Hampel) are too large to distinguish between

Table 2: Mid-points and spreads of the Aumann mean, Huber and Hampel M-estimators, the 1-norm, Hausdorff-type and $d_{\theta=1/3}$ medians of the dissolved oxygen and pH values in Delaware River

		$E[X]$	$K_{\text{Huber}}^M[X]$	$K_{\text{Hampel}}^M[X]$	$\text{Me}[X]$	$M[X]$	$M_{1/3}[X]$
Oxygen	Mid	7.697	7.551	7.487	7.575	7.550	7.551
	Spr	0.963	0.904	0.877	0.725	0.975	0.843
pH	Mid	7.102	7.054	7.030	7.050	7.050	7.053
	Spr	0.127	0.098	0.093	0.050	0.100	0.083

390 ‘normal’ and ‘atypical’ distances between observations in this case, since all the distances would be classified as ‘normal’ and these two M-estimators would coincide with the Aumann mean. On the other hand, Kim and Scott show how to fix the parameters depending on the sample distances. The values used in this example have been 0.558 (Huber), $a = 0.558$, $b = 1.046$ and $c = 1.787$ (Hampel)
 395 for the dissolved oxygen data, and 0.058 (Huber), $a = 0.058$, $b = 0.153$ and $c = 0.183$ (Hampel) for the pH data.

Outliers are not so clearly identified when looking at the dissolved oxygen data, but we can check in Table 2 how the three observations with larger spread and a bit higher mid-point (June, 6 to June, 8 2014) have more influence on the
 400 computation of (both the mid and spr of) the Aumann mean than for the rest of location measures, which present a more robust behavior.

The pH values measured on June, 7 and 8 2014 are clearly identified as outliers taking into account their large spread. This is reflected in the larger spread of the Aumann mean in contrast to the spread of the M-estimators.
 405 Notice that all the spreads are quite small, so the difference in spread between the Aumann mean and the M-estimators is proportionally large.

5. Comparative empirical studies

A comparative simulation is set up to check the robustness of all the M-estimators proposed in this work when dealing with interval-valued data. Four
 410 different studies have been conducted, varying the sample size (small sample size $n = 100$ or large sample size $n = 10000$) and the distributions (the original distribution being symmetric or asymmetric).

For each of the four studies, the comparisons have concerned the following location estimates: Aumann mean, Huber and Hampel M-estimates using the d_{θ}
 415 distance, δ_1 -median, d_H -median and d_{θ} -median, with $\theta \in \{1/3, 1\}$. When $n = 100$, the medians and trimmed means based on the well-known halfspace and simplicial depths (see [44, 45] and [46], respectively) have also been considered. The reason not to include them in the comparisons when $n = 10000$ is that the computation of depths is usually hard when the sample size increases.

For each of the location measures/estimates \mathbf{L} , the Monte Carlo approximation of the estimate, $\mathbf{L}[\mathbf{X}]$, has been computed. This information has been completed with the Monte Carlo approximation of the bias $d(\mathbf{E}[\widehat{\mathbf{L}[\mathbf{X}]_n}], \mathbf{L}[\mathbf{X}])$, the variance $E(d^2(\widehat{\mathbf{L}[\mathbf{X}]_n}, \mathbf{E}[\widehat{\mathbf{L}[\mathbf{X}]_n}]))$ and the mean square error of the estimate $E(d^2(\widehat{\mathbf{L}[\mathbf{X}]_n}, \mathbf{L}[\mathbf{X}]))$ for $d \in \{\delta_1, d_H, d_{1/3}, \delta_2\}$, which have been obtained through the formulas

$$d\left(\frac{\sum_{i=1}^N \widehat{\mathbf{L}[\mathbf{X}]_N^i}}{N}, \mathbf{L}[\mathbf{X}]\right), \frac{1}{N} \sum_{i=1}^N d^2\left(\widehat{\mathbf{L}[\mathbf{X}]_N^i}, \frac{\sum_{i=1}^N \widehat{\mathbf{L}[\mathbf{X}]_N^i}}{N}\right), \frac{1}{N} \sum_{i=1}^N d^2\left(\widehat{\mathbf{L}[\mathbf{X}]_N^i}, \mathbf{L}[\mathbf{X}]\right)$$

420 respectively, where N denotes the number of Monte Carlo iterations and $\widehat{\mathbf{L}[\mathbf{X}]_N^i}$ represents the sample estimate of the i^{th} sample, $i \in \{1, \dots, N\}$.

The general scheme of the four studies has been as follows:

Step 1. A sample of n interval-valued data has been simulated from a compact random interval \mathbf{X} for each of several different situations in such a way that

425

- to generate the interval-valued data, we have considered two real-valued random variables as follows: $X = [X_1 - X_2, X_1 + X_2]$, with $X_1 = \text{mid } X$ and $X_2 = \text{spr } X$ or, alternatively, two order real-valued statistics $X_{(1)}$ and $X_{(2)}$ such that $\mathbf{X} = [X_{(1)}, X_{(2)}]$, i.e., $X_{(1)} = \inf \mathbf{X}$, $X_{(2)} = \sup \mathbf{X}$;

430

- each sample is assumed to be split into a subsample of size $n(1 - c_p)$, where c_p denotes the proportion of contamination and is supposed to range in $\{0, 0.1, 0.2, 0.4\}$, associated with a non-contaminated distribution and a subsample of size $n \cdot c_p$ associated with a contaminated one. C_D plays an additional contamination role, measuring how far the distribution of the contaminated subsample is from the distribution of the non-contaminated one (and ranges in $\{0, 1, 5, 10, 100\}$);

435

- 16 situations for different values of c_p and C_D have been considered and for each of them some cases have been selected, namely, ones in which the random variables X_i (or $X_{(i)}$) are independent (CASES 1 and 3) and others in which they are dependent (CASES 2 and 4).

440

Step 2. $N = 1000$ replications of *Step 1* have been considered for the situation $c_p = C_D = 0$ in order to approximate the population measures by using a Monte Carlo approach.

445

Step 3. $N = 1000$ replications of *Step 1* have been considered for all the situations (c_p, C_D) and the approximated estimates, bias, variance and mean square error have been computed for each location measure.

The following sample sizes and distributions have been used:

Study 1 $n = 100$ and the assumptions for CASE 1 are

- 450
- $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 \sim \chi_1^2$ for the non-contaminated subsample,
 - $X_1 \sim \mathcal{N}(0, 3) + C_D$ and/or $X_2 \sim \chi_4^2 + C_D$ for the contaminated subsample,

whereas CASE 2 assumes that

- 455
- $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 \sim 1/(X_1^2 + 1)^2 + \sqrt{\chi_1^2}$ for the non-contaminated subsample,
 - $X_1 \sim \mathcal{N}(0, 3) + C_D$ and/or $X_2 \sim 1/(X_1^2 + 1)^2 + \sqrt{\chi_1^2} + C_D$ for the contaminated subsample,

being χ_1^2 independent from X_1 and, obviously, $1/(X_1^2 + 1)^2$.

Study 2 $n = 10000$ and same CASES 1 and 2 than in Study 1.

460 **Study 3** $n = 100$ and we vary the distributions considered in Study 1, in order to bound the generated intervals, as commonly happens in real-life applications related to surveys or ratings, and manage asymmetric contamination. CASE 3 assumes that

- 465
- $X_{(1)}, X_{(2)} \sim \text{Beta}(5, 1)$ (they are simply chosen at random and ordered) for the non-contaminated subsample,
 - $X_{(1)}, X_{(2)} \sim \text{Beta}(1, C_D + 1)$ for the contaminated subsample,

whereas CASE 4 assumes that

- 470
- $X_1 \sim \text{Beta}(5, 1)$ and $X_2 \sim \text{Uniform}[0, \min\{X_1, 1 - X_1\}]$ for the non-contaminated subsample,
 - $X_1 \sim \text{Beta}(1, C_D + 1)$ and/or $X_2 \sim \min\{X_1, 1 - X_1\} \cdot \text{Beta}(1, C_D + 1)$ for the contaminated subsample.

Study 4 $n = 10000$ and the same CASES 3 and 4 as in Study 3.

To avoid excessive information in the outputs of the simulations, the details about bias, variance and mean square error can be found in the supplementary material. The conclusions for the four studies have been summarized in Table 3. As a measure of the variation of the choice of the optimal location estimate, the index of qualitative variation (IQV) has been computed for each case as follows (see [47]):

$$\text{IQV} = \frac{k}{k-1} \left(1 - \sum_{i=1}^k f_i^2 \right),$$

475 where k is the number of different categories (i.e., the number of different location estimates considered in each case: $k = 13$ in Studies 1 and 3 and $k = 9$ in Studies 2 and 4) and f_i represents the relative frequency of the i -th category. Therefore, the IQV is equal to 1 for uniformly distributed variables and, on the other hand, coincides with 0 when the variable only takes one value (and, consequently, the relative frequency of the corresponding category is 1).

Table 3: The most suitable location measures/estimates in Studies 1-4. In parenthesis, the optimal value of θ (if any) in most of the corresponding situations and varying the distance used in the computation of the bias, variance and mean square error

	STUDY 1		STUDY 2		STUDY 3		STUDY 4		
	CASE 1	CASE 2	CASE 1	CASE 2	CASE 3	CASE 4	CASE 3	CASE 4	
Bias	$c_p \leq 0.2$	Hampel	Hampel (1)	Hampel	Hampel (1)	Hampel (1/3)	Liu	Hampel (1/3)	Hampel
	$c_p = 0.4$	Hampel (1)	Hampel (1)	Hampel (1)	Hampel (1)	Hampel	Hampel (1)	Hampel	Hampel
	$c_D = 0$	Hampel (1)	δ_1 -median	Hampel (1)	δ_1 -median	Hampel (1/3)	Liu Hampel	Hampel (1/3)	Hampel
	IQV	0.5411	0.5916	0.5400	0.5081	0.3467	0.7758	0.3600	0.7069
Variance	$c_p = 0$	Huber (1)	δ_1 -median	Huber (1)	δ_1 -median	Huber	Hampel (1)	Huber	Hampel (1)
	$c_p \leq 0.2$	Hampel (1) Huber (1)	δ_1 -median	Hampel (1)	Hampel δ_1 -median	Hampel (1)	Hampel (1)	Hampel (1)	Hampel Hausdorff
	$c_p = 0.4$	Hampel (1)	Hampel (1)	Hampel	Hampel (1)	Hampel (1)	Hampel (1/3)	δ_1 -median Hampel	Hampel
	$c_D = 0$	Hampel (1)	δ_1 -median	Hampel (1)	δ_1 -median	Hampel (1)	Hampel (1)	Hampel	Hampel
IQV	0.4618	0.4988	0.5026	0.6652	0.2201	0.3888	0.6443	0.7300	
MSE	$c_p = 0$	Huber (1)	δ_1 -median	Huber (1)	δ_1 -median	Huber	Hampel (1)	Huber	Hampel (1)
	$c_p \leq 0.2$	Hampel (1)	δ_1 -median	Hampel	Hampel (1)	Hampel (1/3)	Hampel (1)	Hampel (1/3)	Hampel
	$c_p = 0.4$	Hampel (1)	Hampel (1)	Hampel (1)	Hampel (1)	Hampel	Hampel (1/3)	Hampel	Hampel
	$c_D = 0$	Hampel (1)	δ_1 -median	Hampel (1)	δ_1 -median	Hampel	Hampel (1)	Hampel (1/3)	Hampel
IQV	0.4761	0.5226	0.6064	0.5609	0.5020	0.3301	0.4510	0.7251	

On the basis of the conclusions gathered in Table 3, one can realize that there
480 is no uniformly most appropriate location estimate and that the outputs seem
to depend more on the considered non-contaminated and contaminated distributions
than on the sample size. However, it is beyond doubt that the Hampel
M-estimator of location usually behaves better than the other location measures
in relation to either the bias, the variance or the mean square error. It is also
485 interesting to highlight that the δ_1 -median is also chosen in several situations
(mostly for smaller contamination proportions) and recall that its computation
is even more straightforward and fast than that for the M-estimators based on
the Representer Theorem. Finally, regarding the qualitative variation, it should
be remarked that, in most of the analyzed situations, the index IQV increases
490 when a larger sample size is considered.

6. Concluding remarks

In this work, several robust alternatives to summarize the location of a compact
random interval have been proposed. Although some (*ad hoc*) extensions
of the concept of median to the interval-valued case had already been stated in
495 the literature, the more general approach of M-estimators of location (of which
such medians are a particular case) has been adopted. M-estimators of location
for compact random intervals have been proven to inherit the main properties
that make them a successful choice in the real settings. Apart from studying
the measurability and consistency of general M-estimators, a particular class of
500 M-estimators - those defined under the Representer Theorem - has been analyzed
in detail. Their existence and uniqueness, their translational and scale
equivariance, their algorithmic solution and their robustness in terms of the finite
sample breakdown point have been shown. Finally, all these measures have
been applied to two real-life examples and have been empirically compared by
505 means of some simulations.

It should be highlighted that the search of robust location measures to summarize
the central tendency of compact random intervals is crucial, in order to avoid
that outliers, atypical observations, errors or data changes invalidate the
statistical conclusions obtained from the analysis of interval-valued data.
510 Nowadays, most of the existing methodology developed to treat such data is
based on the Aumann mean as location measure. Due to its high sensitivity, it
is reasonable to predict that the Aumann mean could be replaced in these techniques
by other location measures with a more robust behavior, like the ones presented
in this work, in order to guarantee the validity of the corresponding
515 statistical conclusions.

When analyzing contaminated datasets, two different approaches could be
followed. The goal of this paper is to remove the outliers' influence on the
outputs and to study the behavior of the majority of the data. Nevertheless,
the opposite view of focusing on the outlier identification also offers an interesting
520 research topic. In that sense, once an internal cross-validation were employed
to evaluate how the removals of individual samples and/or groups of samples
would affect the outputs in Table 2, those results could be reconsidered as an

automated outlier detection tool: the observations that have a large impact on the computation of the Aumann mean, but a small influence on the output of the M-estimator (of any kind), would be classified as outliers.

Finally, some open problems will be now enumerated. First, the robustness has been stated in terms of the finite sample breakdown point due to the lack of realistic parametric families of distributions in the space of compact intervals. Therefore, the use of influence functions has been left as open problem. Secondly, notice that the development of robust scale measures will be required for the extension of the M-estimators of location with unknown dispersion, which present convenient properties like the scale equivariance and are generally used in multivariate spaces. To conclude, hypothesis testing procedures about the proposed location measures could be established and analyzed.

Acknowledgements

The author is grateful to the reviewers and Associate Editor handling this paper, as well as to Prof. M. Á. Gil and Prof. S. Van Aelst, for their insightful comments and suggestions. This research has been partially supported by the Spanish Ministry of Economy and Competitiveness Grant MTM2013-44212-P, the Principality of Asturias/FEDER Grants GRUPIN14-101, Research Contract E-33-2015-0040746, and the Ayuda de Movilidad de Excelencia para Docentes e Investigadores de la Universidad de Oviedo 2015, funded by Banco Santander through Campus de Excelencia Internacional.

Appendix A: Proofs

Proof of Proposition 3.1

It will be shown that the M-estimator of location is a Borel measurable function and, therefore, it is well-defined. The following facts are taken into account:

- $X \equiv \Omega$ is separable and complete since (Ω, \mathcal{A}, P) is a Polish probability space. Recall that either $\boldsymbol{\eta}$ or $\boldsymbol{\nu}$ embed the metric spaces $(\mathcal{K}_c(\mathbb{R}), d_\theta)$, $(\mathcal{K}_c(\mathbb{R}), d_H)$ and $(\mathcal{K}_c(\mathbb{R}), \delta_1)$ isometrically into a cone, U , of a Hilbert (when d_θ is used) or a Banach space (when either d_H or δ_1 are chosen). In case d_θ or d_H are used, $U = \mathbb{R} \times [0, \infty)$, and when δ_1 is chosen, $U = \{(y, z) \in \mathbb{R}^2 : y \leq z\}$. Since all the distances defined in \mathbb{R}^2 by means of such isometries are strongly equivalent among them and w.r.t. the Euclidean distance, they all induce the same topologies, U always being a closed subset. Thus, the completeness of the Hilbert or Banach space is inherited by the cone. On the other hand, the isometrical embedding preserves the separability of the metric spaces $(\mathcal{K}_c(\mathbb{R}), d_\theta)$, $(\mathcal{K}_c(\mathbb{R}), d_H)$ and $(\mathcal{K}_c(\mathbb{R}), \delta_1)$, what has already been commented in Remark 2.3. Consider $Y \equiv U$. Then Y is a complete separable metric space.

- $f : \Omega \times U \rightarrow \mathbb{R}$ such that $f(\omega, (y, z)) = \frac{1}{n} \sum_{i=1}^n \rho(d(X_i(\omega), [y - z, y + z]))$ or alternatively $f(\omega, (y, z)) = \frac{1}{n} \sum_{i=1}^n \rho(d(X_i(\omega), [y, z]))$, depending on the expression of U , is a real-valued measurable function irrespective of the distance $d \in \{d_\theta, d_H, \delta_1\}$.

565

First, $D = \Omega \times U$ is a Borel set, since it is the universal set we are considering and, therefore, it belongs to any σ -algebra. Secondly, the measurability of f derives from the following facts:

- the addition, product and composition of Borel measurable functions are also Borel measurable;
- $\text{mid } X_i, \text{spr } X_i, \text{inf } X_i$ and $\text{sup } X_i, i = 1, \dots, n$, are Borel measurable because of the equivalent definition of compact random intervals. Thus, the mappings $f_i : \Omega \times U \rightarrow U$ with $f_i(\omega, (y, z))$ representing $(\text{mid } X_i(\omega), \text{spr } X_i(\omega))$ or $(\text{inf } X_i(\omega), \text{sup } X_i(\omega))$ are also Borel measurable with respect to the σ -algebra on the product space $\Omega \times U$ and the Borel σ -algebra generated by the topology induced by the corresponding norm on the cone of the Banach (Hilbert for d_θ) space U .
- $f_0 : \Omega \times U \rightarrow U$ such that $f(\omega, (y, z)) = (y, z)$ is also Borel measurable following an analogous reasoning to that for f_i , since the identity is measurable;
- the square root, the square function, the absolute value function and ρ are continuous and, therefore, Borel measurable;
- U is σ -compact, since a locally compact Polish space is equivalent to a locally compact σ -compact metric space.
- For each metric $d \in \{d_\theta, d_H, \delta_1\}$ and each $\omega \in \Omega$, the following function is continuous (and, therefore, lower semi-continuous) with respect to the relative topology on U

570

575

580

585

$$f_\omega : \begin{array}{ccc} U & \longrightarrow & \mathbb{R} \\ (y, z) & \longmapsto & f(\omega, (y, z)). \end{array}$$

For each $\omega \in \Omega$, let (y^*, z^*) be any element of U . It is sufficient to check that, given any sequence $\{(y_n, z_n)\}_{n \in \mathbb{N}} \subset U$ such that $(y_n, z_n) \xrightarrow[n]{\rightarrow} (y^*, z^*)$, i.e.,

$$d([y_n - z_n, y_n + z_n], [y^* - z^*, y^* + z^*]) \xrightarrow[n]{\rightarrow} 0,$$

then $\lim_n f_\omega((y_n, z_n)) = f_\omega((y^*, z^*))$.

For any $n \in \mathbb{N}$ and $i = 1, \dots, n$, the triangular inequality allows us to ensure that $d(X_i(\omega), [y^* - z^*, y^* + z^*]) - d([y_n - z_n, y_n + z_n], [y^* - z^*, y^* + z^*])$

$$\begin{aligned} &\leq d(X_i(\omega), [y_n - z_n, y_n + z_n]) \\ &\leq d(X_i(\omega), [y^* - z^*, y^* + z^*]) + d([y_n - z_n, y_n + z_n], [y^* - z^*, y^* + z^*]). \end{aligned}$$

Since $d([y_n - z_n, y_n + z_n], [y^* - z^*, y^* + z^*]) \xrightarrow{n} 0$, we then conclude that $d(\mathbf{X}_i(\omega), [y_n - z_n, y_n + z_n]) \xrightarrow{n} d(\mathbf{X}_i(\omega), [y^* - z^*, y^* + z^*])$. By the continuity of ρ , $\rho(d(\mathbf{X}_i(\omega), [y_n - z_n, y_n + z_n])) \xrightarrow{n} \rho(d(\mathbf{X}_i(\omega), [y^* - z^*, y^* + z^*]))$ and, obviously, the result follows because the continuity is preserved by the addition and the product of continuous functions.

Applying the result in [48], it is possible to guarantee that the function φ below is Borel-measurable, where

$$\begin{aligned} \varphi : I &\longrightarrow U \\ \omega &\longmapsto (y_0, z_0) \text{ s.t. } f(\omega, (y_0, z_0)) = \inf_{(y,z) \in U} f(\omega, (y, z)) \end{aligned}$$

with $I = \left\{ \omega \in \Omega : \text{for some } (y_0, z_0) \in U, f(\omega, (y_0, z_0)) = \inf_{(y,z) \in U} f(\omega, (y, z)) \right\}$.

Notice that $\varphi(\omega) = (y_0, z_0)$ such that

$$\frac{1}{n} \sum_{i=1}^n \rho(d(\mathbf{X}_i(\omega), [y_0 - z_0, y_0 + z_0])) = \inf_{(y,z) \in U} \frac{1}{n} \sum_{i=1}^n \rho(d(\mathbf{X}_i(\omega), [y - z, y + z])),$$

or, of course, U formalized alternatively when δ_1 is chosen, i.e., φ is the function that assigns to each $\omega \in \Omega$ for which the M-estimate of location exists, the corresponding value of the M-estimate of location.

Sketch of the proof of Theorem 3.2

These assumptions are sufficient to ensure Huber's conditions for consistency (see [49]). Notice that the locally compactness and the second countability required for the space $\mathbb{R} \times [0, \infty)$ in [38] involve d_θ . However, these topological properties are also fulfilled when using d_H or δ_1 because all these metrics are strongly equivalent. The proof follows the same scheme as in [38] (where the loss function was $\rho(x) = |x|$), taking into account that

- When ρ is subadditive apart from non-decreasing, it is possible to preserve the relations and the way of reasoning with distances.
- When ρ is the Huber loss function ρ_a , for any $a > 0$, it can be bounded by means of two lines with the same slope for all $x > 0$

$$ax - a^2/2 \leq \rho_a(x) \leq ax.$$

If ρ has linear upper and lower bounds with the same slope, then it is also possible to preserve the relations and the way of reasoning with distances.

- When ρ is the Hampel loss function $\rho_{a,b,c}$, for any $c > b > a > 0$, it is possible to bound it as follows

$$0 \leq \rho_{a,b,c}(x) \leq a(b + c - a)/2.$$

610 When ρ is bounded, it is also necessary to notice that there is a compact set C in the parameter space such that the sequence of M-estimators of location almost ultimately stays in C , thanks to the parameter space being locally compact and Hausdorff and the Strong Laws of Large Numbers.

Sketch of the proof of Theorem 3.5

615 M-estimators defined through the Representer Theorem are translation equivariant by Proposition 3.4. Therefore, the proof of the upper bound can be carefully extended from the real-valued settings taking care of the semilinearity of $\mathcal{K}_c(\mathbb{R})$. It is possible to prove the equality when ρ has linear upper and lower bounds, following a reasoning like in [37, 50], or when ρ has a finite upper bound C^* satisfying $\rho(\max_{1 \leq i, j \leq n} d_\theta(x_i, x_j)) < \frac{n-2\lfloor \frac{n-1}{2} \rfloor}{n-\lfloor \frac{n-1}{2} \rfloor-1} \cdot C^*$, checking
620 that the M-estimator will always belong to a compact set unless the perturbed sample contains more than $\lfloor (n+1)/2 \rfloor/n$ observations.

Appendix B: Supplementary data

625 Supplementary data to this article, concerning the bias, variance and mean square error of the location measures considered for the simulation study in Section 5, can be found online at <http://dx.doi.org/10.1016/j.chemolab.2016.05.003>

References

- 630 [1] L. Billard, E. Diday, From the statistics of data to the statistics of knowledge: Symbolic data analysis, *J. Amer. Stat. Assoc.* 98 (462) (2003) 470–487.
- [2] P. Giordani, H. A. L. Kiers, Three-way component analysis of interval-valued data, *J. Chemometr.* 18 (5) (2015) 253–264.
- [3] B. Sen, G. Xu, Model based bootstrap methods for interval censored data, *Comput. Stat. Data Anal.* 81 (2015) 121–129.
- 635 [4] Y. Xu, K. F. Lam, B. J. Cowling, Y. B. Cheung, Estimation of intervention effect using paired interval-censored data with clumping below lower detection limit, *Stat. Med.* 34 (2) (2015) 307–316.
- [5] C. Capelli, P. D’Urso, F. D. Lovio, Regime change analysis of interval-valued time series with an application to PM10, *Chemometr. Intell. Lab. Syst.* 146 (2015) 337–346.
640
- [6] S. Kallithraka, I. Arvanitoyannis, P. Kefalas, A. El-Zajouli, E. Soufleros, E. Psarra, Instrumental and sensory analysis of greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin, *Food Chem.* 73 (2001) 501–514.

- 645 [7] P. D'Urso, P. Giordani, A least squares approach to principal component analysis for interval valued data, *Chemometr. Intell. Lab. Syst.* 70 (2004) 179–192.
- [8] H. Minkowski, Volumen und oberfläche, *Mathematische Annalen* 57 (1903) 447–495.
- 650 [9] F. Hausdorff, *Grundzuege der mengenlehre*, Viet, Leipzig, 1914.
- [10] R. A. Vitale, L_p metrics for compact, convex sets, *J. Approx. Theory* 45 (1985) 280–287.
- [11] C. Bertoluzza, N. Corral, A. Salas, On a new class of distances between fuzzy numbers, *Mathware Soft Comput.* 2 (1995) 71–84.
- 655 [12] M. A. Gil, M. A. Lubiano, M. Montenegro, M. T. López-García, Least squares fitting of an affine function and strength of association for interval-valued data, *Metrika* 56 (2002) 97–111.
- [13] G. González-Rodríguez, A. Colubi, M. A. Gil, Fuzzy data treated as functional data: A one-way ANOVA test approach, *Comput. Stat. Data Anal.* 56 (4) (2012) 943–955.
- 660 [14] M. Fréchet, Les Éléments aléatoires de nature quelconque dans un espace distancié, *Ann. Inst. Henri Poincaré* 10 (1948) 215–310.
- [15] G. Matheron, *Random sets and integral geometry*, J. Wiley & Sons, New York, 1975.
- 665 [16] R. J. Aumann, Integrals of set-valued functions, *J. Math. Anal. Appl.* 12 (1965) 1–12.
- [17] I. S. Molchanov, *Theory of random sets. Probability and its applications*, Springer, Berlin, 2005.
- [18] Z. Artstein, R. A. Vitale, A strong law of large numbers for random compact sets, *Ann. Probab.* 3 (1975) 879–882.
- 670 [19] P. J. Huber, Robust estimation of a location parameter, *Ann. Math. Stat.* 35 (1) (1964) 73–101.
- [20] V. J. Yohai, High breakdown-point and high efficiency robust estimates for regression, *Ann. Stat.* 15 (1987) 642–656.
- 675 [21] R. A. Maronna, D. R. Martin, V. J. Yohai, *Robust statistics: Theory and methods*, J. Wiley & Sons, New York, 2006.
- [22] P. J. Huber, *Robust statistics*, Wiley, New York, 1981.
- [23] F. R. Hampel, The influence curve and its role in robust estimation, *J. Amer. Stat. Assoc.* 69 (346) (1974) 383–393.

- 680 [24] J. S. Kim, Kernel methods for classification with irregularly sampled and contaminated data, Ph.D. thesis, University of Michigan (2011).
- [25] J. S. Kim, C. D. Scott, Robust kernel density estimation, *J. Mach. Learn. Res.* 13 (2012) 2529–2565.
- 685 [26] D. L. Donoho, P. J. Huber, The notion of breakdown point, in: P. J. Bickel, K. Doksum, J. J. L. Hodges (Eds.), *A Festschrift for Erich L. Lehmann*, Wadsworth, 1983, pp. 157–184.
- [27] B. Sinova, M. R. Casals, A. Colubi, M. A. Gil, The median of a random interval, in: C. Borgelt, G. González-Rodríguez, W. Trutschnig, M. A. Lubiano, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), *Combining Soft Computing and Statistical Methods in Data Analysis*, Vol. 77 of *Advances in Intelligent and Soft Computing*, Springer, Berlin, 2010, pp. 575–583.
- 690 [28] B. Sinova, S. Van Aelst, Comparing the medians of a random interval defined by means of two different L^1 metrics, in: C. Borgelt, M. A. Gil, J. M. Sousa, M. Verleysen (Eds.), *Towards Advanced Data Analysis by Combining Soft Computing and Statistics*, Vol. 285 of *Studies in Fuzziness and Soft Computing*, Springer, Berlin, 2013, pp. 75–86.
- 695 [29] B. Sinova, M. A. Gil, A. Colubi, S. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach, *Fuzzy Sets Syst.* 200 (2012) 99–115.
- 700 [30] B. Sinova, S. Pérez-Fernández, M. Montenegro, The wabl/ldev/rdev median of a random fuzzy number and statistical properties, in: P. Grzegorzewski, M. Gagolewski, O. Hryniewicz, M. A. Gil (Eds.), *Strengthening Links Between Data Analysis and Soft Computing*, Vol. 315 of *Advances in Intelligent Systems and Computing*, Springer, Heidelberg, 2015, pp. 143–150.
- 705 [31] R. M. C. R. De Souza, F. A. T. De Carvalho, Clustering of interval data based on city-block distances, *Pattern Recogn. Lett.* 25 (2004) 353–365.
- [32] F. A. T. De Carvalho, R. M. C. R. De Souza, M. Chavent, Y. Lechevallier, Adaptive Hausdorff distances and dynamic clustering of symbolic interval data, *Pattern Recogn. Lett.* 27 (2006) 167–179.
- 710 [33] J. C. Gower, Algorithm AS 78: The mediancenter, *Appl. Stat.* 23 (3) (1974) 466–470.
- [34] P. Milasevic, G. R. Ducharme, Uniqueness of the spatial median, *Ann. Stat.* 15 (1987) 1332–1333.
- 715 [35] B. Cadre, Convergent estimators for the L_1 -median of a Banach valued random variable, *Statistics* 35 (4) (2001) 509–521.

- [36] B. Sinova, G. González-Rodríguez, S. Van Aelst, An alternative approach to the median of a random interval using an L^2 metric, in: R. Kruse, M. R. Berthold, C. Moewes, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Sinergies of Soft Computing and Statistics for Intelligent Data Analysis, Vol. 190 of Advances in Intelligent and Soft Computing, Springer, Berlin, 2013, pp. 273–281.
- [37] B. Sinova, S. Van Aelst, A spatial-type interval-valued median for random intervals, Submitted.
- [38] B. Sinova, S. Van Aelst, On the consistency of a spatial-type interval-valued median for random intervals, *Stat. Probab. Lett.* 100 (2015) 130–136.
- [39] Y.-G. Wang, X. Lin, M. Zhu, Z. Bai, Robust estimation using the Huber function with a data-dependent tuning constant, *J. Comput. Graph. Stat.* 16 (2) (2007) 1–14.
- [40] C. Agostinelli, Robust model selection in regression via weighted likelihood methodology, *Stat. Probab. Lett.* 56 (2002) 289–300.
- [41] E. D. Ongley, Environmental quality: changing times for sediment programs, in: J. Bogen, D. E. Walling, T. J. Day (Eds.), *Erosion and Sediment Transport Monitoring Programmes in River Basins*, Vol. 210, IAHS Press: Wallingford, UK, 1992, pp. 379–389.
- [42] T. Voyslavov, S. Tsakovski, V. Simeonov, Surface water quality assessment using self-organizing maps and hasse diagram technique, *Chemometr. Intell. Lab. Syst.* 118 (2012) 280–286.
- [43] E. P. Tao, W. H. Shen, T. L. Liu, X. Q. Chen, Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process, *Chemometr. Intell. Lab. Syst.* 128 (2013) 49–55.
- [44] J. W. Tukey, T6: Order statistics, in: *Mimeographed Notes for Statistics*, Princeton University, 1974.
- [45] J. W. Tukey, Mathematics and the picturing of data, in: *Proceedings of the International Congress of Mathematicians*, 1974, pp. 523–531.
- [46] R. Y. Liu, On a notion of simplicial depth, in: *Proceedings of the National Academy of Sciences of the United States of America*, 1988, pp. 1732–1734.
- [47] J. P. Gibbs, L. D. Poston, The division of labor: Conceptualization and related measures, *Social Forces* 53 (3) (1975) 468–476.
- [48] L. D. Brown, R. Purves, Measurable selections of extrema, *Ann. Stat.* 1 (5) (1973) 902–912.

- 755
- [49] P. J. Huber, The behavior of maximum likelihood estimates under non-standard conditions, in: L. M. Le Cam, J. Neyman (Eds.), Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1967, pp. 221–233.
 - [50] H. P. Lopuhaä, P. J. Rousseeuw, Breakdown points of affine equivariant estimators of multivariate location and covariance matrices, *Ann. Stat.* 19 (1991) 229–248.