

## **M-Estimators and Trimmed Means: From Hilbert-Valued to Fuzzy Set-Valued Data**

**Beatriz Sinova · Stefan Van Aelst ·  
Pedro Terán**

Received: date / Accepted: date

**Abstract** Different approaches to robustly measure the location of data associated with a random experiment have been proposed in the literature, with the aim of avoiding the high sensitivity to outliers or data changes typical for the mean. In particular, M-estimators and trimmed means have been studied in general spaces, and can be used to handle Hilbert-valued data. Both alternatives are of interest due to their success in the classical framework. Since fuzzy set-valued data can be identified with a convex cone of a separable Hilbert space, the previous concepts have been recently applied to the one-dimensional fuzzy case. The aim of this paper is to extend M-estimators and trimmed means to  $p$ -dimensional fuzzy set-valued data, and to theoretically prove that they inherit robustness from the real settings. Some of such theoretical results are more general and directly apply to Hilbert-valued estimators and, in consequence, to functional data. A real-life example will also be included to illustrate the computation and behaviour of these estimators under contamination.

**Keywords** Robust location · Finite sample breakdown point · Functional data · Random fuzzy sets · Random sets

**Mathematics Subject Classification (2010)** 62G35 · 62-07 · 03E72

---

The research of Beatriz Sinova and Pedro Terán was supported by the Spanish Ministry of Economy and Competitiveness under Grant MTM2015-63971-P; and the Principality of Asturias/FEDER Funds under Grants GRUPIN14-101 and GRUPIN-IDI2018-000132. The research of Stefan Van Aelst was supported by Internal Funds KU Leuven (Belgium) under Grant C16/15/068. Their support is gratefully acknowledged.

---

B. Sinova (corresponding author) · P. Terán  
Department of Statistics and O.R. and D.M., University of Oviedo, 33007 Oviedo, Spain  
E-mail: sinovabeatriz@uniovi.es, teranpedro@uniovi.es

S. Van Aelst  
Department of Mathematics, KU Leuven, 3001 Leuven, Belgium  
E-mail: stefan.vanaelst@kuleuven.be

## 1 Introduction

Two successful alternatives in the classical framework, M-estimators and trimmed means, have already been proposed to handle Hilbert-valued data and avoid the high sensitivity to outliers or data changes when summarizing the location. In Sinova et al. (2018), the properties of these M-estimators are studied, but some of the properties these trimmed means satisfy, even those related to their robustness, have not been theoretically proven yet. Due to the perspective of considering functional data as realizations of Hilbert space valued random variables, the proof of such results would be interesting for the functional data framework as well, whose importance is vastly increasing (see, for instance, Valencia et al., 2019, Aneiros et al., 2019 and Hubert et al., 2017, for some recent studies about functional data).

Recently, M-estimators and trimmed means have also been applied to the one-dimensional fuzzy case thanks to the isometrical embedding which identifies fuzzy set-valued data with a convex cone of a separable Hilbert space (see Sinova et al., 2016, and Colubi and González-Rodríguez, 2015, respectively). Since fuzzy set theory was introduced in the sixties (Zadeh, 1965), many studies developing statistical techniques to analyze fuzzy-valued data have been published in the literature. This kind of data is becoming more widely used to model the output of real-life experiments since “paradoxically, one of the principal contributions of fuzzy logic is its high power of precisiation of what is imprecise”, as pointed out by Zadeh (2008).

Formally, fuzzy set-valued data are assumed to be realizations of a random fuzzy set (fuzzy random variable in the sense of Puri and Ralescu, 1986). Many methods for the statistical analysis of random fuzzy sets are based on the well-known Aumann-type mean as measure of the location of the distribution of a random fuzzy set, but its high sensitivity, inherited from that of the mean of random variables, makes it desirable to introduce more robust alternatives, such as location M-estimators and trimmed means.

Properties of location M-estimators for fuzzy number-valued data (i.e., one-dimensional fuzzy set-valued data) have been studied in Sinova et al. (2016), but some of them will be generalized in this paper to cover more choices of the loss function. Moreover, we will show that such results still remain valid for the  $p$ -dimensional case of fuzzy set-valued data (and, in particular, for compact and convex random sets). On the other hand, the analogous properties concerning trimmed means for fuzzy set-valued data have not been tackled yet. Trimming has been proven to be a successful alternative not only for the study of the location (see e.g. Cuesta-Albertos et al., 1997), but also for other applications, like regression and clustering methods (see Rivera-García et al., 2019, Alfons et al., 2013 and García-Escudero et al., 2010, among others). The properties of trimmed means will be analyzed in this paper and, even while they will be formalized for the  $p$ -dimensional fuzzy set-valued data framework, they are still valid when Hilbert-valued or, in particular, functional data are considered. The comparison of location M-estimators and trimmed means under contamination is approached theoretically by means of their finite sample breakdown point.

To illustrate and compare both measures in the fuzzy set-valued framework, a real-life example will be used.

In Section 2, the notions and results related to the space of fuzzy sets and random fuzzy sets used in this work are recalled. In Section 3, fuzzy-valued location M-estimators are generalized to the  $p$ -dimensional case, and their robustness properties are studied. Analogously, trimmed means for fuzzy set-valued data are dealt with in Section 4 and their robustness is proven. Proofs of the theoretical results can be found in the supplementary material (Online Resource). The two approaches are illustrated by means of a real-life example, consisting of 2-dimensional fuzzy set-valued data, in Section 5. Examples which are available in the literature usually involve conical data, whereas our example analyzes more complex data, which is of high importance. Finally, some concluding remarks are given in Section 6.

## 2 The fuzzy setting

### 2.1 The space $\mathcal{F}_c(\mathbb{R}^p)$

In this section, the space  $\mathcal{F}_c(\mathbb{R}^p)$  of fuzzy sets is recalled, in order to give the basic definitions and results required for the understanding of the following sections.

The origin of the concept of fuzzy sets is related to the difficulties attached to the procedure of expressing intrinsically imprecise attributes or measurements, involving some error margins, as fixed values in a precise scale. Fuzzy sets formally introduce gradualness to reflect different levels of imprecision. With this idea in mind, it is easy to understand why fuzzy sets have become so useful in analyzing human opinions, ratings, perceptions..., which are naturally imprecise and subjective.

**Definition 1** A **fuzzy set value** (or **fuzzy vector**) is a normal, upper semi-continuous and quasi-concave mapping  $\tilde{U} : \mathbb{R}^p \rightarrow [0, 1]$ . That is, for all  $\alpha \in (0, 1]$  the  **$\alpha$ -level**

$$\tilde{U}_\alpha = \{\mathbf{x} \in \mathbb{R}^p : \tilde{U}(\mathbf{x}) \geq \alpha\}$$

is a nonempty compact convex set, and the *0-level*,  $\tilde{U}_0 = \text{cl}\{\mathbf{x} \in \mathbb{R}^p : \tilde{U}(\mathbf{x}) > 0\}$  is a nonempty convex set.

The class of all such fuzzy sets is denoted by  $\mathcal{F}_c(\mathbb{R}^p)$ . The interpretation of this notion is the following:  $\tilde{U}(\mathbf{x})$  informs of the ‘degree of compatibility’ of  $\mathbf{x}$  with the quality/property  $\tilde{U}$  represents or, in other words, the ‘degree of truth’ of the assertion “ $\mathbf{x}$  is  $\tilde{U}$ ”.

The most important operations from the statistical point of view, the sum and the product by scalars, can be naturally defined as a level-wise generalization of the usual set arithmetic by means of Zadeh’s extension principle.

**Definition 2** Let  $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$ . The **sum** of  $\tilde{U}$  and  $\tilde{V}$  is defined as the fuzzy set  $\tilde{U} + \tilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$  whose  $\alpha$ -levels are given by

$$(\tilde{U} + \tilde{V})_\alpha = \{\mathbf{y} + \mathbf{z} : \mathbf{y} \in \tilde{U}_\alpha, \mathbf{z} \in \tilde{V}_\alpha\}$$

for all  $\alpha \in [0, 1]$ . More generally, if  $\tilde{U}_i \in \mathcal{F}_c(\mathbb{R}^p)$  for  $i \in \{i_1, \dots, i_m\}$ , with  $m \in \mathbb{N}$ , we denote

$$\sum_{i \in \{i_1, \dots, i_m\}} \tilde{U}_i = \tilde{U}_{i_1} + \dots + \tilde{U}_{i_m}.$$

Let  $\tilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$  and  $\gamma \in \mathbb{R}$ . The **product of  $\tilde{U}$  by the scalar  $\gamma$**  is defined as the fuzzy set  $\gamma \cdot \tilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$  whose  $\alpha$ -levels are given by

$$(\gamma \cdot \tilde{U})_\alpha = \{\gamma \mathbf{y} : \mathbf{y} \in \tilde{U}_\alpha\}$$

for all  $\alpha \in [0, 1]$ .

In order to characterize fuzzy set values, the support function used in the set-valued setting (Minkowski, 1903, see also Castaing and Valadier, 1977, for more details) has been extended to  $\mathcal{F}_c(\mathbb{R}^p)$  by Puri and Ralescu (1985) and Bobylev (1985).

**Definition 3** The **support function** of  $\tilde{U} \in \mathcal{F}_c(\mathbb{R}^p)$  is given by the mapping

$$s_{\tilde{U}} : (0, 1] \times \mathbb{S}^{p-1} \rightarrow \mathbb{R}, \quad s_{\tilde{U}}(\alpha, \mathbf{u}) = \sup_{\mathbf{x} \in \tilde{U}_\alpha} \langle \mathbf{u}, \mathbf{x} \rangle,$$

where  $\langle \cdot, \cdot \rangle$  is the inner product on  $\mathbb{R}^p$  and  $\mathbb{S}^{p-1}$  the unit sphere of  $\mathbb{R}^p$ .

The space  $(\mathcal{F}_c(\mathbb{R}^p), +, \cdot)$  is not linear because of the lack of opposite element for the sum, so metrics between fuzzy sets play a crucial role when analyzing fuzzy set-valued data statistically. In particular, an always well-defined notion of difference between fuzzy sets that preserves the properties it has in connection with the sum in the real settings is lacking. To avoid this drawback, differences will be replaced by distances in the statistical developments. Using the representation of fuzzy sets recalled above, it is possible to define  $L^2$ -type metrics to measure the distance between elements of  $\mathcal{F}_c(\mathbb{R}^p)$ . The choice for  $L^2$ -type metrics is due to the isometrical embedding that can then be established to identify a subspace of  $\mathcal{F}_c(\mathbb{R}^p)$  with a convex cone of a separable Hilbert space in the following way.

Let  $\mathcal{F}_c^2(\mathbb{R}^p) = \{\tilde{U} \in \mathcal{F}_c(\mathbb{R}^p) : s_{\tilde{U}} \in \mathbb{H}_2\}$ , where  $\mathbb{H}_2 = \mathcal{L}^2((0, 1] \times \mathbb{S}^{p-1}, \lambda \otimes \lambda_p)$  is the separable Hilbert space of the  $L^2$ -type real-valued functions defined on the Cartesian product of  $\mathbb{S}^{p-1}$  and the interval  $(0, 1]$ ,  $\lambda$  denotes the Lebesgue measure on  $(0, 1]$  and  $\lambda_p$  denotes the normalized Lebesgue measure on  $\mathbb{S}^{p-1}$ . Along the paper we consider a generic metric  $D$  defined on  $\mathcal{F}_c^2(\mathbb{R}^p) \times \mathcal{F}_c^2(\mathbb{R}^p)$  which satisfies

- $D$  is an  $L^2$ -type metric,
- the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  is separable,

- the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ , endowed with the arithmetic explained above, can be isometrically embedded into a convex cone of a Hilbert space (whence the associated norm through the isometrical embedding,  $\|\cdot\|_D$ , fulfills the parallelogram law).

Examples of metrics fulfilling the preceding conditions are the  $L^2$ -type distance by Klement et al. (1986), which for any fuzzy values  $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$  is given by:

$$\mathbf{d}_2(\tilde{U}, \tilde{V}) = \sqrt{\int_{(0,1]} \sup_{\mathbf{u} \in \mathbb{S}^{p-1}} (s_{\tilde{U}}(\alpha, \mathbf{u}) - s_{\tilde{V}}(\alpha, \mathbf{u}))^2 d\lambda(\alpha)},$$

and the generalized metric by Trutschnig et al. (2009), which for a weighting parameter  $\theta > 0$  and for any fuzzy values  $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$  is given by the following expression:

$$D_\theta(\tilde{U}, \tilde{V}) = \left[ \int_{(0,1] \times \mathbb{S}^{p-1}} ((\text{mid}_{\tilde{U}_\alpha}(\mathbf{u}) - \text{mid}_{\tilde{V}_\alpha}(\mathbf{u}))^2 + \theta (\text{spr}_{\tilde{U}_\alpha}(\mathbf{u}) - \text{spr}_{\tilde{V}_\alpha}(\mathbf{u}))^2) d\lambda_p(\mathbf{u}) d\lambda(\alpha) \right]^{\frac{1}{2}},$$

where  $\text{mid}_{\tilde{U}_\alpha}(\mathbf{u}) = (s_{\tilde{U}}(\alpha, \mathbf{u}) - s_{\tilde{U}}(\alpha, -\mathbf{u}))/2$  and  $\text{spr}_{\tilde{U}_\alpha}(\mathbf{u}) = (s_{\tilde{U}}(\alpha, \mathbf{u}) + s_{\tilde{U}}(\alpha, -\mathbf{u}))/2$  indicate the ‘center’ and the ‘shape’, respectively. Metrics  $\mathbf{d}_2$  and  $D_\theta$  are topologically equivalent. In particular, the space  $(\mathcal{F}_c^2(\mathbb{R}^p), D_\theta)$ , endowed with the above-mentioned arithmetic, can be isometrically embedded into a convex cone of  $\mathbb{H}_2$ .

## 2.2 Random fuzzy sets

The concept of a random fuzzy set (or a fuzzy random variable in the sense of Puri and Ralescu, 1986) mathematically formalizes the random mechanism producing fuzzy set-valued data. It is indeed a well-stated and supported model that combines randomness affecting the generation of the data and imprecision regarding its nature (see, e.g., Gil et al., 2013, for a recent overview).

The notion of a random (compact convex) set, which is involved in the definition of a random fuzzy set, will be now presented.

**Definition 4** Let  $\mathcal{K}_c(\mathbb{R}^p)$  denote the space of nonempty compact convex subsets of  $\mathbb{R}^p$ . Given a probability space  $(\Omega, \mathcal{A}, P)$ , a mapping  $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R}^p)$  is said to be a **random compact convex set** if  $X$  is measurable with respect to the Borel  $\sigma$ -algebra generated by the topology induced by the Hausdorff metric on  $\mathcal{K}_c(\mathbb{R}^p)$ , which is defined as follows:

$$d_H(K, K') = \max \left\{ \sup_{x \in K} \inf_{y \in K'} \|x - y\|, \sup_{y \in K'} \inf_{x \in K} \|x - y\| \right\}$$

for any  $K, K' \in \mathcal{K}_c(\mathbb{R}^p)$ .

The random closed set is defined as follows.

**Definition 5** Let  $Z$  be a separable metric space and let  $X$  be a function on a probability space  $(\Omega, \mathcal{A}, P)$  with values in the family of closed subsets of  $Z$ . The function  $X$  is said to be a **random closed set** or **Effros-measurable** if  $\{\omega : X(\omega) \cap G \neq \emptyset\} \in \mathcal{A}$  for every open  $G \subset Z$ .

**Definition 6** Let  $(\Omega, \mathcal{A}, P)$  be a probability space. A mapping  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R}^p)$  is said to be a **random fuzzy set value** (or **random fuzzy vector**, in short RFV) if it is a Borel-measurable function (i.e., a random element in Fréchet's sense) with respect to the Borel  $\sigma$ -field generated by the topology induced by any metric  $D$  fulfilling the conditions given above.

For the particular cases of the metrics mentioned in this work, it holds that  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R}^p)$  is a random fuzzy set value if for each  $\alpha \in (0, 1]$  the set-valued mapping  $\mathcal{X}_\alpha$ , with  $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$  for all  $\omega \in \Omega$ , is a random compact convex set.

The Borel-measurability of an RFV implies that notions like the *induced distribution* of a random fuzzy set value or the *stochastic independence* of random fuzzy set values follow directly without having to define them explicitly.

Regarding the location or central tendency of the (induced) distribution of an RFV, the best-known measure is the Aumann-type mean introduced by Puri and Ralescu (1986).

**Definition 7** Let  $(\Omega, \mathcal{A}, P)$  be a probability space and  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R}^p)$  an associated random fuzzy set value which is integrably bounded (i.e.,  $E(\sup\{\|\mathbf{x}\| : \mathbf{x} \in \mathcal{X}_0\}) < \infty$ ). The **Aumann-type mean** of  $\mathcal{X}$  is the unique fuzzy set value  $\tilde{E}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R}^p)$  such that for all  $\alpha \in (0, 1]$

$$(\tilde{E}(\mathcal{X}))_\alpha = E[\mathcal{X}_\alpha] = \{E(f) \mid f : \Omega \rightarrow \mathbb{R}^p, f \in L^1(\Omega, \mathcal{A}, P), f \in \mathcal{X}_\alpha \text{ a.s.}[P]\}.$$

The Aumann-type mean preserves all the main properties of the mean of a random variable from both the statistical and probabilistic points of view. For example, it is consistent with the usual fuzzy arithmetic and it is supported by Strong Laws of Large Numbers and by the Fréchet approach (Fréchet, 1948). Unfortunately, as is the case with the mean of a real-valued random variable, the Aumann-type mean is highly sensitive to outliers or data changes. This important drawback motivates the search of more robust location measures by extending successful measures from the real-valued case, such as M-estimators and trimmed means. These extensions will be detailed in the next two sections.

### 3 Location M-estimators for fuzzy set-valued data

M-estimation was introduced by Huber (1964) with the aim to limit the influence of outliers on estimators. Many methods, like least squares or maximum likelihood, evaluate “errors” of the data by means of a squared function, which

assigns outliers an excessive weight. Huber's idea of replacing the squared error by another (usually less rapidly increasing) loss function applied to the errors has proven to be successful in the classical framework. Recently, Kim and Scott (2012) studied M-estimators in the kernel density estimation context. Their results remain valid for Hilbert-valued random elements and also for the one-dimensional RFV (as stated in Sinova et al., 2016).

In this section, the notion of location M-estimator is extended to cover  $p$ -dimensional fuzzy set-valued data. In this paper, M-estimators are based on a non-negative valued loss function  $\rho$  defined on  $\mathbb{R}$  which is assumed to satisfy the following condition.

- (C.1)  $\rho$  vanishes at 0 and it is an even, continuous function which is nondecreasing on the positive real line.

**Definition 8** Let  $(\Omega, \mathcal{A}, P)$  be a probability space,  $\rho$  be a loss function satisfying C.1 and  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  an associated RFV. The **fuzzy M-location value** is any element  $\tilde{g}^M(\mathcal{X})$  in  $\mathcal{F}_c^2(\mathbb{R}^p)$  given by

$$\tilde{g}^M(\mathcal{X}) = \arg \min_{\tilde{g} \in \mathcal{F}_c^2(\mathbb{R}^p)} E(\rho(D(\mathcal{X}, \tilde{g}))),$$

if it exists.

**Definition 9** Let  $(\Omega, \mathcal{A}, P)$  be a probability space and  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  an associated RFV. Moreover, let  $\rho$  be a loss function satisfying C.1, and  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  a simple random sample from  $\mathcal{X}$  (that is,  $\mathcal{X}_1, \dots, \mathcal{X}_n$  are independent RFVs identically distributed as  $\mathcal{X}$ ). Then, the **fuzzy M-estimator of location** is any fuzzy set-valued statistic  $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]$  given by

$$\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)] = \arg \min_{\tilde{g} \in \mathcal{F}_c^2(\mathbb{R}^p)} \frac{1}{n} \sum_{i=1}^n \rho(D(\mathcal{X}_i, \tilde{g})),$$

if it exists.

Notice that neither the fuzzy M-location value nor the fuzzy M-estimator of location are necessarily unique.

The following theorem gives some sufficient conditions which guarantee that fuzzy M-estimators of location are indeed statistics.

**Theorem 1** Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  on a complete probability space  $(\Omega, \mathcal{A}, P)$ . Let  $\rho$  be a loss function satisfying C.1. If  $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]$  is unique (for all  $\omega \in \Omega$ ), then the fuzzy M-estimator of location is well-defined.

*Remark 1* It should be pointed out that Theorem 1 holds for any separable metric space (not only for  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ ), the proof being trivially generalizable.

Since Theorem 1 provides us with some sufficient (and not necessary) conditions for the measurability of M-estimators of location, the completeness of the probability space will not be required in the following results, but it will be assumed that the corresponding M-estimator of location is measurable and, then, well-defined.

The following result is an adaptation of Kim and Scott's Representer Theorem (see Kim and Scott, 2012, Theorems 2 and 3), RT for short, which provides conditions to ensure the existence and uniqueness of M-estimators of location, to the fuzzy set-valued case. The proof is omitted because it is a rather straightforward extension from the corresponding result in Sinova et al. (2016) for the one-dimensional case.

**Theorem 2** Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  on a probability space  $(\Omega, \mathcal{A}, P)$ . Moreover, let  $\rho$  be a loss function which satisfies C.1 and the following assumptions

$$(C.2) \lim_{x \rightarrow 0} \rho(x)/x = 0,$$

(C.3) Set  $\phi(x) = \rho'(x)/x$  and  $\phi(0) \equiv \lim_{x \rightarrow 0} \phi(x)$ , then  $\phi(0)$  exists and it is finite.

Then, the fuzzy M-estimator of location exists and it can be expressed as

$$\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)] = \sum_{i=1}^n u_i \cdot \mathcal{X}_i$$

with  $u_i \geq 0$ ,  $\sum_{i=1}^n u_i = 1$ . Furthermore,  $u_i \propto \phi(D(\mathcal{X}_i, \widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]))$ . Moreover, if the function  $J_n(\widehat{g}) = \frac{1}{n} \sum_{i=1}^n \rho(D(\mathcal{X}_i, \widehat{g}))$  is strictly convex, i.e. whatever  $\lambda \in (0, 1)$  and  $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$  with  $\widetilde{U} \neq \widetilde{V}$  may be,  $J_n(\lambda \widetilde{U} + (1 - \lambda)\widetilde{V}) < \lambda J_n(\widetilde{U}) + (1 - \lambda)J_n(\widetilde{V})$ , then the conditions

- $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)] = \sum_{i=1}^n u_i \cdot \mathcal{X}_i$
- $u_i \propto \phi(D(\mathcal{X}_i, \widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]))$
- $\sum_{i=1}^n u_i = 1$

are sufficient to guarantee that  $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]$  is the unique fuzzy M-estimator of location.

*Remark 2* Theorem 2 guarantees that location M-estimators for fuzzy set-valued data are well-defined, as they can be expressed as weighted linear combinations of the sample observations they will also belong to the space  $\mathcal{F}_c^2(\mathbb{R}^p)$ .

*Remark 3* Conditions C.1-C.3 (in short denoted as RTC, from Representer Theorem's Conditions) are quite general and they are fulfilled by well-known families of loss functions, such as the Huber (1981) family of loss functions

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



with  $a > 0$  a tuning parameter, or the Hampel (1974) family of loss functions, given by

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

with tuning parameters  $c > b > a > 0$ .

### 3.1 Equivariance and symmetry properties for fuzzy M-estimators

The following results show how some of the convenient properties of the (non-fuzzy) M-estimators of location, like their translation equivariance and their behaviour for symmetric distributions, are preserved in the fuzzy-valued setting.

**Proposition 1** (Translation equivariance) *Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  and let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  on a probability space  $(\Omega, \mathcal{A}, P)$ . Moreover, let  $\rho$  be a loss function fulfilling assumptions C.1-C.3, such that the corresponding fuzzy M-estimator of location is unique. Let  $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R}^p)$ . Then, the M-estimator of location based on the simple random sample  $(\mathcal{X}_1 + \tilde{U}, \dots, \mathcal{X}_n + \tilde{U})$ ,  $\widehat{g}^M[(\mathcal{X}_1 + \tilde{U}, \dots, \mathcal{X}_n + \tilde{U})]$ , is equal to  $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)] + \tilde{U}$ , where*

- i)  $\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)] = \sum_{i=1}^n u_i \cdot \mathcal{X}_i$ ,
- ii)  $u_i \propto \phi(D(\mathcal{X}_i, \widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)]))$ ,
- iii)  $\sum_{i=1}^n u_i = 1$ .

*Remark 4* Unluckily, fuzzy M-estimators of location are only scale equivariant when the loss function  $\rho$  is a power function, which is a too strong requirement. This inconvenience is inherited from the classical setting, and a way to overcome it could be to choose the loss function for the computation of M-estimators such that it takes into account the distribution of the distances. More details will be given in Subsection 3.2.

The notion of symmetry of a random fuzzy set value around a vector value is stated as follows: an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  is said to be **symmetrically distributed** about  $\mathbf{c} \in \mathbb{R}^p$  if and only if  $\mathcal{X} - \mathbf{c} := \mathcal{X} + (-\mathbf{c})$  and  $\mathbf{c} - \mathcal{X} := \mathbf{c} + (-1) \cdot \mathcal{X}$  are identically distributed.

**Proposition 2** *Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be an RFV associated with the probability space  $(\Omega, \mathcal{A}, P)$  which is symmetrically distributed for some  $\mathbf{c} \in \mathbb{R}^p$ . Moreover, let  $\rho$  be a loss function fulfilling C.1. If the associated M-location value  $\tilde{g}^M(\mathcal{X})$  is unique, then  $\tilde{g}^M(\mathcal{X})$  is a symmetric fuzzy set about  $\mathbf{c}$  (that is, in fuzzy terminology,  $\tilde{g}^M(\mathcal{X})$  is ‘additively equivalent to’  $\mathbf{c}$ ).*

### 3.2 Algorithmic computation of fuzzy M-estimators under RT conditions

Regarding the practical computation of fuzzy M-estimators of location, the standard iteratively re-weighted least squares algorithm (see, for example, Huber, 1964) can provide us with an approximation, as follows:

**Step 1.** Select initial weights  $\omega_i^{(0)} \in \mathbb{R}$ , for  $i \in \{1, \dots, n\}$ , such that  $\omega_i^{(0)} \geq 0$  and  $\sum_{i=1}^n \omega_i^{(0)} = 1$  (e.g. weights corresponding to a robust estimator of location chosen to initialize the algorithm).

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

$$\tilde{g}_{(k)}^M = \sum_{i=1}^n \omega_i^{(k-1)} \mathcal{X}_i, \quad \omega_i^{(k)} = \frac{\phi(D(\mathcal{X}_i, \tilde{g}_{(k)}^M))}{\sum_{j=1}^n \phi(D(\mathcal{X}_j, \tilde{g}_{(k)}^M))}.$$

**Step 3.** Terminate the algorithm when

$$\frac{|\frac{1}{n} \sum_{i=1}^n \rho(D(\mathcal{X}_i, \tilde{g}_{(k+1)}^M)) - \frac{1}{n} \sum_{i=1}^n \rho(D(\mathcal{X}_i, \tilde{g}_{(k)}^M))|}{\frac{1}{n} \sum_{i=1}^n \rho(D(\mathcal{X}_i, \tilde{g}_{(k)}^M))} < \varepsilon,$$

for some desired tolerance  $\varepsilon > 0$ .

In case the scale equivariance does not hold (see Remark 4), measurement units can have too much influence on the output. In order to avoid this, if the loss function  $\rho$  involves some tuning parameters (as is the case for the well-known Huber and Hampel families of loss functions), they will be selected based on the distribution of the distances from each sample observation to an initial robust estimator of location (e.g., the fuzzy trimmed mean estimator studied in Section 4), as suggested by Kim and Scott (2012).

### 3.3 Consistency and robustness properties of fuzzy M-estimators

The consistency of fuzzy M-estimators can be guaranteed under rather general conditions. First, instead of considering  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  we consider  $(\mathcal{F}_c^2(A), D)$  with  $A$  a non-empty compact convex subset of  $\mathbb{R}^p$  (notice that, in practice, it is very common to work with bounded referentials, like in Hesketh et al., 1988). The proofs of the consistency and finite-sample breakdown point in Theorems 3 and 4 can be rather straightforwardly adapted from those for the one-dimensional case in Sinova et al. (2016) and hence are omitted.

**Theorem 3** Consider a locally compact metric space  $(\mathcal{F}_c^2(A), D)$ , where  $A$  is a non-empty compact convex subset of  $\mathbb{R}^p$ . Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(A)$  be an RFV associated with the probability space  $(\Omega, \mathcal{A}, P)$ . For any  $n \in \mathbb{N}$ , let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$ . Let  $\rho$  be a loss function which satisfies C.1 and either is subadditive and unbounded, or admits upper and lower linear bounds with the same slope, or is bounded. Then, if we assume that the fuzzy M-location value  $\tilde{g}^M(\mathcal{X})$  is unique, the fuzzy M-estimator of location is a strongly consistent estimator of the fuzzy M-location measure, i.e.,

$$\lim_{n \rightarrow \infty} D(\widehat{\tilde{g}^M}[(\mathcal{X}_1, \dots, \mathcal{X}_n)], \tilde{g}^M(\mathcal{X})) = 0 \quad \text{a.s.}[P].$$

*Remark 5* Conditions to ensure the strong consistency are not very restrictive in practice. Indeed, the local compactness of  $(\mathcal{F}_c^2(A), D)$  is fulfilled for many known metrics, like  $D_\theta$  or  $\mathbf{d}_2$ , as will be shown elsewhere. Furthermore, the additional conditions for the loss function are fulfilled for well-known functions like Huber's, Hampel's and Tukey's.

The *finite sample breakdown point* of fuzzy M-estimators of location is derived now. As indicated by Cuevas et al. (2007), this notion can be adapted to estimators taking values in general metric spaces. Following Donoho and Huber (1983), the finite sample breakdown point (denoted fsbp) of the fuzzy M-estimator of location corresponding to a sample of size  $n$  from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$ , where  $(\Omega, \mathcal{A}, P)$  is a probability space and  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  is a metric space such that  $\sup_{\tilde{U}, \tilde{V} \in \mathcal{F}_c^2(\mathbb{R}^p)} D(\tilde{U}, \tilde{V}) = \infty$ , is given by

$$\begin{aligned} & \text{fsbp}(\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)], \tilde{\mathbf{x}}_n, D) \\ &= \frac{1}{n} \min \left\{ k \in \{1, \dots, n\} : \sup_{\tilde{\mathbf{y}}_{n,k}} D(\widehat{g}^M[\tilde{\mathbf{x}}_n], \widehat{g}^M[\tilde{\mathbf{y}}_{n,k}]) = \infty \right\}, \end{aligned}$$

where  $\tilde{\mathbf{x}}_n$  denotes the considered sample of  $n$  observations from  $\mathcal{X}$  and the samples  $\tilde{\mathbf{y}}_{n,k}$  are obtained from  $\tilde{\mathbf{x}}_n$  by perturbing up to  $k$  of its elements.

**Theorem 4** Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be an RFV associated with the probability space  $(\Omega, \mathcal{A}, P)$ , let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$  and  $\tilde{\mathbf{x}}_n = (\tilde{x}_1, \dots, \tilde{x}_n)$  be a realization of the simple random sample. If the loss function  $\rho$  fulfills RT conditions, and the corresponding fuzzy M-estimator of location is unique, then the finite sample breakdown point of the fuzzy M-estimator of location is at most  $\frac{1}{n} \lfloor \frac{n+1}{2} \rfloor$ . Furthermore, under the additional conditions of

- $\rho$  having linear upper and lower bounds with the same slope or
- $\rho$  being bounded by a constant  $C$  such that

$$\rho \left( \max_{1 \leq i, j \leq n} D(\tilde{x}_i, \tilde{x}_j) \right) < \frac{n - 2 \lfloor \frac{n-1}{2} \rfloor}{n - \lfloor \frac{n-1}{2} \rfloor - 1} C,$$

the finite sample breakdown point is exactly given by

$$\text{fsbp}(\widehat{g}^M[(\mathcal{X}_1, \dots, \mathcal{X}_n)], \tilde{\mathbf{x}}_n, D) = \frac{1}{n} \lfloor \frac{n+1}{2} \rfloor.$$

Since the finite sample breakdown point of the Aumann-type mean is  $\frac{1}{n}$ , this proves that fuzzy M-estimators of location are more robust.

#### 4 Fuzzy trimmed means for fuzzy set-valued data

Some robust location measures for general-valued data have already been proposed in the literature, for instance, the usual functional median and the trimmed mean in functional Hilbert spaces (see Cuesta-Albertos and Fraiman, 2006 and López-Pintado and Romo, 2009, among others). Based on Cuesta-Albertos and Fraiman (2006), consider a Hilbert space-valued random element  $X : \Omega \rightarrow \mathbb{H}$ , where  $(\Omega, \mathcal{A}, P)$  is a probability space,  $(\mathbb{H}, \|\cdot\|)$  is a separable Hilbert space and  $P_X$  is the induced probability distribution on the Borel  $\sigma$ -algebra on  $\mathbb{H}$ . For any  $\beta \in (0, 1)$ , the corresponding trimmed mean is any  $\mathbf{g}_{P_X} \in \mathbb{H}$  such that there exists a trimming function  $\tau_{P_X} \in \mathcal{P}_\beta = \{\tau : \mathbb{H} \rightarrow (0, 1) : \tau \text{ measurable, } \int \tau(x) dP_X(x) \geq 1 - \beta\}$  satisfying that

$$\int \|x - \mathbf{g}_{P_X}\|^2 \tau_{P_X}(x) dP_X(x) = \inf_{\mathbf{g} \in \mathbb{H}, \tau \in \mathcal{P}_\beta} \int \|x - \mathbf{g}\|^2 \tau(x) dP_X(x).$$

This definition generalizes the notion of trimmed means based on trimming regions, since that situation is equivalent to using as trimming function only indicator functions of sets, i.e.

$$\min_{A \in \mathcal{E}} \int_A \|x - E(X|A)\|^2 dP_X(x) = \min_{A \in \mathcal{E}, \mathbf{g} \in \mathbb{H}} \int_A \|x - \mathbf{g}\|^2 dP_X(x),$$

where  $\mathcal{E} = \{A \subset \mathbb{H} : P_X(A) = 1 - \beta\}$ . However, it turns out that the best trimming function essentially coincides with the indicator function of a set (as proven in Cuesta-Albertos et al., 1997, for random vectors and in Cuesta-Albertos and Fraiman, 2006, for Hilbert space-valued random elements), so we can restrict ourselves to this case.

**Definition 10** Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be an RFV associated with a probability space  $(\Omega, \mathcal{A}, P)$ . For any  $\beta \in (0, 1)$ , the corresponding **fuzzy trimmed mean**  $E_{[\beta]}(\mathcal{X}) = E(\mathcal{X}|A_{P_X}) \in \mathcal{F}_c^2(\mathbb{R}^p)$  of  $\mathcal{X}$  is defined through

$$A_{P_X} = \arg \min_{\substack{A \subset \mathbb{H}_2 \\ P_{s_{\mathcal{X}}}(A) \geq 1 - \beta}} \int_A (D(\tilde{x}, \tilde{E}(\mathcal{X}|A)))^2 dP_{s_{\mathcal{X}}}(s_{\tilde{x}}),$$

and  $P_{s_{\mathcal{X}}}$  denoting the induced probability distribution on the Borel  $\sigma$ -algebra on  $\mathbb{H}_2$ .

*Remark 6* The particular case of the Hilbert space-valued trimmed means for the fuzzy number-valued case, thanks to the embedding of the space of fuzzy set values into a closed convex cone of a Hilbert space of functions, has been first considered by Colubi and González-Rodríguez (2015) in terms of the metric  $D_\theta$ . Nevertheless, this notion can be extended to deal with the generic metric  $D$  in this paper.

**Definition 11** Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be an RFV associated with a probability space  $(\Omega, \mathcal{A}, P)$  and  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$ . For any  $\beta \in (0, 1)$ , the **fuzzy trimmed mean estimator** is

$$\bar{\mathcal{X}}_{n,\beta} = \frac{1}{h} \cdot \sum_{j \in \widehat{\mathbb{E}}(\mathcal{X}_1, \dots, \mathcal{X}_n)} \mathcal{X}_j,$$

and

$$\widehat{\mathbb{E}}(\mathcal{X}_1, \dots, \mathcal{X}_n) = \arg \min_{\mathbb{E} \subset \{1, \dots, n\}, \#\mathbb{E}=h} \frac{1}{h} \sum_{i \in \mathbb{E}} \left( D \left( \mathcal{X}_i, \frac{1}{h} \cdot \sum_{j \in \mathbb{E}} \mathcal{X}_j \right) \right)^2$$

and  $h = n - \lfloor n\beta \rfloor$  is the cardinality of the trimming region  $\widehat{\mathbb{E}}(\mathcal{X}_1, \dots, \mathcal{X}_n)$ .

Given a sample dataset  $\tilde{\mathbf{x}}_n = (\tilde{x}_1, \dots, \tilde{x}_n)$ , the corresponding **sample trimming region** will be

$$\widehat{\mathbb{E}}_{\tilde{\mathbf{x}}_n} = \arg \min_{\mathbb{E} \subset \{1, \dots, n\}, \#\mathbb{E}=h} \frac{1}{h} \sum_{i \in \mathbb{E}} \left( D \left( \tilde{x}_i, \frac{1}{h} \cdot \sum_{j \in \mathbb{E}} \tilde{x}_j \right) \right)^2,$$

and the set  $\mathcal{E} = \{\mathbb{E} \subset \{1, \dots, n\} : \#\mathbb{E} = h\}$  consists of all the subsets of  $h$  different natural numbers between 1 and the sample size  $n$ . The **sample fuzzy trimmed mean** associated with the sample  $\tilde{\mathbf{x}}_n = (\tilde{x}_1, \dots, \tilde{x}_n)$  corresponds to  $\frac{1}{h} \cdot \sum_{j \in \widehat{\mathbb{E}}_{\tilde{\mathbf{x}}_n}} \tilde{x}_j$ .

Therefore, the trimming region, for a fixed proportion of trimming  $\beta$ , can be seen as the set containing the remaining proportion of sample data with minimum variance. This problem has at least one solution since it is a finite combinatorial problem. Once a trimming region  $\widehat{\mathbb{E}}_{\tilde{\mathbf{x}}_n}$  is determined, the associated trimmed mean and variance are simply the mean and variance of the sample conditioned to  $\widehat{\mathbb{E}}_{\tilde{\mathbf{x}}_n}$ .

Some properties of trimmed mean estimators that have not been analyzed for Hilbert-valued random elements in the literature yet will now be studied. It is very important to remark that, although these properties are formalized for random fuzzy set values, the proofs remain valid when general Hilbert-valued random elements are considered instead, so they are also applicable to, for instance, functional data.

First, Theorem 5 provides us with some sufficient conditions which guarantee that the fuzzy trimmed mean estimator is well-defined.

**Theorem 5** Consider a metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  on a complete probability space  $(\Omega, \mathcal{A}, P)$ . For any trimming proportion  $\beta \in (0, 1)$ , the fuzzy trimmed mean estimator  $\bar{\mathcal{X}}_{n,\beta}$  is well-defined.

As commented for fuzzy M-estimators of location, the completeness of the probability space will not be required in the following results since the conditions in Theorem 5 are only sufficient, but it will be assumed that the corresponding fuzzy trimmed mean estimator is well-defined.

#### 4.1 Equivariance and symmetry properties for fuzzy trimmed means

In this section, the translation and scale equivariance of fuzzy trimmed means estimators is shown, as well as their behaviour when the RFV is symmetrically distributed. As explained before, these results are indeed not only valid for the fuzzy-valued setting, but also for Hilbert space-valued random elements.

**Proposition 3** (*Translation equivariance*) Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  and let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  associated with a probability space  $(\Omega, \mathcal{A}, P)$ . For any trimming proportion  $\beta \in (0, 1)$ , the fuzzy trimmed mean estimator  $\overline{\mathcal{X}}_{n,\beta}$  is translation equivariant, i.e., for any  $\tilde{U} \in \mathcal{F}_c^2(\mathbb{R}^p)$ ,

$$\overline{(\mathcal{X} + \tilde{U})}_{n,\beta} = \overline{\mathcal{X}}_{n,\beta} + \tilde{U}.$$

Unlike fuzzy M-estimators of location, fuzzy trimmed mean estimators are scale equivariant, which presents the advantage of not being influenced by the choice of the measurement units.

**Proposition 4** (*Scale equivariance*) Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  and an RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  on a probability space  $(\Omega, \mathcal{A}, P)$ . Moreover, let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$  and  $\beta \in (0, 1)$  the trimming proportion. The fuzzy trimmed mean estimator  $\overline{\mathcal{X}}_{n,\beta}$  is scale equivariant, i.e., for any  $c \in \mathbb{R}$ ,

$$\overline{(c \cdot \mathcal{X})}_{n,\beta} = c \cdot \overline{\mathcal{X}}_{n,\beta}.$$

With respect to symmetry, fuzzy trimmed mean estimators have the same behaviour as fuzzy M-estimators of location, as shown by the next proposition.

**Proposition 5** Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be a random fuzzy set value associated with the probability space  $(\Omega, \mathcal{A}, P)$  and symmetrically distributed for some  $\mathbf{c} \in \mathbb{R}^p$ . For any trimming proportion  $\beta \in (0, 1)$ , the fuzzy trimmed mean estimator  $\overline{\mathcal{X}}_{n,\beta}$  is a symmetric fuzzy set about  $\mathbf{c}$ .

#### 4.2 Algorithmic computation of fuzzy trimmed means

In Colubi and González-Rodríguez (2015), the Fast Minimum Covariance Determinant (for short, Fast-MCD) algorithm, which is an adaptation of the well-known  $k$ -means algorithm, was rewritten to compute the trimmed mean of a fuzzy-valued sample. We now present an adaptation of this algorithm which tries to avoid, as much as possible, that any local minimum traps the iterative process:

**Step 1.** Set  $h = n - \lfloor n\beta \rfloor \in \{1, \dots, n\}$ , the size of the trimming region, fix  $NS$  the number of starting points and  $nbest$  the number of best trimming regions selected after  $nrep$  initial steps. Initialize  $MSE = \infty$ ;

**Step 2.** Choose at random either 3 observations from the sample (considered as seed) and compute their mean,  $\tilde{f}$ , or an observation considered as an initial mean,  $\tilde{f}$ . It is then possible to build a first region of size  $h$  centered around  $\tilde{f}$ . Since the observation(s) is(are) chosen randomly, this first region will also be random.

**Step 3.** Select the  $h$  observations closest to the mean  $\tilde{f}$  (where closeness refers to the distance corresponding to the norm associated with the inner product of the Hilbert space):

$$\{\tilde{x}_{k_1}, \dots, \tilde{x}_{k_h}\};$$

**Step 4.** Compute the mean  $\tilde{f}^*$  of the  $h$  observations in **Step 3**, and then calculate the corresponding mean squared error given by

$$\tilde{f}^* = \frac{1}{h} \cdot \sum_{i=1}^h \tilde{x}_{k_i}, \quad MSE^* = \frac{1}{h} \sum_{i=1}^h (D(\tilde{x}_{k_i}, \tilde{f}^*))^2;$$

In case  $MSE^*$  is lower than  $MSE$ , update the value of the upper bound, that is, set

$$MSE = MSE^*;$$

**Step 5.** **Steps 3** and **4** are repeated  $nrep$  times.

**Step 6.** Repeat **Steps 2-5**  $NS$  times and choose the  $nbest$  trimming regions with lowest associated  $MSE^*$ .

**Step 7.** For each of the  $nbest$  trimming regions, repeat **Steps 3-4** until convergence.

**Step 8.** The estimate associated with the smallest  $MSE^*$  is the final estimate of the fuzzy trimmed mean. Moreover, the corresponding value  $MSE^*$  is the trimmed Mean Squared Error associated with it.

Notice that, due to **Step 2**, this new algorithm always performs at least as well as Cuesta-Albertos and Fraiman's alternative (see Cuesta-Albertos and Fraiman, 2006). A weakness (fuzzy) trimmed means present is that a selection procedure should be developed to select the optimal trimming proportion  $\beta$  'a priori'.

#### 4.3 Consistency and robustness properties of fuzzy trimmed means

Finally, some properties of the fuzzy trimmed mean estimator will be studied. Recall that, even when these results are stated here for the fuzzy set-valued setting, they are still valid for trimmed means of general Hilbert-valued random elements.

The following result is an immediate application of the strong consistency results by Cuesta-Albertos and Fraiman (2006, 2007) for general Hilbert-valued random elements.

**Theorem 6** Consider a metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$  that can be isometrically embedded into a closed convex cone of a Hilbert space, and the RFV  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  associated with the probability space  $(\Omega, \mathcal{A}, P)$ . For any  $n \in \mathbb{N}$ , let  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$ . Let  $\beta \in (0, 1)$  and assume that the fuzzy trimmed mean  $E_{[\beta]}(\mathcal{X})$  is unique. Then, the fuzzy trimmed mean estimator is a strongly consistent estimator of the fuzzy trimmed mean, i.e.,

$$\lim_{n \rightarrow \infty} D(\overline{\mathcal{X}}_{n,\beta}, E_{[\beta]}(\mathcal{X})) = 0 \quad a.s.[P].$$

The finite sample breakdown point of the fuzzy trimmed mean estimator is now derived to show the robustness of this estimator.

**Theorem 7** Consider the metric space  $(\mathcal{F}_c^2(\mathbb{R}^p), D)$ . Let  $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c^2(\mathbb{R}^p)$  be an RFV associated with the probability space  $(\Omega, \mathcal{A}, P)$ ,  $(\mathcal{X}_1, \dots, \mathcal{X}_n)$  be a simple random sample from  $\mathcal{X}$  and  $\tilde{\mathbf{x}}_n = (\tilde{x}_1, \dots, \tilde{x}_n)$  be a realization of the simple random sample such that all the sample observations are different. For any trimming proportion  $\beta \in (0, 1)$ , the finite sample breakdown point of the fuzzy trimmed mean estimator is exactly given by

$$\text{fsbp}(\overline{\mathcal{X}}_{n,\beta}, \tilde{\mathbf{x}}_n, D) = \begin{cases} \frac{\lfloor n\beta \rfloor + 1}{n} & \text{if } \beta < 0.5 \\ \frac{n - \lfloor n\beta \rfloor}{n} & \text{if } \beta \geq 0.5 \end{cases}$$

Notice that the maximum,  $\frac{1}{n} \lfloor \frac{n+1}{2} \rfloor$ , will be attained when  $\beta = 0.5$ . For any other value of  $\beta$ , the finite sample breakdown point of fuzzy trimmed means is smaller and, therefore, their behaviour is less robust. Anyway, both location M-estimators and trimmed means present a more robust performance than the Aumann-type mean in terms of finite sample breakdown points.

## 5 Illustrative example

In this section, a two-dimensional real-life example is considered to illustrate fuzzy M-estimators and fuzzy trimmed means. Both proposals will also be compared to the location measure used in the literature, the Aumann-type mean, in order to show their advantages. Examples which are available in the literature for  $p = 2$  usually involve conical data (see e.g. Sugano, 2011, Salski, 2007, and Celmiņš, 1987, among others), but we consider more complex data, the Cartesian product of two trapezoidal fuzzy numbers, which models the natural fuzzy relation between the answers given in terms of a fuzzy rating scale (Hesketh et al., 1988) to two items of a questionnaire.

*Example.* 68 fourth grade students from Colegio San Ignacio school (Oviedo, Spain) have been asked to answer some questions selected from the joint Student questionnaire TIMSS (Trends in International Mathematics and Science Study) - PIRLS (Progress in International Reading Literacy Study). Items have been designed to use a fuzzy rating scale with referential  $[0, 10]$ , instead of a standard 4-point Likert scale.



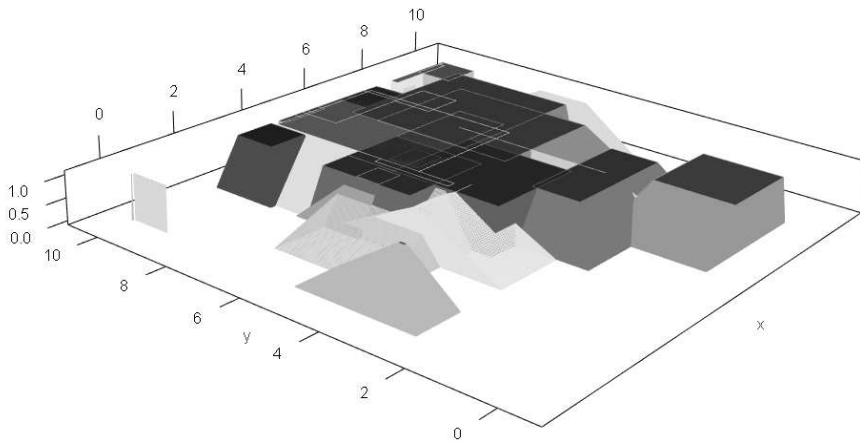
Statistical conclusions have been empirically shown to frequently differ depending on the considered psychometric rating scale (see, for instance, Lubiano et al., 2016). To simplify the instructions given to the 9-10 year-old students, only trapezoidal fuzzy numbers have been considered, as it has been empirically shown that the shape is not statistically relevant (see, for instance, Lubiano et al., 2017). A trapezoidal fuzzy number will be denoted as  $\text{Tra}(a, b, c, d)$  with  $a \leq b \leq c \leq d$ , where for each  $\alpha \in [0, 1]$ ,

$$(\text{Tra}(a, b, c, d))_\alpha = [\alpha b + (1 - \alpha)a, \alpha c + (1 - \alpha)d].$$

For this study, we have considered the responses given by the 31 female students about their degree of agreement with the following statements:

- “I like Mathematics” (question M1),
- “My Maths teacher is easy to understand” (question M2).

The joint/linked responses to both items can be properly modelled as two-dimensional fuzzy set values by considering a natural fuzzy relation, namely, the Cartesian product of the two trapezoidal fuzzy scores (see Zadeh, 1975, for which the  $\alpha$ -level corresponds to the cartesian product of the ‘individual’  $\alpha$ -levels), as displayed in Figure 1.



**Fig. 1** The sample of 31 fuzzy set values built from the (joint/linked) fuzzy scores of the girls to items M1 and M2. The  $z$ -axis represents the degree of compatibility (between 0 and 1) of the values  $x$  and  $y$  in  $[0, 10]$  with the students’ opinion regarding M1 and M2 items, respectively.

The Aumann-type mean, the fuzzy trimmed mean with two common choices for the trimming proportion, i.e.  $\beta = 0.15$  and  $\beta = 0.2$ , and the fuzzy M-estimator of location based on the Hampel loss function are compared by means of the metric  $D_{\theta=1/3}$ . The tuning parameters  $0 < a < b < c$  involved in

the Hampel loss function allow us to control the degree of suppression of the outliers, since the smaller their values the greater such a degree. They have been selected following the suggestion in Kim and Scott (2012), that is, by considering the median, the 75<sup>th</sup> and the 85<sup>th</sup> percentiles of the distances between each observation and a robust seed (which, in this case, has been chosen to be the fuzzy trimmed mean with  $\beta = 0.15$ ). For this dataset, this yields  $a = 5.633175$ ,  $b = 7.278062$  and  $c = 8.329507$ .

The computation of the considered location measures is based on the support function of the fuzzy sets. The outcome of each Cartesian product of two trapezoidal fuzzy numbers with referential  $[0, 10]$ ,  $\tilde{V}_1$  and  $\tilde{V}_2$ , is a truncated square pyramid  $\tilde{U}$  with support function given by  $s_{\tilde{U}} : (0, 1] \times \mathbb{S}^1 \rightarrow \mathbb{R}$  with

$$s_{\tilde{U}}(\alpha, \mathbf{u}) = \begin{cases} -\varrho_{\mathbf{v}_{21}} \cos(\pi + \vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{21}}) & \text{if } \vartheta_{\mathbf{u}} \in \left(-\frac{\pi}{2}, -\vartheta_{\mathbf{w}_{21}}\right), u_1 > 0 \\ \varrho_{\mathbf{v}_{21}} \cos(\vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{21}}) & \text{if } \vartheta_{\mathbf{u}} \in [-\vartheta_{\mathbf{w}_{21}}, 0), u_1 > 0 \\ \varrho_{\mathbf{v}_{22}} \cos(\vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{22}}) & \text{if } \vartheta_{\mathbf{u}} \in \left[0, \frac{\pi}{2}\right), u_1 > 0 \\ \varrho_{\mathbf{v}_{12}} \cos(\pi + \vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{12}}) & \text{if } \vartheta_{\mathbf{u}} \in \left(-\frac{\pi}{2}, -\vartheta_{\mathbf{w}_{21}}\right), u_1 < 0 \\ -\varrho_{\mathbf{v}_{12}} \cos(\vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{12}}) & \text{if } \vartheta_{\mathbf{u}} \in [-\vartheta_{\mathbf{w}_{21}}, 0), u_1 < 0 \\ -\varrho_{\mathbf{v}_{11}} \cos(\vartheta_{\mathbf{u}} - \vartheta_{\mathbf{v}_{11}}) & \text{if } \vartheta_{\mathbf{u}} \in \left[0, \frac{\pi}{2}\right), u_1 < 0, \end{cases}$$

where  $\varrho_{\mathbf{u}} = \varrho_{(u_1, u_2)} = \sqrt{u_1^2 + u_2^2}$ ,  $\mathbf{v}_{ij} = (x_i(\alpha), y_j(\alpha))$ ,  $\mathbf{w}_{ij} = (y_j(\alpha), x_i(\alpha))$ ,

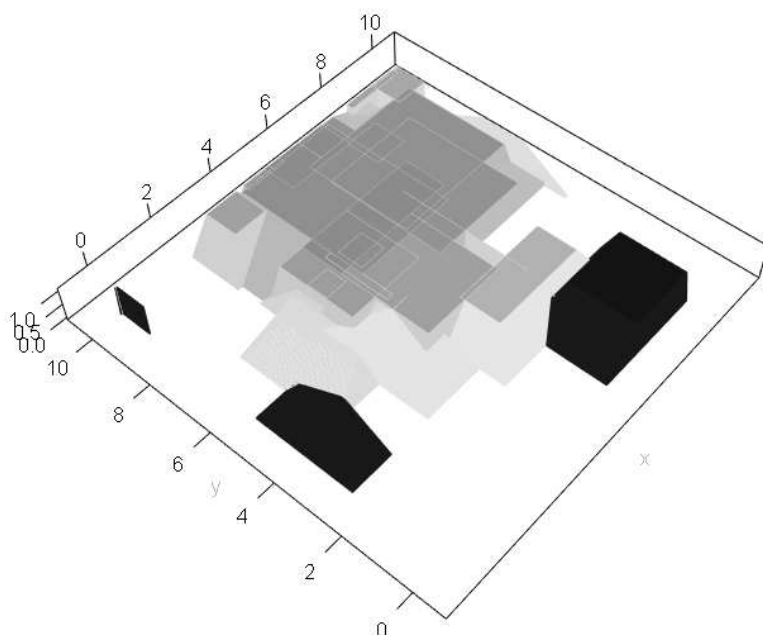
$$\vartheta_{\mathbf{u}} = \vartheta_{(u_1, u_2)} = \begin{cases} \arctan(u_2/u_1), & \text{if } u_1 \neq 0 \\ \frac{\pi}{2}, & \text{if } u_1 = 0, u_2 > 0 \\ -\frac{\pi}{2}, & \text{if } u_1 = 0, u_2 < 0, \end{cases}$$

$x_1(\alpha) = \inf(\tilde{V}_1)_\alpha$ ,  $x_2(\alpha) = \sup(\tilde{V}_1)_\alpha$ ,  $y_1(\alpha) = \inf(\tilde{V}_2)_\alpha$  and  $y_2(\alpha) = \sup(\tilde{V}_2)_\alpha$ .

The  $D_{1/3}$  distance between  $(0, 0)$  and each of the considered location measures has been computed: this distance equals 17.117 for the Aumann-type mean, 18.023 for the fuzzy trimmed mean with trimming proportion  $\beta = 0.15$ , 18.599 for the fuzzy trimmed mean with  $\beta = 0.2$ , and 18.260 for the fuzzy M-estimator of location based on Hampel's loss function with the above mentioned tuning parameters. The reason for the difference between the Aumann-type mean and the robust location measures is now explained. First, the fuzzy trimmed mean with  $\beta = 0.15$  excludes the four observations in Table 1 from the computations due to their higher distance to the 'centre' of the data.

**Table 1** Trimmed observations for the computation of the fuzzy trimmed mean ( $\beta = 0.15$ ).

Response to item M1	Response to item M2
Tra(5.850, 7.025, 9.050, 9.100)	Tra(0.000, 0.1250, 2.050, 2.550)
Tra(0.000, 0.025, 0.025, 0.025)	Tra(9.975, 9.975, 9.975, 10.000)
Tra(0.000, 1.125, 1.200, 1.275)	Tra(2.500, 3.750, 3.900, 5.450)
Tra(0.050, 0.050, 0.075, 0.075)	Tra(9.025, 9.025, 9.950, 9.950)



**Fig. 2** The sample of 31 fuzzy sets built from the answers of the girls to items M1 and M2. In black, the observations trimmed for the computation of the fuzzy trimmed mean.

As said before, these fuzzy sets correspond to the students with the lowest degrees of agreement with respect to at least one of the two items, which can be clearly seen from Figure 2. When  $\beta = 0.2$ , the fuzzy trimmed mean excludes the two extra observations in Table 2.

**Table 2** Trimmed observations for the computation of the fuzzy trimmed mean with the choice  $\beta = 0.2$ , but not with  $\beta = 0.15$ .

Response to item M1	Response to item M2
Tra(4.325, 5.025, 7.925, 8.500)	Tra(1.750, 2.500, 3.675, 3.675)
Tra(2.500, 2.550, 4.275, 4.300)	Tra(2.500, 4.625, 4.625, 6.900)

On the other hand, the weights assigned to the observations in Table 1 when computing the fuzzy M-estimator of location based on Hampel's loss function were 0 in all cases. Hence, the lowest degrees of agreement have a larger impact on the computation of the Aumann-type mean than on the other location measures, which present a more robust behaviour. Regarding the comparison of the fuzzy trimmed means and the fuzzy M-estimator of location, the latter

‘removes’ the 4 observations in Table 1 and the first observation in Table 2, and assigns the second smallest weight (0.011 out of the maximum weight 0.043) to the second observation in Table 2. Hence, in this case, it provides an intermediate solution and detects 5 outliers (it can be checked that the computation of the estimator using as a seed the fuzzy trimmed mean with  $\beta = 0.2$  also assigns a null weight to only 5 observations).

In Table 3, the distances between the Aumann-type mean, the fuzzy trimmed means and the fuzzy M-estimator of location based on Hampel’s loss function (with  $a = 5.633$ ,  $b = 7.278$  and  $c = 8.330$ ) computed for the whole dataset and three reduced datasets are shown. The reduced datasets have been obtained by removing either the four observations in Table 1, the five observations contained in Table 1 and the first row of Table 2 or the six observations in Tables 1-2.

**Table 3** Distances between the location measures computed for the whole dataset and three reduced datasets without 4, 5 or 6 observations.

Measure	4 observations	5 observations	6 observations
Aumann-type mean	1.035	1.202	1.516
Fuzzy trimmed mean ( $\beta = 0.15$ )	0.876	0.876	0.799
Fuzzy trimmed mean ( $\beta = 0.2$ )	1.477	1.826	1.559
Fuzzy M-estimator	0.820	0.820	0.283

Naturally, the removal of the fifth observation from the dataset does not have any impact on the computation of the fuzzy M-estimator (with respect to its performance when only the other four observations are deleted) because, as said before, it detects the five outliers from the beginning. On the other hand, the fuzzy trimmed mean with  $\beta = 0.15$  discards the 15% of the observations for the computation of the mean and that implies that the number of observations removed when the sample size is 27 (reduced dataset without 4 outliers) or 26 (reduced dataset without 5 outliers) is 4 and 3, respectively. Therefore, the size of the trimming region coincides in both situations (23 observations in each case). For this reason, only the Aumann-type mean and the fuzzy trimmed mean with  $\beta = 0.2$  present different values in the first two columns of Table 3.

It can be noticed that the fuzzy M-estimator is the location measure which achieves the lowest distance, followed by the fuzzy trimmed mean with  $\beta = 0.15$ . Contrarily, the Aumann-type mean computed for the whole dataset is clearly further from the Aumann-type mean computed for the reduced datasets (indeed, the higher number of outliers removed, the higher this distance) and something similar happens to the fuzzy trimmed mean with  $\beta = 0.2$  (the trimming proportion is too high for this sample and non-outlying observations are trimmed, so some important information about the ‘centre’ of the data is lost).

Table 4 completes the study with the computation of the distances between the four estimates (the Aumann-type mean, the two fuzzy trimmed means and

the fuzzy M-estimator of location based on Hampel's loss function) on the full dataset and the Aumann-type mean of the reduced datasets.

**Table 4** Distances between the location measures computed for the whole dataset and the Aumann-type mean of a reduced dataset without 4, 5 or 6 observations.

Measure	4 observations	5 observations	6 observations
Aumann-type mean	1.035	1.202	1.516
Fuzzy trimmed mean ( $\beta = 0.15$ )	0	0.343	0.619
Fuzzy trimmed mean ( $\beta = 0.2$ )	0.619	0.320	0
Fuzzy M-estimator	0.482	0.217	0.418

Obviously, the first row of Tables 3 and 4 coincide and, as expected, the robust estimators on the full dataset are close to the Aumann-type mean on the reduced datasets. The fact that all distances but the one for the fuzzy trimmed mean with  $\beta = 0.2$  increase when passing from 5 deleted observations to 6 is an indication that there are indeed 5 outliers in this dataset.

## 6 Concluding Remarks

Fuzzy M-estimators of location and trimmed means have been extended to  $\mathcal{F}_c^2(\mathbb{R}^p)$  and their properties have been analyzed. It has been proven, by means of their finite sample breakdown point, that the behaviour of both alternatives is more robust than for the Aumann-type mean. In terms of their finite sample breakdown point, fuzzy M-estimators of location are a more robust approach than fuzzy trimmed means when the trimming proportion is less than 0.5 (when the trimming proportion is exactly 0.5 both alternatives are equally robust). Both estimators share the translation equivariance, symmetry with respect to symmetrically distributed random fuzzy set values and strong consistency, but only fuzzy trimmed means are always scale equivariant. Finally, a real-life example has empirically shown the robustness of fuzzy M-estimators of location and fuzzy trimmed means as an alternative to the Aumann-type mean.

As future research lines, it would be interesting to use other tools from robust statistics, such as the influence function, so realistic parametric families of distributions on  $\mathcal{F}_c^2(\mathbb{R}^p)$  should be proposed first. Furthermore, scale equivariant fuzzy M-estimators could be defined by means of robust scale measures for fuzzy set-valued data (see de la Rosa de Sáa et al., 2017, for some robust scale estimators introduced for the one-dimensional case). Finally, hypothesis testing procedures related to these measures could be established.

## SUPPLEMENTARY MATERIAL

Online Resource: Proofs of the theoretical results stated in the article. (.pdf)

**Acknowledgements** The authors are grateful to the Editor and reviewers, as well as to their colleagues Prof. M. A. Gil and Prof. G. González-Rodríguez, for their insightful comments and suggestions.

### Conflict of interest

The authors declare that they have no conflict of interest.

### References

- Alfons A, Croux C, Gelper S (2013) Sparse least trimmed squares regression for analyzing high-dimensional large data sets. *Ann Appl Stat* 7(1):226–248
- Aneiros G, Cao R, Fraiman R, Genest C, Vieu P (2019) Recent advances in functional data analysis and high-dimensional statistics. *J Multivar Anal* 170:3–9
- Bobylev VN (1985) Support function of a fuzzy set and its characteristic properties. *Math Notes (USSR)* 37(4):281–285
- Castaing C, Valadier M (1977) *Convex Analysis and Measurable Multifunctions*, Lect. Notes Math., vol 580. Springer
- Celmiņš A (1987) Least squares model fitting to fuzzy vector data. *Fuzzy Sets and Systems* 22:245–269
- Colubi A, González-Rodríguez G (2015) Fuzziness in data analysis: Towards accuracy and robustness. *Fuzzy Sets and Systems* 281:260–271
- Cuesta-Albertos JA, Fraiman R (2006) Impartial trimmed means for functional data. In: Liu RY, Serfling R, Souvaine DL (eds) *Data Depth: Robust Multivariate Statistical Analysis, Computational Geometry and Applications*, DIMACS Series, vol 72, Amer. Math. Soc., pp 121–145
- Cuesta-Albertos JA, Fraiman R (2007) Impartial trimmed  $k$ -means for functional data. *Comput Statist Data Anal* 51(10):4864–4877
- Cuesta-Albertos JA, Gordaliza A, Matrán C (1997) Trimmed  $k$ -means: an attempt to robustify quantizers. *Ann Statist* 25(2):553–576
- Cuevas A, Febrero M, Fraiman R (2007) Robust estimation and classification for functional data via projection-based depth notions. *Comput Statist* 22(3):481–496
- Donoho DL, Huber PJ (1983) The notion of breakdown point. In: Bickel PJ, Doksum K, Jr Hodges JL (eds) *A Festschrift for Eric L. Lehmann*, Wadsworth, pp 157–184
- Fréchet M (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann I H Poincaré* 10:215–310
- García-Escudero LA, Gordaliza A, Mayo-Iscar A, Martín RS (2010) Robust clusterwise linear regression through trimming. *Comput Statist Data Anal* 54:3057–3069
- Gil MA, Colubi A, Terán P (2013) Random fuzzy sets: why, when, how. *BEIO* 30(1):5–29

- Hampel FR (1974) The influence curve and its role in robust estimation. *J Amer Statist Assoc* 69:383–393
- Hesketh T, Pryor R, Hesketh B (1988) An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences. *Int J Man-Mach Stud* 29:21–35
- Huber PJ (1964) Robust estimation of a location parameter. *Ann Math Statist* 35:73–101
- Huber PJ (1981) *Robust Statistics*. J. Wiley & Sons
- Hubert M, Rousseeuw P, Segaert P (2017) Multivariate and functional classification using depth and distance. *Adv Data Anal Classif* 11:445–466
- Kim JS, Scott CD (2012) Robust kernel density estimation. *J Mach Learn Res* 13:2529–2565
- Klement EP, Puri ML, Ralescu DA (1986) Limit theorems for fuzzy random variables. *Proc R Soc Lond Ser A Math Phys Eng Sci* 407:171–182
- López-Pintado S, Romo J (2009) On the concept of depth for functional data. *J Amer Statist Assoc* 104(486):718–734
- Lubiano MA, Montenegro M, Sinova B, de la Rosa de Sáa S, Gil MA (2016) Hypothesis testing for means in connection with fuzzy rating scale-based data: algorithms and applications. *European J Oper Res* 251:918–929
- Lubiano MA, Salas A, Gil MA (2017) A hypothesis testing-based discussion on the sensitivity of means of fuzzy data with respect to data shape. *Fuzzy Sets and Systems* 328:54–69
- Minkowski H (1903) Volumen und oberfläche. *Math Ann* 57:447–495
- Puri ML, Ralescu DA (1985) The concept of normality for fuzzy random variables. *Ann Probab* 13:1373–1379
- Puri ML, Ralescu DA (1986) Fuzzy random variables. *J Math Anal Appl* 114:409–422
- Rivera-García D, García-Escudero LA, Mayo-Iscar A, Ortega J (2019) Robust clustering for functional data based on trimming and constraints. *Adv Data Anal Classif* 13:201–225
- de la Rosa de Sáa S, Lubiano MA, Sinova B, Filzmoser P (2017) Robust scale estimators for fuzzy data. *Adv Data Anal Classif* 11(4):731–758
- Salski A (2007) Fuzzy clustering of fuzzy ecological data. *Ecol Inform* 2:262–269
- Sinova B, Gil MA, Van Aelst S (2016) M-estimates of location for the robust central tendency of fuzzy data. *IEEE Trans Fuzzy Syst* 24(4):945–956
- Sinova B, González-Rodríguez G, Van Aelst S (2018) M-estimators of location for functional data. *Bernoulli* 24(3):2328–2357
- Sugano N (2011) Fuzzy set theoretical approach to the tone triangular system. *J Comput* 6(11):2345–2356
- Trutschnig W, González-Rodríguez G, Colubi A, Gil MA (2009) A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. *Inform Sci* 179(23):3964–3972
- Valencia D, Lillo RE, Romo J (2019) A kendall correlation coefficient between functional data. *Adv Data Anal Classif* DOI 10.1007/s11634-019-00360-z, URL <https://doi.org/10.1007/s11634-019-00360-z>

- 
- Zadeh LA (1965) Fuzzy sets. *Inform Control* 8(3):338–353
- Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning–i. *Inform Sci* 8(3):199–249
- Zadeh LA (2008) Is there a need for fuzzy logic? *Inform Sci* 178:2751–2779