Advantages of M-estimators of location for fuzzy numbers based on Tukey's biweight loss function

Beatriz Sinova[a,*], Stefan Van Aelst[b,c]

[a]*Department of Statistics and O.R. and D.M., University of Oviedo, 33007 Oviedo, Spain*
[b]*Department of Mathematics, KU Leuven, 3001 Leuven, Belgium*
[c]*Department of Applied Mathematics, Computer Science and Statistics, Ghent University, 9000 Gent, Belgium*

Abstract

M-estimators of location have been adapted to summarize the central tendency of random fuzzy numbers in a robust way. Under mild conditions on the loss function, which include the well-known Huber and Hampel families of loss functions, fuzzy number-valued location M-estimators exist and can be expressed as weighted means of the observations. Huber and Hampel loss functions depend on one and three tuning parameters, respectively. Some empirical analyses have been developed to compare the finite-sample behavior of the corresponding location M-estimators when these tuning parameters are well-chosen quantiles of the distribution of distances from an initial estimate to each observation. In that sense, it has been shown that the flexibility of the three tuning parameters in the Hampel loss function makes the corresponding M-estimator more accurate than the M-estimator based on the Huber loss function in many situations. More recently, Tukey's biweight (or bisquare) loss function has also been used to compute M-estimators of location for random fuzzy numbers. The robustness of all these estimators has been proven by their finite sample breakdown point, but a simulation study to compare the finite-sample behavior of the M-estimators defined in terms of Hampel and Tukey loss functions is still lacking. This paper aims to develop such simulation results and analyze the advantages of choosing

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

the Tukey loss function.

Keywords: M-estimator of location, random fuzzy number, Hampel loss function, Tukey biweight loss function

2010 MSC: 62G35, 62–07

1. Introduction

The interest of the statistical analysis of fuzzy number-valued data mainly lies in the wide variety of real-life experiments, characterized by an underlying imprecision, that this type of data can model mathematically. Indeed, fuzzy numbers are especially useful to describe ratings, opinions, judgements, perceptions and other data often in connection with human valuations in a natural and very expressive way. As discussed in De la Rosa de Sáa *et al.* [1], in these situations, responses cannot usually be expected to be expressible in terms of values on a precise scale, since they are essentially imprecise. For this reason, numerous methodologies have already been adapted to cover the fuzzy-valued setting. Most of these procedures are based on the use of the Aumann-type mean as location measure. The Aumann-type mean (see [2]) is a generalization of the concept of mean for real-valued random variables, from which it inherits not only very convenient statistical and probabilistic properties, but also the high sensitivity to outliers or data changes. Unfortunately, this means that the statistical conclusions of the developed methods involving the Aumann-type mean may be invalidated under data contamination, which is very frequent in real-life experiments.

In order to limit the influence of outliers on the location estimate, some robust measures have already been proposed in the literature. Among them, we could mention different extensions of the concept of median to the fuzzy number-valued case (see e.g. [3] and [4]) or, more recently, the more general M-estimation approach ([5]).

In the classical framework, M-estimators of location were introduced by Huber [6] as a way to overcome the lack of robustness of the standard least squares and maximum likelihood estimators. The key aspect is the involvement of a loss function applied to the errors of the data that is selected to be less rapidly increasing than the square loss function that is used in the least squares or maximum likelihood procedures. There exist several well-known families of loss functions, such as Huber, Hampel and Tukey biweight (or bisquare), that can be used for the computation of M-estimators.

Regarding the adaptation of this notion to the fuzzy number-valued case, Sinova *et al.* [5] proved that under mild conditions on the loss function, fuzzy number-valued location M-estimators exist and can be expressed as weighted means of the observations. Although only Huber and Hampel families of loss functions were explicitly mentioned in [5], it was commented in [7] that the Tukey biweight loss function could be used as well to compute M-estimators of location for random fuzzy numbers. That is to say, the three families of loss functions fulfill the required conditions to guarantee the existence of the corresponding fuzzy number-valued M-estimators of location and express them as convex linear combinations of the sample observations. Moreover, the robustness of the M-estimators of location based on these three choices has already been proven by means of their finite sample breakdown point. However, once M-estimators of location based on Huber and Hampel loss functions have already been studied, is there still some interest in introducing the ones based on the Tukey biweight loss function? In other words, does the Tukey biweight loss function present any advantage in contrast to the previously studied loss functions?

Huber and Hampel loss functions depend on one and three nonnegative tuning parameters, respectively. Although different values for these parameters could be proposed, a quite convenient way to take into account the measurement units of the problem and avoid the consequences of the lack of scale equivariance of M-estimators of location consists of choosing the values through quantiles of the distribution of distances from an initial location estimate to each observation. This approach was suggested by Kim and Scott in [8] and followed by Sinova *et al.* [5]. In the latter paper, an empirical comparison of the use of

Huber and Hampel loss functions with tuning parameters selected in this way (the parameter for Huber's being the median of the distribution of the distances to the initial estimate and the three parameters for Hampel's being the median, the 75th and the 85th percentiles) showed that Hampel's was, in general, a more accurate choice because of the flexibility allowed by the larger number of tuning parameters.

Tukey biweight loss function has a shape more similar to the Hampel loss function since it is not convex, contrary to the Huber loss function, and it is bounded. For that reason, we find their comparison particularly interesting. Indeed, the common aspects of the two loss functions allow us to develop a deeper empirical study than the one presented in [5]. Once we have chosen the three tuning parameters involved in Hampel's loss function as detailed above, our aim will be to conclude whether it is possible to select the tuning parameter for the Tukey biweight loss function in such a way that the finite-sample performance of the fuzzy number-valued M-estimator of location based on Tukey's loss function improves the results achieved by the M-estimator based on Hampel's loss function.

The rest of the paper is structured as follows. In Section 2, the preliminaries on the space of fuzzy numbers and their main features are recalled. Location M-estimators for random fuzzy numbers are introduced in Section 3 and their main properties, including strong consistency and robustness, are summarized. The empirical comparison of the finite-sample behaviour of the M-estimators of location for random fuzzy numbers, which is the main focus of this paper, is presented in Section 4 and illustrated by means of a real-life example in Section 5. Finally, some concluding remarks are provided in Section 6.

2. Preliminaries on the space of fuzzy numbers

Let $\mathcal{F}_c(\mathbb{R})$ denote the class of (bounded) fuzzy numbers, that is, the mappings $\widetilde{U} : \mathbb{R} \to [0,1]$ such that their α-levels

$$\widetilde{U}_\alpha = \begin{cases} \{x \in \mathbb{R} : \widetilde{U}(x) \geq \alpha\} & \text{if } \alpha \in (0,1] \\ \mathrm{cl}\{x \in \mathbb{R} : \widetilde{U}(x) > 0\} & \text{if } \alpha = 0, \end{cases}$$

are nonempty compact intervals. $\widetilde{U}(x)$ can be interpreted as the 'degree of compatibility' of x with $\widetilde{U}$' (or 'degree of truth' of the assertion "x is $\widetilde{U}$").

As can be deduced from this definition, the levels involved in fuzzy number-valued data add a certain gradualness to the imprecision of interval-valued data. This fact makes fuzzy data very convenient to describe ratings, opinions and other imprecise human assessments.

With respect to the arithmetic to handle these data, Zadeh's extension principle, which extends level-wise the usual interval arithmetic, is generally considered.

Definition 2.1. *Let $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$. The* ***sum*** *of $\widetilde{U}$ and $\widetilde{V}$ is defined as the fuzzy number $\widetilde{U} + \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ given for each $\alpha \in [0,1]$ by*

$$(\widetilde{U} + \widetilde{V})_\alpha = \textit{Minkowski sum of } \widetilde{U}_\alpha \textit{ and } \widetilde{V}_\alpha = \left[\inf \widetilde{U}_\alpha + \inf \widetilde{V}_\alpha, \sup \widetilde{U}_\alpha + \sup \widetilde{V}_\alpha\right].$$

Let $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma \in \mathbb{R}$. The ***product*** *of $\widetilde{U}$* ***by the scalar*** *γ is defined as the fuzzy number $\gamma \cdot \widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ given for each $\alpha \in [0,1]$ by*

$$(\gamma \cdot \widetilde{U})_\alpha = \gamma \cdot \widetilde{U}_\alpha = \begin{cases} \left[\gamma \cdot \inf \widetilde{U}_\alpha, \gamma \cdot \sup \widetilde{U}_\alpha\right] & \textit{if } \gamma \geq 0 \\ \left[\gamma \cdot \sup \widetilde{U}_\alpha, \gamma \cdot \inf \widetilde{U}_\alpha\right] & \textit{otherwise.} \end{cases}$$

It should be remarked that the space $(\mathcal{F}_c(\mathbb{R}), +, \cdot)$ is not linear, so there is no "difference operation" between fuzzy numbers that is well-defined and, at the same time, keeps the main properties of the difference between real values in connection with the sum. For that reason, differences involved in statistical developments for real numbers are often replaced by distances between fuzzy numbers whenever this is possible (see [9] and [10] for more details about the distance-based analysis of fuzzy number-valued data).

The following family of distances between fuzzy numbers has been chosen due to its versatility. It was introduced by Montenegro *et al.* [11] (and, later, generalized to fuzzy vector-valued data in [12]) as an extension of the family of metrics proposed by Bertoluzza *et al.* [13].

Definition 2.2. *Let $\theta \in (0, +\infty)$ and let φ be an absolutely continuous probability measure on $([0,1], \mathcal{B}_{[0,1]})$ with the mass function being positive on $(0,1)$. The* ***mid/spr-based*** L^2 ***distance*** *between any two fuzzy numbers $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ is defined as*

$$D_\theta^\varphi(\widetilde{U}, \widetilde{V}) = \left[\int_{[0,1]} \left(\operatorname{mid} \widetilde{U}_\alpha - \operatorname{mid} \widetilde{V}_\alpha \right)^2 d\varphi(\alpha) \right.$$

$$\left. + \theta \int_{[0,1]} \left(\operatorname{spr} \widetilde{U}_\alpha - \operatorname{spr} \widetilde{V}_\alpha \right)^2 d\varphi(\alpha) \right]^{1/2},$$

where $\operatorname{mid} \widetilde{U}_\alpha = (\inf \widetilde{U}_\alpha + \sup \widetilde{U}_\alpha)/2$ *and* $\operatorname{spr} \widetilde{U}_\alpha = (\sup \widetilde{U}_\alpha - \inf \widetilde{U}_\alpha)/2$.

The nonnegative parameter θ included in the definition of the mid/spr-based L^2 distance weighs the relative importance assigned to the 'deviation in shape' (represented by spr) in contrast to the 'deviation in center' (represented by mid). Neither θ nor the measure φ have a stochastic meaning. Some interesting choices for θ are 1 (D_1^φ generalizes the well-known distance ρ_2 by Diamond and Kloeden [14]) and 1/3 ($D_{1/3}^\varphi$ assigns the same relevance to all the points in the intervals, since the metric can be rewritten as

$$D_{1/3}^\varphi(\widetilde{U}, \widetilde{V}) = \sqrt{\int_{[0,1]} \int_{[0,1]} \left(\widetilde{U}_\alpha^{[\eta]} - \widetilde{V}_\alpha^{[\eta]} \right)^2 d\ell(\eta)\, d\varphi(\alpha)},$$

with $\ell =$ Lebesgue measure on $([0,1], \mathcal{B}_{[0,1]})$ and $\widetilde{U}_\alpha^{[\eta]} = \eta \cdot \sup \widetilde{U}_\alpha + (1-\eta) \cdot \inf \widetilde{U}_\alpha$ for all $\eta \in [0,1]$).

The fact that the space $(\mathcal{F}_c(\mathbb{R}), D_\theta^\varphi)$ can be isometrically embedded into a convex cone of a certain Hilbert space by means of the so-called support function (in Puri and Ralescu's sense [15]) will be crucial for the adaptation of M-estimators of location. For more details about this identification, please check, among others, [12, 16, 17].

Regarding the mathematical modelling of the random mechanism generating fuzzy data, the notion of random fuzzy number in Puri and Ralescu's sense [15, 18] will be considered. Notice that randomness affects the generation of the data, whereas fuzziness is assumed to affect the nature of the data.

Definition 2.3. *Let $(\Omega, \mathcal{A}, P)$ be a probability space modeling a random experiment. A mapping $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ is said to be a* ***random fuzzy number*** *associated with the random experiment if, and only if, for each $\alpha \in [0,1]$ the interval-valued mapping $\mathcal{X}_\alpha$ (where $\mathcal{X}_\alpha(\omega) = \big(\mathcal{X}(\omega)\big)_\alpha$ for all $\omega \in \Omega$) is a random compact interval or equivalently, the real-valued functions* $\inf \mathcal{X}_\alpha$ *and* $\sup \mathcal{X}_\alpha$ *are random variables.*

Indeed, a random fuzzy number is Borel-measurable with respect to the Borel σ-field associated with the mid/spr-based L^2 distance, among other metrics (see e.g. [17]), which guarantees that notions like the *induced distribution of a random fuzzy number* or the *stochastic independence of random fuzzy numbers* (and, hence, the notion of simple random sample from a random fuzzy number) can be immediately stated.

Certainly, the best-known measure to summarize the central tendency of a random fuzzy number is the Aumann-type mean given by

Definition 2.4. *Let $\mathcal{X}$ be a random fuzzy number and assume that the expected values of the random variables* $\inf \mathcal{X}_0$ *and* $\sup \mathcal{X}_0$ *are finite. The* ***Aumann-type mean*** *of $\mathcal{X}$ is the fuzzy number $\widetilde{E}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ such that for each $\alpha \in [0,1]$*

$$\big(\widetilde{E}(\mathcal{X})\big)_\alpha = [E(\inf \mathcal{X}_\alpha), E(\sup \mathcal{X}_\alpha)] \,.$$

The Aumann-type mean of a random fuzzy number fulfills very nice properties inherited from the classical settings, such as Strong Laws of Large Numbers (see [19]) and the Fréchet approach [20] using the D_θ^φ metric. Unfortunately, also the lack of robustness has been inherited, which has motivated the search for location measures with a more robust behavior. Several extensions of the concept of median have been proposed (see [3, 4] for more details) and, more recently, M-estimators of location have been adapted to cover the fuzzy number-valued setting. Such an adaptation will be carefully reviewed in the next section.

3. Location M-estimators for random fuzzy numbers

In this section, the adaptation of location M-estimators to the fuzzy number-valued setting will be reviewed. As already said in the introduction, M-estimators of location were first introduced by Huber [6] as a way to overcome the high sensitivity of the standard least squares and maximum likelihood estimators to outliers or data changes, by replacing the square loss function by a less rapidly increasing function. We assume that the loss function, denoted by ρ, satisfies the conditions

C.1 $\rho : \mathbb{R}^+ \to \mathbb{R}$ is continuous and non-decreasing, and $\rho(0) = 0$.

The M-estimators of location for fuzzy number-valued data are then defined as follows.

Definition 3.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be an associated random fuzzy number. Moreover, let $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ be a simple random sample from $\mathcal{X}$. Then, the* ***fuzzy M-estimator of location*** *is the fuzzy number-valued statistic $\widehat{\widetilde{g}^M}[(\mathcal{X}_1, \ldots, \mathcal{X}_n)]$, given by*

$$\widehat{\widetilde{g}^M}[(\mathcal{X}_1, \ldots, \mathcal{X}_n)] = \arg \min_{\widetilde{U} \in \mathcal{F}_c(\mathbb{R})} \frac{1}{n} \sum_{i=1}^{n} \rho(D_\theta^\varphi(\mathcal{X}_i, \widetilde{U})),$$

if it exists.

The Representer Theorem (Theorem 3.1) is a consequence of Kim and Scott's theory on robust kernel density estimation (see [8]). Although they studied Hilbert spaces in this particular context, their results remain valid for any Hilbert space. Taking into account that, as pointed out, fuzzy numbers can be identified with a convex cone of a Hilbert space thanks to an isometrical embedding, the theorem has been adapted to cover fuzzy M-estimators of location. Note that the adaptation of any methodology from Hilbert Space-Valued

Data Analysis to the fuzzy-valued framework must guarantee that the result does not move out of the convex cone. In this case, this is easy to ensure since the Representer Theorem below expresses fuzzy M-estimators of location as a weighted mean of the observations and such an operator is closed in $\mathcal{F}_c(\mathbb{R})$.

Theorem 3.1. [5] *Let $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ be a simple random sample from a random fuzzy number $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ on a probability space $(\Omega, \mathcal{A}, P)$. Moreover, let the loss function ρ satisfy the assumptions*

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

C.3 Let $\phi(x) = \rho'(x)/x$ and $\phi(0) \equiv \lim_{x\to 0} \phi(x)$. Assume that $\phi(0)$ exists and is finite.

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

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

with $\omega_i \geq 0$, $\sum_{i=1}^{n} \omega_i = 1$. Furthermore, $\omega_i \propto \phi(D_\theta^\varphi(\mathcal{X}_i, \widehat{g}^M[(\mathcal{X}_1, \ldots, \mathcal{X}_n)]))$.

The conditions on ρ assumed in the Representer Theorem are not too strong, since well-known families such as Huber, Tukey and Hampel fulfill them (see [5] for the study of Huber and Hampel families of loss functions and [7] for the analysis of Tukey's family). These families of loss functions are now recalled.

The *Huber loss function* [21] given by

$$\rho_a^H(x) = \begin{cases} x^2/2 & \text{if } 0 \leq x \leq a \\ a(x - a/2) & \text{otherwise,} \end{cases}$$

with $a > 0$ a tuning parameter, is a convex function and puts less emphasis on large errors compared to the squared error loss. On the other hand, the *Tukey biweight or bisquare* [22] family of loss functions is given by

$$\rho_d^T(x) = \begin{cases} d^2/6 \cdot (1 - (1 - (x/d)^2)^3) & \text{if } 0 \leq x \leq d \\ d^2/6 & \text{otherwise,} \end{cases}$$

with tuning parameter $d > 0$, and the *Hampel loss function* [23] corresponds to

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

where the nonnegative parameters $a < b < c$ allow us to control the degree of suppression of large errors. The smaller their values, the higher this degree. Both the Tukey biweight and Hampel families of loss functions are not convex anymore and can better cope with extreme outliers, since observations far from the center ($x \geq d$ or $x \geq c$) all contribute equally to the loss.

Even while the Tukey biweight and Hampel loss functions share key aspects, and the Hampel loss function had already been used for the computation of fuzzy M-estimators of location, it has also been interesting to study fuzzy M-estimators of location based on the Tukey biweight loss function. As Sinova and Van Aelst comment in [7], "the benefit of the Tukey loss function is to combine the better performance of Hampel's loss function regarding extreme outliers with the simplicity of an expression depending on just one tuning parameter, like the Huber loss function".

Notice that the expression provided by the Representer Theorem, which presents fuzzy M-estimators of location as a weighted mean of the observations, is implicit, since the fuzzy M-estimator of location itself is involved in the computation of the weights. For that reason, an algorithm is required to approximate the value of the fuzzy M-estimator of location in practice. A natural choice is the standard iteratively re-weighted least squares algorithm:

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

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

$$\widetilde{g}^M_{(k)} = \sum_{i=1}^n \omega_i^{(k-1)}\,\mathcal{X}_i, \quad \omega_i^{(k)} = \frac{\phi(D_\theta^\varphi(\mathcal{X}_i, \widetilde{g}^M_{(k)}))}{\sum_{j=1}^n \phi(D_\theta^\varphi(\mathcal{X}_j, \widetilde{g}^M_{(k)}))}.$$

Step 3. Terminate the algorithm when

$$\frac{|\frac{1}{n}\sum_{i=1}^n \rho(D_\theta^\varphi(\mathcal{X}_i, \widetilde{g}^M_{(k+1)})) - \frac{1}{n}\sum_{i=1}^n \rho(D_\theta^\varphi(\mathcal{X}_i, \widetilde{g}^M_{(k)}))|}{\frac{1}{n}\sum_{i=1}^n \rho(D_\theta^\varphi(\mathcal{X}_i, \widetilde{g}^M_{(k)}))} < \varepsilon,$$

for some desired tolerance $\varepsilon > 0$.

It has been shown that Kim and Scott's suggestion of selecting the tuning parameters involved in the loss functions on the basis of the distribution of the distances to the center provides us with good estimates (see [8, 5]). More precisely, first an initial robust estimator of location is computed (e.g., the 1-norm median in [4]) and then, we calculate the distance between each observation and this initial estimate. Following Kim and Scott's suggestion, a, b and c are taken to be the median, the 75th and the 85th percentiles of those distances.

Several properties of fuzzy M-estimators of location have been theoretically studied in [5, 7]. In particular, all fuzzy M-estimators of location based on a loss function fulfilling the conditions required for the Representer Theorem have been proven to be translation equivariant. Scale equivariance does not hold unless ρ is a power function, but the selection procedure chosen for the tuning parameters and described above avoids the influence of the measurement units on the estimate.

The following theorem provides us with a set of sufficient conditions to guarantee the strong consistency of fuzzy M-estimators of location. Note that the space is limited to fuzzy numbers defined on a bounded referential, which is very common in practice (see e.g. [24]). The sufficient assumptions on ρ include the Huber, Tukey and Hampel families of loss functions, among others.

Theorem 3.2. [7] *Consider the metric space $(\mathcal{F}_c([q,r]), D_\theta^\ell)$, with $q < r$. Let $\mathcal{X} : \Omega \to \mathcal{F}_c([q,r])$ be a random fuzzy number associated with a probability space $(\Omega, \mathcal{A}, P)$. Under any of the following assumptions:*

- *ρ is subadditive and unbounded,*
- *ρ has linear upper and lower bounds with the same slope,*
- *ρ is bounded,*

and whenever the associated M-location value

$$\tilde{g}^M(\mathcal{X}) = \arg \min_{\widetilde{U} \in \mathcal{F}_c([q,r])} E\left[\rho\left(D_\theta^\ell(\mathcal{X}, \widetilde{U})\right)\right]$$

exists and is unique, the M-estimator of location is a strongly consistent estimator of $\tilde{g}^M(\mathcal{X})$, i.e.,

$$\lim_{n\to\infty} D_\theta^\ell(\widehat{\widetilde{g}^M}[(\mathcal{X}_1, \dots, \mathcal{X}_n)], \tilde{g}^M(\mathcal{X})) = 0 \quad a.s.\ [P].$$

Finally, the robustness of the fuzzy M-estimation approach can be theoretically proven by means of the finite sample breakdown point (fsbp). This measure (see [25, 26]) indicates the minimum proportion of sample observations that have to be perturbed to make the distance between the M-estimates of the original data and contaminated data arbitrarily large.

Theorem 3.3. [5, 7] *Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$ be a random fuzzy number associated with a probability space $(\Omega, \mathcal{A}, P)$ and let $(\widetilde{x}_1, \dots, \widetilde{x}_n)$ be a sample of observations obtained from $\mathcal{X}$. Moreover, let ρ be a loss function fulfilling the assumptions in Theorem 3.1, such that the corresponding sample M-estimate of location is unique. Then, the finite sample breakdown point of the corresponding location M-estimator is at most $\frac{1}{n}\lfloor\frac{n+1}{2}\rfloor$, where $\lfloor\cdot\rfloor$ denotes the floor function.*

Moreover, under any of these additional assumptions,

- *ρ has linear upper and lower bounds with the same slope,*
- *ρ is bounded by a constant $C < \infty$ and satisfies*

$$\rho\left(\max_{1\le i,j\le n} D_\theta^\ell(\widetilde{x}_i, \widetilde{x}_j)\right) < \frac{n - 2\lfloor\frac{n-1}{2}\rfloor}{n - \lfloor\frac{n-1}{2}\rfloor - 1} \cdot C,$$

the finite sample breakdown point is exactly $\frac{1}{n}\lfloor\frac{n+1}{2}\rfloor$.

4. Empirical comparison

In Section 3, it has been recalled that the Huber, Tukey and Hampel families of loss functions fulfill the sufficient conditions stated by the Representer Theorem to guarantee the existence of the corresponding fuzzy number-valued M-estimators of location and express them as convex linear combinations of the sample observations. Regarding the robustness of these fuzzy M-estimators of location, the three measures based on the Huber, Tukey and Hampel loss functions share the same finite sample breakdown point, which means that the impact of global data contamination is the same. Under some assumptions, it has been proven that they can even withstand up to 50% data contamination, which is the maximum for translation equivariant location estimators.

Fuzzy M-estimators of location were first computed using the Huber and Hampel families of loss functions (see [5]), whereas Tukey's biweight loss function has been considered more recently [7]. Taking into account the good theoretical and empirical performance of fuzzy M-estimators of location based on the Huber and Hampel loss functions, we could wonder about the necessity of introducing M-estimators based on Tukey's biweight loss function.

The empirical studies in [5] showed that in general the Hampel loss function is a more suitable choice than the Huber loss function when the tuning parameters are chosen as explained in Section 3. It should be clarified that the procedure suggested by Kim and Scott in [8] to select the values for the tuning parameters seems to be specifically designed for the Hampel loss function and the procedure was adapted to cover the Huber loss function as well. Naturally, other choices for the value of the tuning parameter in the Huber loss function could be explored depending on each dataset, as will be done for the Tukey biweight loss function in this section.

The aim is to analyze the potential advantages of using the Tukey loss function instead of the Hampel loss function for the definition and computation of fuzzy M-estimators of location. We have not included Huber's loss function in the computations of this section for the following two reasons. On the one hand, the flexibility allowed by the three tuning parameters in the Hampel loss function makes the comparison Hampel versus Tukey more interesting than Huber versus Tukey (since the latter families of loss functions depend on only one tuning parameter). On the other hand, the Hampel and the Tukey loss functions are closer in shape than the Huber and the Tukey loss functions.

The simulation study developed in this section completes the variety of situations considered in [5]. However, the focus of the simulations is not centered on the comparison of the Hampel M-estimator and the Tukey M-estimator for a certain value of the tuning parameter that appears in the Tukey biweight loss function, but on the search for a value of such a tuning parameter to improve the results from the Hampel M-estimator. That is, once the tuning parameters for the Hampel loss function have been selected following Kim and Scott's suggestion and the Hampel M-estimator of location has been computed, our interest lies in the search for a value of the tuning parameter involved in the Tukey loss function such that the corresponding Tukey M-estimator provides a better estimate than the Hampel M-estimator.

4.1. Comparison of Tukey's and Hampel's loss functions

Since the Tukey biweight loss function has a shape similar to that of the Hampel loss function, as has already been mentioned, we first compare both loss functions. Figure 1 shows the Tukey loss function with tuning parameter $d = 4.685$ and the Hampel loss function with $a = 2$, $b = 4$ and $c = 8$. Both are not convex (contrary to the Huber loss function) and they are bounded.

As can be seen in Figure 1, both loss functions assign a constant value to all the observations that lie far from the center ($x \geq d$ or $x \geq c$), but this value, which coincides with the upper bound of ρ, differs from one loss function to another. Tukey's biweight loss function is bounded by $d^2/6$ and Hampel's loss function by $a(b+c-a)/2$. For the comparison, it would be reasonable to bound both loss functions by the same constant. That is, once the tuning parameters a, b and c have been selected, d would be chosen to satisfy $d^2/6 = a(b+c-a)/2$. Solving the equation yields $d = \sqrt{3a(b+c-a)}$.

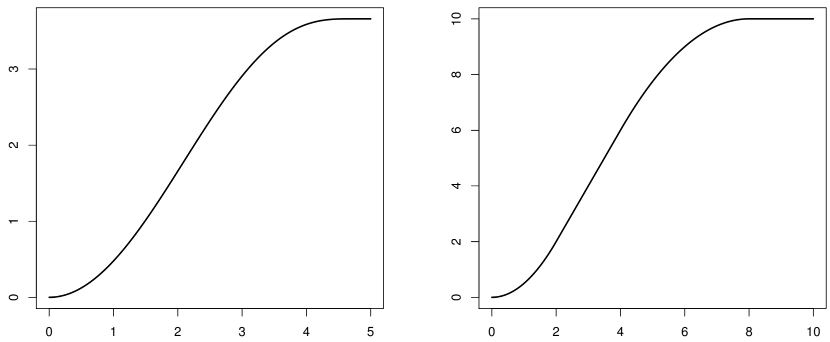

Figure 1: Tukey loss function with $d = 4.685$ (left) and Hampel loss function with $a = 2$, $b = 4$ and $c = 8$ (right)

For example, if a, b and c have been selected as in Figure 1, then correspondingly $d = 7.745967$. Loss functions $\rho_{2,4,8}$ and $\rho^T_{7.745967}$ have been displayed in Figure 2 (left) to easily check how close they are. Functions $\phi_{2,4,8}$ and $\phi^T_{7.745967}$ in Figure 2 (right) allow us to easily compare the weights that the M-estimators of location based on the loss functions $\rho_{2,4,8}$ and $\rho^T_{7.745967}$ assign to the outlying observations, since the weights from the Representer Theorem are proportional to the ϕ function evaluated on the corresponding distances.

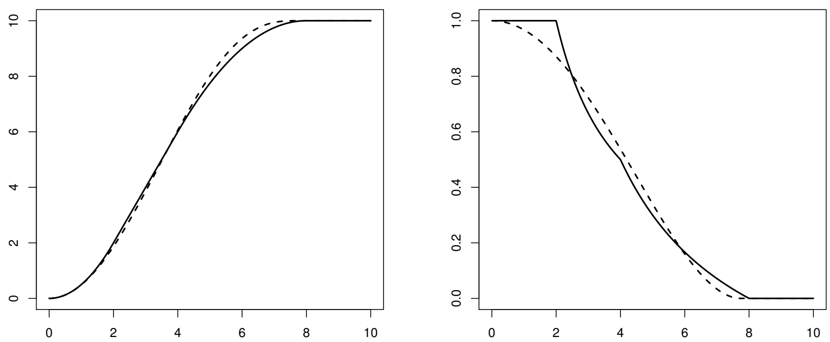

Figure 2: Tukey loss function with $d = 7.745967$ (dashed line) and Hampel loss function with $a = 2$, $b = 4$ and $c = 8$ (left) and their ϕ function (right)

In this case, $\sqrt{3a(b+c-a)} < c$ and thus the Tukey biweight loss function classifies as extreme more observations than the Hampel loss function. This means that if there is more data contamination than the one handled by Hampel's loss function (those observations with $x \geq c$, with c the 85th percentile as explained before), the Tukey biweight loss function better copes with the remaining outliers (the outlying observations with $\sqrt{3a(b+c-a)} \leq x < c$). However, the final performance of the Tukey biweight loss function also depends on the importance it assigns to small errors (i.e., whether it focuses more or less than the Hampel loss function on those errors). On the other hand, if the Hampel loss function already treats all the outliers as extreme observations, the Tukey ρ_d^T loses some valid information when considering some non-contaminated observations as outliers.

Analogously, the case $\sqrt{3a(b+c-a)} > c$ can be studied. If all the outliers are identified correctly by the Tukey loss function, then the Hampel loss function may classify some observations as outliers when they are not. The final choice between the Tukey and Hampel loss functions depends on their focus on small errors. On the other hand, in case not all data contamination is handled correctly by Tukey (since there exist some outlying observations with $x < \sqrt{3a(b+c-a)}$), the Hampel loss function may better cope with the remaining outlying observations (at least with those which fulfill $x \geq c$).

4.2. A proposal for the tuning parameter in the Tukey loss function

We could now try to slightly modify the proposal $d = \sqrt{3a(b+c-a)}$ in order to improve these results, i.e., to get an M-estimator of location based on the Tukey loss function with a better finite-sample performance than the one based on the Hampel loss function in as many situations as possible.

- First, when the Hampel loss function identifies all the outliers correctly, and Tukey's loss function may lose some valid information ($d \leq c \leq x$), we could think of choosing the loss function ρ_c^T instead of $\rho^T_{\sqrt{3a(b+c-a)}}$.
- Analogously, when the Hampel loss function does not necessarily identify

all the outliers, but at least identifies more extreme observations than the Tukey loss function (i.e., $d > c$ and some outliers satisfy $0 \leq x < d$), we could also choose the loss function ρ_c^T.

- Recall that, as it has already been commented, the choice $\rho^T_{\sqrt{3a(b+c-a)}}$ in the other two situations ($d \leq c$ and some outliers fulfilling $0 \leq x < c$, and $d > c$ and all outliers with $x \geq d$) does not always guarantee that the Tukey fuzzy M-estimator of location will present a better performance than the Hampel fuzzy M-estimator of location. Notice that in case both of them identify the outliers correctly, the final decision will depend on the weights assigned to small errors, which have not been discussed above.

With respect to the choice of ρ_c^T in some cases, the comparison between ρ_c^T and the Hampel loss function $\rho_{a,b,c}$, as shown in Figure 3, is addressed by multiplying ρ_c^T by the constant $3a(b+c-a)/c^2$, so both loss functions have the same upper bound.

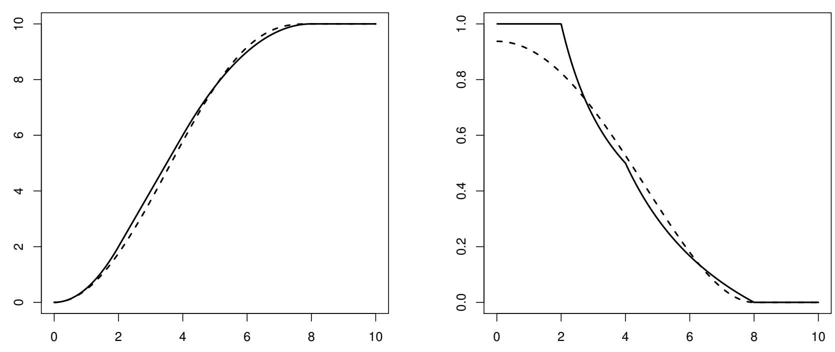

Figure 3: $0.9375 \cdot \rho_8^T$ (dashed line) and $\rho_{2,4,8}$ loss functions (left) and their ϕ function (right)

It can be proven using the derivatives of the loss functions $3a(b+c-a)/c^2 \cdot \rho_c^T$ and $\rho_{a,b,c}$ that the first one assigns a bit less importance to large errors, as Figure 3 shows. Notice that the derivative of the loss function ρ_c^T is given by

$$\frac{d}{dx}\rho_c^T(x) = \begin{cases} x - 2\dfrac{x^3}{c^2} + \dfrac{x^5}{c^4} & \text{if } 0 \le x \le c \\ 0 & \text{otherwise,} \end{cases}$$

and the derivative of $\rho_{a,b,c}$ is

$$\frac{d}{dx}\rho_{a,b,c}(x) = \begin{cases} x & \text{if } 0 \le x < a \\ a & \text{if } a \le x < b \\ \dfrac{a(x-c)}{b-c} & \text{if } b \le x < c \\ 0 & \text{if } c \le x. \end{cases}$$

Proposition 4.1. *Let $0 < a < b < c$ be the tuning parameters in the Hampel loss function. Given any $0 < \varepsilon < \min\left\{\dfrac{ac^3}{27a(c-b)(b+c-a)}, c-b\right\}$, it holds that*

$$\left[\frac{d}{dx}\left(\frac{3a(b+c-a)}{c^2}\rho_c^T(x)\right)\right]_{x=c-\varepsilon} < \left[\frac{d}{dx}\rho_{a,b,c}(x)\right]_{x=c-\varepsilon},$$

and, consequently, $\dfrac{3a(b+c-a)}{c^2}\phi_c^T(c-\varepsilon) < \phi_{a,b,c}(c-\varepsilon)$.

Proof First, simplifying the expression of $\dfrac{d}{dx}\left(\dfrac{3a(b+c-a)}{c^2}\rho_c^T(x)\right)$ evaluated at $c-\varepsilon$, we get

$$\begin{aligned}\left[\frac{d}{dx}\left(\frac{3a(b+c-a)}{c^2}\rho_c^T(x)\right)\right]_{x=c-\varepsilon} &= \frac{3a(b+c-a)}{c^2}\left(c-\varepsilon-2\frac{(c-\varepsilon)^3}{c^2}+\frac{(c-\varepsilon)^5}{c^4}\right) \\ &= \frac{3a(b+c-a)}{c^2}\left(4\frac{\varepsilon^2}{c}-8\frac{\varepsilon^3}{c^2}+5\frac{\varepsilon^4}{c^3}-\frac{\varepsilon^5}{c^4}\right).\end{aligned}$$

Since $0 < \varepsilon < c$, we know that $\dfrac{\varepsilon^3}{c^3} \le \dfrac{\varepsilon}{c}$ and

$$\left[4\frac{\varepsilon}{c} - 8\left(\frac{\varepsilon}{c}\right)^2 + 5\left(\frac{\varepsilon}{c}\right)^3 - \left(\frac{\varepsilon}{c}\right)^4\right] \le 9\frac{\varepsilon}{c}.$$

Note that $\varepsilon < \dfrac{ac^3}{27a(c-b)(b+c-a)}$ implies that $a > 9\dfrac{\varepsilon}{c}(c-b)\dfrac{3a(b+c-a)}{c^2}$, so the following inequality holds

$$\frac{3a(b+c-a)}{c^2}\left(4\frac{\varepsilon^2}{c}-8\frac{\varepsilon^3}{c^2}+5\frac{\varepsilon^4}{c^3}-\frac{\varepsilon^5}{c^4}\right) \le \frac{3a(b+c-a)}{c^2}9\frac{\varepsilon}{c}\varepsilon$$

$$< \frac{a}{c-b}\varepsilon = \left[\frac{d}{dx}\rho_{a,b,c}(x)\right]_{x=c-\varepsilon},$$

which completes the proof and allows us to conclude that $\frac{3a(b+c-a)}{c^2}\rho_c^T$ assigns a bit less importance to large errors than $\rho_{a,b,c}$. ■

In summary, in order to try to improve the results achieved by the Hampel fuzzy M-estimator of location, we could choose the value for the tuning parameter in the Tukey loss function as specified in Table 1.

Table 1: Choice for the tuning parameter in the Tukey biweight loss function ρ_d^T

	$x \geq \max\{\sqrt{3a(b+c-a)}, c\}$ holds for	
Case	all outliers	some outliers
$\sqrt{3a(b+c-a)} \leq c$	ρ_c^T	$\rho_{\sqrt{3a(b+c-a)}}^T$
$\sqrt{3a(b+c-a)} > c$	$\rho_{\sqrt{3a(b+c-a)}}^T$	ρ_c^T

The problem with the proposal in Table 1 is that, in practice, we do not know how much contamination is present in the dataset and whether the outlying observations fulfill the condition $x \geq \max\{\sqrt{3a(b+c-a)}, c\}$ or not, while the suggested choice for the tuning parameter is exactly the opposite in these two situations.

The proposal could be summarized as

$$d = \lambda \min\{\sqrt{3a(b+c-a)}, c\} + (1-\lambda)\max\{\sqrt{3a(b+c-a)}, c\},$$

with λ equal to 0 when all the outliers fulfill $x > \max\{\sqrt{3a(b+c-a)}, c\}$ and equal to 1 if it holds $x \leq \max\{\sqrt{3a(b+c-a)}, c\}$ for some outliers.

In order to distinguish between the two contamination situations ($\lambda = 0$ or $\lambda = 1$), we can use the information given by the tuning parameters in the Hampel loss function. Recall that a, b and c represent the median, 75th and 85th percentiles of the distances between the observations and an initial estimate (throughout this paper, this initial estimate is chosen to be the 1-norm median [4]) following Kim and Scott's procedure. As the considered contamination proportion in the simulation design from Section 4.3 is at most 0.4, note that if not all outliers are farther than c from the initial estimate, then at least one of the ratios a/b and b/c becomes quite small. We could then estimate λ by $1 - \min\{a/b, b/c\}$. For this reason, in the simulation study developed in this section, we select the value for the tuning parameter in the Tukey loss function as follows

$$d = \left(1 - \min\left\{\frac{a}{b}, \frac{b}{c}\right\}\right) \min\{\sqrt{3a(b+c-a)}, c\} + \min\left\{\frac{a}{b}, \frac{b}{c}\right\} \max\{\sqrt{3a(b+c-a)}, c\}.$$

4.3. Simulation design

The general scheme of the simulations completes the designs used in [5] (note that CASE 2 has been slightly modified by replacing the term $0.1 \cdot \chi_1^2$ by $\sqrt{\chi_1^2}$ inspired by the relation of the χ_{n-1}^2 distribution with the variance estimator).

Trapezoidal fuzzy data are generated according to four real-valued random variables. In CASES 1-4, 6 and 11-14, $\mathcal{X} = \text{Tra}(X_1 - X_2 - X_3, X_1 - X_2, X_1 + X_2, X_1 + X_2 + X_4)$, so $\inf \mathcal{X}_0 = X_1 - X_2 - X_3$, $\inf \mathcal{X}_1 = X_1 - X_2$, $\sup \mathcal{X}_1 = X_1 + X_2$ and $\sup \mathcal{X}_0 = X_1 + X_2 + X_4$, whereas in CASES 5 and 7-10, the trapezoidal fuzzy data are generated according to $\mathcal{X} = \text{Tra}(X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)})$, with $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq X_{(4)}$ the order statistics of X_1, X_2, X_3, X_4.

A contamination proportion equal to $c_p \in \{0, 0.1, 0.2, 0.4\}$ is introduced in each sample. Although it is not explicitly detailed in this next section, any kind of outlier is allowed in these simulation studies: the four random variables detailed above can follow the corresponding distributions for the contaminated observations or just some (at least one) of them. This means that we deal with outliers in location, outliers in shape and/or outliers in both location and shape.

STUDY 1 (mid-points of the 1-levels generated from a symmetric distribution)

A second parameter, $C_D \in \{0, 1, 5, 10, 100\}$, determines the distance between the distribution of the regular and contaminated observations.

In CASE 1 the variables X_i are independent. In particular,

- $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim \chi^2_1$ for the regular observations.
- $X_1 \sim \mathcal{N}(0, 3) + C_D$ and $X_2, X_3, X_4 \sim \chi^2_4 + C_D$ for the contaminated observations.

In CASE 2 dependence between the variables X_i is introduced as follows.

- $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2+\sqrt{\chi^2_1}$ for the non-contaminated subsample (with χ^2_1 independent of X_1),
- $X_1 \sim \mathcal{N}(0, 3) + C_D$ and $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + \sqrt{\chi^2_1} + C_D$ for the contaminated subsample (with χ^2_1 independent of X_1).

STUDY 2 (mid-points of the 1-levels generated from an asymmetric distribution)

The parameter C_D again determines the distance between the distribution of the regular and contaminated observations.

In CASE 3 the variables X_i are independent and

- $X_1 \sim \mathcal{W}(2, 3)$ (Weibull distribution of parameters of shape and scale 2 and 3, respectively) and $X_2, X_3, X_4 \sim \chi^2_1$ for the regular observations.
- $X_1 \sim \mathcal{W}(6, 3) + C_D$ and $X_2, X_3, X_4 \sim \chi^2_4 + C_D$ for the contaminated observations.

In CASE 4 dependence between the variables X_i is introduced as follows.

- $X_1 \sim \mathcal{W}(2, 3)$ and $X_2, X_3, X_4 \sim 1/(X_1^2+1)^2+\sqrt{\chi^2_1}$ for the non-contaminated subsample (with χ^2_1 independent of X_1),
- $X_1 \sim \mathcal{W}(6, 3) + C_D$ and $X_2, X_3, X_4 \sim 1/(X_1^2 + 1)^2 + \sqrt{\chi^2_1} + C_D$ for the contaminated subsample (with χ^2_1 independent of X_1).

STUDY 3 (bounded referential [0,1] and mid-points of the 1-levels generated from a symmetric distribution)

In this study, the parameter $C_D \in \{0, 1, 5, 10, 100\}$ does not measure the distance between the distribution of the regular and contaminated observations anymore, since the 0-level of all fuzzy numbers is assumed to belong to the interval $[0, 1]$ and, therefore, regular and contaminated observations do not lie far from each other. In this situation, C_D determines the difference in shape of regular and contaminated observations.

In CASE 5 the variables X_i are distributed as

- $X_1, X_2, X_3, X_4 \sim \text{Beta}(5, 5)$ for the regular observations.
- $X_1, X_2, X_3, X_4 \sim \text{Beta}\left(0.5, \frac{5}{\max\{0.5, C_D\}}\right)$ for the contaminated observations.

In CASE 6 the variables X_i are distributed as

- $X_1 \sim \text{Beta}(5, 5)$, $X_2 \sim \text{Uniform}[0, \min\{X_1, 1-X_1\}]$, $X_3 \sim \text{Uniform}[0, X_1 - X_2]$ and $X_4 \sim \text{Uniform}[0, 1 - X_1 - X_2]$ for the regular observations.
- $X_1 \sim \text{Beta}\left(0.5, \frac{5}{\max\{0.5, C_D\}}\right)$, $X_2 \sim \min\{X_1, 1-X_1\} \cdot \text{Beta}\left(0.5, \frac{5}{\max\{0.5, C_D\}}\right)$, $X_3 \sim (X_1 - X_2) \cdot \text{Beta}\left(0.5, \frac{5}{\max\{0.5, C_D\}}\right)$ and $X_4 \sim (1 - X_1 - X_2) \cdot \text{Beta}\left(0.5, \frac{5}{\max\{0.5, C_D\}}\right)$ for the contaminated observations.

STUDY 4 (bounded referential [0,1] and mid-points of the 1-levels generated from an asymmetric distribution)

As in STUDY 3, the parameter C_D determines the difference in shape of regular and contaminated observations.

For CASES 7-10, the variables X_i are distributed as

- $X_1, X_2, X_3, X_4 \sim \text{Beta}(p, q)$ for the regular observations.
- $X_1, X_2, X_3, X_4 \sim \text{Beta}(q, C_D + 1)$ if $p > q$ or

  $X_1, X_2, X_3, X_4 \sim \text{Beta}(C_D + 1, p)$ if $p < q$ for the contaminated observations.

The choices for the parameters p and q in the beta distribution are specified in Table 2.

$CASES$	p	q
7 & 11	5	1
8 & 12	4	2
9 & 13	2	4
10 & 14	1	5

Table 2: Choices for p and q in CASES 7-14 of STUDY 4

For CASES 11-14, the variables X_i are distributed as

- $X_1 \sim \text{Beta}(p, q)$, $X_2 \sim \text{Uniform}[0, \min\{X_1, 1-X_1\}]$, $X_3 \sim \text{Uniform}[0, X_1 - X_2]$ and $X_4 \sim \text{Uniform}[0, 1 - X_1 - X_2]$ for the regular observations.

- $X_1 \sim \text{Beta}(q, C_D + 1)$, $X_2 \sim \min\{X_1, 1 - X_1\} \cdot \text{Beta}(q, C_D + 1)$, $X_3 \sim (X_1 - X_2) \cdot \text{Beta}(q, C_D + 1)$ and $X_4 \sim (1 - X_1 - X_2) \cdot \text{Beta}(q, C_D + 1)$ if $p > q$ or

  $X_1 \sim \text{Beta}(C_D + 1, p)$, $X_2 \sim \min\{X_1, 1 - X_1\} \cdot \text{Beta}(C_D + 1, p)$, $X_3 \sim (X_1 - X_2) \cdot \text{Beta}(C_D + 1, p)$ and $X_4 \sim (1 - X_1 - X_2) \cdot \text{Beta}(C_D + 1, p)$ if $p < q$ for the contaminated observations.

Again, the choices for the parameters p and q in the beta distribution are specified in Table 2.

For each of the four cases above, the steps to follow are now explained in detail. The considered metric is $D^{\ell}_{1/3}$ and the tuning parameters in the Hampel and Tukey loss functions are fixed as in Section 4.2, taking the 1-norm median as initial estimate.

Step 1. The population targets associated with the M-estimates based on Hampel and Tukey loss functions are approximated by Monte Carlo using $N = 10000$ samples consisting of $n = 100$ regular fuzzy number-valued observations ($c_p = C_D = 0$).

Step 2. For each case and each choice (c_p, C_D), we generated $N = 1000$ random samples of size $n = 100$ and calculated the corresponding M-estimates based on Hampel and Tukey loss functions. Their performance has been compared by means of the corresponding mean squared error.

4.4. Results

Tables 3-7 show the results of the comparative analysis described in Section 4.3. The practical calculation of M-estimates of location as in Section 3 has been implemented in R [27] and the functions are available in

http://bellman.ciencias.uniovi.es/smire/Archivos/FuzzyMestimators.R.

For each situation, the lowest estimated mean squared error has been highlighted in bold. Indeed, a t-test has been applied to check if the lowest mean square error in each situation is significantly smaller than the other one. Furthermore, the variance associated to the fuzzy M-estimates of location based on Hampel and Tukey loss functions has been computed.

It can be seen that the proposal for the choice of the tuning parameter in the Tukey loss function does not provide us with a uniformly best fuzzy M-estimator of location. However, these empirical studies show that the corresponding fuzzy M-estimator is competitive and improves the results achieved by the Hampel fuzzy M-estimator of location with tuning parameters selected following Kim and Scott's procedure in most of the studied situations. The results could be summarized as follows. In general, the Tukey fuzzy M-estimator of location behaves better

- for moderate amounts of contamination ($c_p \leq 0.1$ and also $c_p = 0.2$ with $C_D < 5$) in all cases except from CASE 2 with $C_D = 5, 10, 100$;
- for intermediate amounts of contamination ($c_p = 0.2$ and $C_D \geq 5$) in CASES 5-9 and 11-13. Note that this includes all cases concerning the bounded referential (STUDIES 3 and 4) except CASES 10 and 14, which are, indeed, the most similar situations to the simulation procedure used when the referential is unbounded;

Table 3: CASES 1-3: Mean squared error and variance (both multiplied by 100) associated to fuzzy M-estimates of location based on Hampel and Tukey loss functions with tuning parameters fixed as in Section 4.2. A t-test has been applied to check if the lowest mean square error in each situation (in bold) is significantly smaller than the other one, and, in case there is no significant difference (at significant level 0.05), such a situation has been distinguished in grey colour.

		STUDY 1 - CASE 1					STUDY 1 - CASE 2					STUDY 2- CASE 3				
		TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST
c_p	C_D	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value
0	0	**2.6157**	2.6144	2.8727	2.8714	$1.50{\cdot}10^{-13}$	**4.0925**	2.2892	4.3831	2.7220	$1.90{\cdot}10^{-15}$	**5.5997**	5.5987	6.1656	6.1651	$1.24{\cdot}10^{-10}$
0.1	0	**2.9977**	2.7922	3.2317	3.0099	$<2.2{\cdot}10^{-16}$	**3.8895**	2.3682	4.1096	2.7536	$5.64{\cdot}10^{-8}$	**5.5960**	5.1089	6.1760	5.6531	$3.76{\cdot}10^{-15}$
0.1	1	**2.9636**	2.7252	3.2305	2.9721	$<2.2{\cdot}10^{-16}$	**4.3790**	2.2339	4.6151	2.6419	$3.90{\cdot}10^{-10}$	**6.5221**	5.3313	7.2127	5.9216	$1.62{\cdot}10^{-12}$
0.1	5	**2.8791**	2.6579	2.9736	2.7416	$1.28{\cdot}10^{-6}$	3.4926	2.0876	**3.4137**	2.2798	0.0004	**4.8249**	4.3409	5.0416	4.4175	$2.24{\cdot}10^{-6}$
0.1	10	**2.5937**	2.3612	2.6824	2.4407	$1.60{\cdot}10^{-7}$	3.4132	2.3137	**3.3370**	2.4666	$7.44{\cdot}10^{-5}$	**4.8918**	4.3709	5.0939	4.3823	$5.90{\cdot}10^{-6}$
0.1	100	**2.7117**	2.4736	2.7479	2.4986	0.04784	3.1375	2.0568	**3.0774**	2.2210	0.0027	**4.8532**	4.1680	5.1570	4.2380	$4.77{\cdot}10^{-11}$
0.2	0	**4.0221**	3.0956	4.2861	3.2855	$<2.2{\cdot}10^{-16}$	**3.8828**	2.5353	4.1622	3.0020	$4.34{\cdot}10^{-12}$	**6.6272**	4.9124	7.1932	5.3646	$6.93{\cdot}10^{-13}$
0.2	1	**4.4077**	3.0964	4.6602	3.2568	$4.51{\cdot}10^{-16}$	**5.2680**	2.5492	5.4307	2.9154	$8.18{\cdot}10^{-6}$	**11.3933**	6.2881	12.1995	6.8560	$3.36{\cdot}10^{-12}$
0.2	5	4.7338	2.7454	**4.5900**	2.7708	$4.53{\cdot}10^{-11}$	3.5835	2.2396	**3.4675**	2.3362	$1.42{\cdot}10^{-9}$	**8.4816**	4.6093	9.0706	4.6129	$< 2.2{\cdot}10^{-16}$
0.2	10	5.9077	2.7116	**4.6859**	2.6752	$<2.2{\cdot}10^{-16}$	**2.1459**	1.8035	2.3182	1.8591	$< 2.2{\cdot}10^{-16}$	10.2730	3.8425	**9.6049**	3.8252	$< 2.2{\cdot}10^{-16}$
0.2	100	8.1551	2.6064	**4.6307**	2.5906	$<2.2{\cdot}10^{-16}$	**1.7404**	1.6692	2.1991	1.8056	$< 2.2{\cdot}10^{-16}$	15.1263	3.8554	**10.5225**	3.9183	$< 2.2{\cdot}10^{-16}$
0.4	0	**10.3164**	4.3191	10.8703	4.4841	$<2.2{\cdot}10^{-16}$	**3.4834**	2.5817	3.5338	2.8018	0.0491	**13.6835**	6.6060	14.5373	7.0396	$< 2.2{\cdot}10^{-16}$
0.4	1	**13.8694**	4.7834	14.6758	4.9372	$<2.2{\cdot}10^{-16}$	**7.9043**	3.2041	7.9223	3.4848	0.3211	**38.6145**	11.8231	40.3808	12.5889	$< 2.2{\cdot}10^{-16}$
0.4	5	**29.6953**	7.6437	34.8549	8.6801	$<2.2{\cdot}10^{-16}$	**15.0096**	5.0720	17.3709	5.9372	$< 2.2{\cdot}10^{-16}$	**55.2370**	14.8144	59.7958	14.9210	$4.16{\cdot}10^{-10}$
0.4	10	**44.8112**	11.8664	53.3491	12.9101	$<2.2{\cdot}10^{-16}$	**20.8983**	8.0349	27.7716	9.3791	$< 2.2{\cdot}10^{-16}$	**58.1400**	15.0637	79.9412	18.3472	$< 2.2{\cdot}10^{-16}$
0.4	100	**24.5023**	17.0867	82.3317	31.3990	$<2.2{\cdot}10^{-16}$	**25.0629**	24.8888	58.7913	34.9118	$< 2.2{\cdot}10^{-16}$	**37.5671**	18.7257	118.4374	34.1891	$< 2.2{\cdot}10^{-16}$

Table 4: CASES 4-6: Mean squared error and variance (both multiplied by 100) associated to fuzzy M-estimates of location based on Hampel and Tukey loss functions with tuning parameters fixed as in Section 4.2. A t-test has been applied to check if the lowest mean square error in each situation (in bold) is significantly smaller than the other one, and, in case there is no significant difference (at significant level 0.05), such a situation has been distinguished in grey colour.

		STUDY 2 - CASE 4					STUDY 3 - CASE 5					STUDY 3- CASE 6				
		TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST
c_p	C_D	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value
0	0	**4.9822**	4.9361	5.5423	5.4935	$7.72 \cdot 10^{-13}$	**0.0167**	0.0167	0.0210	0.0210	$< 2.2 \cdot 10^{-16}$	**0.0504**	0.0503	0.0595	0.0594	$< 2.2 \cdot 10^{-16}$
0.1	0	**5.7313**	5.1098	6.4991	5.7361	$1.12 \cdot 10^{-15}$	**0.0164**	0.0158	0.0184	0.0179	$< 2.2 \cdot 10^{-16}$	**0.0457**	0.0453	0.0502	0.0498	$< 2.2 \cdot 10^{-16}$
0.1	1	**8.6874**	6.6169	10.5960	8.2461	$< 2.2 \cdot 10^{-16}$	**0.0158**	0.0153	0.0179	0.0175	$< 2.2 \cdot 10^{-16}$	**0.0470**	0.0463	0.0521	0.0516	$< 2.2 \cdot 10^{-16}$
0.1	5	**4.9733**	4.4412	5.2503	4.5375	$8.39 \cdot 10^{-7}$	**0.0159**	0.0155	0.0190	0.0186	$< 2.2 \cdot 10^{-16}$	**0.0460**	0.0459	0.0537	0.0536	$< 2.2 \cdot 10^{-16}$
0.1	10	**4.5576**	4.0488	4.8021	4.0884	$6.89 \cdot 10^{-7}$	**0.0172**	0.0169	0.0204	0.0201	$< 2.2 \cdot 10^{-16}$	**0.0466**	0.0466	0.0544	0.0543	$< 2.2 \cdot 10^{-16}$
0.1	100	**4.8090**	4.2800	5.0210	4.2679	$9.03 \cdot 10^{-5}$	**0.0155**	0.0153	0.0170	0.0169	$2.05 \cdot 10^{-13}$	**0.0450**	0.0446	0.0499	0.0496	$< 2.2 \cdot 10^{-16}$
0.2	0	**6.4002**	4.6650	7.3476	5.2929	$1.09 \cdot 10^{-9}$	**0.0219**	0.0161	0.0231	0.0176	$1.31 \cdot 10^{-6}$	**0.0462**	0.0433	0.0491	0.0462	$3.64 \cdot 10^{-8}$
0.2	1	**18.0175**	9.4674	21.7114	11.8630	$< 2.2 \cdot 10^{-16}$	**0.0209**	0.0154	0.0219	0.0171	$1.32 \cdot 10^{-5}$	**0.0491**	0.0457	0.0522	0.0491	$6.65 \cdot 10^{-10}$
0.2	5	**7.0356**	3.8855	7.8717	3.9147	$< 2.2 \cdot 10^{-16}$	**0.0180**	0.0158	0.0206	0.0185	$< 2.2 \cdot 10^{-16}$	**0.0483**	0.0474	0.0528	0.0520	$< 2.2 \cdot 10^{-16}$
0.2	10	**8.7300**	3.1452	8.7429	3.1606	0.4076	**0.0183**	0.0168	0.0207	0.0192	$< 2.2 \cdot 10^{-16}$	**0.0467**	0.0467	0.0516	0.0516	$< 2.2 \cdot 10^{-16}$
0.2	100	13.3225	2.8159	**10.1470**	3.2899	$< 2.2 \cdot 10^{-16}$	**0.0186**	0.0147	0.0193	0.0159	0.0002	**0.0483**	0.0463	0.0506	0.0488	$5.22 \cdot 10^{-7}$
0.4	0	**9.0811**	4.2936	10.4543	4.8743	$< 2.2 \cdot 10^{-16}$	0.1435	0.0328	**0.1366**	0.0366	$7.69 \cdot 10^{-6}$	0.0977	0.0691	**0.0855**	0.0621	0.0062
0.4	1	**56.8851**	18.2626	64.0791	21.0338	$< 2.2 \cdot 10^{-16}$	0.1335	0.0270	**0.1303**	0.0297	0.0023	0.1013	0.0705	**0.0891**	0.0635	$4.45 \cdot 10^{-9}$
0.4	5	**34.3435**	11.0423	37.7986	11.0676	$2.32 \cdot 10^{-8}$	**0.0439**	0.0214	0.0471	0.0240	$7.28 \cdot 10^{-11}$	**0.0560**	0.0504	0.0584	0.0533	$5.54 \cdot 10^{-6}$
0.4	10	**41.1221**	11.9462	58.2582	15.1092	$< 2.2 \cdot 10^{-16}$	**0.0318**	0.0201	0.0360	0.0229	$< 2.2 \cdot 10^{-16}$	**0.0517**	0.0515	0.0549	0.0547	$1.00 \cdot -9$
0.4	100	**40.6934**	25.8834	94.0305	39.1042	$< 2.2 \cdot 10^{-16}$	0.1166	0.0261	**0.1154**	0.0290	0.1155	0.0908	0.0731	**0.0862**	0.0716	0.0086

Table 5: CASES 7-9: Mean squared error and variance (both multiplied by 100) associated to fuzzy M-estimates of location based on Hampel and Tukey loss functions with tuning parameters fixed as in Section 4.2. A t-test has been applied to check if the lowest mean square error in each situation (in bold) is significantly smaller than the other one, and, in case there is no significant difference (at significant level 0.05), such a situation has been distinguished in grey colour.

		STUDY 4 - CASE 7					STUDY 4 - CASE 8					STUDY 4- CASE 9				
		TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST
c_p	C_D	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value
0	0	**0.0123**	0.0123	0.0141	0.0141	$1.36 \cdot 10^{-14}$	**0.0223**	0.0223	0.0272	0.0272	$< 2.2 \cdot 10^{-16}$	**0.0229**	0.0228	0.0277	0.0276	$< 2.2 \cdot 10^{-16}$
0.1	0	**0.0152**	0.0136	0.0159	0.0142	0.0005	**0.0242**	0.0241	0.0288	0.0288	$< 2.2 \cdot 10^{-16}$	**0.0236**	0.0235	0.0288	0.0287	$< 2.2 \cdot 10^{-16}$
0.1	1	**0.0133**	0.0113	0.0139	0.0117	0.0001	**0.0260**	0.0245	0.0298	0.0282	$< 2.2 \cdot 10^{-16}$	**0.0301**	0.0263	0.0353	0.0317	$< 2.2 \cdot 10^{-16}$
0.1	5	**0.0129**	0.0112	0.0135	0.0115	$1.98 \cdot 10^{-6}$	**0.0244**	0.0217	0.0272	0.0246	$1.53 \cdot 10^{-14}$	**0.0247**	0.0222	0.0271	0.0249	$3.85 \cdot 10^{-12}$
0.1	10	**0.0129**	0.0113	0.0136	0.0116	$3.92 \cdot 10^{-8}$	**0.0245**	0.0219	0.0273	0.0246	$2.67 \cdot 10^{-14}$	**0.0249**	0.0232	0.0276	0.0260	$1.94 \cdot 10^{-12}$
0.1	100	**0.0127**	0.0106	0.0133	0.0108	$5.34 \cdot 10^{-7}$	**0.0225**	0.0211	0.0248	0.0236	$9.42 \cdot 10^{-11}$	**0.0200**	0.0191	0.0216	0.0208	$7.66 \cdot 10^{-10}$
0.2	0	**0.0224**	0.0124	0.0230	0.0126	0.0004	**0.0268**	0.0263	0.0322	0.0316	$< 2.2 \cdot 10^{-16}$	**0.0242**	0.0240	0.0285	0.0282	$< 2.2 \cdot 10^{-16}$
0.2	1	**0.0250**	0.0116	0.0260	0.0115	$5.30 \cdot 10^{-11}$	**0.0335**	0.0255	0.0384	0.0305	$< 2.2 \cdot 10^{-16}$	**0.0378**	0.0244	0.0416	0.0288	$4.34 \cdot 10^{-12}$
0.2	5	**0.0282**	0.0099	0.0291	0.0098	$1.12 \cdot 10^{-9}$	**0.0415**	0.0244	0.0432	0.0267	0.0003	**0.0484**	0.0267	0.0496	0.0296	0.0212
0.2	10	**0.0277**	0.0106	0.0284	0.0107	0.0016	**0.0408**	0.0230	0.0421	0.0247	0.0162	**0.0400**	0.0239	0.0414	0.0258	0.0076
0.2	100	**0.0290**	0.0110	0.0293	0.0110	0.0886	**0.0371**	0.0227	0.0401	0.0255	$1.22 \cdot 10^{-8}$	**0.0323**	0.0212	0.0330	0.0226	0.0147
0.4	0	**0.0860**	0.0178	0.0868	0.0175	0.0394	**0.0329**	0.0306	0.0378	0.0354	$< 2.2 \cdot 10^{-16}$	**0.0305**	0.0295	0.0361	0.0351	$< 2.2 \cdot 10^{-16}$
0.4	1	**0.1586**	0.0195	0.1619	0.0193	$1.04 \cdot 10^{-5}$	**0.0777**	0.0348	0.0802	0.0391	0.0037	**0.1024**	0.0366	0.1071	0.0424	$4.61 \cdot 10^{-7}$
0.4	5	**0.2338**	0.0212	0.2526	0.0240	$< 2.2 \cdot 10^{-16}$	0.2215	0.0403	**0.2143**	0.0435	$6.44 \cdot 10^{-5}$	0.3044	0.0496	**0.2939**	0.0581	$8.50 \cdot 10^{-5}$
0.4	10	**0.2518**	0.0291	0.2738	0.0350	$1.55 \cdot 10^{-14}$	0.2479	0.0402	**0.2387**	0.0432	0.0003	0.3049	0.0484	**0.2876**	0.0565	$8.07 \cdot 10^{-8}$
0.4	100	**0.2779**	0.0295	0.2917	0.0418	$2.66 \cdot 10^{-6}$	**0.2406**	0.0402	0.2498	0.0443	$3.10 \cdot 10^{-7}$	**0.2665**	0.0449	0.2680	0.0522	0.3331

Table 6: CASES 10-12: Mean squared error and variance (both multiplied by 100) associated to fuzzy M-estimates of location based on Hampel and Tukey loss functions with tuning parameters fixed as in Section 4.2. A t-test has been applied to check if the lowest mean square error in each situation (in bold) is significantly smaller than the other one, and, in case there is no significant difference (at significant level 0.05), such a situation has been distinguished in grey colour.

		STUDY 4 - CASE 10					STUDY 4 - CASE 11					STUDY 4- CASE 12				
		TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST
c_p	C_D	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value	MSE	VAR	MSE	VAR	p-value
0	0	**0.0138**	0.0138	0.0154	0.0154	$4.02{\cdot}10^{-11}$	**0.0442**	0.0442	0.0499	0.0498	$< 2.2{\cdot}10^{-16}$	**0.0640**	0.0639	0.0735	0.0734	$< 2.2{\cdot}10^{-16}$
0.1	0	**0.0145**	0.0129	0.0154	0.0137	$8.88{\cdot}10^{-5}$	**0.0405**	0.0403	0.0441	0.0440	$1.46{\cdot}10^{-10}$	**0.0657**	0.0653	0.0748	0.0744	$< 2.2{\cdot}10^{-16}$
0.1	1	**0.0125**	0.0108	0.0133	0.0114	$1.68{\cdot}10^{-8}$	**0.0400**	0.0397	0.0441	0.0438	$6.14{\cdot}10^{-15}$	**0.0669**	0.0667	0.0767	0.0766	$< 2.2{\cdot}10^{-16}$
0.1	5	**0.0131**	0.0112	0.0136	0.0113	$3.27{\cdot}10^{-5}$	**0.0414**	0.0406	0.0448	0.0436	$2.27{\cdot}10^{-9}$	**0.0600**	0.0591	0.0656	0.0648	$8.49{\cdot}10^{-9}$
0.1	10	**0.0131**	0.0115	0.0139	0.0118	$6.14{\cdot}10^{-10}$	**0.0397**	0.0391	0.0425	0.0417	$3.82{\cdot}10^{-6}$	**0.0628**	0.0625	0.0679	0.0676	$2.15{\cdot}10^{-10}$
0.1	100	**0.0125**	0.0109	0.0130	0.0111	$1.38{\cdot}10^{-5}$	**0.0430**	0.0423	0.0464	0.0455	$9.43{\cdot}10^{-10}$	**0.0582**	0.0574	0.0634	0.0624	$1.60{\cdot}10^{-11}$
0.2	0	**0.0219**	0.0120	0.0227	0.0124	$5.88{\cdot}10^{-5}$	**0.0417**	0.0397	0.0447	0.0431	$1.05{\cdot}10^{-8}$	**0.0606**	0.0590	0.0682	0.0667	$1.41{\cdot}10^{-13}$
0.2	1	**0.0266**	0.0122	0.0276	0.0122	$2.45{\cdot}10^{-12}$	**0.0416**	0.0389	0.0434	0.0408	$7.98{\cdot}10^{-5}$	**0.0685**	0.0652	0.0741	0.0715	$1.59{\cdot}10^{-8}$
0.2	5	**0.0283**	0.0100	0.0288	0.0100	$9.04{\cdot}10^{-5}$	**0.0496**	0.0455	0.0511	0.0471	0.0006	**0.0736**	0.0673	0.0768	0.0708	0.0002
0.2	10	**0.0274**	0.0112	0.0276	0.0112	0.0785	**0.0491**	0.0446	0.0502	0.0460	0.0278	**0.0628**	0.0590	0.0669	0.0622	$1.67{\cdot}10^{-7}$
0.2	100	0.0289	0.0097	**0.0285**	0.0096	0.0315	**0.0529**	0.0472	0.0546	0.0495	0.0217	**0.0586**	0.0536	0.0638	0.0582	$2.30{\cdot}10^{-13}$
0.4	0	**0.0943**	0.0184	0.0946	0.0182	0.2826	0.0615	0.0460	**0.0610**	0.0468	0.1184	**0.0742**	0.0704	0.0832	0.0794	$< 2.2{\cdot}10^{-16}$
0.4	1	**0.1596**	0.0190	0.1639	0.0188	$1.17{\cdot}10^{-7}$	**0.1059**	0.0760	0.1066	0.0779	0.1136	**0.1064**	0.0840	0.1119	0.0914	$8.80{\cdot}10^{-5}$
0.4	5	**0.2317**	0.0268	0.2539	0.0280	$< 2.2{\cdot}10^{-16}$	**0.1308**	0.1049	0.1532	0.1248	$< 2.2{\cdot}10^{-16}$	0.2407	0.1545	**0.2160**	0.1416	$5.46{\cdot}10^{-6}$
0.4	10	**0.2436**	0.0268	0.2724	0.0306	$< 2.2{\cdot}10^{-16}$	**0.1092**	0.0859	0.1270	0.1023	$< 2.2{\cdot}10^{-16}$	0.1934	0.1458	**0.1793**	0.1391	0.0791
0.4	100	**0.2655**	0.0357	0.2785	0.0416	$1.79{\cdot}10^{-5}$	**0.1055**	0.0820	0.1098	0.0876	$1.78{\cdot}10^{-6}$	0.1074	0.0823	**0.1050**	0.0803	0.3849

Table 7: CASES 13-14: Mean squared error and variance (both multiplied by 100) associated to fuzzy M-estimates of location based on Hampel and Tukey loss functions with tuning parameters fixed as in Section 4.2. A t-test has been applied to check if the lowest mean square error in each situation (in bold) is significantly smaller than the other one, and, in case there is no significant difference (at significant level 0.05), such a situation has been distinguished in grey colour.

		STUDY 4 - CASE 13					STUDY 4 - CASE 14				
		TUKEY		HAMPEL		T-TEST	TUKEY		HAMPEL		T-TEST
0	0	**0.0651**	0.0650	0.0743	0.0742	$3.67 \cdot 10^{-16}$	**0.0441**	0.0440	0.0508	0.0508	$< 2.2 \cdot 10^{-16}$
0.1	0	**0.0671**	0.0667	0.0788	0.0784	$< 2.2 \cdot 10^{-16}$	**0.0425**	0.0418	0.0451	0.0446	$8.45 \cdot 10^{-8}$
0.1	1	**0.0637**	0.0629	0.0726	0.0720	$1.47 \cdot 10^{-15}$	**0.0429**	0.0410	0.0457	0.0444	$7.48 \cdot 10^{-7}$
0.1	5	**0.0674**	0.0655	0.0716	0.0700	$3.40 \cdot 10^{-6}$	**0.0401**	0.0383	0.0428	0.0414	$5.39 \cdot 10^{-9}$
0.1	10	**0.0631**	0.0609	0.0687	0.0664	$1.19 \cdot 10^{-7}$	**0.0418**	0.0398	0.0444	0.0428	$1.31 \cdot 10^{-8}$
0.1	100	**0.0571**	0.0554	0.0625	0.0607	$1.35 \cdot 10^{-12}$	**0.0428**	0.0400	0.0454	0.0431	$2.68 \cdot 10^{-7}$
0.2	0	**0.0670**	0.0653	0.0738	0.0721	$2.88 \cdot 10^{-10}$	**0.0460**	0.0436	0.0478	0.0459	0.0014
0.2	1	**0.0705**	0.0668	0.0780	0.0746	$1.91 \cdot 10^{-9}$	**0.0458**	0.0385	0.0459	0.0400	0.4047
0.2	5	**0.0660**	0.0562	0.0687	0.0598	0.0002	0.0475	0.0351	**0.0460**	0.0354	$6.70 \cdot 10^{-5}$
0.2	10	**0.0662**	0.0577	0.0691	0.0606	$5.23 \cdot 10^{-6}$	0.0562	0.0397	**0.0538**	0.0397	$8.84 \cdot 10^{-8}$
0.2	100	**0.0612**	0.0514	0.0655	0.0545	$7.32 \cdot 10^{-12}$	0.0598	0.0415	**0.0583**	0.0420	0.0020
0.4	0	**0.0923**	0.0796	0.1010	0.0889	$1.05 \cdot 10^{-11}$	0.0612	0.0458	**0.0603**	0.0466	0.0373
0.4	1	**0.0959**	0.0798	0.0995	0.0850	0.0043	0.1247	0.0609	**0.1238**	0.0623	0.0674
0.4	5	0.2501	0.1375	**0.2308**	0.1313	$7.14 \cdot 10^{-7}$	**0.1476**	0.0571	0.1564	0.0612	$1.41 \cdot 10^{-9}$
0.4	10	0.2417	0.1362	**0.2219**	0.1325	0.0009	**0.1492**	0.0566	0.1639	0.0610	$2.07 \cdot 10^{-15}$
0.4	100	**0.1875**	0.1007	0.1943	0.1113	0.2080	**0.1708**	0.0661	0.1832	0.0692	$5.90 \cdot 10^{-11}$

- for high amounts of contamination ($c_p = 0.4$) in CASES 1-4 (that is, those considering an unbounded referential), 7 and 10 (among the cases contained in STUDY 4, those with a larger distance between the regular and the contaminated observations, that is, more similar to the unbounded referential situation).

In summary, the referential set being unbounded or bounded seems to have much more influence on the performance of fuzzy M-estimators of location than the symmetry of the distribution generating the mid-points of the 1-levels of the fuzzy numbers. The results indicate that the Tukey biweight loss function can also be a very useful choice for the computation of fuzzy M-estimators of location for both an unbounded (mainly low or high contamination rates with this choice of the tuning parameter) and a bounded referential set (mainly low or intermediate contamination rates with this tuning parameter).

In relation to those cases in which the Tukey fuzzy M-estimator of location has not improved the results achieved by the Hampel fuzzy M-estimator, the following comments should be taken into account.

- Due to the proposed choice of the tuning parameter in Tukey's loss function, the Tukey fuzzy M-estimator of location mainly detects low or high amounts of contamination, whereas it can experience some problems when dealing with an intermediate amount of contamination. For example, in STUDIES 1 and 2, when $c_p = 0.2$ and C_D is large, the proposal correctly guesses that not all the outliers are larger than $\max\{\sqrt{3a(b+c-a)}, c\}$. Therefore, the estimate of λ is close to 1 and d is approximately equal to $\min\{\sqrt{3a(b+c-a)}, c\}$. However, if c is so large in comparison with a and b (which are not affected by the 20% of outliers), the computation $\sqrt{3a(b+c-a)}$ may lead to a value greater than c and, in consequence, the Tukey fuzzy M-estimator would be even less suitable than the Hampel version.

- Even while the Tukey fuzzy M-estimator of location is design to mainly detect low or high amounts of contamination, this is not always achieved, as it can be seen in STUDIES 3 and 4 (see Tables 4-7). The reason is that in those situations fuzzy numbers are generated within a small referential set and the regular and contaminated observations are not far away, so some of the numerous contaminated observations may mislead the final estimate.

5. Real-life example

The comparison of the fuzzy M-estimates based on Hampel and Tukey loss functions is now illustrated by means of a real-life example, concerning the well-known TIMSS-PIRLS assessments. In 2011, TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading Literacy Study) joined to assess the fourth grade students in three fundamental curricular areas: mathematics, science, and reading.

We have adapted some items for the Student questionnaire TIMSS/PIRLS (see http://timss.bc.edu/ timss2011/downloads/T11_StuQ_4.pdf) to work with a fuzzy rating scale [24] instead of a 4-point Likert scale (namely DISAGREE A LOT, DISAGREE A LITTLE, AGREE A LITTLE and AGREE A LOT), which is the standard format. The reason is that the fuzzy rating scale allows us to combine a free-response format with a fuzzy valuation, in such a way that human valuations are captured more expressively and richly (see e.g. [1]). More details about the data collection can be found in [5].

Sixty-eight fourth grade students from Colegio San Ignacio in Asturias (Spain) answered this questionnaire after receiving some instructions. Only trapezoidal fuzzy numbers have been considered to ease their understanding of the task. The results are shown in Figure 4. We now consider the information about the item "My teacher of Mathematics is easy to understand" collected among the male students.

The output for the location M-estimates based on Hampel and Tukey loss functions is displayed in Figure 5.

It can be seen that Hampel and Tukey M-estimates, which are based on the

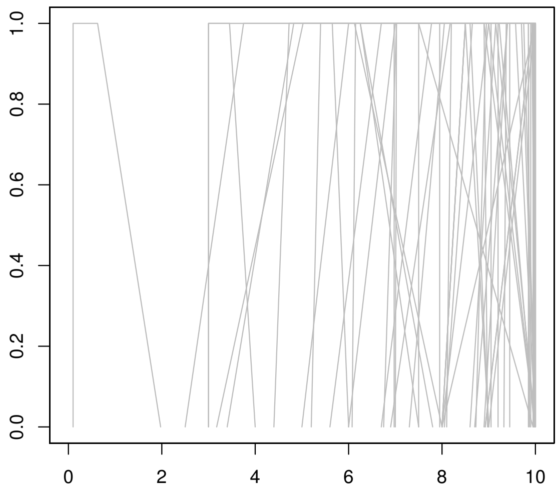

Figure 4: Sample fuzzy data of the 68 fuzzy rating scale-based responses

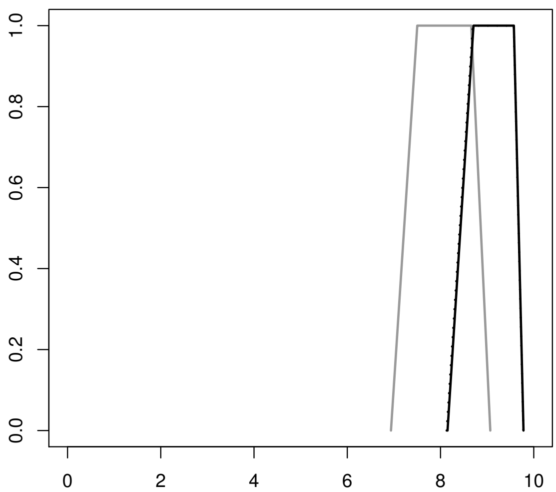

Figure 5: Aumann-type mean (grey), Hampel (−−) and Tukey (black) location M-estimates of the 68 fuzzy rating scale-based responses

$D^{\ell}_{1/3}$ metric and use the 1-norm median as initial solution, almost coincide in this real-life example when the parameter of the Tukey loss function is chosen as explained in Section 4.2. Naturally, both of them improve the behaviour of the

Aumann-type mean and, as explained in [5, 7], they preserve the trapezoidal shape of the sample data. It is quite remarkable that the Tukey M-estimate, which depends on just one tuning parameter, performs as well as the Hampel M-estimate, whose three tuning parameters provide much more flexibility.

6. Concluding remarks

In this paper, the relevance of the Tukey biweight or bisquare family of loss functions for the computation of fuzzy number-valued M-estimators of location has been analyzed. Their use has been recently introduced in the literature and a comparative analysis with the Hampel fuzzy M-estimators of location, which had been previously proposed, was of interest for practical reasons.

First, the concept of fuzzy M-estimators of location has been recalled, as well as their main properties. To complement these theoretical results, some simulations have been designed in order to compare the fuzzy M-estimators based on two different families of loss functions: Hampel and Tukey biweight. For Hampel's loss function, Kim and Scott proposed a selection procedure for its tuning parameters in [8], which had been used in the adaptation of M-estimators to the fuzzy-valued setting (see [5]). We have now shown that the Tukey loss function can be a very competitive alternative when a careful choice of its tuning parameter is used. Its advantages with respect to the Hampel fuzzy M-estimator seem to depend more on the bounded or unbounded referential than on the distribution of the mid-points of the 1-levels being symmetric or not. It is interesting to notice that even when our motivation to propose the choice for the tuning parameter in the Tukey loss function comes from the Hampel estimate (since, as it has already been said, it provided us with the best results in general), the Hampel estimate does not have to be calculated first in order to calculate the newly tuned Tukey estimate and the inspiration could have come from anywhere else. For instance, some fixed specific quantiles of the distribution of distances to the initial estimate could also have been considered as possible values for the tuning parameter of the Tukey loss function, namely, the median, the 75th and the 85th percentiles of those distances (that is, the previous choice for the three parameters involved in the Hampel loss function). In principle, this option seems to be more rigid and not to cover as many different situations as our proposal, but the comparison in terms of mean squared error could be tackled in the future.

In future research, it would be interesting to develop a more thorough study of those situations in which the Hampel fuzzy M-estimator of location still presents a better finite-sample behaviour than the M-estimator based on Tukey loss function. Another problem of interest would be the analysis of other families of loss functions for which the Representer Theorem still holds.

Acknowledgments

The authors are grateful to the reviewers and Editor handling this paper, as well as to Prof. M. Á. Gil, for their insightful comments and suggestions.This research has been partially supported by the Spanish Ministry of Economy and Competitiveness through the Grant MTM2013-44212-P, the Principality of Asturias/FEDER Grant GRUPIN14-101, Grant C16/15/068 of International Funds KU Leuven and IAP research network grant nr. P7/06 of the Belgian government. Their support is gratefully acknowledged.

References

[1] S. De la Rosa de Sáa, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, Fuzzy rating scale-based questionnaires and their statistical analysis, IEEE Trans. Fuzzy Syst. 23 (1) (2015) 111–126.

[2] M.L. Puri, D.A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[3] 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.

[4] 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.

[5] B. Sinova, M. A. Gil, S. Van Aelst, M-estimates of location for the robust central tendency of fuzzy data, IEEE Trans. Fuzzy Syst. 24 (4) (2016) 945–956.

[6] P. J. Huber, Robust estimation of a location parameter, Ann. Math. Statist. 35 (1964) 73–101.

[7] B. Sinova, S. Van Aelst, Tukey's biweight loss function for fuzzy set-valued M-estimators of location, in: M. B. Ferraro, P. Giordani, B. Vantaggi, M. Gagolewski, M. A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Soft Methods for Data Science, Vol. 456 of Advances in Intelligent Systems and Computing, Springer, Switzerland, 2017, pp. 447–454.

[8] J. S. Kim, C. D. Scott, Robust kernel density estimation, J. Mach. Learn. Res. 13 (2012) 2529–2565.

[9] A. Blanco-Fernández, R. M. Casals, A. Colubi, N. Corral, M. García-Bárzana, M. A. Gil, G. González-Rodríguez, M. T. López, M. A. Lubiano, M. Montenegro, A. B. Ramos-Guajardo, S. de la Rosa de Sáa, B. Sinova, A distance-based statistical analysis of fuzzy number-valued data, Int. J. Approx. Reason. 55 (2014) 1487–1501.

[10] A. Blanco-Fernández, R. M. Casals, A. Colubi, N. Corral, M. García-Bárzana, M. A. Gil, G. González-Rodríguez, M. T. López, M. A. Lubiano, M. Montenegro, A. B. Ramos-Guajardo, S. de la Rosa de Sáa, B. Sinova, Rejoinder on "a distance-based statistical analysis of fuzzy number-valued data", Int. J. Approx. Reason. 55 (2014) 1601–1605.

[11] M. Montenegro, M. R. Casals, M. A. Lubiano, M. A. Gil, Two-sample hypothesis tests of means of a fuzzy random variable, Inf. Sci. 133 (2001) 89–100.

[12] W. Trutschnig, G. González-Rodríguez, A. Colubi, M. A. Gil, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread, Inf. Sci. 179 (2009) 3964–3972.

[13] C. Bertoluzza, N. Corral, A. Salas, On a new class of distances between fuzzy numbers, Math & Soft Comput. 2 (1995) 71–84.

[14] P. Diamond, P. Kloeden, Metric spaces of fuzzy sets, Fuzzy Sets Syst. 35 (1990) 241–249.

[15] M. L. Puri, D. A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[16] A. Blanco-Fernández, M. R. Casals, A. Colubi, N. Corral, M. Garca-Bárzana, M. A. Gil, G. González-Rodríguez, M. T. López, M. A. Lubiano, M. Montenegro, A. B. Ramos-Guajardo, S. de la Rosa de Sáa, B. Sinova, Random fuzzy sets: A mathematical tool to develop statistical fuzzy data analysis, Iran. J. Fuzzy Syst. 10 (2) (2013) 1–28.

[17] M. A. Gil, A. Colubi, P. Terán, Random fuzzy sets: why, when, how, BEIO 30 (1) (2013) 5–29.

[18] M. L. Puri, D. A. Ralescu, The concept of normality for fuzzy random variables, Ann. Probab. 11 (1985) 1373–1379.

[19] A. Colubi, M. López-Díaz, J. S. Domínguez-Menchero, M. A. Gil, A generalized strong law of large numbers, Probab. Theory Relat. Fields 114 (1999) 401–417.

[20] M. Fréchet, Les éléments aléatoires de nature quelconque dans un espace distancié, Ann. L'Inst. H. Poincaré 10 (1948) 215–310.

[21] P. J. Huber, Robust Statistics, J. Wiley & Sons, New York, 1981.

[22] A. E. Beaton, J. W. Tukey, The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data, Technometrics 16 (1974) 147–185.

[23] F. R. Hampel, The influence curve and its role in robust estimation, J. Amer. Stat. Assoc. 69 (1974) 383–393.

[24] T. Hesketh, R. Pryor, B. Hesketh, An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences, Int. J. Man-Mach. Stud. 29 (1988) 21–35.

[25] F. R. Hampel, A general qualitative definition of robustness, Ann. Math. Stat. 42 (1971) 1887–1896.

[26] D. L. Donoho, P. J. Huber, The notion of breakdown point, in: P. J. Bickel, K. Doksum, J. L. J. Hodges (Eds.), A Festschrift for Erich L. Lehmann, Wadsworth, Belmont, 1983.

[27] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2017).
URL `http://www.R-project.org/`