Fuzzy *c*-ordered medoids clustering for interval-valued data

Pierpaolo D'Urso [a,*], Jacek M. Leski [b,c]

[a] *Department of Social Sciences and Economics, Sapienza University of Rome, P.le Aldo Moro 5, Rome, Italy*
[b] *Institute of Electronics, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland*
[c] *Department of Computer Medical Systems, Institute of Medical Technology and Equipment, Roosevelt St. 118, 41-800 Zabrze, Poland*

ARTICLE INFO

ABSTRACT

Fuzzy clustering for interval-valued data helps us to find natural vague boundaries in such data. The Fuzzy *c*-Medoids Clustering (FcMdC) method is one of the most popular clustering methods based on a partitioning around medoids approach. However, one of the greatest disadvantages of this method is its sensitivity to the presence of outliers in data. This paper introduces a new robust fuzzy clustering method named Fuzzy *c*-Ordered-Medoids clustering for interval-valued data (FcOMdC-ID). The Huber's M-estimators and the Yager's Ordered Weighted Averaging (OWA) operators are used in the method proposed to make it robust to outliers. The described algorithm is compared with the fuzzy *c*-medoids method in the experiments performed on synthetic data with different types of outliers. A real application of the FcOMdC-ID is also provided.

1. Introduction

In many real applications, empirical information represented by data cannot be expressed in terms of single valued data but in form of interval-valued data. We can distinguish the following situations:

- The phenomena cannot be explained by using single-valued data, i.e. the data are intrinsically interval-valued. Examples are the concepts of monthly temperature in meteorological stations or the daily rate of exchange between euro and dollar or euro and sterling. For example, for daily temperatures or daily pollution levels registered in different places or for mineral concentrations of food items, it could be more interesting to consider the minimum and maximum values registered than the average ones because they offer more detailed information about the examined phenomenon taking into account the variability of the features involved. Data in which each observation is given as an interval of values—indicated by a minimum and a maximum—are called interval-valued data. Notice that the intervals need not pertain to the actually observed maxima and minima, but could also pertain to, for instance, interquartile

intervals, or intervals pertaining to the middle 90% of the scores [44].

- Interval data may occur due to a lack of knowledge, that is when the true value of a variable is unknown and only an interval of values including the true value is available. Thus, the available information is imprecise and therefore cannot be revealed exactly by single numerical data.
- Interval-valued data may also arise as a result of aggregating huge databases, which are impossible to analyze in the original form.

In all these cases, standard statistical methods for single-valued data (numerical data) are unable to properly take into account the nature of the empirical information formalized in an interval manner (interval-valued data).

In the literature, there are many examples of real applications with interval-valued data. For instance, "researchers often have only interval data on variables that can, in principle, be measured more precisely [i.e.] the interval data on wealth in the Health and Retirement Study (HRS) provide a ready illustration [61]. Let v denote a person's wealth. Under the HRS questionnaire protocol, a respondent is asked to report v. If he does not comply, the respondent is then asked to report if wealth falls within a sequence of brackets. The HRS thus yields a wealth interval $[v_0, v_1]$ for each respondent. The interval is degenerate when a respondent provides a point value of wealth, is an informative interval of positive width when the respondent answers the subsequent

* Corresponding author.
 E-mail addresses: pierpaolo.durso@uniroma1.it (P. D'Urso),
jleski@polsl.pl (J.M. Leski).

bracket questions, and is the uninformative interval $[-\infty, +\infty]$ otherwise" [76].

In database protection and security intelligence, for protecting confidential data, e.g. health information of different patients, it is necessary that statistical programs do not have full access to the data because otherwise, by computing sufficiently many different statistics, it is possible to uniquely reconstruct the actual values; in these cases for preventing this form of happening is useful to supply the data processing programs not with exact data, but with an interval of possible values; in environmetrics, we may refer to the minimum and maximum values registered in the considered period of pollutant concentration data recorded at various places; in finance, we may examine the minimum–maximum daily rate of exchange data between euro-dollar or euro-sterling; in medicine, we may make reference to minimum and maximum values of the daily systolic and diastolic pressure, pulse rate, temperature of patients; in meteorology, we can consider the minimum–maximum daily temperature, humidity and wind speed registered in different places; in chemometrics, we may study the range of mineral concentrations of food products analyzed at different times or in different experimental situations and so on.

In the literature, examples of fields in which interval data-based information is analyzed are: pattern recognition [25], web [33], telecommunications [31], meteorology [51,71,33], environmetrics [7], chemometrics [59,34–36], medicine [3,81,85], physics [47], education [22,85], economics and finance [45].

From a methodological point of view, interesting contributions on interval-valued data analysis have been proposed in principal component analysis [34,25,70], multidimensional scaling [24,47], regression analysis [76,28,79,80,45], regression trees [71,7], time series analysis [2,86], discriminant analysis [26,27], self-organizing maps [31,5,51]; cluster analysis [9,17–20,23,50,52,63].

Recently a great attention is focused on fuzzy clustering of interval-valued data (see, e.g., [35,36,14,21,33]). In this regard, an interesting line of research is that of the managing of fuzzy clustering in the presence of outlier interval-valued data, i.e. robust fuzzy clustering [35,33].

According to Huber [56], a robust method should have the following properties:

- it should have a reasonably good accuracy at the assumed model;
- small deviations from the model assumptions should impair the performance only by a small amount;
- larger deviations from the model assumptions should not cause a catastrophe.

In the literature, in a fuzzy clustering framework for single valued data, different robust approaches can be considered, i.e.:

- *Noise approach*: outliers are assigned to the so-called noise cluster (see, e.g., [11–13,82]).
- *Semifuzzy approach*: the problem of outliers is avoided by setting a maximum number of clusters that a datum can belong to, setting membership degrees to zero if a predefined maximal distance is exceeded, or by defining a minimum threshold for membership degrees [84].
- *Metric approach*: metrics with robust properties are incorporated in the objective functions of the clustering problem; e.g., L_p norm $(0 < p < 1)$ [53], L_1 norm [60,65,72], exponential metric [92]. In the metric approach, we can consider the sub-approach *ordered statistics approach*. In this case the concept of weight functions in robust statistics is related to the concept of membership functions in fuzzy set theory or to possibility distributions in possibility theory. Different types of robust methods (such as the M, R, L estimators, and the least median of squares method) can be adopted [56]. Frigui and Krishnapuram [38] proposed a robust

fuzzy clustering algorithm, termed robust *C*-prototypes algorithm, based on a generalization of the M-estimator to estimate the *C* prototypes and on the application of a loss function ρ to squared distances to reduce the effect of outliers. Davé and Krishnapuram [12] discussed the M-estimator and the connection between robust statistics (i.e. the M-estimator) and fuzzy clustering. The M-estimator uses a suitable symmetric positive-definite function (called the robust-loss function) and forms the objective function by summing the loss over all points. Winkler et al. [91] proposed a robust fuzzy clustering with polynomial fuzzifier function in connection with M-estimators in which for each data point its distances to the prototypes are ordered. In particular, Leski [72] proposed a robust fuzzy clustering method using the Vapnik's ε -insensitive estimator in which the fuzzy *c*-medians can be obtained as a particular case of the suggested method. In order to improve the robustness of the fuzzy clustering, Leski [74] used together the Huber's M-estimators and the Yager's OWA (Ordered Weighted Averaging) operators [93]. In particular, Leski [74] combined the fuzzy *c*-means clustering with the robust ordered statistics using Huber's M-estimator to develop a new fuzzy clustering method called as Fuzzy *C*-Ordered-Means clustering.

- *Trimmed approach*: the clustering procedure is applied to the data remaining after a fixed fraction of outlying data objects are eliminated (see, e.g., [39–42,67]).
- *Possibilistic approach*: following the Possibilistic Theory membership degrees are expressed as degrees of compatibility (possibility degrees); in this approach, to avoid the problem of outliers, outliers are included in all clusters with small membership (see, e.g., [68,69,87,94,83]).
- *Evidential approach*: new clustering techniques based on the concept of belief functions [77].
- *Influence weighting approach*: a suitable weighting system for objects is used in the clustering process and low weights are objectively assigned to outliers [64,29].

In a non-single valued data analysis framework—i.e. fuzzy clustering for imprecise data (fuzzy and interval data)—only some of the previous methodological approaches have been explored. In particular, Butkiewicz [4] proposed a transformation of fuzzy data into crisp data and uses an *ordered statistics approach*-based classical robust fuzzy clustering model for single valued data. Hung and Yang [58] suggested a metric approach-based robust fuzzy clustering for univariate fuzzy data by considering an exponential distance. Zarandi et al. [96], suggested an *influence weighting approach*-based robust fuzzy clustering model for fuzzy data using a Wasserstein-type distance. Coppi et al. [10] proposed a *possibilistic approach*-based robust clustering method for multivariate fuzzy data. D'Urso and De Giovanni [32] proposed robust fuzzy clustering methods for fuzzy data based on different methodological approaches, i.e. based on *noise*, *metric* and *trimmed* approaches. For interval-valued data according to our knowledge only two robust fuzzy clustering methods have been proposed in the literature, i.e. the noise approach-based fuzzy clustering proposed by D'Urso and Giordani [35] and the trimmed approach-based fuzzy clustering suggested by D'Urso et al. [33].

In this paper, by adopting a Partitioning Around Medoids (PAM) procedure we propose a fuzzy clustering using Huber's M-estimators and Yager's OWA (Ordered Weighted Averaging) operators. The proposed method belongs to *the ordered statistics approach* and inherits the advantages connected to PAM clustering strategy, fuzzy theory and robust methodological approach; i.e.:

- By using a PAM-based clustering approach, each cluster is represented by an observed object (medoid) and not by a fictitious object (centroid) (as in the fuzzy *c*-means clustering),

which is very appealing and useful in a wide range of applications. This is very important for the interpretation of the selected clusters, as well.

- The concept of partial membership underlying the fuzzy approach appears to be more appealing and flexible than that of the classical clustering procedures [78,89]; furthermore, the fuzzy clustering methods have been shown to be computationally more efficient because dramatic changes in the value of cluster membership are less likely to occur in estimation procedures [78] and less afflicted by local optima problems [54]; in addition, the memberships for any given set of respondents indicate whether there is a second-best cluster almost as good as the best cluster—a result which traditional clustering models cannot uncover [37].

- Notice that PAM-based fuzzy clustering represents a robustification of the fuzzy c-means clustering; however, it provides only a "timid robustification" of the fuzzy c-means clustering, because a single outlier still serves to breakdown the clustering [39]. However, as remarked by Garcia-Escudero et al. [43] the clustering model based on medoids "resists to the presence of 1 outlier in a remote position, but it breaks down when we increase to 3 the number of outliers". For this reason, for neutralizing the disruptive effects of more outlier interval-valued data in the clustering process we propose a robust fuzzy c-medoids clustering method using a strategy based on the combination of the Huber's M-estimators and the Yager's OWA (Ordered Weighted Averaging) operators [93].

The paper is organized as follows. In Section 2, we describe the interval-valued data and a metric for comparing interval-valued data; successively, we propose a robust fuzzy clustering for interval-valued data belonging to ordered statistics sub-approach, by considering simultaneously the Huber's M-estimators and the Yager's Ordered Weighted Averaging (OWA) operators [93]. In particular, we combine the fuzzy c-medoids clustering with the robust ordered statistics using Huber's M-estimator to develop a new fuzzy clustering method called Fuzzy c-Ordered-Medoids clustering for interval-valued data (FcOMdC-ID). In Section 3, we present a simulation study and in Section 4 we apply our method to a real world case. In Section 5 some remarks conclude the paper.

2. A robust fuzzy clustering method for interval-valued data: FcOMdC-ID (Fuzzy c-Ordered Medoids Clustering for Interval-valued Data)

2.1. Interval-valued data

Algebrically an interval valued datum can be formalized as $x_{ij} = [\underline{x}_{kj}, \overline{x}_{kj}], k = 1, ..., N; j = 1, ..., J$ where x_{kj} represents the jth interval-valued variable observed on the kth object; $\underline{x}_{kj}$ and $\overline{x}_{kj}$ denote, respectively, the lower and upper bounds of the interval; in particular, they represent the minimum and maximum values

registered for the jth interval valued variable with respect to the kth object. Each object is represented geometrically by a *hyper-rectangle* in $\Re^J$ having 2^J vertices. The 2^J vertices correspond to all the possible (lower bound, upper bound) combinations. In particular, in $\Re(J = 1)$ the generic object is represented by a segment; in $\Re^2(J = 2)$, it is represented by a rectangle with $2^2 = 4$ vertices, and so on [8].

Let us assume J interval-valued variables are observed on N objects. Then, the data can be stored in the so-called *interval-valued matrix* as follows:

$$\mathbf{X} \equiv \{x_{kj} = [\underline{x}_{kj}, \overline{x}_{kj}] : \quad k = 1, ..., N; j = 1, ..., J\} \tag{1}$$

where x_{kj} represents the jth interval-valued variable observed on the kth object and $\underline{x}_{kj}$ and $\overline{x}_{kj}$ indicate, respectively, the lower bound (min) and upper bound (max) of the interval-valued datum x_{kj}.

If we deal with J single valued variables, each object is represented as a point in the reference space $\Re^J$. Instead, in the case of interval valued data, each object is represented as a hyperrectangle (in $\Re^J$) having 2^J vertices (a rectangle with $2^J = 4$ vertices if $J=2$). Then, the ith row of $\mathbf{X}$ refers to the ith object that can be seen as a hyperrectangle in $\Re^J$.

To this purpose, we indicate $\mathbf{M}$ as the matrix whose generic element is the *midpoint* (*center*) of the associated interval. In particular, the *midpoint matrix* (*center matrix*) is

$$\mathbf{M} \equiv \left\{ m_{kj} = \frac{\overline{x}_{kj} + \underline{x}_{kj}}{2} : \ k = 1, ..., N; j = 1, ..., J \right\}, \tag{2}$$

where m_{kj} is the midpoint (center) of the associated interval value for $k = 1, ..., N$ and $j = 1, ..., J$.

Furthermore, we define the *radius matrix* (*spread matrix*) $\mathbf{R}$ of order $(N \times J)$ in the following way:

$$\mathbf{R} \equiv \left\{ r_{kj} = \frac{\overline{x}_{kj} - \underline{x}_{kj}}{2} : \ k = 1, ..., N; j = 1, ..., J \right\}, \tag{3}$$

where r_{kj} is the radius (spread) of the associated interval value for $k = 1, ..., N$ and $j = 1, ..., J$.

Then, we can denote the generic interval-valued datum pertaining to the kth object with respect to the jth interval-valued variable as the ordered couple (m_{kj}, r_{kj}), where m_{kj} denotes the midpoint (center) and r_{kj} the radius (spread). In this way, the lower and upper bounds of the interval-valued datum can be obtained as $m_{kj} - r_{kj}$ and $m_{kj} + r_{kj}$, respectively.

By considering the previous reformulation of the interval-valued data, the *interval-valued matrix* can be reformalized as follows:

$$\mathbf{X} \equiv \left\{ x_{kj} = (m_{kj}, r_{kj}); k = 1, ..., N; j = 1, ..., J \right\}. \tag{4}$$

Notice that as for standard datasets (single valued dataset), also in the interval case, the data can be corrupted by noise and outliers. In particular, we can distinguish three possible types of outlier interval-valued data (see Fig. 1; inlier data=no outlier data; [29]):
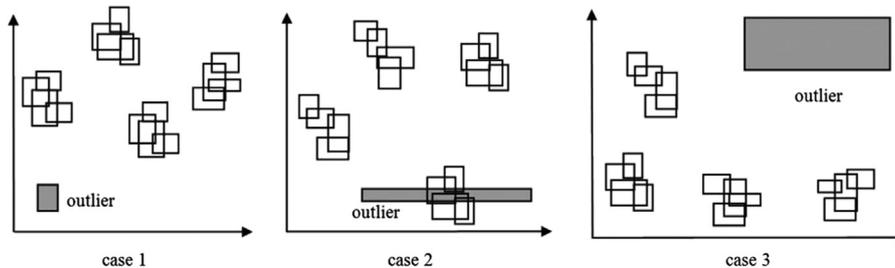
Fig. 1. Examples of outlier interval-valued data in $\Re^2$.

1. outlier midpoint (center)-inlier radius (spread);
2. inlier center-outlier spreads;
3. outlier center-outlier spreads.

2.2. Interval-valued data pre-processing.

Usually, in classical statistical analysis for single valued data, data standardization is recommended to prevent the results from being strongly affected by the measurement scale. Also in interval-valued data analysis, a suitable data pre-processing, centering, normalization and standardization is recommended. In the literature, there are several measures of standardization of interval data. For instance, de Carvalho et al. [15,18] proposed some alternative standardization methods for interval data. Guo et al. [48] proposed a standardization of interval data based on the empirical descriptive statistics. See, also Billard and Diday [3], D'Urso and Giordani [35], D'Urso and De Giovanni [31] and D'Urso et al. [33].

2.3. Distance measures for interval-valued data

In the literature, several distance measures for interval-valued data have been proposed; see, e.g., Ichino and Yaguchi [59], Gowda and Ravi [46], de Carvalho and de Souza [16] D'Urso and Giordani [34], Guru et al. [50] and Kim and Billard [66].

2.3.1. A distance measure based on the midpoint and radius of the interval data

A distance used for clustering interval-valued data is the distance measure proposed by D'Urso and Giordani [34] and successively adopted by D'Urso and De Giovanni [31] and D'Urso et al. [33].

In particular, in order to compare two generic objects k' and k'' described by J interval-valued variables, we consider all the vertices of the two hyperrectangles pertaining to the objects involved.

Thus, we have the following Euclidean distance measure for interval-valued data:

$$D(\mathbf{x}_{k'}, \mathbf{x}_{k''}) = \left[\sum_{k=1}^{2^J} \| (\mathbf{m}_{k'} + \mathbf{r}_{k'} * \mathbf{h}_\nu) - (\mathbf{m}_{k''} + \mathbf{r}_{k''} * \mathbf{h}_\nu) \|^2 \right]^{\frac{1}{2}} \quad (5)$$

in which $\mathbf{x}_{k'}, \mathbf{x}_{k''}, \mathbf{m}_{k'}, \mathbf{m}_{k''}, \mathbf{r}_{k'}$ and $\mathbf{r}_{k''}$ denote, respectively, the k'th and k''th rows of $\mathbf{X}, \mathbf{M}$ and $\mathbf{R}$, where the symbol $*$ is the Hadamard product, which is the elementwise product of two matrices (vectors) of the same order. Furthermore, the vectors $\mathbf{h}_\nu, \nu = 1, ..., 2^J$, help us to define every vertex of the hyperrectangle associated to each object separately. In fact, their elements are equal to ± 1 in order to refer exactly to every vertex. The elements of $\mathbf{h}_\nu, \nu = 1, ..., 2^J$, are the row of a new matrix, say $\mathbf{H}$, of order $(2^J \times J)$. If $J=3$, we get:

$$\mathbf{H} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}. \quad (6)$$

Using the vector $\mathbf{h}_1$ (first row of $\mathbf{H}$) we have the vector of the lower bounds pertaining to the kth object:

$$\left(\underline{x}_{k1} \ \underline{x}_{k2} \ \underline{x}_{k3} \right) = \mathbf{m}_k + \mathbf{r}_k * (-1 \ -1 \ -1) = \mathbf{m}_k - \mathbf{r}_k. \quad (7)$$

Analogous to (7), by means of $\mathbf{h}_5$, we can obtain the vector of the upper bounds:

$$(\overline{x}_{k1} \ \overline{x}_{k2} \ \overline{x}_{k3}) = \mathbf{m}_k + \mathbf{r}_k * (1 \ 1 \ 1) = \mathbf{m}_k + \mathbf{r}_k. \quad (8)$$

Thus, it can be shown that the Euclidean distance (5) can be simplified as

$$\begin{aligned} D(\mathbf{x}_{k'}, \mathbf{x}_{k''}) &= \left[2^J \| \mathbf{m}_{k'} - \mathbf{m}_{k''} \|^2 + 2^J \| \mathbf{r}_{k'} - \mathbf{r}_{k''} \|^2 \right]^{\frac{1}{2}} \\ &\approx \left[\| \mathbf{m}_{k'} - \mathbf{m}_{k''} \|^2 + \| \mathbf{r}_{k'} - \mathbf{r}_{k''} \|^2 \right]^{\frac{1}{2}} \\ &= \left[\sum_{j=1}^{J} \left[(m_{k'j} - m_{k''j})^2 + (r_{k'j} - r_{k''j})^2 \right] \right]^{\frac{1}{2}}. \end{aligned} \quad (9)$$

Notice that the distance measure (9) is the same distance used in de Carvalho et al. [15], which is a special case of the one in de Carvalho and de Souza [16].

2.3.2. Robust distance measures for interval data

The distance measure recalled in the previous subsection uses a quadratic loss function ($\mathcal{L}_{\text{SQR}}(e) = e^2$, where the so-called model residuals are $e = m_{k'j} - m_{k''j}$ or $e = r_{k'j} - r_{k''j}$) as a dissimilarity measure between the interval-valued data and the medoids. The reason for using this measure is mathematical, that is, for simplicity and low computational burden. However, this approach is sensitive to noise and outliers. In the literature there are many proposals of robust loss functions. For example, the Huber's one is of special interest [56]:

$$\mathcal{L}_{\text{HUB}}(e) = \begin{cases} e^2/\delta^2, & |e| \le \delta, \\ |e|/\delta, & |e| > \delta, \end{cases} \quad (10)$$

where $\delta > 0$ denotes a parameter. Many other robust loss functions may be taken into account:

- LINear (LIN)

$$\mathcal{L}_{\text{LIN}}(e) = |e|, \quad (11)$$

- SIGmoidal (SIG) with parameters $\alpha, \beta > 0$

$$\mathcal{L}_{\text{SIG}}(e) = 1/(1 + \exp(-\alpha(|e| - \beta))), \quad (12)$$

- LOGarithmic (LOG)

$$\mathcal{L}_{\text{LOG}}(e) = \log(1 + e^2). \quad (13)$$

For a vector argument $\mathbf{e} = [e_1, e_2, ..., e_J]$ a loss function takes the form:

$$\mathcal{L}(\mathbf{e}) = \sum_{j=i}^{J} \mathcal{L}(e_j), \quad (14)$$

2.4. The clustering method

The different robust loss functions are connected to robustness of the distance measure. Obviously, in this case we consider the distance between pair of objects (and then not between object and medoid). Using $\mathcal{D}(\mathbf{x}_k, \tilde{\mathbf{x}}_i) = \mathcal{L}(\mathbf{m}_k - \tilde{\mathbf{m}}_i) + \mathcal{L}(\mathbf{r}_k - \tilde{\mathbf{r}}_i)$ as a dissimilarity measure between the kth interval-valued datum and the ith medoid, and additional weighting, the fuzzy c-ordered medoids criterion function takes the form:

$$J(\mathbf{U}, \tilde{\mathbf{X}}) = \sum_{i=1}^{c} \sum_{k=1}^{N} \beta_k (u_{ik})^\lambda \mathcal{D}(\mathbf{x}_k, \tilde{\mathbf{x}}_i), \quad (15)$$

where $\tilde{\mathbf{X}} = \{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, ..., \tilde{\mathbf{x}}_c\}$. The ith medoid is taken from dataset, i.e., $\tilde{\mathbf{x}}_i \in \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N\}$. λ is a weighting exponent in $[1, \infty)$. This parameter influences a fuzziness of the clusters. $\beta_k \in [0, 1]$ denotes the typicality of the kth datum with respect to the clusters; smaller β_k results in a more atypical data. These parameters are derived based on the ordering of the distances of data from medoids [74,75].

The set of all possible fuzzy partitions of N vectors into c clusters is defined by

$$\mathcal{J}_{gfc} = \left\{ \mathbf{U} \in \mathfrak{R}^{c \times N} \Big|_{\substack{1 \leq i \leq c \\ 1 \leq k \leq N}} u_{ik} \in [0, 1]; \sum_{i=1}^{c} u_{ik} = 1; 0 < \sum_{k=1}^{N} u_{ik} < N \right\}, \quad (16)$$

The overall assessment of the typicality of the kth datum is obtained using s-norm S ($\star_S$) [74]:

$$\underset{1 \leq k \leq N}{\forall} \beta_k = \beta_{1k} \star_S \beta_{2k} \star_S \cdots \star_S \beta_{ck}, \quad (17)$$

which may be linguistically interpreted as the following sentence: "The kth datum is typical IF AND ONLY IF the kth datum is typical with respect to the first cluster OR the kth datum is typical with respect to the second cluster OR ⋯ OR the kth datum is typical with respect to the cth cluster". $\beta_{ik} \in [0, 1]$ denotes the typicality of the kth datum with respect to the ith cluster. Instead of s-norm, the maximum operation is chosen

$$\underset{1 \leq k \leq N}{\forall} \beta_k = \beta_{1k} \vee \beta_{2k} \vee \cdots \vee \beta_{ck}. \quad (18)$$

The necessary conditions for the minimization of (15) with respect to the elements of the partition matrix can be described as

$$\underset{\substack{1 \leq k \leq N \\ 1 \leq s \leq c}}{\forall} u_{sk} = \mathcal{D}(\mathbf{x}_k, \tilde{\mathbf{x}}_s)^{\frac{1}{1-\lambda}} \Bigg/ \left[\sum_{j=1}^{c} \mathcal{D}(\mathbf{x}_k, \tilde{\mathbf{x}}_j)^{\frac{1}{1-\lambda}} \right]. \quad (19)$$

A local optimal solution with respect to the medoids can be obtained as follows [33]:

$$\underset{1 \leq i \leq c}{\forall} \tilde{\mathbf{x}}_i = \mathbf{x}_{q_i}, \quad (20)$$

where

$$q_i = \underset{1 \leq \ell \leq N}{\arg \min} \sum_{k=1}^{N} \beta_k (u_{ik})^{\lambda} \mathcal{D}(\mathbf{x}_k, \mathbf{x}_{\ell}). \quad (21)$$

Let $\pi_i : \{1, 2, ..., N\} \rightarrow \{1, 2, ..., N\}$ be the permutation function for ith medoid. The rank-ordered dissimilarities between ith medoid and data satisfy the following conditions:

$$\mathcal{D}(\mathbf{x}_{\pi_i(1)}, \tilde{\mathbf{x}}_i) \leq \mathcal{D}(\mathbf{x}_{\pi_i(2)}, \tilde{\mathbf{x}}_i) \leq \mathcal{D}(\mathbf{x}_{\pi_i(3)}, \tilde{\mathbf{x}}_i) \leq \cdots \leq \mathcal{D}(\mathbf{x}_{\pi_i(N)}, \tilde{\mathbf{x}}_i). \quad (22)$$

If β_{ik} parameters fulfill $\beta_{i\pi_i(1)} \geq \beta_{i\pi_i(2)} \geq \cdots \geq \beta_{i\pi_i(N)}$, then the impact of outliers is reduced by down-weighting the respective dissimilarities. The form of parameters β_{ik} may be piecewise-linear [74,75]

$$\beta_{i\pi_i(n)} = \left\{ \left[(p_c N - n)/(2 p_l N) + 0.5 \right] \wedge 1 \right\} \vee 0 \quad (23)$$

or sigmoidal

$$\beta_{i\pi_i(n)} = 1 \Big/ \left\{ 1 + \exp \left[\frac{2.944}{p_a N} (n - p_c N) \right] \right\}, \quad (24)$$

where $\wedge$ and $\vee$ denote min and max operations, respectively. The functions defined by (23) and (24) are called Piecewise-Linearly-weighted OWA (PLOWA) and Sigmoidally-weighted OWA (SOWA), respectively. Both functions are nonincreasing with respect to argument $n \in \{1, 2, ..., N\}$. For $n = p_c N$ both functions are equal to 0.5. Parameters $p_l > 0$ and $p_a > 0$ influence their slope. In the case of the piecewise-linear function, for $n \in [p_c N - p_l N, p_c N + p_l N]$ its value linearly decreases from 1 to 0 [74]. For the sigmoidal function, a value of 2.944 is chosen to obtain that for $n \in [p_c N - p_a N, p_c N + p_a N]$ its value decreases from 0.95 to 0.05 [74]. The following

values: $p_c \in [0.7, 1.0]$, $p_l = 0.1$, $p_a = 0.1$ were used. If ordering of dissimilarities from (22) is not applied, which is equivalent to using Uniformly weighting function for OWA–UOWA ($\beta_{i\pi_i(n)} = 1$ for all i, n), then we call this case as clustering without ordering (or with no weighting function).

The final form of the proposed Fuzzy c-Ordered-Medoids Clustering for Interval-valued Data (FcOMdC-ID) can be described

```
Algorithm FcOMdC-ID
```

1. Fix $c(1 < c < N)$, $\lambda \in (1, \infty)$. Choose dissimilarity measure. Initialize $\tilde{\mathbf{X}}^{(0)} \subset \mathbf{X}$, $\beta_{ik} = 1$ and set the iteration index $\ell = 1$,
2. Calculate the fuzzy partition matrix $\mathbf{U}^{(\ell)}$ for the ℓth iteration using (19),
3. Rank-order the dissimilarities between the ith medoid and data (see (22)) obtaining the permutation function $\pi_i(k)$,
4. Calculate $\beta_{i\pi_i(k)}$ using (23) or (24) or uniform weighting,
5. Update overall typicality parameters β_k using (18),
6. Update the medoids for the ℓth iteration $\tilde{\mathbf{X}}^{(\ell)}$ using $\mathbf{U}^{(\ell)}$ and (20) and (21),
7. If $\|\tilde{\mathbf{X}}^{(\ell+1)} - \tilde{\mathbf{X}}^{(\ell)}\|_F > \xi$ then $j \leftarrow j + 1$ and go to Step 2 else stop.

Notice that the fuzziness parameter λ plays an important role in fuzzy clustering. It should be suitably chosen in advance. Although $1 < \lambda < \infty$, values too close to 1 will result in a partition with all memberships close to 0 or 1. Excessively large values will lead to disproportionate overlap with all memberships close to $1/c$ [90]. Consequently, neither of these types of λ is recommended [1]. Although there have been some empirical heuristic procedures to determine the value of λ (e.g., [78,90,32]), there seems to exist no theoretically justifiable manner of selecting λ. Wang et al. [88] and Yang et al. [95] suggest $\lambda = 2$ and Leski [73] uses $\lambda = 1.1$. Since the medoid always has membership equal to 1 in the cluster, raising its membership to the power λ has no effect. Thus, when λ is high, the mobility of the medoids may be lost [30]. For this reason, a value between 1 and 1.5 for λ is recommended by Kamdar and Joshi [62]. For a discussion and a detailed list of references on the choice of λ see D'Urso and De Giovanni [32].

Notice that, with respect to the clustering methods for interval data suggested in the literature, our clustering method has different advantages; i.e.:

- *Our method vs standard (non fuzzy) clustering methods* (e.g., the de Souza-de Carvalho's method [23] and the de Carvalho et al.'s method [18]) based on, respectively, adaptive City-block and Hausdorff distances; de Carvalho-Lechevallier's methods [19,20] based on, respectively, single Hausdorff and City-block adaptive distances and adaptive quadratic distances; de Carvalho–de Souza's method ([21] based on single and adaptive distances): conversely to standard clustering methods for interval data, our method inherits all the benefits of the fuzzy approach to cluster analysis and of the PAM-based clustering approach (see, Section 1); in addition our method is robust to the presence of one or more outliers in the datasets (see Section 1).
- *Our method vs. non-robust fuzzy clustering methods* (i.e. the de Carvalho's method [14] and the de Carvalho–Tenorio's method [21] based on the k-means algorithm): with respect to non-robust fuzzy clustering methods proposed in the literature based on the k-means algorithm our method inherits the benefits connected to the PAM-based clustering approach (see, Section 1) and is resistant to the presence of outliers neutralizing suitably the disruptive effects of possible anomalous data in the clustering process (see Sections 1 and 2).
- *Our method vs robust fuzzy clustering methods* (i.e. the D'Urso-Giordani's method [35]) based on the k-means algorithm and the noise approach; the D'Urso et al.'s method ([33] based on the PAM and the trimmed approach): with respect to D'Urso

et al.'s method [33] our method is more informative on the intensity of the outliers. In particular, since the D'Urso et al.'s method is based on the trimmed approach it neutralizes the outliers and suitably removes them from the clustering process. Then this method achieves its robustness with respect to outliers by trimming away a certain fraction of the data objects and requires the specification of the "trimming ratio", that is the percentage of the data objects that has to be trimmed, i.e., the percentage of objects discarded in the clustering process and, thus, not considered in the optimization problem [33]. In this way, the method neutralizes the negative effect of the outliers, but no additional information about the intensity of the anomaly (atypicality) of the outliers is provided with respect to the entire dataset. Conversely, our proposed method neutralizes the disruptive effect of the outliers adopting the PAM approach and combining the Huber's M-estimators and the Yager's OWA operators; in this way by means of the typicality parameters $\beta_k \in [0, 1]$ our method provides suitable information on the atypicality level of the data: if outliers are present in the dataset, our method tends to give them typicalities very low or close to 0. Then, typicality is an important means for alleviating the undesirable effects of outliers. In fact, by means of the typicality our method tunes suitably the influence of the outlier interval data in the clustering process. With respect to D'Urso–Giordani's method [35], our method inherits the benefits connected to the typicality information and to the PAM-based clustering approach (see Section 1).

3. Simulation study

In all experiments for FcOMdC-ID the weighting exponent $\lambda = 1.5$ was used. The iterations were stopped as soon as the Frobenius norm of the successive $\tilde{\mathbf{X}}$ matrices difference was less than 10^{-4}. All experiments were run in the MATLAB environment. The purpose of this experiment was to investigate the sensitivity to various types and number of outliers for the traditional FcMdC-ID and the FcOMdC-ID methods. The two-dimensional (two features vectors) interval-valued data sets, presented in Fig. 2, consist of three well-separated groups (each of 25 points), and a varying number of outliers. For each group interval-valued data were generated randomly. Both the centers and the spreads were from Gaussian distribution with mean $\mathfrak{m}$ and standard deviation $\mathfrak{s}$; $\mathcal{N}(\mathfrak{m}, \mathfrak{s})$. The centers (two features) and the spreads of the three groups were as follows:

- first group:
  centers $\mathcal{N}(10, 2)$, $\mathcal{N}(10, 1)$ and spreads $\mathcal{N}(2, 0.05)$, $\mathcal{N}(1, 0.05)$,
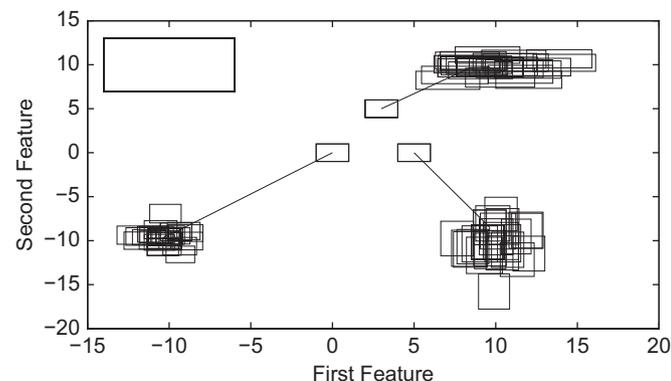
- second group:
  centers $\mathcal{N}(10, 1)$, $\mathcal{N}(-10, 2)$ and spreads $\mathcal{N}(1, 0.05)$, $\mathcal{N}(2, 0.05)$,
- third group:
  centers $\mathcal{N}(-10, 1)$, $\mathcal{N}(-10, 1)$ and spreads $\mathcal{N}(1, 0.05)$, $\mathcal{N}(1, 0.05)$.

The tested methods were initialized using medoids with centers $(5, 0)$, $(0, 0)$ and $(3, 5)$, and spreads $(1, 1)$, $(1, 1)$ and $(1, 1)$, marked on the figure. For each experiment all outliers are identical. This is the least favorable case, since if in the place of such outliers a medoid would be placed, it would cause the greatest possible reduction criterion function. When outliers are dispersed reducing the values of the criterion would not be so big. In other words, 'force' attracting prototypes for such outliers is the largest. Thus, it is a drastic method for the verification of resistance of clustering methods. The number of outliers varies from 0 (no outliers) to 25. There are three types of outliers:

- with respect to the centers and spreads: centers $(-10, 10)$ and spreads $(4, 3)$ (see Fig. 2),
- with respect to the spreads: centers $(-10, 10)$ and spreads $(1, 1)$ (see Fig. 3),
- with respect to the centers: centers $(-10, -10)$ and spreads $(0.25, 9)$ (see Fig. 4).

Fig. 2 illustrates the performance of the FcOMdC-ID method for 20 centers and spreads outliers, and the SIG loss function. In this figure, we can observe the traces of the medoids centers (lines) and spreads (boldfaced rectangles) calculated in the successive iterations. We can see that the medoids terminate near the expected clusters centers and spreads. Fig. 3 illustrates the performance of the FcOMdC-ID method for 20 spreads outliers and the SIG loss function. Fig. 4 illustrates the performance of the FcOMdC-ID method for 20 center outliers and the SIG loss function. We can notice that despite the outliers type the medoids terminate near the expected clusters centers.

The effects of the FcMdC and the FcOMdC-ID methods investigation for varying number of outliers and its type are presented in Figs. 5–7. For the computed terminal medoids, we measured the performance of clustering by the Frobenius norm of the difference between the expected value of centers/spreads matrix and the terminal medoids matrix. Figs. 5 and 7 show that for few outliers (from 1 to 4 center or center/spread outliers type) the terminal medoids determined by all methods are close to the true centers/spreads. But for a greater number of outliers the terminal prototypes errors are smaller for FcOMdC-ID than for the FcMdC method. We can conclude that the best performance of the FcOMdC-ID method is obtained for the SIG loss function.

Fig. 2. Performance of the FcOMdC-ID method with the SIG loss function for 20 center and spread outliers.
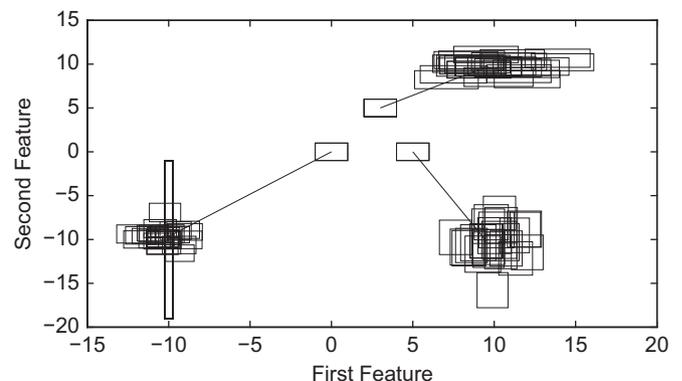
Fig. 3. Performance of the FcOMdC-ID method with the SIG loss function for 20 spread outliers.

Fig. 6 shows that that despite the large number of outliers (spread outliers type) the terminal medoids determined by all methods are close to the true centers/spreads. But the best results were archived for the FcOMdC-ID method with SIG and HUB loss functions.

Fig. 7 shows that the FcMdC method performance is catastrophically deteriorated for 3 outliers. For the FcOMdC-ID method the terminal medoids determined for all loss functions are close to the true centers/spreads. The performance is catastrophically deteriorated for the number of outliers equals to 18, for the LOG loss function only. However, even for smaller number of outliers and for the FcOMdC-ID method with SQR loss function rather serious increase of errors is visible. In the case of the center outliers, the best results were archived for the FcOMdC-ID with SIG loss function.
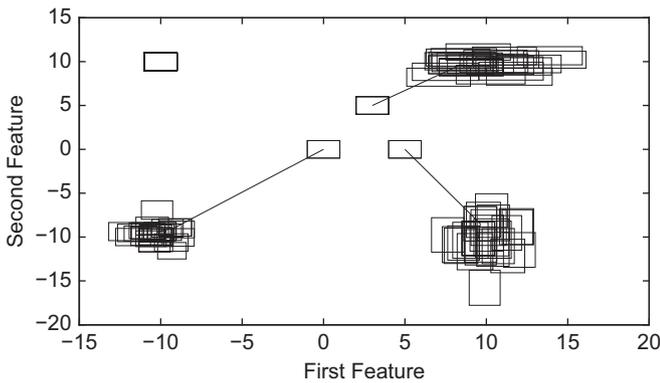
In all the above experiments, the groups are of the same cardinality (equal to 25). When groups have different cardinality (unbalance data), as it seems, the fewest cardinality determines the resistance of the method. In order to verify this hypothesis, we conducted the following experiment. The cardinality of one of the group (located in the upper right corner of Fig. 2) is reduced from 25 to 12 elements. Other groups remained unchanged. The centers and spread outliers are added to data. The effects of the FcMdC and the FcOMdC-ID methods for such data and for varying number of outliers are presented in Fig. 8. Comparing this figure to the corresponding experiment on balanced groups cardinality, we see that the supposition proved true. In fact, the behavior of the FcOMdC-ID method for individual loss functions is similar, but scaled. Similar values for errors occur for twice fewer outliers. It is also noted that for SQR loss function we get similar errors as for the FcMdC method. As before, the best results were obtained for LIN, SIG and HUB loss functions.

Another experiment was devoted to checking what are the effects of the FcOMdC-ID method, when overlapping groups are present in the data. For this purpose, processed data are shown in Fig. 2. There have been "move up" about 14 units of the group located in the bottom right corner, to give data presented graphically in Fig. 9. Other groups remained unchanged. The center and spread outliers are added to data. The effects of the FcMdC and FcOMdC-ID methods for such data and for varying number of outliers are presented in Fig. 10. Comparing this figure to the corresponding experiment on well separated groups, we see that in most cases the results are very similar to those obtained for well-separated groups of data. However, a surprise can be better performance of the FcMdC method and worse performance of the FcOMdC-ID method with the LOG loss function.

The next experiment was conducted to compare the developed method to the trimmed fuzzy c-medoids for interval-valued data

Fig. 4. Performance of the FcOMdC-ID method with the SIG loss function for 20 center outliers.
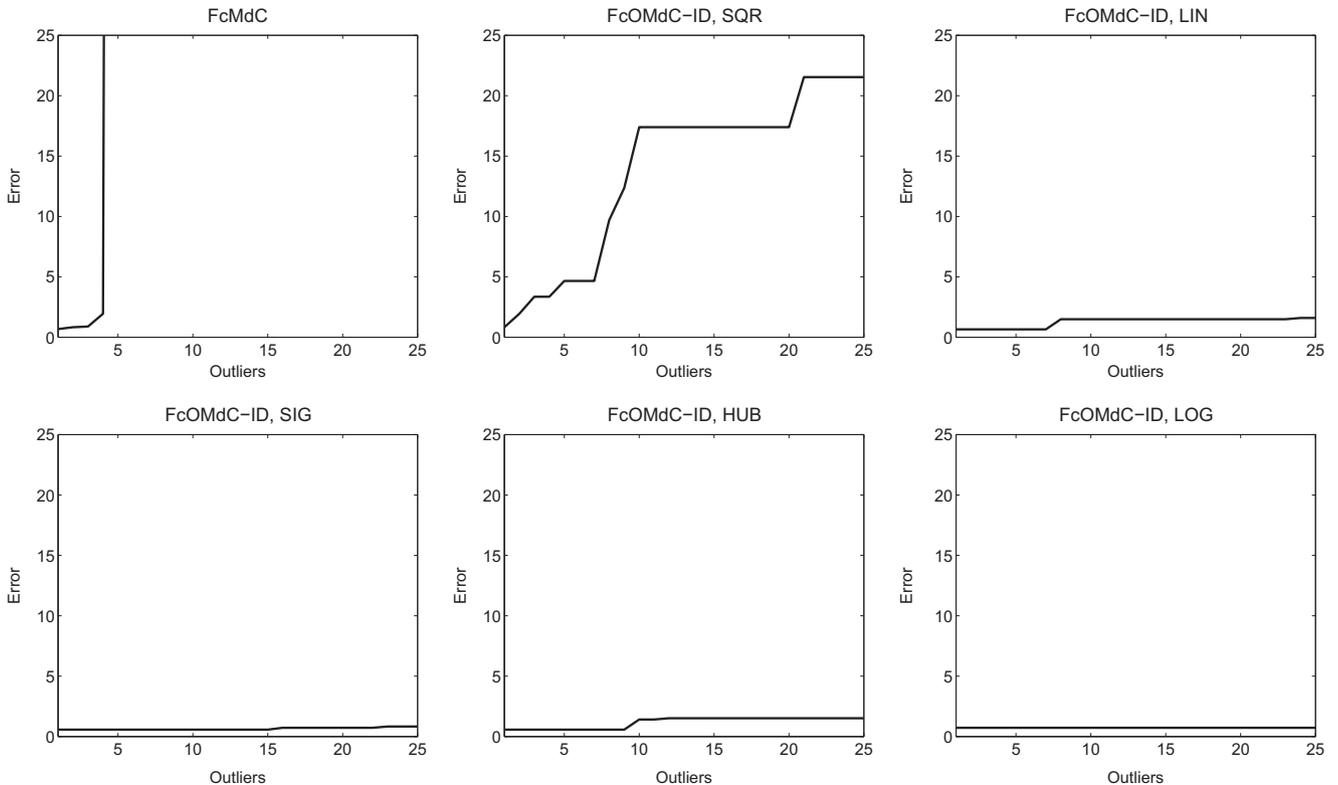
Fig. 5. The Frobenius norm of the clusters centers errors for various numbers of center and spread outliers in the synthetic data from Fig. 2. The traditional FcMdC and the FcOMdC-ID with various loss functions are presented in subplots.

method (TrFCMd-ID). Experiments were performed as before for three types of outliers: (1) center and spread outliers, (2) spread outliers and (3) center outliers. For the TrFCMd-ID method, two values of the trimming parameter $\alpha' = \{0.75, 0.95\}$ were used. The effects of the investigation for varying number of outliers and its type are presented in Fig. 11. The scales of individual subplots
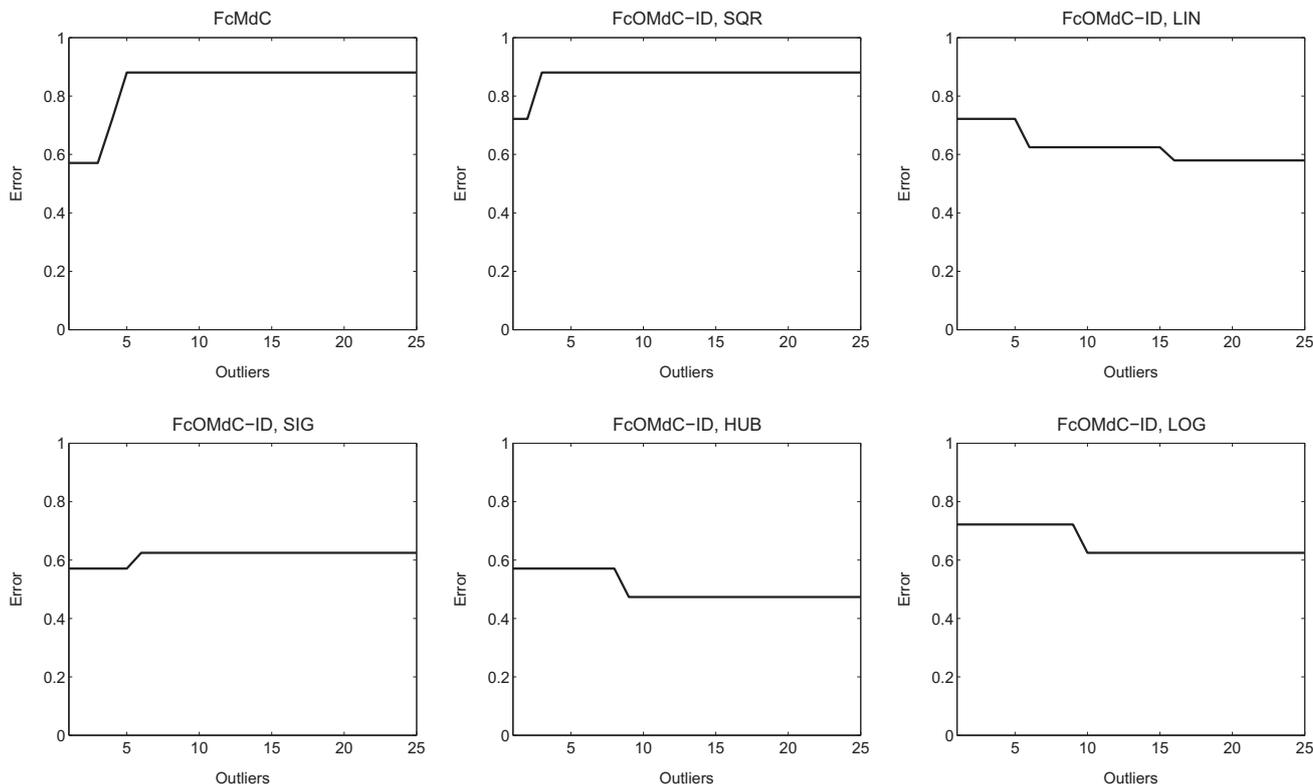
Fig. 6. The Frobenius norm of the cluster centers errors for various numbers of spread outliers in the synthetic data from Fig. 3. The traditional FcMdC and the FcOMdC-ID with various loss functions are presented in subplots.
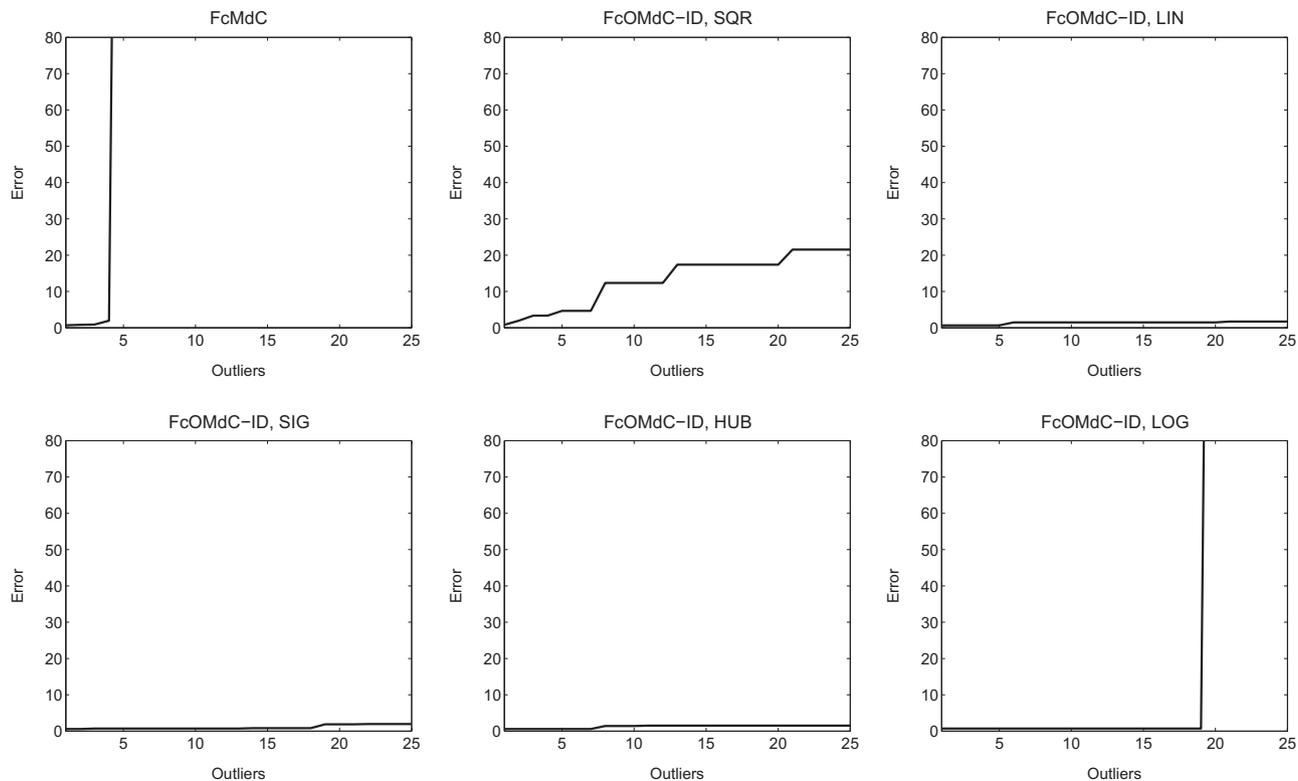
Fig. 7. The Frobenius norm of the cluster centers errors for various numbers of center outliers in the synthetic data from Fig. 4. The traditional FcMdC and the FcOMdC-ID with various loss functions are presented in subplots.
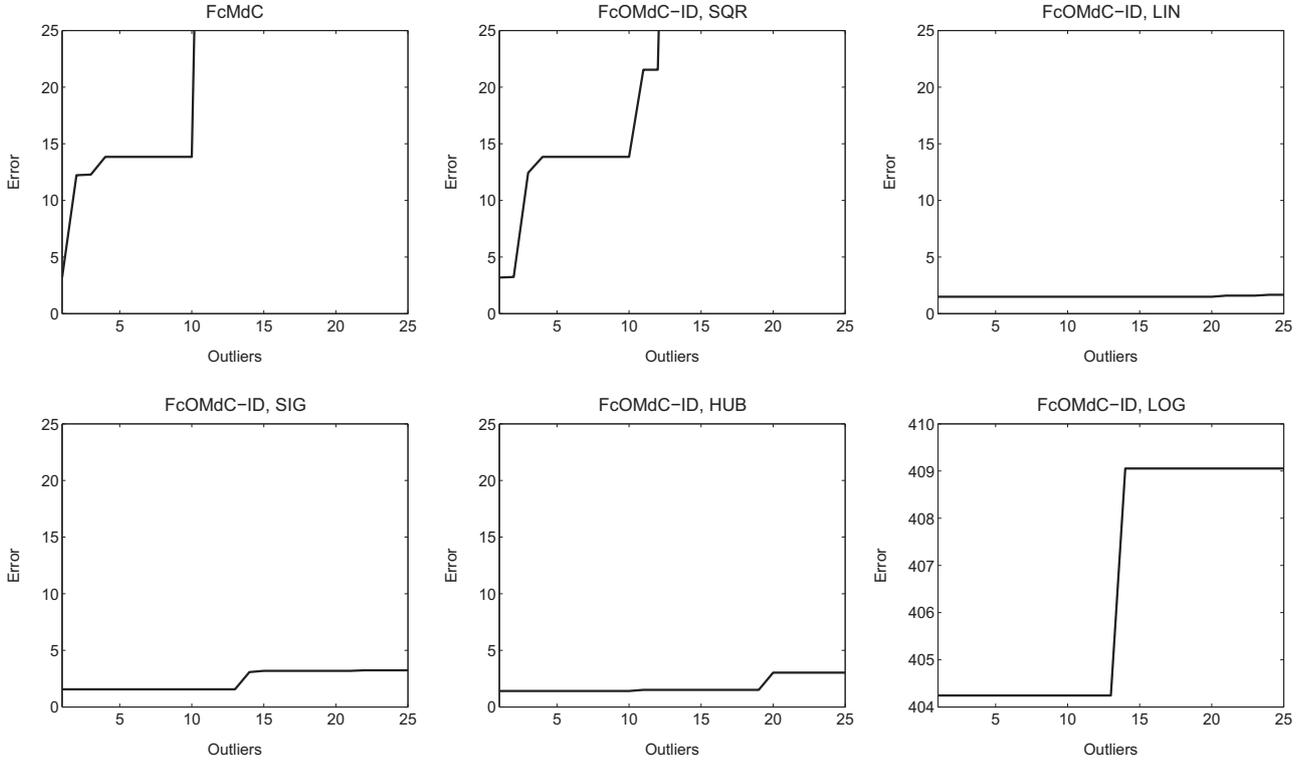
Fig. 8. The Frobenius norm of the cluster center errors for various numbers of center and spread outliers in the unbalanced synthetic data from Fig. 2. The traditional FcMdC and the FcOMdC-ID with various loss functions are presented in subplots.
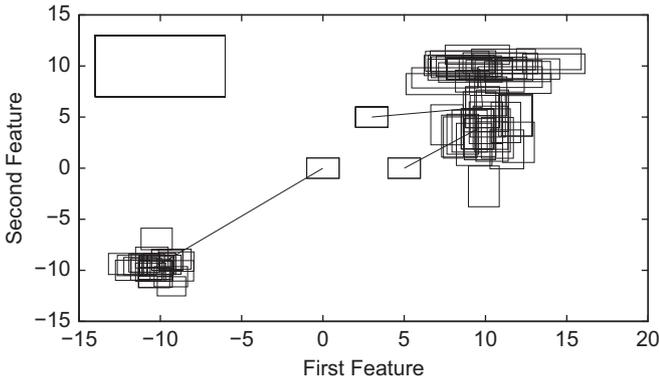
Fig. 9. Performance of the FcOMdC-ID method with the SIG loss function for overlapping groups of data and 20 center and spread outliers.

correspond to those used in Figs. 5–7, for easy comparison. Taking these figures into account, several observations can be made. First of all, it should be noted that, the less value of $\alpha' = 0.75$ causes the method lead to a very small errors for a small number of outliers (for the center, and the center and spread outliers). However, when the specified number of outliers is exceeded then errors of the method grow dramatically. For the spread outliers, behavior of the TrFCMd-ID and FcOMdC-IDs methods are similar. On the other hand, greater value $\alpha' = 0.95$ causes, the TrFCMd-ID method leads to errors comparable to the FcOMdC-ID method for SQR function loss. In conclusion, the experiment shows that for SIG and HUB loss functions, method FcOMdC-ID leads to smaller errors compared to the TrFCMd-ID method. This is due to the simultaneous use of two methods to obtain resistant behavior to outliers: the loss function and the OWA operation.

4. Application: identification of climatic areas

4.1. Temperature data (Guru's data, 2004, 2005)

In this subsection, an application of the proposed algorithm FcOMdC-ID with various loss functions to real dataset is shown. We use the minimum and maximum temperature in degree centigrade (interval temperature data) over 12 months observed in 37 cities. Data are drawn from Guru et al. [50] and Guru and Kiranagi [49]. The dataset is shown in Table 1. Notice that, with respect to Guru's dataset, in Table 1 we have added another city (Ojmjakon) considered in the next application (Section 4.2). The location map of the 37 cities and Ojmjakon is shown in Fig. 12, while the latitudes and longitudes are reported in Table 2. In the experiments we consider the number of clusters ranging from 2 to 6, and parameter of SOWA ordering function $p_c \in \{0.7, 0.8, 0.9, 1.0\}$. To obtain the best number of clusters and the value of p_c the Fuzzy Silhouette (FS) is used [6]

$$FS = \frac{\sum_{k=1}^{N} (u_{rk} - u_{qk})^{\gamma} s_k}{\sum_{k=1}^{N} (u_{rk} - u_{qk})^{\gamma}}, \tag{25}$$

where u_{rk}, u_{qk} are the first and second largest elements of the kth column of fuzzy partition matrix **U**, γ is a weighting coefficient (in our experiments equals 1.0), and

$$s_k = \frac{b_k - a_k}{b_k \vee a_k}, \tag{26}$$

where b_k is the minimum of the distances (using dissimilarity $\mathcal{D}$) of kth datum to all remaining data from other clusters, a_k is the average distance of kth datum to remaining data belonging to its highest membership cluster.

Tables 3–5 show, for different values of λ ($\lambda = 1.5, \lambda = 1.8, \lambda = 2.0$, respectively) the fuzzy silhouette value obtained for climatic dataset using the FcOMdC-ID algorithm for various number

of clusters, the parameter of the SOWA function and the type of loss function. Each element of the tables give the greatest value of FS obtained for 100 independent trials over different random choices of initial medoids. The higher the value of FS the better the assignment of data to the clusters (clusters are more compact and more well-separated). It should be noted that one should not
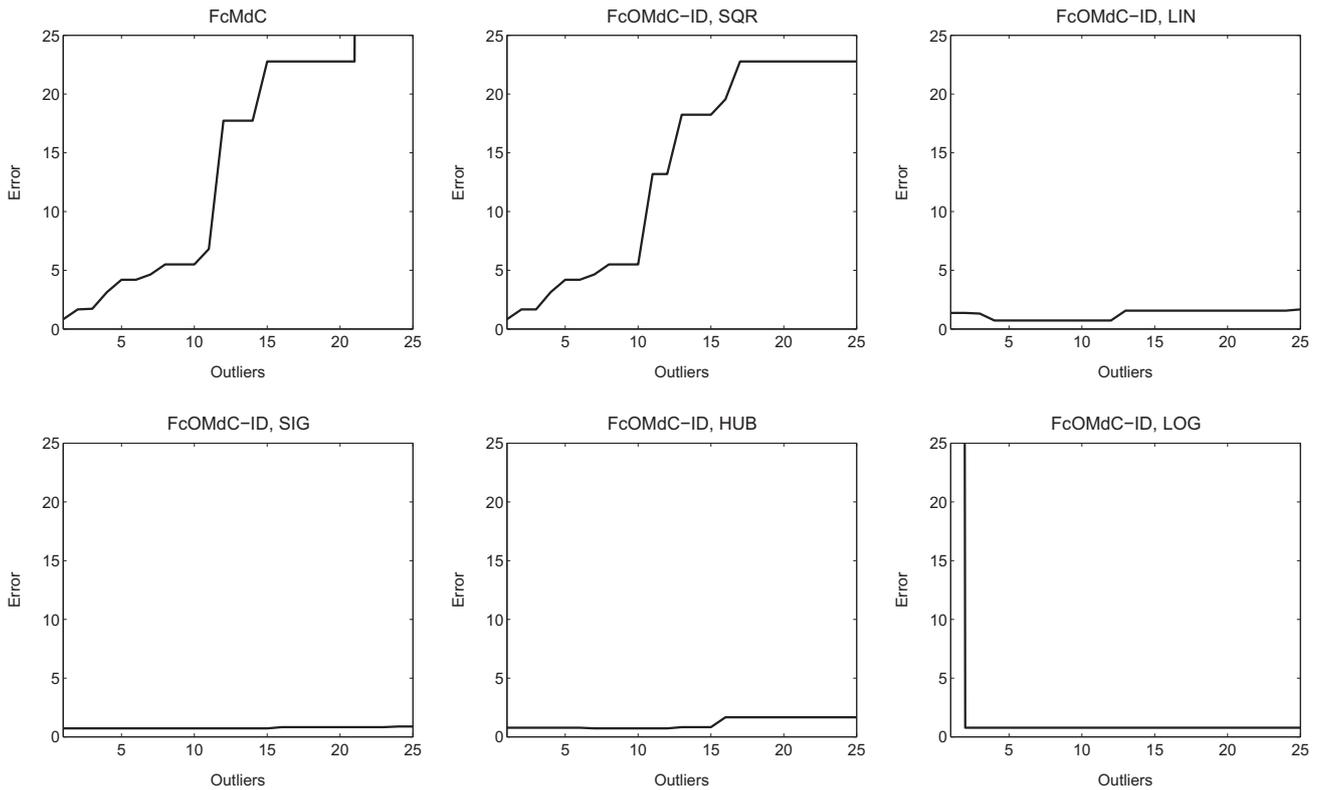
Fig. 10. The Frobenius norm of the clusters centers errors for various numbers of center and spread outliers in the synthetic data (with overlapping groups) from Fig. 9. The traditional FcMdC and the FcOMdC-ID with various loss functions are presented in subplots.
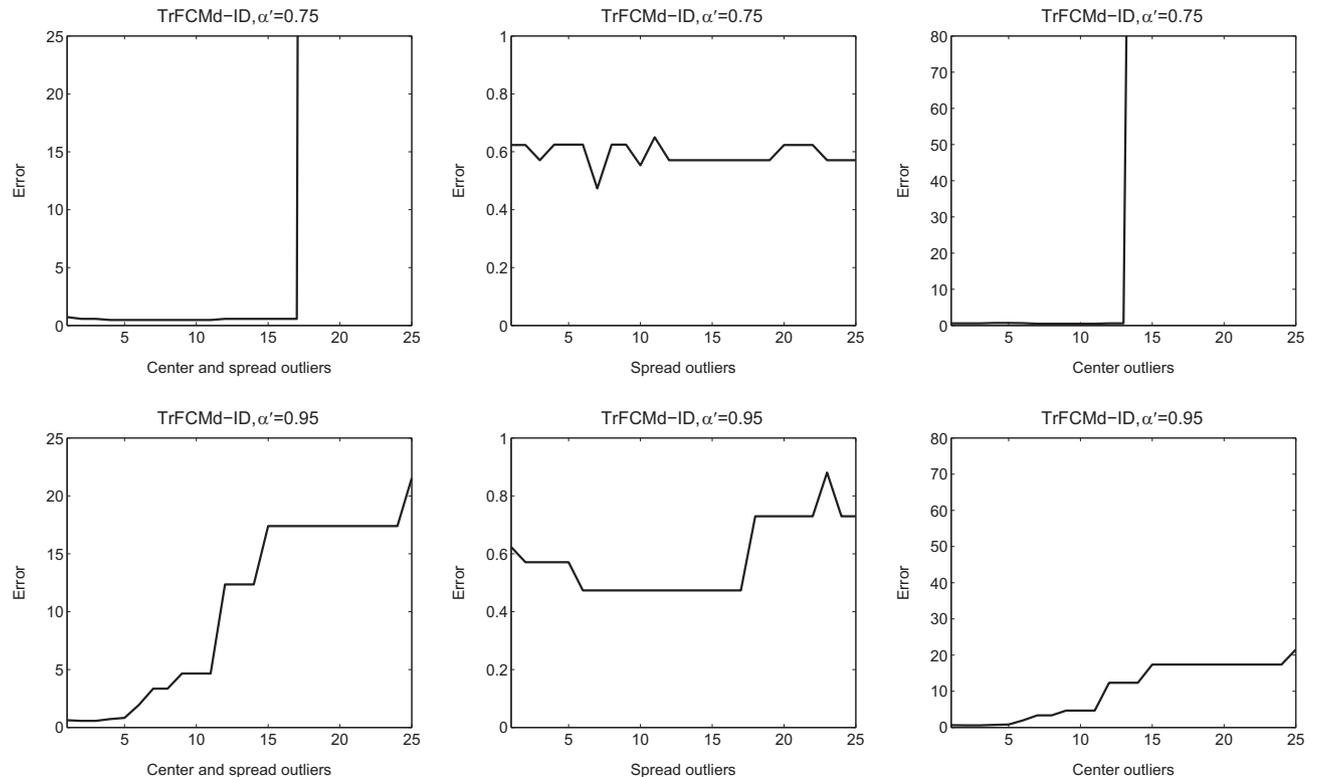
Fig. 11. The Frobenius norm of the cluster center errors for various numbers of outliers in the synthetic data from Figs. 2 to 4. The TrFCMd-ID method with various trimming value and types of outliers is presented in subplots.

Table 1
City temperature interval dataset.

No.	City	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug	Sept	Oct	Nov	Dec.
0	Amsterdam	[−4,4]	[−5,3]	[2,12]	[5,15]	[7,17]	[10,20]	[10,20]	[12,23]	[10,20]	[5,15]	[1,10]	[−1,4]
1	Athens	[6,12]	[6,12]	[8,16]	[11,19]	[16,25]	[19,29]	[22,32]	[22,32]	[19,28]	[16,23]	[11,18]	[8,14]
2	Bahrain	[13,19]	[14,19]	[17,23]	[21,27]	[25,32]	[28,34]	[29,36]	[30,36]	[28,34]	[24,31]	[20,26]	[15,21]
3	Bombay	[19,28]	[19,28]	[22,30]	[24,32]	[27,33]	[26,32]	[25,30]	[25,30]	[24,30]	[24,32]	[23,32]	[20,30]
4	Cairo	[8,20]	[9,22]	[11,25]	[14,29]	[17,33]	[20,35]	[22,36]	[22,35]	[20,33]	[18,31]	[14,26]	[10,20]
5	Calcutta	[13,27]	[16,29]	[21,34]	[24,36]	[26,36]	[26,33]	[26,32]	[26,32]	[26,32]	[24,32]	[23,32]	[20,30]
6	Colombo	[22,30]	[22,30]	[23,31]	[24,31]	[25,31]	[25,30]	[25,29]	[25,29]	[25,30]	[24,29]	[23,29]	[22,30]
7	Copenhagen	[−2,2]	[−3,2]	[−1,5]	[3,10]	[8,16]	[11,20]	[14,22]	[14,21]	[11,18]	[7,12]	[3,7]	[1,4]
8	Dubai	[13,23]	[14,24]	[17,28]	[19,31]	[22,34]	[25,36]	[28,39]	[28,39]	[25,37]	[21,34]	[17,30]	[14,26]
9	Frankfurt	[−10,9]	[−8,10]	[−4,17]	[0,24]	[3,27]	[7,30]	[8,32]	[8,31]	[5,27]	[0,22]	[−3,14]	[−8,10]
10	Geneva	[−3,5]	[−6,6]	[3,9]	[7,13]	[10,17]	[15,17]	[16,24]	[16,23]	[11,19]	[6,13]	[3,8]	[−2,6]
11	Hong Kong	[13,17]	[12,16]	[15,19]	[19,23]	[22,27]	[25,29]	[25,30]	[25,30]	[25,29]	[22,27]	[18,23]	[14,19]
12	Kuala Lumpur	[22,31]	[23,32]	[23,33]	[23,33]	[23,32]	[23,32]	[23,31]	[23,32]	[23,32]	[23,31]	[23,31]	[23,31]
13	Lisbon	[8,13]	[8,14]	[9,16]	[11,18]	[13,21]	[16,24]	[17,26]	[18,27]	[17,24]	[14,21]	[11,17]	[8,14]
14	London	[2,6]	[2,7]	[3,10]	[5,13]	[8,17]	[11,20]	[13,22]	[13,21]	[11,19]	[8,14]	[5,10]	[3,7]
15	Madras	[20,30]	[20,31]	[22,33]	[26,35]	[28,39]	[27,38]	[26,36]	[26,35]	[25,34]	[24,32]	[22,30]	[21,29]
16	Madrid	[1,9]	[1,12]	[3,16]	[6,19]	[9,24]	[13,29]	[16,34]	[16,33]	[13,28]	[8,20]	[4,14]	[1,9]
17	Manila	[21,27]	[22,27]	[24,29]	[24,31]	[25,31]	[25,31]	[23,29]	[24,28]	[25,28]	[24,29]	[22,28]	[22,27]
18	Mauritius	[22,28]	[22,29]	[22,29]	[21,28]	[19,25]	[18,24]	[17,23]	[17,23]	[17,24]	[18,25]	[19,27]	[21,28]
19	Mexico City	[6,22]	[15,23]	[17,25]	[18,27]	[18,27]	[18,27]	[18,27]	[18,26]	[18,26]	[16,25]	[19,27]	[21,28]
20	Moscow	[−13,−6]	[−12,−5]	[−8,0]	[0,8]	[7,18]	[11,23]	[13,24]	[11,22]	[6,16]	[1,8]	[−5,0]	[−11,−5]
21	Munich	[−6,1]	[−5,3]	[−2,9]	[3,14]	[7,18]	[10,21]	[12,23]	[11,23]	[8,20]	[4,13]	[0,7]	[−4,3]
22	Nairobi	[12,25]	[13,26]	[14,25]	[14,24]	[13,22]	[12,21]	[11,21]	[11,21]	[11,24]	[13,24]	[13,23]	[13,23]
23	New Delhi	[−2,4]	[−3,4]	[1,9]	[6,15]	[12,22]	[17,27]	[21,29]	[20,28]	[16,24]	[11,19]	[5,12]	[−2,6]
24	New York	[−2,4]	[−3,4]	[1,9]	[6,15]	[12,22]	[17,27]	[21,29]	[20,28]	[16,24]	[11,19]	[5,12]	[−2,6]
25	Paris	[1,7]	[1,7]	[2,12]	[5,16]	[8,19]	[12,22]	[14,24]	[13,24]	[11,21]	[7,16]	[4,10]	[1,6]
26	Rome	[4,11]	[5,13]	[7,16]	[10,19]	[13,23]	[17,18]	[20,31]	[20,31]	[17,27]	[13,21]	[9,16]	[5,12]
27	San Francisco	[6,13]	[6,14]	[7,17]	[8,18]	[10,19]	[11,21]	[12,22]	[12,22]	[12,23]	[11,22]	[8,18]	[6,14]
28	Seoul	[0,7]	[1,6]	[1,8]	[6,16]	[12,22]	[16,25]	[18,31]	[16,30]	[9,28]	[3,24]	[7,19]	[1,8]
29	Singapore	[23,30]	[23,30]	[24,31]	[24,31]	[24,30]	[25,30]	[25,30]	[25,30]	[24,30]	[24,30]	[24,30]	[23,30]
30	Stockholm	[−9,−5]	[−9,−6]	[−4,2]	[1,8]	[6,15]	[11,19]	[14,22]	[13,20]	[9,15]	[5,9]	[1,4]	[−2,2]
31	Sydney	[20,30]	[20,30]	[18,26]	[16,23]	[12,20]	[5,17]	[8,16]	[9,17]	[11,20]	[13,22]	[16,26]	[20,30]
32	Tehran	[0,5]	[5,8]	[10,15]	[15,18]	[20,25]	[28,30]	[36,38]	[38,40]	[29,30]	[18,20]	[9,12]	[−5,0]
33	Tokyo	[0,9]	[0,10]	[3,13]	[9,18]	[14,23]	[18,25]	[22,29]	[13,31]	[20,27]	[13,21]	[8,16]	[2,21]
34	Toronto	[−8,−1]	[−8,−1]	[−4,4]	[−2,11]	[−8,18]	[13,24]	[16,27]	[16,26]	[12,22]	[6,14]	[−1,17]	[−5,1]
35	Vienna	[−2,1]	[−1,3]	[1,8]	[5,14]	[10,19]	[13,22]	[15,24]	[14,23]	[11,19]	[7,13]	[2,7]	[1,3]
36	Zurich	[−11,9]	[−8,15]	[−7,18]	[−1,21]	[2,27]	[6,30]	[10,31]	[8,25]	[5,23]	[3,22]	[0,19]	[−11,8]
37	Ojmjakon	[−51,−42]	[−48,−36]	[−41,−22]	[−25,−5]	[−6,7]	[2,18]	[5,21]	[1,18]	[−4,8]	[−21,−10]	[−41,−32]	[−49,−40]
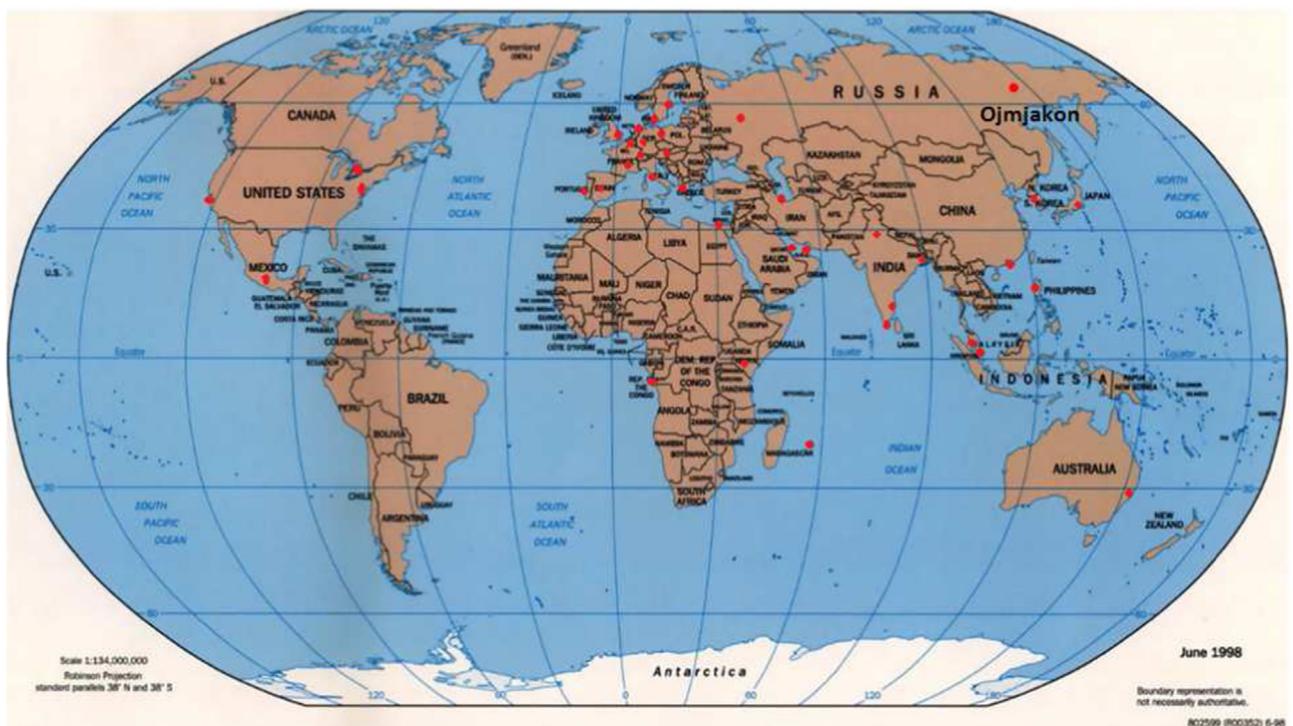
Fig. 12. Location map for 37 cities (Guru's dataset) and Ojmjakon.

Table 2
Latitude and longitude for 37 Cities (Guru's dataset) and Ojmjakon.

No.	City	Latitude	Longitude	No.	City	Latitude	Longitude
0	Amsterdam	52°30′N	4°52′E	19	Mexico City	19°24′N	99°13′W
1	Athens	37°59′N	23°44′E	20	Moscow	55°45′N	37°37′E
2	Bahrain	26°12′N	50°36′E	21	Munich	48°08′N	11°35′E
3	Bombay	18°56′N	72°51′E	22	Nairobi	1°17′S	36°49′E
4	Cairo	30°0′N	31°77′E	23	New Delhi	28°37′N	77°13′E
5	Calcutta	22°35′N	88°21′E	24	New York	40°44′N	73°55′W
6	Colombo	6°0′N	79°52′E	25	Paris	48°51′N	2°20′E
7	Copenhagen	55°42′N	12°34′E	26	Rome	41°52′N	12°37′E
8	Dubai	25°13′N	55°17′E	27	San Francisco	37°46′N	122°26′W
9	Frankfurt	52°21′N	14°33′E	28	Seoul	37°55′N	127°3′E
10	Geneva	46°14′N	6°4′E	29	Singapore	1°22′N	103°55′E
11	Hong Kong	22°17′N	114°8′E	30	Stockholm	59°23′N	18°00′E
12	Kuala Lumpur	3°8′N	101°42′E	31	Sydney	33°52′S	151°12′E
13	Lisbon	38°42′N	9°5′W	32	Tehran	35°40′N	51°26′E
14	London	51°30′N	0°9′W	33	Tokyo	35°41′N	139°44′E
15	Madras	13°5′N	80°18′E	34	Toronto	43°40′N	79°22′W
16	Madrid	40°26′N	3°42′W	35	Vienna	48°13′N	16°22′E
17	Manila	14°37′N	121°0′E	36	Zurich	47°22′N	8°32′E
18	Mauritius	20°37′S	57°42′E	37	Ojmjakon	63°27′N	142°46′E

Table 3
Fuzzy silhouette index for various number of clusters, the parameter of the SOWA function and type of loss function. $\lambda = 1.5$.

		Number of clusters				
Method	p_c	2	3	4	5	6
FcOMdC-ID, SQR	0.7	0.4129	0.3960	0.4182	0.3147	0.2874
	0.8	0.4129	0.3960	0.4182	0.3024	0.2451
	0.9	0.4129	0.3960	0.4182	0.3053	0.2874
	1.0	0.4129	0.3960	0.4182	0.3312	0.3034
FcOMdC-ID,LIN	0.7	0.3172	0.2904	0.2985	0.2661	0.2196
	0.8	0.3172	0.2927	0.2985	0.2412	0.2309
	0.9	0.3172	0.2904	0.2700	0.2585	0.2486
	1.0	0.3172	0.2927	0.2931	0.2535	0.2458
FcOMdC-ID,SIG	0.7	0.3233	0.3017	0.3041	0.2930	0.2702
	0.8	0.3233	0.3017	0.3087	0.2844	0.2764
	0.9	0.3233	0.3017	0.3041	0.2886	0.2721
	1.0	0.3233	0.3017	0.3036	0.2470	0.3009
FcOMdC-ID,HUB	0.7	0.3185	0.2951	0.3016	0.2475	0.2497
	0.8	0.3185	0.2951	0.2960	0.2923	0.2589
	0.9	0.3185	0.2935	0.3016	0.2480	0.2334
	1.0	0.3185	0.2951	0.3016	0.2700	0.2837
FcOMdC-ID,LOG	0.7	0.1948	0.2419	0.2323	0.2281	0.1975
	0.8	0.1948	0.2419	0.2422	0.2017	0.2052
	0.9	0.1948	0.2419	0.2509	0.2045	0.2126
	1.0	0.1948	0.2405	0.2504	0.2130	0.2180

Table 4
Fuzzy silhouette index for various number of clusters, the parameter of the SOWA function and type of loss function. $\lambda = 1.8$.

		Number of clusters				
Method	p_c	2	3	4	5	6
FcOMdC-ID, SQR	0.7	0.4506	0.4185	0.4432	0.4017	0.3358
	0.8	0.4506	0.4185	0.4123	0.3176	0.3650
	0.9	0.4506	0.4185	0.5025	0.2836	0.2729
	1.0	0.4506	0.4185	0.4471	0.4067	0.3602
FcOMdC-ID,LIN	0.7	0.3409	0.3085	0.3015	0.2424	0.2370
	0.8	0.3409	0.3085	0.3099	0.2498	0.2159
	0.9	0.3409	0.3086	0.3099	0.2557	0.2362
	1.0	0.3409	0.3086	0.2692	0.2892	0.2634
FcOMdC-ID,SIG	0.7	0.3493	0.3244	0.2954	0.2914	0.2804
	0.8	0.3493	0.3244	0.3177	0.3258	0.2136
	0.9	0.3493	0.3244	0.3228	0.2757	0.2978
	1.0	0.3493	0.3236	0.3362	0.2813	0.2322
FcOMdC-ID,HUB	0.7	0.3429	0.3306	0.3006	0.2843	0.2608
	0.8	0.3429	0.3281	0.2780	0.2438	0.2324
	0.9	0.3429	0.3280	0.2744	0.2492	0.2544
	1.0	0.3429	0.3281	0.3101	0.2965	0.2557
FcOMdC-ID,LOG	0.7	0.2079	0.2574	0.2458	0.1911	0.2271
	0.8	0.2079	0.2526	0.2564	0.2227	0.2545
	0.9	0.2079	0.2574	0.2492	0.2066	0.1929
	1.0	0.2079	0.2577	0.2576	0.1974	0.2223

compare the results obtained for different loss functions because they use different dissimilarity measure. Analysis of these tables allow drawing the following conclusions. First, most often we get the greatest values of the FS for two or four clusters. Second, impact of parameter p_c on the quality of obtained clusters is minimal. This effect is due to the elastic influence of the ordering function SOWA to reduce the outliers impact (by its soft down-weighting) on the result clusters.

From Tables 3–5, we deduce that an optimal partition for our FcOMdC-ID clustering method is obtained for SIG loss function, $\lambda = 2$ and $p_c = 0.7$. The membership degree matrix is shown in Table 6.

In order to compare the results obtained by our FcOMdC-ID method vs the results obtained by some methods proposed in the literature, we also show in Table 6 the membership degree

matrices obtained by the following methods: the partition obtained by a panel of human experts [50,49], the MSV (Mutual Similarity Value) agglomerative based clustering [50], the FcMd-ID (Fuzzy c-Medoids for Interval Data) method [33], the NcFcM-ID (Fuzzy c-Means for Interval Data) method D'Urso and Giordani [35], the TrFcMd-ID (Trimmed Fuzzy c-Medoids for Interval Data) method D'Urso et al. [33]. In addition, in order to compare the different partitions we report in Table 7 the Fuzzy Rand Index matrix. Notice that, the Fuzzy Rand Index [57] is a fuzzy extension of the original Rand index based on agreements and disagreements in two partitions: $\omega = (a+b+c+d)^{-1}(a+d)$ (a, d numbers of pairs of units belonging to the same/different clusters in the two partitions; b, c number of pairs of units belonging to the same clusters in one partition and to different clusters in the other).

Table 5
Fuzzy silhouette index for various number of clusters, the parameter of the SOWA function and type of loss function. $\lambda = 2.0$.

Method	p_c	Number of clusters				
		2	3	4	5	6
FcOMdC-ID, SQR	0.7	0.4690	0.4260	0.5115	0.2978	0.3321
	0.8	0.4690	0.4244	0.4463	0.3353	0.3282
	0.9	0.4690	0.4342	0.4309	0.3722	0.3361
	1.0	0.4690	0.4342	0.4340	0.3910	0.3615
FcOMdC-ID,LIN	0.7	0.3499	0.3331	0.3023	0.2587	0.2657
	0.8	0.3499	0.3210	0.3397	0.2956	0.2445
	0.9	0.3499	0.3329	0.2912	0.2639	0.2469
	1.0	0.3499	0.3358	0.3217	0.2607	0.2302
FcOMdC-ID,SIG	0.7	0.3602	0.3530	0.3232	0.3230	0.2619
	0.8	0.3602	0.3383	0.3284	0.3118	0.2571
	0.9	0.3602	0.3383	0.3438	0.2421	0.2482
	1.0	0.3602	0.3502	0.3381	0.2662	0.2747
FcOMdC-ID,HUB	0.7	0.3522	0.3413	0.2894	0.2339	0.2079
	0.8	0.3522	0.3413	0.3322	0.2897	0.2565
	0.9	0.3522	0.3411	0.3199	0.2368	0.2354
	1.0	0.3522	0.3384	0.3100	0.2855	0.2343
FcOMdC-ID,LOG	0.7	0.2193	0.2588	0.2382	0.2107	0.2134
	0.8	0.2129	0.2360	0.2678	0.1798	0.1863
	0.9	0.2129	0.2648	0.2437	0.2052	0.2089
	1.0	0.2129	0.2392	0.2352	0.2114	0.1600

In the partition obtained by a panel of human experts [50,49] (see Table 6) the cities belonging to cluster 1 are located in between 0° and 40° latitudes and the cities which are classified under cluster 2 are located in between 40° and 60° latitudes. However, the observers classified some cities (Athens, Lisbon, San Francisco, Seoul and Tokyo) which are closer to the sea coast and are located in between the latitudes 0° and 40° as members of cluster 2, because such cities bear low temperature which is similar to that of the cities which are located in between 40° and 60° latitudes. The cities nearer to sea coast bear relatively low temperature because of the cool breeze from the sea coast and also due to high humidity present in the atmosphere. Mauritius being the only island (in the considered data set) that is classified as a member of a singleton cluster and Tehran as a member of the other singleton because It is characterized by a singular climate behavior (small fluctuation in temperature; see Fig. 11). The partition obtained through the MSV (Mutual Similarity Value) agglomerative based clustering [50] is in accordance with the realistic clusters provided by the panel of human observers. By applying the FcOMdC-ID method (with the SIG loss function, $\lambda = 2$ and $p_c = 0.7$) and the other fuzzy clustering methods, i.e. the timidly robust FcMd-ID method [32] (with $\lambda = 2$) and the robust methods NcFcM-ID method [35] (with $\lambda = 2$ and noise distance δ equal to $\sum_{i=1}^{c} \sum_{k=1}^{N} D_{ik}^2 / cN$, where D_{ik}^2 indicates the squared Euclidean distance between the kth observations and the ith centroid) and TrFcMd-ID method [33] (with $\lambda = 2$ and trimming parameter $\alpha = 0$), we obtain results consistent with the two non-singleton clusters indicated by the panel of experts. These evidences are corroborated by the Fuzzy Rand Index matrix shown in Table 7. Furthermore, notice that in the Guru's dataset there are no outliers, i.e. cities with minimum and maximum temperature that lie an abnormal distance from other minimum and maximum temperature values. The results obtained by all robust clustering methods confirm this evidence. In fact, as expected, the obtained results are consistent with the empirical evidences (data without

outliers) and then show a stable behavior of our robust model and of the other robust ones in the examined real context.

4.2. Temperature data with outlier (modified Guru's data)

In order to show the performance, effectiveness and robustness of our clustering method, we contaminate the Guru's dataset with an outlier, i.e. we consider the monthly minimum and maximum temperature of a very cold village: Ojmjakon. Ojmjakon is a rural locality in Ojmjakonsky District of the Sakha Republic, Russia, located along the Indigirka River, 30 km northwest of Tomtor on the Kolyma Highway. Ojmjakon is one of the coldest permanently inhabited locales on the planet. With an extreme subarctic climate, Ojmjakon is known as one of the candidates for the Northern Pole of Cold. The ground there is permanently frozen (continuous permafrost). Although winters in Ojmjakon are long and excessively cold, summers are mild, sometimes with hot, and very hot, days. The climate is quite dry, but as average monthly temperatures are below freezing for seven months of the year, substantial evaporation occurs only in summer months. Summers are much wetter than winters.

The considered monthly minimum and maximum temperatures, the location and the latitude and longitude of Ojmjakon are shown together to the other cities in Tables 1 and 2 and Fig. 12.

To show the different monthly climatic features of the cities during the year, in Fig. 13a we represent the violin plots in the different months for the entire interval dataset for the midpoints of the interval data.

The violin plot synergistically combines the box plot and the density trace (or smoothed histogram) into a single display that reveals structure found within the data [55]. In particular, it includes a box plot with two slight modifications. First, a circle replaces the median line which facilitates quick comparisons when viewing multiple groups. Second, outside points which are traditionally classified as mild and severe outliers are not identified by individual symbols. The density trace supplements traditional summary statistics by graphically showing the distributional characteristics of batches of data [55]. The density trace is plotted symmetrically to the left and the right of the (vertical) box plot. There is no difference in these density traces other than the direction in which they extend. Adding two density traces gives a symmetric plot which makes it easier to see the magnitude of the density. This hybrid of the density trace and the box plot allows quick and insightful comparison of several distributions [55]. With the addition of the density trace to the box plot, violin plots provide a better indication of the shape of the distribution. This includes showing the existence of clusters in data. The density trace highlights the peaks, valleys, and bumps in the distribution [55]. About that, in the application, as we can see below, we select $\lambda = 2$ by means of the fuzzy silhouette index (see Fig. 14). This result is corroborated by the bimodality of the density traces (in particular in the summer months) of midpoint temperatures shown in the violin plots in Fig. 13a.

In addition, from the violin plots shown in Fig. 13a we can see the seasonal pattern of the temperatures, the changes in temperature during the months and clearly identify the presence of an outlier.

We apply the modified Guru's dataset (Table 1) to our FcOMdC-ID method (with SIG loss function, $\lambda = 2$ and $p_c = 0.7$) and other robust fuzzy clustering methods, i.e. the timidly robust FcMd-ID method [33] (with $\lambda = 2$) and the robust methods NcFcM-ID method [35] (with $\lambda = 2$ and noise distance δ equal to $\sum_{i=1}^{c} \sum_{k=1}^{N} D_{ik}^2 / cN$, where D_{ik}^2 indicates the squared Euclidean distance between the kth observations and the ith centroid) and

Table 6
Membership degrees (Guru's dataset).

No.	City	FcOMdC-ID			MSV, panel of human experts				FcMd-ID		NcFcM-ID			TrFcMd-ID	
		Cluster 1	Cluster 2	(β_k)	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Noise cluster	Cluster 1	Cluster 2
0	Amsterdam	0.0114	0.9886	0.9999	0	1	0	0	0.0187	0.9813	0.0222	0.8915	0.0863	0.0187	0.9813
1	Athens	0.3094	0.6906	0.9809	0	1	0	0	0.2915	0.7085	0.2635	0.4547	0.2817	0.2915	0.7085
2	Bahrain	0.8974	0.1026	0.9988	1	0	0	0	0.8841	0.1159	0.7688	0.0611	0.1701	0.8841	0.1159
3	Bombay	1.0000	0.0000	1.0000	1	0	0	0	1.0000	0.0000	0.9235	0.0156	0.0609	1.0000	0.0000
4	Cairo	0.7333	0.2667	0.9913	1	0	0	0	0.7265	0.2735	0.6399	0.1295	0.2306	0.7265	0.2735
5	Calcutta	0.9726	0.0274	0.9996	1	0	0	0	0.9677	0.0323	0.8912	0.0235	0.0853	0.9677	0.0323
6	Colombo	0.9944	0.0056	0.9999	1	0	0	0	0.9902	0.0098	0.8842	0.0230	0.0928	0.9902	0.0098
7	Copenaghen	0.0281	0.9719	0.9995	0	1	0	0	0.0265	0.9735	0.0277	0.8566	0.1157	0.0265	0.9735
8	Dubai	0.8909	0.1091	0.9982	1	0	0	0	0.9117	0.0883	0.8072	0.0464	0.1464	0.9117	0.0883
9	Frankfurt	0.0970	0.9030	0.9396	0	1	0	0	0.1083	0.8917	0.0805	0.5996	0.3200	0.1083	0.8917
10	Geneva	0.0237	0.9763	0.9998	0	1	0	0	0.0224	0.9776	0.0203	0.9069	0.0728	0.0224	0.9776
11	Hong Kong	0.8034	0.1966	0.9973	1	0	0	0	0.7537	0.2463	0.6667	0.1210	0.2123	0.7537	0.2463
12	Kuala Lampur	0.9854	0.0146	0.9998	1	0	0	0	0.9837	0.0163	0.8518	0.0281	0.1202	0.9837	0.0163
13	Lisbon	0.1905	0.8095	0.9871	0	1	0	0	0.1867	0.8133	0.1989	0.5518	0.2492	0.1867	0.8133
14	London	0.0053	0.9947	0.9999	0	1	0	0	0.0079	0.9921	0.0216	0.9097	0.0686	0.0079	0.9921
15	Madras	0.9709	0.0291	0.9995	1	0	0	0	0.9755	0.0245	0.8052	0.0347	0.1601	0.9755	0.0245
16	Madrid	0.0958	0.9042	0.9973	0	1	0	0	0.0763	0.9237	0.0717	0.7718	0.1565	0.0763	0.9237
17	Manila	0.9886	0.0114	0.9998	1	0	0	0	0.9864	0.0136	0.9116	0.0189	0.0695	0.9864	0.0136
18	Mauritius	0.8641	0.1359	0.9992	0	0	1	0	0.8597	0.1403	0.6741	0.0831	0.2428	0.8597	0.1403
19	Mexico City	0.6593	0.3407	0.9941	1	0	0	0	0.6859	0.3141	0.6132	0.1519	0.2349	0.6859	0.3141
20	Moscow	0.0771	0.9229	0.9718	0	1	0	0	0.0995	0.9005	0.0612	0.5303	0.4085	0.0995	0.9005
21	Munich	0.0208	0.9792	0.9996	0	1	0	0	0.0289	0.9711	0.0277	0.8489	0.1235	0.0289	0.9711
22	Nairobi	0.4228	0.5772	0.9718	1	0	0	0	0.5065	0.4935	0.3898	0.2463	0.3639	0.5065	0.4935
23	New Deli	0.8213	0.1787	0.9960	1	0	0	0	0.8340	0.1660	0.6561	0.0900	0.2539	0.8340	0.1660
24	New York	0.0546	0.9454	0.9992	0	1	0	0	0.0516	0.9484	0.0332	0.8728	0.0940	0.0516	0.9484
25	Paris	0.0000	1.0000	1.0000	0	1	0	0	0.0000	1.0000	0.0075	0.9701	0.0225	0.0000	1.0000
26	Rome	0.1667	0.8333	0.9941	0	1	0	0	0.1646	0.8354	0.1623	0.6116	0.2261	0.1646	0.8354
27	San Francisco	0.0714	0.9286	0.9988	0	1	0	0	0.0947	0.9053	0.1194	0.6613	0.2193	0.0947	0.9053
28	Seoul	0.0586	0.9414	0.9982	0	1	0	0	0.0662	0.9338	0.0535	0.8092	0.1373	0.0662	0.9338
29	Singapore	0.9926	0.0074	0.9999	1	0	0	0	0.9877	0.0123	0.8621	0.0267	0.1113	0.9877	0.0123
30	Stockholm	0.0732	0.9268	0.9809	0	1	0	0	0.0737	0.9263	0.0540	0.6344	0.3116	0.0737	0.9263
31	Sydney	0.6137	0.3863	0.9871	1	0	0	0	0.5982	0.4018	0.3461	0.1801	0.4738	0.5982	0.4018
32	Tehran	0.3919	0.6081	0.8251	0	0	0	1	0.3658	0.6342	0.2095	0.3152	0.4753	0.3658	0.6342
33	Tokyo	0.1634	0.8366	0.9960	0	1	0	0	0.1301	0.8699	0.1109	0.6883	0.2008	0.1301	0.8699
34	Toronto	0.0641	0.9359	0.9913	0	1	0	0	0.0778	0.9222	0.0547	0.6678	0.2774	0.0778	0.9222
35	Vienna	0.0123	0.9877	0.9998	0	1	0	0	0.0159	0.9841	0.0165	0.9224	0.0611	0.0159	0.9841
36	Zurich	0.1008	0.8992	0.9127	0	1	0	0	0.1134	0.8866	0.0827	0.5836	0.3337	0.1134	0.8866

Table 7
Fuzzy Rand Index matrix (Guru's dataset).

	FcOMdC-ID	MSV	FcMd-ID	NcFcM-ID	TrFcMd-ID
FcOMdC-ID	1.0000	0.7830	0.9796	0.9079	0.9796
MSV	0.7830	1.0000	0.7860	0.7302	0.7860
FcMd-ID	0.9796	0.7860	1.0000	0.9129	1.0000
NcFcM-ID	0.9079	0.7302	0.9129	1.0000	0.9129
TrFcMd-ID	0.9796	0.7860	1.0000	0.9129	1.0000
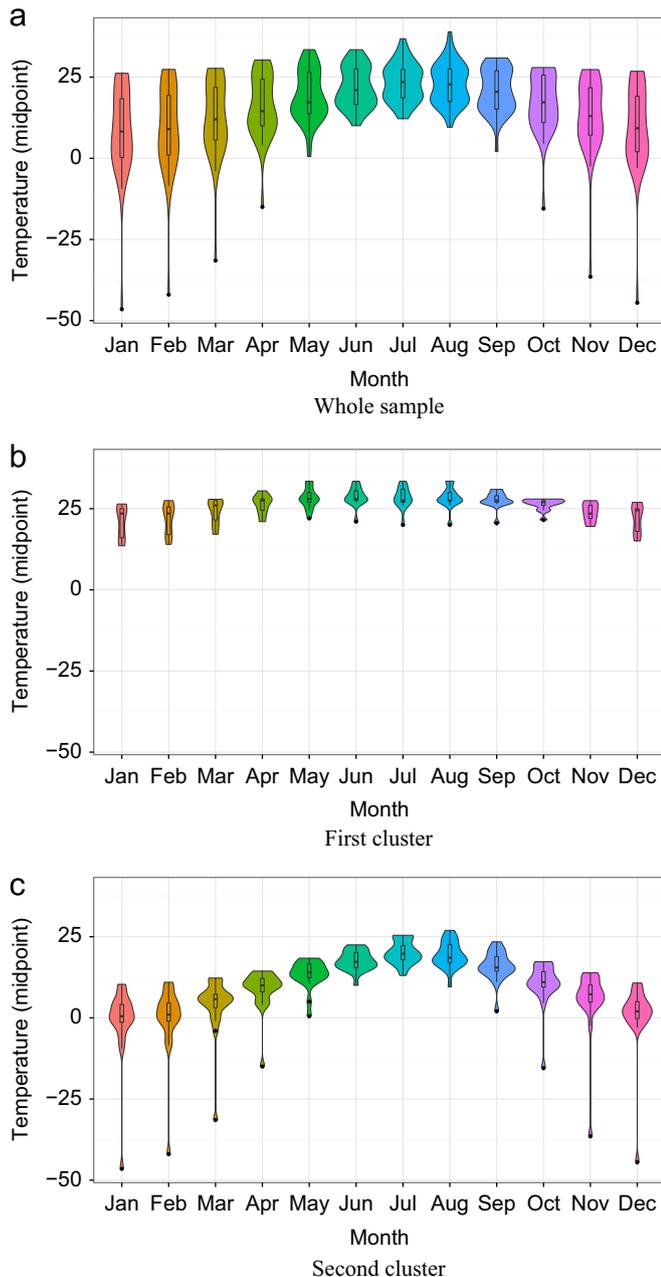
Fig. 13. Violin plots—midpoints. (a) Whole sample. (b) First cluster. (c) Second cluster.

the TrFcMd-ID method [33] (with $\lambda = 2$ and trimming parameter $\alpha = 0.025$). All results are shown in Table 8.

As we can see, the fuzzy partitions obtained by means of the different robust and timidly robust methods are similar, as can also be seen from the Fuzzy Rand Index matrix (Table 9). In fact, all methods identify and neutralize in a different manner but

effectively the outlier. In particular, the timid robust FcMd-ID method neutralizes the negative effects in the clustering process of the outlier because the structure of the cluster is not altered but it does not identify clearly the outlier. However, as emphasized by Garcíia-Escudero et al. [43], the clustering method based on medoids resists the presence of one outlier in a remote position, but it breaks down when we increase the number of outliers to three. The TrFcMd-ID method and the NcFcM-ID method identify the outlier and neutralize its negative effect, respectively, removing it from the analyzed dataset and putting the outlier in the noise cluster. Our method is more informative than the other ones because by means of the typicality parameter, which represents an important means for alleviating the undesirable effects of outliers, we have a measure of the intensity level of the outlier. In particular, since Ojmjakon presents a very low level of typicality, i.e. $\beta = 0.0119$, it is represented in the set of the cities as a village with very atypical levels of temperatures. Then, our method not only is able to neutralize and tune suitably the influence of the anomalous interval data in the clustering process, but it also provides a suitable measure of the typicality/atypicality level of the outlier interval data.

With reference to the results obtained by means of our method, in Fig. 15 we show the radar plots (zoom stars in the interval data analysis) for the 2 medoid cities, for the fuzzy cities and for the outlier city.

As we can see, the prototype (medoid) of the cluster 1, Colombo, represents the cities (belonging to cluster 1) characterized by warm temperatures in all the months of the year, while the medoid of the cluster 2, Paris, represents a cluster with cities characterized by temperate climate. The "fuzzy" cities Athens and Mexico City are warm in summer and cold in the other months; vice versa for the "fuzzy" cities Nairobi and Sydney that are located in the opposite hemisphere. The fuzziness of Tehran is connected to its climatic conditions: it is very cold in winter and very hot in summer with small month temperature changes. As we can see, the outlier city, Ojmjakon, is characterized by very cold temperatures. These empirical findings are corroborated by the results shown in Figs. 13b and c, in which we represent the violin plots for the midpoint temperatures obtained by considering the cities with membership degrees to cluster 1 and cluster 2 greater than 0.60 (cut-off=0.60).

5. Conclusions

The real interval-valued data can contain outliers. Therefore the clustering methods need to be robust. This paper combines the fuzzy c-medoids clustering with the robust ordered statistics using Huber's M-estimator. The developed FcOMdC-ID clustering method is based on various dissimilarity measures (as squared, linear, Huber, sigmoidal and logarithmic) and an ordering of models residuals. The method is introduced as the problem of a constrained minimization of the criterion function. The necessary conditions for obtaining local minimum of the criterion function with respect to the elements of the partition matrix are shown. The existing fuzzy c-medoids clustering method can be obtained as special cases of the method developed. The study of the FcOMdC-ID with the traditional fuzzy c-medoids as the reference methods is included. These numerical examples show the usefulness of the method proposed when applied to clustering data with different types of outliers.

Furthermore, in order to show the effectiveness of our method in an empirical context, the FcOMdC-ID method is applied and compared with other robust fuzzy clustering methods.
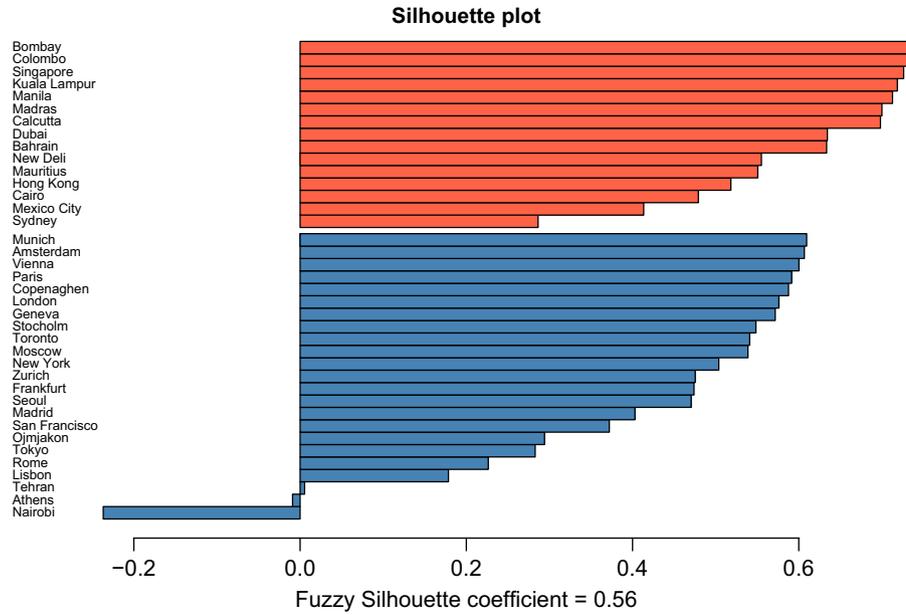
Fig. 14. Silhouette plot.

Table 8
Membership degrees (modified Guru's dataset).

No.	City	FcOMdC-ID			FcMd-ID		NcFcM-ID			TrFcMd-ID	
		Cluster 1	Cluster 2	(β_k)	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Noise cluster	Cluster 1	Cluster 2
0	Amsterdam	0.0954	0.9046	0.9998	0.0198	0.9802	0.0234	0.9166	0.0599	0.0187	0.9813
1	Athens	0.4025	0.5975	0.9763	0.3061	0.6939	0.3021	0.4906	0.2072	0.2915	0.7085
2	Bahrain	0.7380	0.2620	0.9989	0.8729	0.1271	0.8177	0.0631	0.1192	0.8841	0.1159
3	Bombay	0.9302	0.0698	1.0000	0.9874	0.0126	0.9313	0.0188	0.0499	1.0000	0.0000
4	Cairo	0.6023	0.3977	0.9925	0.7124	0.2876	0.7079	0.1324	0.1597	0.7265	0.2735
5	Calcutta	0.7979	0.2021	0.9995	0.9457	0.0543	0.9106	0.0258	0.0636	0.9677	0.0323
6	Colombo	1.0000	0.0000	0.9999	0.9933	0.0067	0.9004	0.0266	0.0730	0.9902	0.0098
7	Copenaghen	0.1471	0.8529	0.9997	0.0283	0.9717	0.0298	0.8884	0.0818	0.0265	0.9735
8	Dubai	0.7045	0.2955	0.9993	0.8816	0.1184	0.8483	0.0483	0.1033	0.9117	0.0883
9	Frankfurt	0.2423	0.7577	0.9763	0.1099	0.8901	0.0907	0.6722	0.2371	0.1083	0.8917
10	Geneva	0.1351	0.8649	0.9998	0.0238	0.9762	0.0217	0.9272	0.0511	0.0224	0.9776
11	Hong Kong	0.6911	0.3089	0.9976	0.7767	0.2233	0.7318	0.1223	0.1460	0.7537	0.2463
12	Kuala Lampur	0.9232	0.0768	0.9998	0.9798	0.0202	0.8743	0.0322	0.0935	0.9837	0.0163
13	Lisbon	0.3299	0.6701	0.9925	0.2050	0.7950	0.2274	0.5894	0.1832	0.1867	0.8133
14	London	0.0685	0.9315	0.9999	0.0086	0.9914	0.0240	0.9261	0.0499	0.0079	0.9921
15	Madras	0.8329	0.1671	0.9997	0.9563	0.0437	0.8354	0.0398	0.1248	0.9755	0.0245
16	Madrid	0.2399	0.7601	0.9984	0.0799	0.9201	0.0778	0.8118	0.1104	0.0763	0.9237
17	Manila	0.9407	0.0593	0.9999	1.0000	0.0000	0.9245	0.0216	0.0539	0.9864	0.0136
18	Mauritius	0.7504	0.2496	0.9984	0.9046	0.0954	0.7364	0.0882	0.1754	0.8597	0.1403
19	Mexico City	0.5615	0.4385	0.9949	0.7097	0.2903	0.6829	0.1545	0.1626	0.6859	0.3141
20	Moscow	0.2242	0.7758	0.9280	0.1032	0.8968	0.0715	0.6131	0.3154	0.0995	0.9005
21	Munich	0.1264	0.8736	0.9995	0.0305	0.9695	0.0291	0.8855	0.0854	0.0289	0.9711
22	Nairobi	0.4468	0.5532	0.9889	0.5422	0.4578	0.4544	0.2718	0.2738	0.5065	0.4935
23	New Deli	0.6480	0.3520	0.9965	0.7984	0.2016	0.7206	0.0960	0.1834	0.8340	0.1660
24	New York	0.1924	0.8076	0.9993	0.0539	0.9461	0.0350	0.9002	0.0648	0.0516	0.9484
25	Paris	0.0000	1.0000	1.0000	0.0000	1.0000	0.0085	0.9749	0.0166	0.0000	1.0000
26	Rome	0.3077	0.6923	0.9949	0.1748	0.8252	0.1833	0.6521	0.1646	0.1646	0.8354
27	San Francisco	0.2122	0.7878	0.9976	0.1031	0.8969	0.1352	0.7038	0.1610	0.0947	0.9053
28	Seoul	0.1972	0.8028	0.9989	0.0685	0.9315	0.0576	0.8459	0.0965	0.0662	0.9338
29	Singapore	0.9935	0.0065	0.9998	0.9909	0.0091	0.8825	0.0307	0.0869	0.9877	0.0123
30	Stockholm	0.2217	0.7783	0.9838	0.0774	0.9226	0.0611	0.7061	0.2329	0.0737	0.9263
31	Sydney	0.5582	0.4418	0.9838	0.6383	0.3617	0.4151	0.2096	0.3753	0.5982	0.4018
32	Tehran	0.4530	0.5470	0.8559	0.3721	0.6279	0.2526	0.3700	0.3774	0.3658	0.6342
33	Tokyo	0.2982	0.7018	0.9965	0.1352	0.8648	0.1234	0.7315	0.1451	0.1301	0.8699
34	Toronto	0.2059	0.7941	0.9889	0.0808	0.9192	0.0606	0.7367	0.2027	0.0778	0.9222
35	Vienna	0.1011	0.8989	0.9999	0.0171	0.9829	0.0176	0.9396	0.0428	0.0159	0.9841
36	Zurich	0.2462	0.7538	0.9655	0.1148	0.8852	0.0937	0.6575	0.2488	0.1134	0.8866
37	Ojmjakon	0.3874	0.6126	0.0119	0.3118	0.6882	0.0436	0.0918	0.8646	trimmed	

Table 9
Fuzzy Rand Index matrix (modified Guru's dataset).

	FcOMdC-ID	FcMd-ID	NcFcM-ID	TrFcMd-ID
FcOMdC-ID	1.0000	0.8722	0.8463	0.8672
FcMd-ID	0.8722	1.0000	0.9127	0.9825
NcFcM-ID	0.8463	0.9127	1.0000	0.9257
TrFcMd-ID	0.8672	0.9825	0.9257	1.0000
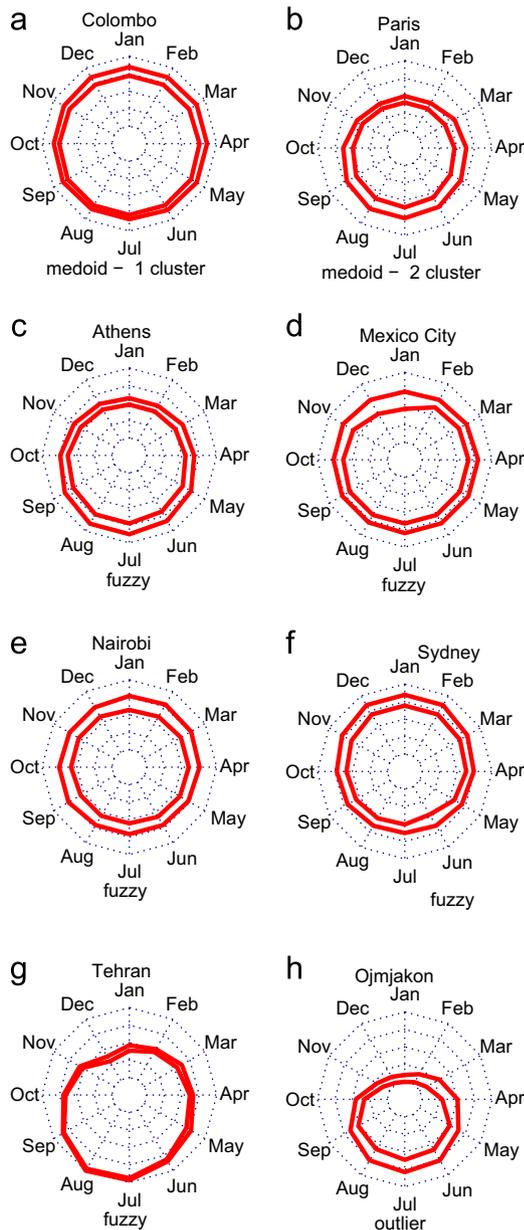
Fig. 15. Radar plots (zoom stars) for medoid, fuzzy and outlier cities.

Conflict of Interest

None declared.

Acknowledgments

References

[1] P. Arabie, J.D. Carroll, W. DeSarbo, J. Wind, Overlapping clustering: a new method for product positioning, J. Market. Res. (1981) 310–317.

[2] J. Arroyo, R. Espínola, C. Maté, Different approaches to forecast interval time series: a comparison in finance, Comput. Econ. 37 (2) (2011) 169–191.

[3] L. Billard, E. Diday, Symbolic Data Analysis: Conceptual Statistics and Data Mining, Wiley, England, 2006.

[4] B. Butkiewicz, Robust fuzzy clustering with fuzzy data, in: Proceedings of Advances in Web Intelligence, Third International Atlantic Web Intelligence Conference, AWIC 2005, Lecture Notes in Computer Science, vol. 352, Springer, Berlin, Heidelberg, 2005, pp. 76–82.

[5] G. Cabanes, Y. Bennani, R. Destenay, A. Hardy, A new topological clustering algorithm for interval data, Pattern Recognit. 46 (11) (2013) 3030–3039.

[6] R. Campello, E. Hruschka, A fuzzy extension of the silhouette width criterion for cluster analysis, Fuzzy Sets Syst. 157 (21) (2006) 2858–2875.

[7] C. Cappelli, P. D'Urso, F. Di Iorio, Regime change analysis of interval-valued time series with an application to pm10, Chemom. Intell. Lab. Systems 146 (2015) 337–346.

[8] P. Cazes, A. Chouakria, E. Diday, Y. Schektrman, Extension de l'analyse en composantes principales à des données de type intervalle, Rev. Stat. Appl. 45 (3) (1997) 5–24.

[9] R. Coppi, P. D'Urso, Fuzzy k-means clustering models for triangular fuzzy time trajectories, Stat. Methods Appl. 11 (1) (2002) 21–40.

[10] R. Coppi, P. D'Urso, P. Giordani, Fuzzy and possibilistic clustering for fuzzy data, Comput. Stat. Data Anal. 56 (4) (2012) 915–927.

[11] R.N. Davé, Characterization and detection of noise in clustering, Pattern Recognit. Lett. 12 (11) (1991) 657–664.

[12] R.N. Davé, R. Krishnapuram, Robust clustering methods: a unified view, IEEE Trans. Fuzzy Syst. 5 (2) (1997) 270–293.

[13] R.N. Davé, S. Sen, Robust fuzzy clustering of relational data, IEEE Trans. Fuzzy Syst. 10 (6) (2002) 713–727.

[14] F.d.A. de Carvalho, Fuzzy c-means clustering methods for symbolic interval data, Pattern Recognit. Lett. 28 (2007) 423–437.

[15] F.d.A. de Carvalho, P. Brito, H.-H. Bock, Dynamic clustering for interval data based on l 2 distance, Comput. Stat. 21 (2) (2006) 231–250.

[16] F.d.A. de Carvalho, R.M. de Souza, New metrics for constrained boolean symbolic objects, in: Proceedings of the Conference on Knowledge Extraction and Symbolic Data Analysis (KESDA '98), Luxembourg, Office for Official Publications of the European Communities, 1998, pp. 175–187.

[17] F.d.A. de Carvalho, R.M. de Souza, Unsupervised pattern recognition models for mixed feature-type symbolic data, Pattern Recognit. Lett. 31 (5) (2010) 430–443.

[18] F.d.A. de Carvalho, R.M. de Souza, M. Chavent, Y. Lechevallier, Adaptive Hausdorff distances and dynamic clustering of symbolic interval data, Pattern Recognit. Lett. 27 (3) (2006) 167–179.

[19] F.d.A. de Carvalho, Y. Lechevallier, Dynamic clustering of interval-valued data based on adaptive quadratic distances, IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum. 39 (6) (2009) 1295–1306.

[20] F.d.A. de Carvalho, Y. Lechevallier, Partitional clustering algorithms for symbolic interval data based on single adaptive distances, Pattern Recognit. 42 (7) (2009) 1223–1236.

[21] F.d.A. de Carvalho, C. Tenorio, Fuzzy k-means clustering algorithms for inter-valalued data based on adaptive quadratic distances, Fuzzy Sets Syst. 161 (2) (2010) 2978–2999.

[22] J.V. de Oliveira, W. Pedrycz, Advances in Fuzzy Clustering and its Applications, Wiley, Chichester, 2007.

[23] R.M.C.R. de Souza, F.d.A. de Carvalho, Clustering of interval data based on city–block distances, Pattern Recognit. Lett. 25 (3) (2004) 353–365.

[24] T. Denoeux, M. Masson, Multidimensional scaling of interval-valued dissimilarity data, Pattern Recognit. Lett. 21 (1) (2000) 83–92.

[25] A. Douzal-Chouakria, L. Billard, E. Diday, Principal component analysis for interval-valued observations, Stat. Anal. Data Mining: ASA Data Sci. J. 4 (2) (2011) 229–246.

[26] Silva A. Duarte, P. Brito, Linear discriminant analysis for interval data, Comput. Stat. 21 (2) (2006) 289–308.

[27] Silva A. Duarte, P. Brito, Discriminant analysis of interval data: an assessment of parametric and distance-based approaches, J. Classif. available online. http://dx.doi.org/10.1007/s00357-015-9189-8, 2015.

[28] P. D'Urso, Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, Comput. Stat. Data Anal. 42 (1) (2003) 47–72.

[29] P. D'Urso, Fuzzy clustering for data time arrays with inlier and outlier time trajectories, IEEE Trans. Fuzzy Syst. 13 (5) (2005) 583–604.

[30] P. D'Urso, Fuzzy clustering, in: C. Hennig, M. Meila, F. Murtagh, R. Rocci (Eds.), Handbook of Cluster Analysis, 2015, pp. 545–574.

[31] P. D'Urso, L. De Giovanni, Midpoint radius self-organizing maps for interval-valued data with telecommunications application, Appl. Soft Comput. 11 (5) (2011) 3877–3886.

[32] P. D'Urso, L. De Giovanni, Robust clustering of imprecise data, Chemom. Intell. Lab. Syst. 136 (2014) 58–80.

[33] P. D'Urso, L. De Giovanni, R. Massari, Trimmed fuzzy clustering for interval-valued data, Adv. Data Anal. Classif. 9 (1) (2015) 21–40.

[34] P. D'Urso, P. Giordani, A least squares approach to principal component analysis for interval valued data, Chemom. Intell. Lab. Syst. 70 (2) (2004) 179–192.

[35] P. D'Urso, P. Giordani, A robust fuzzy k-means clustering model for interval valued data, Comput. Stat. 21 (2) (2006) 251–269.

[36] P. D'Urso, P. Giordani, A weighted fuzzy c-means clustering model for fuzzy data, Comput. Stat. Data Anal. 50 (6) (2006) 1496–1523.

[37] B.S. Everitt, S. Landau, M. Leese, Cluster Analysis, 4th ed., Arnold Press, London, 2001.

[38] H. Frigui, R. Krishnapuram, A robust algorithm for automatic extraction of an unknown number of clusters from noisy data, Pattern Recognit. Lett. 17 (12) (1996) 1223–1232.

[39] L.A. García-Escudero, A. Gordaliza, Robustness properties of k-means and trimmed k-means, J. Am. Stat. Assoc. 94 (447) (1999) 956–969.

[40] L.A. García-Escudero, A. Gordaliza, A proposal for robust curve clustering, J. Classif. 22 (2) (2005) 185–201.

[41] L.A. García-Escudero, A. Gordaliza, C. Matrán, A central limit theorem for multivariate generalized trimmed k-means, Ann. Stat. 27 (1999) 1061–1079.

[42] L.A. García-Escudero, A. Gordaliza, C. Matrán, Trimming tools in exploratory data analysis, J. Comput. Graph. Stat. 12 (2) (2003).

[43] L.A. García-Escudero, A. Gordaliza, C. Matrán, A. Mayo-Iscar, A review of robust clustering methods, Adv. Data Anal. Classif. 4 (2) (2010) 89–109.

[44] P. Giordani, H.A. Kiers, Three-way component analysis of interval-valued data, J. Chemom. 18 (5) (2004) 253–264.

[45] G. González-Rivera, W. Lin, Constrained regression for interval-valued data, J. Bus. Econ. Stat. 31 (4) (2013) 473–490.

[46] K.C. Gowda, T. Ravi, Agglomerative clustering of symbolic objects using the concepts of both similarity and dissimilarity, Pattern Recognit. Lett. 16 (6) (1995) 647–652.

[47] P.J. Groenen, S. Winsberg, O. Rodriguez, E. Diday, I-Scal: multidimensional scaling of interval dissimilarities, Comput. Stat. Data Anal. 51 (1) (2006) 360–378.

[48] J. Guo, W. Li, C. Li, S. Gao, Standardization of interval symbolic data based on the empirical descriptive statistics, Comput. Stat. Data Anal. 56 (3) (2012) 602–610.

[49] D.S. Guru, B.B. Kiranagi, Multivalued type dissimilarity measure and concept of mutual dissimilarity value for clustering symbolic patterns, Pattern Recognit. 38 (1) (2005) 151–156.

[50] D.S. Guru, B.B. Kiranagi, P. Nagabhushan, Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns, Pattern Recognit. Lett. 25 (10) (2004) 1203–1213.

[51] C. Hajjar, H. Hamdan, Interval data clustering using self-organizing maps based on adaptive Mahalanobis distances, Neural Netw. 46 (2013) 124–132.

[52] A. Hardy, N. Kasaro, A new clustering method for interval data, Math. Sci. Hum. 187 (2009) 79–91.

[53] R.J. Hathaway, J.C. Bezdek, Y. Hu, Generalized fuzzy c-means clustering strategies using L_p norm distances, IEEE Trans. Fuzzy Syst. 8 (5) (2000) 576–582.

[54] W.J. Heiser, P.J. Groenen, Cluster differences scaling with a within-clusters loss component and a fuzzy successive approximation strategy to avoid local minima, Psychometrika 62 (1) (1997) 63–83.

[55] J.L. Hintze, R.D. Nelson, Violin plots: a box plot-density trace synergism, Am. Stat. 52 (2) (1998) 181–184.

[56] P.J. Huber, Robust Statistics, Wiley, New York, 1981.

[57] E. Hüllermeier, M. Rifqi, S. Henzgen, R. Senge, Comparing fuzzy partitions: a generalization of the rand index and related measures, IEEE Trans. Fuzzy Syst. 20 (3) (2012) 546–556.

[58] W.-L. Hung, M.-S. Yang, Fuzzy clustering on lr-type fuzzy numbers with an application in Taiwanese tea evaluation, Fuzzy Sets Syst. 150 (3) (2005) 561–577.

[59] M. Ichino, H. Yaguchi, Generalized Minkowski metrics for mixed feature-type data analysis, IEEE Trans. Syst. Man Cybern. 24 (4) (1994) 698–708.

[60] K. Jajuga, L_1-norm based fuzzy clustering, Fuzzy Sets Syst. 39 (1) (1991) 43–50.

[61] F.T. Juster, R. Suzman, An overview of the health and retirement study, J. Hum. Resour. (1995) S7–S56.

[62] T. Kamdar, A. Joshi, On creating adaptive Web servers using Weblog Mining, Technical Report TR-CS-00-05, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, 2000.

[63] C.-H. Kao, J. Nakano, S.-H. Shieh, Y.-J. Tien, H.-M. Wu, C.-K. Yang, C.-H. Chen, Exploratory data analysis of interval-valued symbolic data with matrix visualization, Comput. Stat. Data Anal. 79 (2014) 14–29.

[64] A. Keller, Fuzzy clustering with outliers, in: Fuzzy Information Processing Society, 2000. NAFIPS, 19th International Conference of the North American, IEEE, Piscataway, 2000, pp. 143–147.

[65] P.R. Kersten, Fuzzy order statistics and their application to fuzzy clustering, IEEE Trans. Fuzzy Syst. 7 (6) (1999) 708–712.

[66] J. Kim, L. Billard, Dissimilarity measures for histogram-valued observations, Commun. Stat. Theory Methods 42 (2) (2013) 283–303.

[67] R. Krishnapuram, A. Joshi, O. Nasraoui, L. Yi, Low-complexity fuzzy relational clustering algorithms for web mining, IEEE Trans. Fuzzy Syst. 9 (4) (2001) 595–607.

[68] R. Krishnapuram, J.M. Keller, A possibilistic approach to clustering, IEEE Trans. Fuzzy Syst. 1 (2) (1993) 98–110.

[69] R. Krishnapuram, J.M. Keller, The possibilistic c-means algorithm: insights and recommendations, IEEE Trans. Fuzzy Syst. 4 (3) (1996) 385–393.

[70] J. Le-Rademacher, L. Billard, Symbolic covariance principal component analysis and visualization for interval-valued data, J. Comput. Graph. Stat. 21 (2) (2012) 413–432.

[71] Y. Lertworaprachaya, Y. Yang, R. John, Interval-valued fuzzy decision trees with optimal neighbourhood perimeter, Appl. Soft Comput. 24 (2014) 851–866.

[72] J.M. Leski, Towards a robust fuzzy clustering, Fuzzy Sets Syst. 137 (2) (2003) 215–233.

[73] J.M. Leski, Fuzzy (c+p)-means clustering and its application to a fuzzy rule-based classifier: towards good generalization and good interpretability, IEEE Trans. Fuzzy Syst. 23 (4) (2015) 802–812.

[74] J.M. Leski, Fuzzy c-ordered-means clustering, Fuzzy Sets Syst. 286 (1) (2016) 114–133.

[75] J.M. Leski, M. Kotas, On robust fuzzy c-regression models, Fuzzy Sets Syst. 279 (1) (2015) 112–129.

[76] C.F. Manski, E. Tamer, Inference on regressions with interval data on a regressor or outcome, Econometrica 70 (2) (2002) 519–546.

[77] M.-H. Masson, T. Denoeux, Ecm: an evidential version of the fuzzy c-means algorithm, Pattern Recognit. 41 (4) (2008) 1384–1397.

[78] A.B. McBratney, A.W. Moore, Application of fuzzy sets to climatic classification, Agric. For. Meteorol. 35 (1) (1985) 165–185.

[79] E. Neto, F.d.A. de Carvalho, Centre and range method for fitting a linear regression model to symbolic interval data, Comput. Stat. Data Anal. 52 (3) (2008) 1500–1515.

[80] E. Neto, F.d.A. de Carvalho, Constrained linear regression models for symbolic interval-valued variables, Comput. Stat. Data Anal. 54 (2) (2010) 333–347.

[81] S.R.K. Raju, Symbolic data analysis in cardiology, in: E. Diday, K.C. Gowda (Eds.), Symbolic Data Analysis and Its Applications, CEREMADE, Paris, 1997, pp. 245–249.

[82] F. Rehm, F. Klawonn, R. Kruse, A novel approach to noise clustering for outlier detection, Soft Comput. 11 (5) (2007) 489–494.

[83] F. Rehm, R. Winkler, R. Kruse, Fuzzy clustering with repulsive prototypes, in: A. Laurent, M. Lesot (Eds.), Scalable Fuzzy Algorithms for Data Management and Analysis, IGI Global, 2010, pp. 332–336.

[84] S.Z. Selim, M.A. Ismail, Soft clustering of multidimensional data: a semi-fuzzy approach, Pattern Recognit. 17 (5) (1984) 559–568.

[85] S.-F. Su, C.-C. Chuang, C.-W. Tao, J.-T. Jeng, C.-C. Hsiao, Radial basis function networks with linear interval regression weights for symbolic interval data, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 42 (1) (2012) 69–80.

[86] P. Teles, P. Brito, Modelling interval time series with space-time processes, Commun. Stat.: Theory Methods 44 (17) (2015) 3599–3627.

[87] H. Timm, C. Borgelt, C. Döring, R. Kruse, An extension to possibilistic fuzzy cluster analysis, Fuzzy Sets Syst. 147 (1) (2004) 3–16.

[88] X.-Y. Wang, J.M. Garibaldi, B. Bird, M.W. George, Novel developments in fuzzy clustering for the classification of cancerous cells using FTIR spectroscopy, in: J.V. DeOliveira, W. Pedrycz (Eds.), Advances in Fuzzy Clustering and its Applications, Wiley, Chichester, 2007, pp. 404–425.

[89] M. Wedel, W.A. Kamakura, Market Segmentation: Conceptual and Methodological Foundations, Kluwer Academic, Boston, 1998.

[90] M. Wedel, J.-B.E. Steenkamp, A fuzzy clusterwise regression approach to benefit segmentation, Int. J. Res. Market. 6 (4) (1989) 241–258.

[91] R. Winkler, F. Klawonn, R. Kruse, Fuzzy clustering with Polynomial Fuzzifier function in connection with M-Estimators, Appl. Comput. Math. 10 (1) (2011) 146–163.

[92] K.-L. Wu, M.-S. Yang, Alternative c-means clustering algorithms, Pattern Recognit. 35 (10) (2002) 2267–2278.

[93] R.R. Yager, On ordered weighted averaging aggregation operators in multi-criteria decision making, IEEE Trans. Syst. Man Cybern. 18 (1) (1988) 183–190.

[94] M.-S. Yang, K.-L. Wu, Unsupervised possibilistic clustering, Pattern Recognit. 39 (1) (2006) 5–21.

[95] M.-S. Yang, K.-L. Wu, J.-N. Hsieh, J. Yu, Alpha-cut implemented fuzzy clustering algorithms and switching regressions, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 38 (3) (2008) 588–603.

[96] M.F. Zarandi, Z.S. Razaee, M. Karbasian, A fuzzy case based reasoning approach to value engineering, Expert Syst. Appl. 38 (8) (2011) 9334–9339.

Pierpaolo D'Urso is currently a professor of Statistics at Sapienza University of Rome. Currently he teaches statistics and multivariate statistics at Sapienza University. He received his Ph.D. in Statistics and his bachelor's degree in Statistics both from Sapienza. He is an Associate Editor of several journals, like Advances in Data Analysis and Classification, Information Sciences, Metron—International Journal of Statistics, Advances in Computational Research. He is a Member of the Editorial Board of several international journals. He is a member of the Evaluation board of the LUISS University, and served as a member of the Advisory Committee on Statistics at Presidency of the

Council of Ministers of Italy. He has served as a referee of several international journals and conferences in statistics, data mining and applied mathematics. He has authored over 100 articles published in international journals, conference proceedings, volumes. His recent research activity focuses on fuzzy clustering, clustering and classification of time series, fuzzy regression, and statistical methods in web marketing, neuromarketing, geomarketing, consumer path analysis and tourism.

Jacek M. Leski was born in Gliwice, Poland and received the M.Sc., Ph.D. and D.Sc. degrees in electronics from Silesian University of Technology in 1987, 1989 and 1995, respectively. He is a Professor in Biomedical Information Processing at the Silesian University of Technology and at the present moment, the Head of its Division of Biomedical Electronics. He is also a Professor at the Institute of Medical Technology and Equipment, Zabrze, Poland. His research interests include digital processing of biomedical signals, fuzzy and neuro-fuzzy modeling, pattern recognition and learning theory. In these subjects he has authored and coauthored 5 monographs, 12 chapters in monographs and more than 170 international journal and conference papers. He is a Senior Member of the IEEE, the EMBS, the IEEE Signal Processing Society, the IEEE System, Man and Cybernetics Society, the IEEE Neural Networks Society, IEEE Communication Society, IEEE Control Systems Society, the Polish Biomedical Engineering Society and the Polish Society of Theoretical and Applied Electrotechnics.