Spatial Statistics 57 (2023) 100764

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Fuzzy clustering of spatial interval-valued data

Pierpaolo D'Urso [a,*], Livia De Giovanni [b], Lorenzo Federico [b], Vincenzina Vitale [a]

[a] Department of Social Sciences and Economics, Sapienza - University of Rome, P.za Aldo Moro, 5-00185, Rome, Italy

[b] Department of Political Sciences and Data Lab, Luiss university - Viale Romania, 32-00197, Rome, Italy

ARTICLE INFO

Article history:
Received 18 April 2023
Received in revised form 3 July 2023
Accepted 17 July 2023
Available online 2 August 2023

Keywords:
Spatial imprecise data
Fuzzy clustering
Partitioning around medoids
Entropy
Environmental data
Networks

ABSTRACT

In this paper, two fuzzy clustering methods for spatial interval-valued data are proposed, i.e. the fuzzy C-Medoids clustering of spatial interval-valued data with and without entropy regularization. Both methods are based on the Partitioning Around Medoids (PAM) algorithm, inheriting the great advantage of obtaining non-fictitious representative units for each cluster.

In both methods, the units are endowed with a relation of contiguity, represented by a symmetric binary matrix. This can be intended both as contiguity in a physical space and as a more abstract notion of contiguity. The performances of the methods are proved by simulation, testing the methods with different contiguity matrices associated to natural clusters of units. In order to show the effectiveness of the methods in empirical studies, three applications are presented: the clustering of municipalities based on interval-valued pollutants levels, the clustering of European fact-checkers based on interval-valued data on the average number of impressions received by their tweets and the clustering of the residential zones of the city of Rome based on the interval of price values.

* Corresponding author.
E-mail addresses: pierpaolo.durso@uniroma1.it (P. D'Urso), ldegiovanni@luiss.it (L. De Giovanni), lfederico@luiss.it (L. Federico), vincenzina.vitale@uniroma1.it (V. Vitale).

https://doi.org/10.1016/j.spasta.2023.100764

1. Introduction

In the literature on data analysis, a great deal of attention is paid to statistical methods to treat interval-valued data, in different research areas (Coppi et al., 2006; Denoeux and Masson, 2000; D'Urso and Giordani, 2005; D'Urso and De Giovanni, 2014; D'Urso and Leski, 2016; Giordani and Kiers, 2004).

Interesting methods have been suggested in a classical cluster analysis framework. In particular, Gowda and Diday (1991) hinted a clustering method for symbolic data; (Guru et al., 2004) proposed a similarity measure to compare interval-valued data and a modified agglomerative method for clustering symbolic data. De Carvalho and Lechevallier (2009b) proposed a partitional dynamic clustering method for interval data based on adaptive Hausdorff distances.

An interesting line of research has focused on the clustering of interval-valued data based on fuzzy approaches, where the weighting exponent m controls the extent of membership sharing between fuzzy clusters (De Carvalho and Tenório, 2010; D'Urso et al., 2015; D'Urso and Giordani, 2006a; D'Urso et al., 2017). Li and Mukaidono (1995) remarked that this unusual parameter is unnatural and does not have a physical meaning.

As an alternative approach to methods based on weighting exponent m (Fuzzy Partitioning Around Centroids or Medoids), in the literature many authors proposed fuzzy clustering techniques based on an entropy-regularization term. In particular, considering standard data, Yao et al. (2000) proposed an entropy-based fuzzy clustering method which automatically identifies the number and initial locations of cluster centers. Successively, it removes all data points having dissimilarity larger than a threshold with the chosen cluster center; the procedure is repeated until all data points are removed. Ichihashi (2000) and Miyagishi et al. (2000) suggested a generalized objective function with additional variables. These authors consider a covariance matrix and show an equivalence between their Kullback–Leibler (KL) fuzzy clustering and the Gaussian mixture model. The method of fuzzy clustering using the KL information is called entropy-based method of FCM. Coppi and D'Urso (2006) suggested fuzzy unsupervised clustering methods based on Shannon entropy regularization to classify time-varying data. Zarinbal et al. (2014) proposed a new fuzzy clustering method based on FCM and the relative entropy is added to the objective function as a regularization function to maximize the dissimilarity between clusters. Kahali et al. (2019) presented an entropy-based FCM segmentation method that incorporates the uncertainty of classification of individual pixels within the classical framework of FCM. Gao et al. (2019) showed a novel method considering noise intelligently based on the existing FCM approach, called adaptive-FCM and its extended version (adaptive-REFCM) in combination with relative entropy. We remark that the cited methods use entropy-based regularization regard standard data in the clustering problem.

An additional source of complexity in the clustering problems is due to the spatial nature of data. In this regard, different methods have been suggested in the clustering literature to discover spatial patterns for different kind of spatial units, e.g., urban areas or image pixels. The main challenge these methods deal with is the identification of an appropriate algorithm to capture both spatial dependence and spatial heterogeneity.

Fouedjio (2016) classifies clustering of spatial data into four main approaches: (1) non-spatial clustering with geographical coordinates as additional variables (D'Urso and Vitale 2020, Fouedjio (2016)); (2) non-spatial clustering based on a spatial dissimilarity measure; (3) spatially constrained clustering; (4) model-based clustering (Durante and Sempi 2016, D'Urso et al. (2021)). A fifth approach worth of notice consists in including a spatial penalty term in the objective function of the clustering method, as suggested by Pham (2001). Examples of applications for image pixels segmentation can be found in Tolias and Panas (1998a,b), Pham and Prince (1999), Liew et al. (2000, 2003), Pham (2001), Liew et al. (2003), Chuang et al. (2006). While this proposal has been introduced for solving image segmentation problem, the idea beyond can be easily extended (López-Oriona et al. 2022, D'Urso et al. (2022a), D'Urso et al. (2022b), D'Urso et al. (2019a,b) , Disegna et al. (2017), Coppi et al. (2010)).

A comparison of some clustering methods for interval-valued data proposed in the literature is presented in Table 1. To the best of our knowledge, there are not methods of fuzzy clustering based on PAM and/or entropy regularization for spatial interval-valued data.

Table 1
Comparison of some clustering methods for interval-valued data proposed in the literature.

Author	Methodological approach	Metric	Uncertainty managing	Robustness	Spatial data
Guru et al. (2004)	Agglomerative	Similarity measure	crisp	No	No
D'Urso and Giordani (2006a)	c-means	Squared Euclidean distance	fuzzy	Yes, noise cluster approach	No
De Carvalho (2007)	c-means	Squared Euclidean distance	fuzzy	No	No
De Carvalho and Lechevallier (2009a)	Partition dynamic	Adaptive Hausdorff distance	fuzzy	No	No
De Carvalho and Tenório (2010)	c-means	Adaptive quadratic distance	fuzzy	No	No
D'Urso et al. (2015)	c-medoids clustering	Squared Euclidean distance	fuzzy	Yes, trimmed approach	No
D'Urso and Leski (2016)	c-medoids	Squared Euclidean distance and Loss functions	fuzzy	Yes, Huber's M-estimators, Ordered weighted averaging, Typicality degrees	No
D'Urso et al. (2017)	c-medoids	Exponential transformation of squared Euclidean distance	fuzzy	Yes, exponential metric approach	No
Maharaj et al. (2019)	hierarchical, non-hierarchical, c-means, c-medoids	Frobenius distance between the autocorrelation and cross-correlation functions of the interval time series upper and lower bound	crisp	No	No
D'Urso et al. (2023a)	Entropy regularization	Exponential transformation of squared Euclidean distance	fuzzy	Yes, exponential metric approach	No
D'Urso et al. (2023b)	c-medoids	Distance between the wavelet variance–covariance matrices of centers and radius at each scale of two time series	fuzzy	No	No

Following this line of research, in this paper two fuzzy clustering methods for interval-valued data and spatial penalty term are proposed. The methods are the fuzzy clustering of spatial interval-valued data with and without entropy regularization, henceforth denoted with EFMd-SpID and EFMd-SpID respectively. Both methods are based on the fuzzy Partitioning Around Medoids (PAM) algorithm, inheriting the great advantage of obtaining non-fictitious representative units for each cluster. The main difference lies in the way fuzziness is introduced, i.e. in the EFMd-SpID linearly by subtracting from the clustering criterion the entropy of the fuzzy partition weighted by a suitable positive factor, in the EFMd-SpID by raising each membership degree to a power greater than 1.

The paper is structured as follows: in Section 2 the clustering methods are described; in Section 3 a simulation study is presented; in Section 4 the methodology is illustrated by analyzing real data; Section 5 concludes.

2. Fuzzy clustering of spatial interval-valued data

2.1. Dissimilarity measure for interval-valued data

An interval-valued datum can be formalized as $x_{ij} = [\underline{x}_{ij}, \overline{x}_{ij}],\ i = 1, \ldots, I;\ j = 1, \ldots, J$, where x_{ij} indicates the jth interval-valued variable observed on the ith unit; $\underline{x}_{ij}$ and $\overline{x}_{ij}$ denote, respectively, the lower and upper bounds of the interval, i.e., they represent the minimum and maximum values of the jth interval-valued variable with respect to the ith unit. Each unit is represented geometrically by a hyper-rectangle in $\Re^J$ having 2^J vertices. All the 2^J vertices correspond to all the possible (lower bound, upper bound) combinations. In particular, in $\Re$ ($J = 1$) the generic unit is represented by a segment; in $\Re^2$ ($J = 2$), it is represented by a rectangle with $2^2 = 4$ vertices, and so on Cazes et al. (1997).

Then, assuming J interval-valued variables are observed on I units, the entire dataset can be stored in the so-called *interval-valued matrix* as follows:

$$\mathbf{X} \equiv \{x_{ij} = [\underline{x}_{ij}, \overline{x}_{ij}] : i = 1, \ldots, I;\ j = 1, \ldots, J\}. \tag{1}$$

By denoting with

$$\mathbf{M} \equiv \left\{ m_{ij} = \frac{\overline{x}_{ij} + \underline{x}_{ij}}{2} : i = 1, \ldots, I;\ j = 1, \ldots, J \right\}, \tag{2}$$

the *midpoint matrix* (*center matrix*), where m_{ij} is the midpoint (center) of the associated interval value for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, and with

$$\mathbf{R} \equiv \left\{ r_{ij} = \frac{\overline{x}_{ij} - \underline{x}_{ij}}{2} : i = 1, \ldots, I;\ j = 1, \ldots, J \right\}, \tag{3}$$

the *radius matrix*, where r_{ij} is the radius (spread) of the associated interval for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, we can reformulate the interval-valued matrix (1) as follows:

$$\tilde{\mathbf{X}} \equiv \{\tilde{x}_{ij} = [m_{ij}, r_{ij}] : i = 1, \ldots, I;\ j = 1, \ldots, J\} = \{\tilde{\mathbf{x}}_i = [\mathbf{m}_i, \mathbf{r}_i] : i = 1, \ldots, I\}, \tag{4}$$

where $\mathbf{m}_i$ and $\mathbf{r}_i$ denote, respectively, the ith row of $\mathbf{M}$ and $\mathbf{R}$.

Then, $\tilde{x}_{ij} = [m_{ij}, r_{ij}]$ represents an alternative formalization of the interval-valued datum $x_{ij} = [\underline{x}_{ij}, \overline{x}_{ij}]$. In this way, the lower and upper bounds of the interval-valued datum can be obtained as $\underline{x}_{ij} = m_{ij} - r_{ij}$ and $\overline{x}_{ij} = m_{ij} + r_{ij}$, respectively.

In the literature, several metrics have been suggested for interval-valued data (D'Urso and Giordani 2004, Guru et al. (2004), Ichino and Yaguchi (1994), Kabir et al. (2017)). In this paper, we adopt a weighted distance measure (D'Urso et al. 2014, D'Urso et al. (2016)).

The distance measure is weighted as the dissimilarity between each pair of units is measured by separately considering the midpoints and the radii of the interval-valued data and using a suitable weighting system for such components (D'Urso and Giordani, 2006b).

In formula, the weighted distance measure between units i and i' is:

$$d_w(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i'}) = [w_m^2 d^2(\mathbf{m}_i, \mathbf{m}_{i'}) + w_r^2 d^2(\mathbf{r}_i, \mathbf{r}_{i'})]^{\frac{1}{2}} \tag{5}$$

where $d^2(\mathbf{m}_i, \mathbf{m}_{i'}) = \|\mathbf{m}_i - \mathbf{m}_{i'}\|^2$ is the squared Euclidean distance between the midpoints and $d^2(\mathbf{r}_i, \mathbf{r}_{i'}) = \|\mathbf{r}_i - \mathbf{r}_{i'}\|^2$ is the squared Euclidean distance between the radii, while w_m and w_r are the weights for the midpoint component and the radius component, respectively. For practical purposes we will set $w_r = w$, $w_m = 1 - w$, under the constraint $0 \le w \le 0.5$ (*coherence* condition).

The distance measure (5) has the following features: (i) $d_w(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i'})$ is a metric. The properties of identity, non negativity, symmetry are easily verified. Also the triangular inequality property is satisfied, as it can be verified after some algebra. (ii) $d_w(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_{i'})$ is computationally easy and theoretically intuitive.

2.2. Dealing with space: Contiguity matrix

When dealing with spatial data the within group dispersion has to be minimized and the spatial autocorrelation between contiguous spatial units has to be taken into consideration. In the literature, there are different ways of defining neighborhood and consequently there are different ways of constructing contiguity matrices among spatial units (Páez and Scott, 2005). We want to use a very general definition of contiguity that can be applied to a vast variety of settings and that it is easy to formalize mathematically. This can be intended both as contiguity in a physical space, such as spatial units being adjacent (sharing a border) or belonging to the same macro-area, even if they are not adjacent and as a more abstract notion of contiguity, such as friendship on social networks, written texts having the same topic, readers, or authors or businesses having financial transactions. When we have a set of I units we can abstract their contiguity (adjacency) relations as an $I \times I$ matrix $\mathbf{A}$ such that $a_{ii'} = 1$ if the unit i is contiguous to the unit i' and $a_{ii'} = 0$ otherwise. All these different definitions of connectivity and contiguity can be embedded in the clustering procedure through their matrix $\mathbf{A}$. In this paper we will deal only with *undirected* contiguity relations, that is, we will always assume the matrix $\mathbf{A}$ to be symmetric.

2.3. Fuzzy clustering of spatial interval-valued data

In this paper, two fuzzy clustering methods of spatial interval-valued data are proposed, the fuzzy clustering of spatial interval-valued data with entropy regularization (Section 2.3.1) and the fuzzy clustering of spatial interval-valued data without entropy regularization (Section 2.3.2). For both methods, the PAM algorithm, also known as Fuzzy C-Medoids (FCMd), is adopted thanks of its great advantage of obtaining non-fictitious representative units as final result. This allows for more appealing and easy to interpret results of the final partition (Kaufman and Rousseeuw, 2005). From a computational perspective, fuzzy clustering algorithms are generally more efficient (dramatic changes in the value of cluster membership are less likely to occur in estimation procedures) and they are less affected by both local optima and convergence problems (Everitt et al., 2001; Hwang et al., 2007).

2.3.1. Fuzzy C-medoids clustering of spatial interval-valued data with entropy regularization

Given a set of I units, of which we know both some attributes which are represented in an interval-valued form according to (4) and a contiguity structure represented by a matrix $\mathbf{A}$, we want to find a fuzzy partition into C clusters that avoids clustering together units with very different attribute values and clustering separately units which are contiguous. More formally given the matrices $\tilde{\mathbf{X}}$ and $\mathbf{A}$, we want to build a membership matrix $\mathbf{U} = \{u_{ic}, i \leq I, c \leq C\}$ in which u_{ic} is the membership degree of the unit i to the cluster c that solves the following optimization problem:

$$
\begin{aligned}
\min: & \sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic} d_w^2(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_c) + p\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic} ln(u_{ic}) + \frac{\gamma}{2}\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic}\sum_{i'=1}^{I}\sum_{c'\in C_c} a_{ii'}u_{i'c'} = \\
& = \sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic}[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c)] + p\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic} ln(u_{ic}) + \\
& + \frac{\gamma}{2}\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic}\sum_{i'=1}^{I}\sum_{c'\in C_c} a_{ii'}u_{i'c'} \\
& s.t. \sum_{c=1}^{C} u_{ic} = 1,\ u_{ic} \geq 0,\ 0 \leq w \leq 0.5
\end{aligned}
\tag{6}
$$

where $p\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic} ln(u_{ic})$ is the entropy term; $\frac{\gamma}{2}\sum_{i=1}^{I}\sum_{c=1}^{C} u_{ic}\sum_{i'=1}^{I}\sum_{c'\in C_c} a_{ii'}u_{i'c'}^{m}$ is the spatial penalty term; $\gamma \geq 0$ is the tuning parameter of the spatial information (spatial coefficient); $a_{ii'}$ is the generic element of the $(I \times I)$ "contiguity" matrix $\mathbf{A}$; C_c is the set of the C clusters, with the exclusion of cluster c; u_{ic} is the membership degree of the unit i to the cluster c. For each spatial unit i and each cluster c, the sum of the membership degrees of the contiguous (neighboring, adjacent) spatial units (as indicated in the matrix $\mathbf{A}$) in all the clusters except cluster c (summarized C_c) is constrained to be as small as possible. We can observe that the spatial coefficient γ tunes the trade-off between internal cohesion based on the feature vectors and the spatial homogeneity of the clusters. For $\gamma = 0$ the spatial regularization is not taken into account.

By solving the constrained quadratic minimization problem shown in Eq. (6) *via* the Lagrangian multiplier method, we obtain the optimal solutions u_{ic} and w:

$$u_{ic} = \frac{exp\left\{-\frac{1}{p}\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right]\right\}}{\sum_{c'} exp\left\{-\frac{1}{p}\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_{c'}) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_{c'}) + \gamma \sum_{i'=1}^{I} \sum_{c'' \in C_{c'}} a_{ii'} u_{i'c''}\right]\right\}} \tag{7}$$

$$w_0 = \frac{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} d^2(\mathbf{m}_i, \mathbf{m}_c)}{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}(d^2(\mathbf{m}_i, \mathbf{m}_c) + d^2(\mathbf{r}_i, \mathbf{r}_c))}. \tag{8}$$

$$w = min\{w_0, 0.5\}$$

In particular, by considering the following Lagrangian function:

$$\begin{aligned} L_m(u_{ic}, \lambda, w) = &\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} d_w^2(\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{x}}_c) + p \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} ln(u_{ic}) + \\ &+ \frac{\gamma}{2} \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'} - \lambda \left(\sum_{c=1}^{C} u_{ic} - 1 \right) \end{aligned} \tag{9}$$

and setting the first partial derivatives with respect to u_{ic} and λ equal to zero, we obtain:

$$\begin{aligned} \frac{\partial L_m(u_{ic}, \lambda, w)}{\partial u_{ic}} = 0 \Leftrightarrow &\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c)\right] + \\ &+ p(ln(u_{ic}) + 1) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'} - \lambda = 0 \end{aligned} \tag{10}$$

$$\frac{\partial L_m(u_{ic}, \lambda w)}{\partial \lambda} = 0 \Leftrightarrow \sum_{c=1}^{C} u_{ic} - 1 = 0. \tag{11}$$

From Eq. (10), we obtain:

$$ln(u_{ic}) = \frac{1}{p}\left[\lambda - [(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c)] - \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right] - 1 \tag{12}$$

Notice that in (10) the coefficient $\frac{1}{2}$ pertaining to γ vanishes because the derivative results in a term corresponding to the product of u_{ic} and its neighbors, plus a term corresponding to the inverse product of the neighbors and u_{ic}. Then:

$$u_{ic} = \exp\left\{\frac{\lambda}{p} - \frac{1}{p}\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right] - 1\right\}. \tag{13}$$

By considering Eq. (11):

$$exp\left(\frac{\lambda}{p} - 1\right) = \frac{1}{\sum_{c=1}^{C} \exp\left\{-\frac{1}{p}\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right]\right\}} \tag{14}$$

and by replacing Eq. (14) in Eq. (13), we obtain (7).

The normalization condition for w is implicitly satisfied. To take into account the *coherence condition*, we derive with respect to w and select the minimum between the obtained value and 0.5:

$$\frac{\partial L_m(u_{ic}, \lambda, w)}{\partial w} = 0$$

$$w = \frac{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} d^2(\mathbf{m}_i, \tilde{\mathbf{m}}_c)}{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}(d^2(\mathbf{m}_i, \tilde{\mathbf{m}}_c) + d^2(\mathbf{r}_i, \tilde{\mathbf{r}}_c))}. \tag{15}$$

The inclusion of the spatial penalty term in the iterative solution ensures that the final membership degrees of spatial units contiguous to i are as high as possible with the cluster to which i belongs. In order to ensure a certain degree of spatial homogeneity within the clusters, γ tunes the membership degrees. Notice that this does not imply that a cluster is formed by solely areas that are contiguous in one or more senses. Let i and c be a generic spatial unit and a generic cluster, respectively. The aim of the spatial penalty term is to reduce the membership degrees of all units contiguous to i computed in all clusters but c (Coppi et al., 2010).

The fuzzy clustering procedure is illustrated in Algorithm 1.

Algorithm 1 Fuzzy C-Medoids clustering for Spatial Interval-Value Data with entropy regularization (EFMd-SpID) algorithm

```
Fix C, max.iter and generate randomly the membership degree matrix U;
Set iter = 0;
Set medoids= {x̃_1, ..., x̃_C} arbitrarily;
repeat
    Set old.medoids=medoids;
    Update U by using (17);
    Compute w by using (24);
    Update medoids as follows:
    for c = 1 to C do
        q = arg min_{1≤i'≤I} Σ_{i''=1}^{I} u_{i''c}[(1 − w)^2 d^2(m_{i'}, m_{i''}) + w^2 d^2(r_{i'}, r_{i''})] +
        + p Σ_{i''=1}^{I} Σ_{c=1}^{C} u_{i''c} ln(u_{i''c})
        Set ⇒ x̃_c = x̃_q
    end for
    iter ← iter + 1;
until medoids=old.medoids or iter = max.iter
```

The fuzzy clustering algorithm that minimizes (6) is built by adopting an estimation strategy based on the Fu and Albus heuristic algorithm (Fu and Albus 1977, Krishnapuram et al. (1999, 2001)). Indeed, the alternating optimization estimation procedure cannot be adopted because the necessary conditions cannot be derived by differentiating the objective function in (6) with respect to the medoids.

2.3.2. Fuzzy C-medoids clustering of spatial interval-valued data without entropy regularization

The clustering method can be formalized as follows:

$$\begin{aligned}
\min : & \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m d_w^2(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_c) + \frac{\gamma}{2} \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m = \\
& = \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m [(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c)] + \frac{\gamma}{2} \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m \\
& s.t. \sum_{c=1}^{C} u_{ic} = 1, \; u_{ic} \geq 0, \; 0 \leq w \leq 0.5
\end{aligned} \tag{16}$$

where m is the fuzziness parameter; the other terms are defined as in (6). By solving the constrained quadratic minimization problem shown in Eq. (16) *via* the Lagrangian multiplier method, we obtain the optimal solutions u_{ic} and w:

$$u_{ic} = \frac{\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right]^{-\frac{1}{m-1}}}{\sum_{c'=1}^{C} \left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_{c'}) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_{c'}) + \gamma \sum_{i'=1}^{I} \sum_{c'' \in C_{c'}} a_{ii'} u_{i'c''}\right]^{-\frac{1}{m-1}}} \tag{17}$$

$$w_0 = \frac{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m d^2(\mathbf{m}_i, \mathbf{m}_c)}{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m (d^2(\mathbf{m}_i, \mathbf{m}_c) + d^2(\mathbf{r}_i, \mathbf{r}_c))}. \tag{18}$$

$$w = min\{w_0, 0.5\}$$

In particular, by considering the following Lagrangian function:

$$L_m(u_{ic}, \lambda, w) = \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m d_w^2(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_c) + \frac{\gamma}{2} \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m - \lambda \left(\sum_{c=1}^{C} u_{ic} - 1 \right) \tag{19}$$

and setting the first partial derivatives with respect u_{ic} and λ equal to zero, we obtain:

$$\begin{aligned} \frac{\partial L_m(u_{ic}, \lambda, w)}{\partial u_{ic}} = 0 \Leftrightarrow & m u_{ic}^{m-1} \left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c)\right] + \\ & + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'} - \lambda = 0 \end{aligned} \tag{20}$$

$$\frac{\partial L_m(u_{ic}, \lambda w)}{\partial \lambda} = 0 \Leftrightarrow \sum_{c=1}^{C} u_{ic} - 1 = 0. \tag{21}$$

From Eq. (20), we obtain:

$$u_{ic} = \left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \frac{1}{\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right]^{\frac{1}{m-1}}} \tag{22}$$

By considering Eq. (21):

$$1 = \left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} \sum_{c=1}^{C} \frac{1}{\left[(1-w)^2 d^2(\mathbf{m}_i, \mathbf{m}_c) + w^2 d^2(\mathbf{r}_i, \mathbf{r}_c) + \gamma \sum_{i'=1}^{I} \sum_{c' \in C_c} a_{ii'} u_{i'c'}\right]^{\frac{1}{m-1}}} \tag{23}$$

and by replacing Eq. (23) in Eq. (22), we obtain (17).

The normalization condition for w is implicitly satisfied. To take into account the *coherence condition*, we derive with respect to w and select the minimum between the obtained value and 0.5:

$$\frac{\partial L_m(u_{ic}, \lambda, w)}{\partial w} = 0$$

$$w = \frac{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} d^2(\mathbf{m}_i, \mathbf{m}_c)}{\sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic} (d^2(\mathbf{m}_i, \mathbf{m}_c) + d^2(\mathbf{r}_i, \mathbf{r}_c))}. \tag{24}$$

The fuzzy clustering procedure is illustrated in Algorithm 2.

In summary, some remarks on the proposed methods. It can be noted that:

Algorithm 2 Fuzzy C-Medoids clustering for Spatial Interval-Value Data without entropy regularization (FMd-SpID) algorithm

```
Fix $C$, $max.iter$ and generate randomly the membership degree matrix $\mathbf{U}$;
Set $iter = 0$;
Set $medoids = \{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_C\}$ arbitrarily;
repeat
    Set old.medoids=medoids;
    Update U by using (17);
    Compute $w$ by using (24);
    Update medoids as follows:
    for $c = 1$ to $C$ do
        $q = \arg\min_{1 \le i' \le I} \sum_{i''=1}^{I} u_{i''c}^{m} [(1-w)^2 d^2(\mathbf{m}_{i'}, \mathbf{m}_{i''}) + w^2 d^2(\mathbf{r}_{i'}, \mathbf{r}_{i''})]$
        Set $\Rightarrow \tilde{\mathbf{x}}_c = \tilde{\mathbf{x}}_q$
    end for
    $iter \leftarrow iter + 1$;
until medoids=old.medoids or $iter = max.iter$
```

- interval-valued data are considered;
- spatial contiguity relations are included;
- the methods inherit all the peculiarities of the distance (5) for interval-valued data;
- the PAM approach makes easier the interpretation: the medoids represent their clusters with respect to the entity of the values of the centers (by means of the centers of the medoids) and with respect to the uncertainty levels around the values of the centers (represented by the radii of the medoids);
- the complexity of the methods is $O(I^2)$ being the complexity linear with respect to J and C;
- the iterative algorithm converges to, at least, a local optimum; to limit the risk of hitting local optima, more than one random start is recommended.

2.4. Validity measure

In general, internal validity measures provide useful guidelines in the identification of the best partition (as suggested by Handl et al., 2005; D'Urso, 2015). A suitable measure for fuzzy clustering algorithm has been suggested by Campello and Hruschka (2006).

The Fuzzy Silhouette (FS) index (Campello and Hruschka, 2006) is a popular measure computed as the weighted average of the individual silhouette widths, s_i, as follows:

$$FS = \frac{\sum_{i=1}^{I} (u_{ip} - u_{iq})^{\alpha} \cdot s_i}{\sum_{i=1}^{I} (u_{ip} - u_{iq})^{\alpha}}, \qquad s_i = \frac{(b_{ip} - a_{ip})}{max(b_{ip}, a_{ip})} \tag{25}$$

where a_{ip} is the average distance of unit i to all other units belonging to cluster p; b_{ip} is the minimum (over clusters) of the average distances d_{iq} of unit i to all units belonging to another cluster q, $q \neq p$; u_{ip} and u_{iq} are the first and second largest elements of the ith row of the fuzzy partition matrix $\mathbf{U}$; $\alpha \geq 0$ is an optional user defined weighting coefficient. The distance used in (25) is the weighted distance (5). Note that the traditional (crisp) Silhouette measure is obtained by setting $\alpha = 0$.

The higher the value of FS, the better the assignment of the units to the clusters simultaneously obtaining the minimization of the intra-cluster distance and the maximization of the inter-cluster distance.

2.5. Spatial autocorrelation

The Fuzzy Moran (FM)'s index is a measure of spatial autocorrelation to assess the post-cluster autocorrelation between units. It is a generalization of the spatial autocorrelation measure introduced by Coppi et al. (2010). The suggested spatial autocorrelation of the final partition is a multivariate fuzzy generalization of the Moran's index (Gittleman and Kot, 1990). The idea of the *FM* index is to compute the spatial autocorrelation between classified units in which both the matrix of membership degrees and the spatial proximity matrices are considered. The *FM* index is defined as follows:

$$FM = \frac{tr\left[\bar{\tilde{\mathbf{X}}}'\mathbf{U}_c^{\frac{1}{2}}\mathbf{A}\mathbf{U}_c^{\frac{1}{2}}\bar{\tilde{\mathbf{X}}}\right]}{tr\left[\bar{\tilde{\mathbf{X}}}'\mathbf{U}_c^{\frac{1}{2}}\,diag(\mathbf{A}'\mathbf{A})\mathbf{U}_c^{\frac{1}{2}}\bar{\tilde{\mathbf{X}}}\right]} \tag{26}$$

where $\mathbf{U}_c$ is the square diagonal matrix of order I of the membership degrees of cluster c; $\bar{\tilde{\mathbf{X}}}$ is the centered matrix; $\mathbf{A}$ is the spatial matrix.

The *FM* index ranges between -1 and 1. A value of 1 indicates perfect positive autocorrelation, 0 indicates no autocorrelation and -1 indicates perfect negative autocorrelation. Thus higher the value of *FM*, the better the geographical assignment of the units to the clusters.

3. Simulation

The aim of the simulation study is to assess the sensitivity of the clustering process to the contiguity matrices. The fuzzy clustering method of spatial interval-valued data with entropy regularization EFMd-SpID is used. In the simulation study $I = 16$, $J = 2$. The generation process either for the midpoint or for the radius of the $J = 2$ variables is presented in Table 3. For each unit 6 values of the $J = 2$ variables are generated, and the minimum and maximum and related midpoint and radius over the period are computed (Fig. 1). The natural clusters are units (1, 2, 3, *4*), (*5*, 6, 7, 8), (9, 10, 11, *12*), (*13*, 14, 15, 16), numbering the units clockwise starting from the bottom left. The four units 4, 5, 12, 13 show "soft" membership to the natural clusters (in the middle in Fig. 1, top).

Two contiguity matrices are considered. The contiguous units are (1, 2, 3, 4), (5, 6, 7, 8), (9, 10, 11, 12), (13, 14, 15, 16) in P_1; (1, 6, 9, 14), (2, 7, 10, 15), (3, 8, 11, 16), (4, 5, 12, 13) in P_2 (Fig. 2). We choose these two different contiguity matrices to showcase how the behavior of the clustering changes depending on how the two kinds of data (attributes and contiguity structure) taken as input by the algorithm are related to each other. In particular, we defined P_1 in such a way that units who are contiguous also have similar values of their attributes and P_2 so that instead units who are contiguous are more likely to have very distant values in the attributes. The degree of agreement between contiguity structure and attribute similarity is called assortativity and is measured by the *attribute assortativity coefficient* of the contiguity network, originally defined in Newman (2003). For each of the two variables j we compute the assortativity coefficients of the centers:

$$\rho(A, j) = \frac{\sum_{(i,i'):a_{ii'}=1}(m_{ij} - \overline{m}_j)(m_{i'j} - \overline{m}_j)}{\sum_{(i,i'):a_{ii'}=1}(m_{ij} - \overline{m}_j)^2}. \tag{27}$$

By definition $\rho(A, j) \in [-1, 1]$ and the higher it is, the more the contiguity structure agrees with the attribute similarity. When $\rho(A, j) > 0$ we say that the contiguity structure is assortative, when $\rho(A, j) < 0$ that it is disassortative. As shown in Table 2, the contiguity matrix P_1 is strongly assortative, while P_2 is disassortative. We underline that the assortativity coefficient is an intrinsic property of the data, which is neither taken explicitly as an input nor derived from the output of the model. That said, the final clustering looks very different on assortative and disassortative contiguity structures. We will see in the analysis how on assortative structures the algorithm usually finds more coherent clusters, as the same membership assignments will minimize both the spatial penalty and the attribute-based objective function, while on disassortative structures the algorithm has to find a balance between contrasting optimizations and usually finds membership degrees that are less crisp and sometimes harder to label properly.

The value of the entropy parameter $p = 1$ is selected by the fuzzy silhouette without spatial penalty term for C=4. The parameter γ ranges in [0.5; 2] with step 0.5.

The performances – fuzzy silhouette (25) and values of the objective function (6) – and the values of the spatial autocorrelation (26) are presented in Tables 4, 5, and 6.

Table 2
Attribute assortativity coefficients in the simulated data.

	$\rho(A, 1)$	$\rho(A, 2)$
P_1	0.846	0.823
P_2	−0.332	−0.333

Table 3
Data generation process.

	Units 1,2,3	Unit 4	Unit 5	Units 6,7,8	Units 9,10,11	Unit 12	Unit 13	Units 14,15,16
$j = 1$	$U[0.0, 0.5]$	$U[0.7, 0.8]$	$U[0.7, 0.8]$	$U[0.0, 0.5]$	$U[1.5, 2.0]$	$U[1.3, 1.4]$	$U[1.3, 1.4]$	$U[1.5, 2.0]$
$j = 2$	$U[0.0, 0.5]$	$U[0.7, 0.8]$	$U[1.3, 1.4]$	$U[1.5, 2.0]$	$U[1.5, 2.0]$	$U[1.3, 1.4]$	$U[0.7, 0.8]$	$U[0.0, 0.5]$

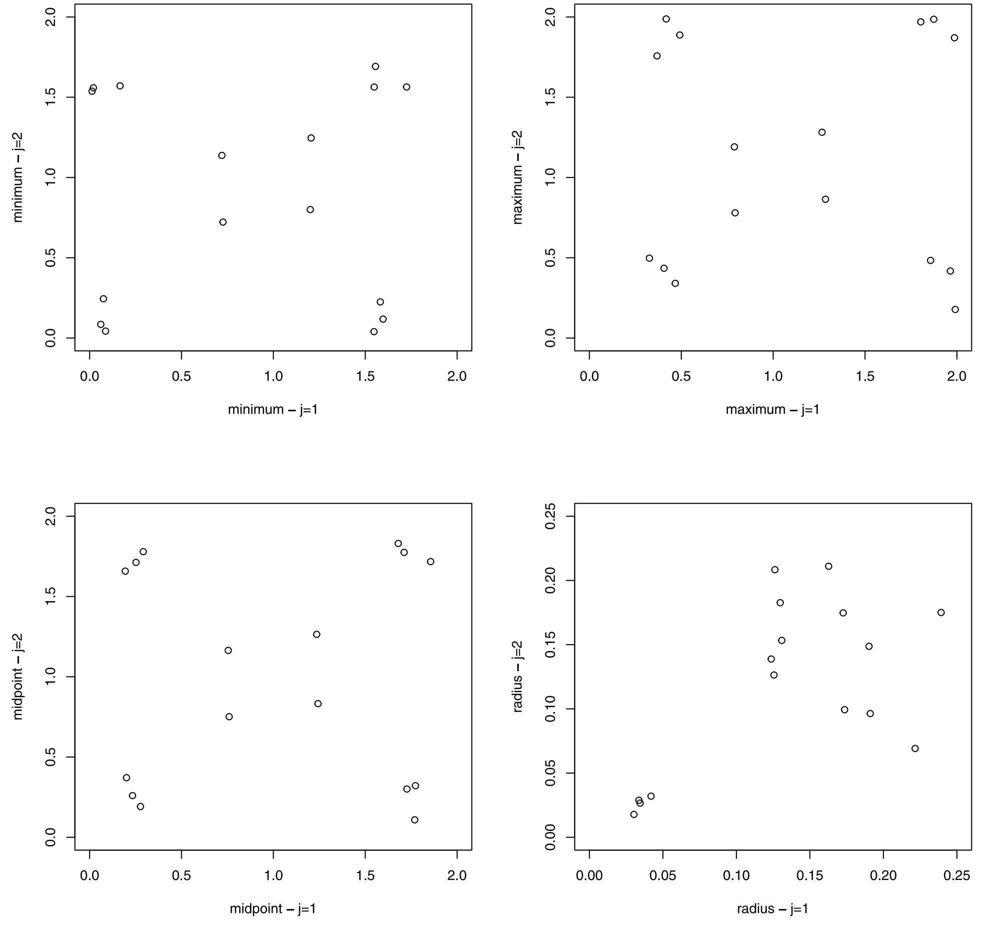

Fig. 1. Data generation process - $C = 4$.

The values of the fuzzy silhouette are higher in the presence of contiguities consistent with the natural clusters (adjacency matrix P_1, Table 4).

As expected, the values of the objective function are higher in the presence of the adjacency matrix P_2 (Table 5).

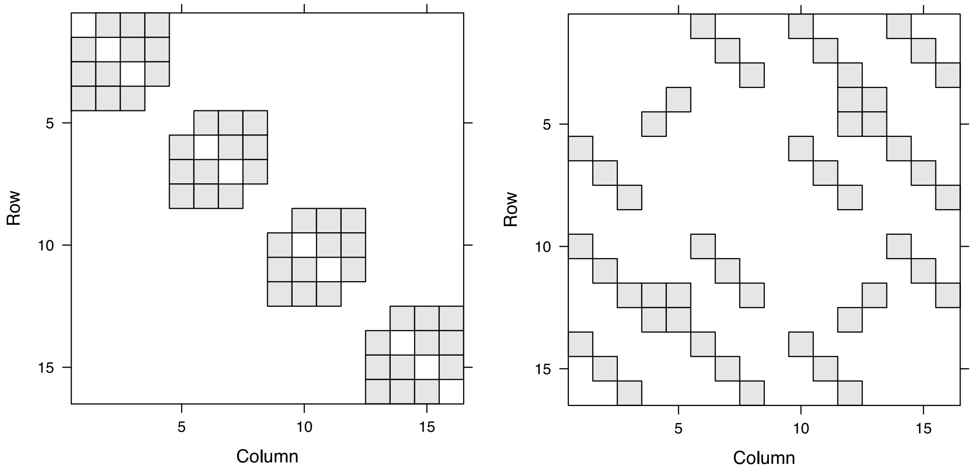

Fig. 2. Contiguity matrices P_1 (left), P_2 (right).

Table 4
Fuzzy silhouette (25).

Adjacency	γ				
	0.0	0.5	1.0	1.5	2.0
P_1	0.42	0.61	0.76	0.76	0.76
P_2	0.42	0.42	−0.13	−0.52	−0.58

Table 5
Values of the objective function (6).

Adjacency	γ				
	0.0	0.5	1.0	1.5	2.0
P_1	−12.91	−4.81	−0.54	0.37	0.66
P_2	−12.91	−3.15	5.76	10.57	15.63

Table 6
Values of the spatial autocorrelation (26).

Adjacency	γ				
	0.0	0.5	1.0	1.5	2.0
P_1	0.74	0.77	0.80	0.82	0.82
P_2	−0.32	−0.32	−0.30	−0.31	−0.29

As expected, the values of the spatial correlation are higher in the presence of the adjacency matrix P_1 (Table 6).

The highest fuzzy membership and the related cluster, alongside with the medoids, are illustrated in Tables 7 and 8. The clusters are (**1**, **2**, **3**, **4**), (**5**, **6**, **7**, **8**), (**9**, **10**, **11**, **12**), (**13**, **14**, **15**, **16**) with spatial constraint P_1; (**1**, 5, **6**, **9**, **14**), (**2**, *3*, **7**, *8*, **10**, *11*, 13, **15**, *16*), (4), (12) with spatial constraint P_2, where the contiguous units in each cluster are given in bold or italic. Units 4, 12 form a single cluster, units 5 and 13 are assigned to the other clusters. We recall that the four "soft" units have radius different from the other units (Fig. 1).

In the presence of the adjacency matrices P_1 the clustering method forces the "soft" units 4, 5, 12, 13 in the natural clusters while increasing the value of the spatial parameter γ. The medoids

Table 7
Highest fuzzy membership - $C = 4$, adjacency matrix P_1.

γ	0.0		0.5		1.0		1.5		2.0	
Unit	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster
1	0.294	1	0.485	1	0.906	1	0.982	1	0.996	1
2	0.298	1	0.489	1	0.910	1	0.982	1	0.996	1
3	0.297	1	0.488	1	**1.000**	1	**1.000**	1	**1.000**	1
4	**1.000**	1	**1.000**	1	0.868	1	0.972	1	0.984	1
5	**1.000**	2	**1.000**	2	0.865	2	0.972	2	0.984	2
6	0.290	2	0.477	2	**1.000**	2	**1.000**	2	**1.000**	2
7	0.290	2	0.476	2	0.906	2	0.982	2	0.996	2
8	0.290	2	0.476	2	0.906	2	0.982	2	0.996	2
9	0.298	3	0.490	3	**1.000**	3	**1.000**	3	**1.000**	3
10	0.297	3	0.489	3	0.910	3	0.982	3	0.996	3
11	0.299	3	0.492	3	0.912	3	0.983	3	0.996	3
12	**1.000**	3	**1.000**	3	0.870	3	0.973	3	0.984	3
13	**1.000**	4	**1.000**	4	0.869	4	0.972	4	0.984	4
14	0.296	4	0.484	4	0.918	4	0.984	4	0.997	4
15	0.291	4	0.479	4	**1.000**	4	**1.000**	4	**1.000**	4
16	0.290	4	0.478	4	0.911	4	0.982	4	0.996	4

Table 8
Highest fuzzy membership - $C = 4$, adjacency matrix P_2.

γ	0.0		0.5		1.0		1.5		2.0	
Unit	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster	Highest membership	Cluster
1	0.294	1	0.485	1	0.906	1	0.982	1	0.996	1
1	0.294	2	0.289	3	**1.000**	1	0.963	1	0.993	1
2	0.298	2	0.291	3	0.272	4	0.959	3	0.992	2
3	0.297	2	0.291	3	0.271	4	0.945	4	0.992	2
4	**1.000**	2	**1.000**	3	0.308	4	**1.000**	2	**1.000**	4
5	**1.000**	3	**1.000**	1	**1.000**	4	**1.000**	1	**1.000**	1
6	0.290	3	0.287	1	0.667	1	0.972	1	0.994	1
7	0.290	3	0.286	1	0.311	4	0.959	3	0.990	2
8	0.290	3	0.285	1	0.307	4	0.921	4	0.990	2
9	0.298	4	0.290	4	0.538	1	0.964	1	0.992	1
10	0.297	4	0.291	4	0.324	2	0.965	3	0.992	2
11	0.299	4	0.293	4	0.325	2	0.954	4	0.993	2
12	**1.000**	4	**1.000**	4	**1.000**	2	0.324	3	**1.000**	3
13	**1.000**	1	**1.000**	2	**1.000**	3	**1.000**	3	**1.000**	2
14	0.296	1	0.290	2	0.643	1	0.945	1	0.990	1
15	0.291	1	0.287	2	0.323	3	0.963	3	0.994	2
16	0.290	1	0.286	2	0.326	3	**1.000**	4	0.994	2

are not the "soft" units (as in the case of $\gamma = 0$) and the memberships of the "soft" units to the clusters is lower than the memberships of the other units to the same cluster.

In the presence of the adjacency matrix P_2 the clustering method forces the contiguous units in the same cluster while increasing the value of the spatial parameter γ (Fig. 3)

4. Applications to real data

4.1. Application to environmental data

The atmosphere is composed mainly of two gaseous elements: nitrogen and oxygen. These two elements are joined by a long list of other chemical species, gaseous and non-gaseous, present in smaller or trace amounts. Multiple anthropogenic activities result in the release of a significant amount of chemicals into the atmosphere, some of which are already present in the atmosphere in lower quantities and some of which are completely absent, potentially harmful to humans, flora and fauna. Such substances can be the cause of atmospheric pollution. The problem then arises of defining the ambient air quality of an area and assessing its status in order to be able to quantify whether it is at risk.

Considering a well-defined location at a certain instant in time, the air quality is the composition of the air at that location, and once the distribution in space and time of the concentration in the air of the various pollutants and their limits are known, it is possible to make a judgment on the state of the air at various points in the territory, based on measured or reconstructed data.

To assess whether the state of the air, and thus the possible presence of pollutants, is of concern or not, it is necessary to define concentration limits for the various pollutants and to measure

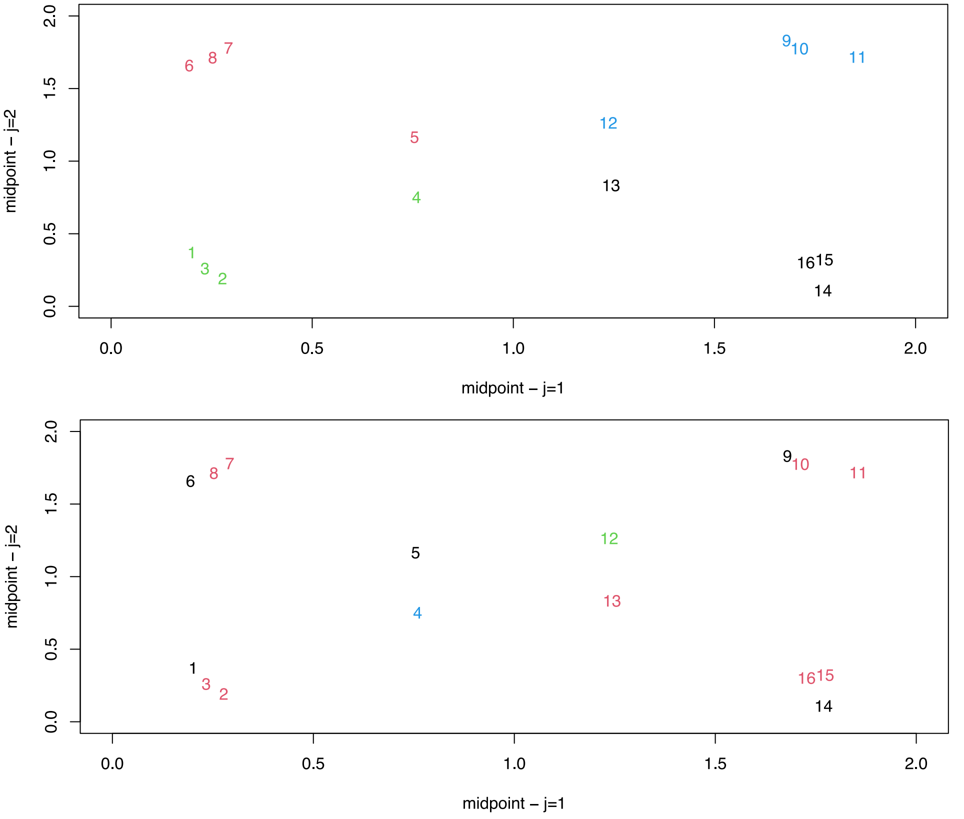

Fig. 3. EFMd-SpID clustering, p=2, adjacency matrix P_1 (top), adjacency matrix P_2, bottom.

the distribution in space and time of the ground concentration of the various pollutant species. Currently available technology allows only to measure with temporal continuity the concentration of some of the pollutants at some well-defined points of the territory, through fixed measuring stations, usually organized in monitoring networks. It is inevitable that in the territory of interest, even if there is a monitoring network composed of numerous measuring points, the state of the air is sampled at a discrete and small number of points, and this implies the need to integrate the measurements of the monitoring network with model reconstructions. The assimilation of measurements into model tools means that these measurements are extended to the entire area under consideration (spatialization) in an objective, repeatable and physically consistent manner. Air quality status monitoring therefore aims to reconstruct the spatial and temporal distribution of the concentration of the pollutants; verify the exceeding of the critical levels provided for the pollutants; identify the areas and/or agglomerations where limits are exceeded; predict future events well in advance, assess their criticality, and define their spatial and temporal consistency.

ARPA Lazio (Regional Environmental Protection Agency) verifies the state of air quality through a specific monitoring system.

Two pollutants are considered in the application. The "Particulate Matter (PM)" refers to a very heterogeneous collection of solid or liquid particles that, due to their small characteristic size, remain suspended in the lower part of the troposphere for longer or shorter periods of time. Its constituent particles vary in size and contain various substances such as: sand, ash, dust, soot, siliceous substances, plant substances, metal compounds, salts, elements such as lead and other heavy metals, inorganic chemical compounds and organic chemical compounds. Since its presence

Table 9
Fuzzy silhouette - EFMd-SpID.

$C \backslash p$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0.47	0.47	0.45	0.46	0.48	0.49	0.47	0.51	0.50	0.50
3	0.48	0.49	0.50	0.50	0.51	0.52	0.43	0.47	0.43	0.52
4	0.42	0.41	0.43	0.42	0.42	0.40	0.40	0.33	0.40	0.40
5	0.36	0.40	0.41	0.37	0.37	0.35	0.35	0.28	0.30	0.20
6	0.36	0.39	0.38	0.30	0.34	0.30	0.31	0.12	0.25	0.22

in the air can adversely affect human health, it is considered a pollutant (regardless of its chemical constitution) and, along with nitrogen dioxide and ozone, is a major source of concern for air quality status. It should be noted that the presence of particulate matter in the atmosphere, especially in urban areas, has now reached alarming levels, especially in winter when periods of limited atmospheric ventilation are more frequent.

One classification of Particulate Matter is based on a descriptive simplification of the particle size that characterizes it. The PM 2.5 is the fine fraction consisting of the subset of particles less than 2.5 μm in diameter. They are mainly derived from emissions from vehicular traffic, industrial activities, power plants. Particles generated in atmospheric chemistry processes go in this class.

The two most important nitrogen oxides from the point of view of air pollution are nitrogen oxide NO and nitrogen dioxide NO2, whose primary origin in the lower layers of the atmosphere is combustion processes and, in urban areas, motor vehicle exhaust and domestic heating. The NO2 dioxide is one of the most important secondary pollutants, i.e., resulting from the reaction of different chemical species in the atmosphere and therefore not emitted directly.

Pollutant concentration values of PM 2.5 and NO2 calculated through modeling tools and monitoring network measurements were downloaded by the ARPA website.[1] The analysis considers the 121 municipalities in the province of Rome. The year 2021 has been considered. The minimum and maximum values of the 2 pollutants over the year are presented in Table 11 with related midpoints and radii. The contiguity matrix among municipalities registers the administrative borders and was downloaded from the National Institute of Statistics (Istat).

The Fuzzy C-Medoids clustering for Spatial Interval-Value Data with entropy regularization (EFMd-SpID) and the Fuzzy C-Medoids clustering for Spatial Interval-Value Data without entropy regularization (FMd-SpID) have been applied to the standardized values of the midpoints and radii of the $J = 2$ variables. To determine the best clustering of the municipalities, the optimal iterative solutions were obtained by solving the EFMd-SpID and FMd-SpID methods with the Lagrangian multipliers method where:

1. the entropy parameter p and the number of clusters C for the EFMd-SpID method and the fuzziness parameter m and the number of clusters C for the FMd-SpID method have been identified by means of the fuzzy cluster validity measure (25), fixing the spatial coefficient γ equal to 0;
2. given the selected values of the parameters (p, C) or (m, C), the value of the spatial penalty coefficient γ has been selected in order to maximize the multivariate spatial autocorrelation (26) considering the contiguity matrix in the methods (6), (16).

The results of the fuzzy silhouette in the methods without penalty term are presented in Tables 9 and 10. The selected values of the parameters are $p = 0.30$, $C = 3$ and $m = 1.5$, $C = 3$. The selected value of the spatial coefficient γ by using the spatial correlation coefficient is 1.0 for the method EFMd-SpID and 0.5 for the method FMd-SpID (Fig. 4). The value of the weight (of the radius) w is 0.36 for EFMd-SpID and 0.35 for the FMd-SpID. The memberships to the 3 clusters and the highest membership cluster are presented in Table 12.

The two partitions are very similar; the value of the Fuzzy Rand Index (Campello, 2007) between the two partitions is 0.76. The comments refer to the EFMd-SpID clustering method. The clustering

[1] ARPA Lazio

Table 10
Fuzzy silhouette - FMd-SpID.

$C\backslash m$	1.0	1.5	1.7
2	0.46	0.48	0.49
3	0.48	0.52	0.52
4	0.37	0.44	0.46
5	0.37	0.41	0.43
6	0.36	0.40	0.30

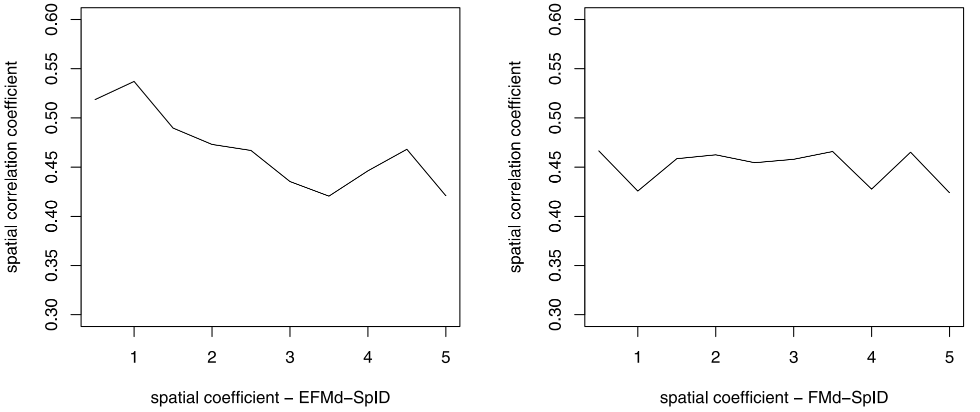

Fig. 4. Multivariate spatial autocorrelation for different values of the spatial coefficient (26).

results of the EFMd-SpID are presented in Fig. 7; for comparison, the results without spatial penalty term ($\gamma = 0$) are presented in Fig. 8.

The medoids are Fiano Romano, Sambuci and Sant'Oreste. Fiano Romano has orographic type Inner Hill and altitude 97 meters asl. It is located north of Rome. It is charaterized by high midpoints of PM 2.5 and NO2, and high radii, in particular NO2. Members of this cluster are Rome and municipalites close to Rome, in the north as Monte Libretti, Monte Rotondo and in the south as Frascati, Monte Compatri, Monte Porzio Catone, Rocca di Papa, Rocca Priora, San Cesareo, and Fiumicino.

Sant'Oreste has orographic type Inner Hill and altitude 420 meters asl. It is located north of Rome. It is charaterized by middle midpoints of PM 2.5 and NO2, and middle radii. Members of this cluster are municipalites located north of Rome as Morlupo, Torrita Tiberina.

Sambuci has orographic type Mountain Interior and Altitude 434 meters asl. It is located west of Rome toward the interior of the peninsula. It is charaterized by low midpoints of PM 2.5 and NO2, and low radii. Members of this cluster are municipalities located close to Abruzzo, in the inner part of the peninsula, as Mandela, Rocca Santo Stefano, Roccagiovine, Rolate.

The municipality Civitavecchia is "fuzzy" according to the membership treshold 0.6 (Maharaj and D'Urso, 2011). Its membership is shared between the high pollutant cluster with medoid Fiano Romano mainly for the high NO2 radius and the low pollutant cluster with medoid Sambuci mainly for the low value of the PM 2.5 midpoint. Civitavecchia is located on the sea. Genazzano also shows shared membership between the same two clusters, the cluster with medoid Fiano Romano mainly for the high pollutants midpoints and the low pollutant cluster with medoid Sambuci mainly for the low values of the radii. It is shown in the ternary plot (Fig. 5).

The spatial constraint reinforces in the same cluster municipalities similar with respect to pollutant concentration and contiguous, whereas splits into different clusters municipalities similar but not contiguous (Fig. 6).

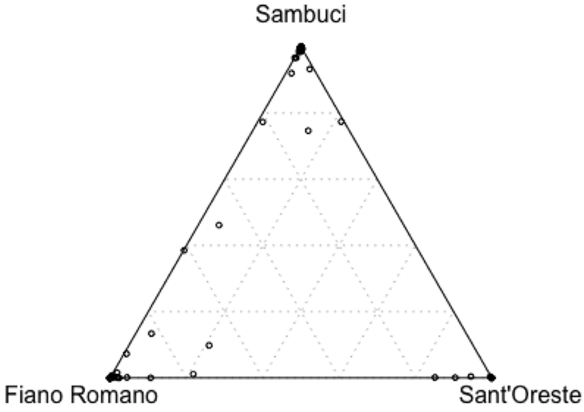

Fig. 5. Ternary plot EFMd-SpID.

4.2. Application to the european fact-checking network

This application considers a looser definition of contiguity, based on social media interactions. We select as units 45 European fact checking outlets and their presence on Twitter from December 16th 2022 to January 27th 2023. Here, as is standard for the application of the method, we require two different kinds of data. The interval-value data $\mathbf{X}$ are the average number of impressions received by their tweets posted each day of the selected time interval on a logarithmic scale. On the other hand the contiguity (adjacency) matrix $\mathbf{A}$ measures whether or not in the same timeframe there has been some user that has either replied to or retweeted at least one tweet from both unit i and i'.

The data are presented in Table 15. The adjacency matrix and the network structure are presented in Fig. 9.

The Fuzzy C-Medoids clustering for Spatial Interval-Value Data with entropy regularization (EFMd-SpID) and the Fuzzy C-Medoids clustering for Spatial Interval-Value Data without entropy regularization (FMd-SpID) algorithm have been applied.

The results of the fuzzy silhouette in the methods without penalty term are presented in Tables 13 and 14.

The selected values of the parameters by using the fuzzy silhouette FS are $p = 0.50$, $C = 3$ and $m = 1.7$, $C = 3$. The selected value of the spatial coefficient γ by using the spatial correlation coefficient FM is 0.5 for the method EFMd-SpID and 1.5 for the method FMd-SpID (Fig. 10).

The size of the clusters is 9, 8, 28, and 6, 8, 31 respectively for the methods EFMd-SpID and FMd-SpID. The memberships to the 3 clusters and the highest membership cluster are presented in Table 16. The two partitions are very similar; the value of the Fuzzy Rand Index (Campello, 2007) between the two partitions is 0.75. The medoids are vrtnws, ScienceFeedback and ellinikahoaxes for the method EFMd-SpID; vrtnws, ScienceFeedback and lavoceinfo for the method FMd-SpID. What we observe is that the 3 groups found by the algorithm can be interpreted as follows: (1) Popular fact-checkers with isolated audiences, (2) Fact-checkers with little to no outreach on Twitter, (3) Most popular fact-checkers which have a shared audience. There is one "fuzzy" fact-checker according to the membership threshold 0.6 (Maharaj and D'Urso, 2011): Faktabaari with the highest membership in cluster 1 shared with cluster 3 (Fig. 11).

The partitions are visualized in Fig. 12. The ternary plots in Fig. 11 show the "fuzzy" fact-checkers. From the first plot in Fig. 12 we observe that the EFMd-SpID classifies units 4 and 44 in cluster 1 despite the fact that they are closer to vrtnews, the medoid of cluster 3, and further from ellinikahoaxes, the medoid of cluster 1, than units 9, 24, 30, and 39, which are classified into cluster 3. This is because 4 and 44 are isolated in the contiguity network, while 9, 24, 30, and 39

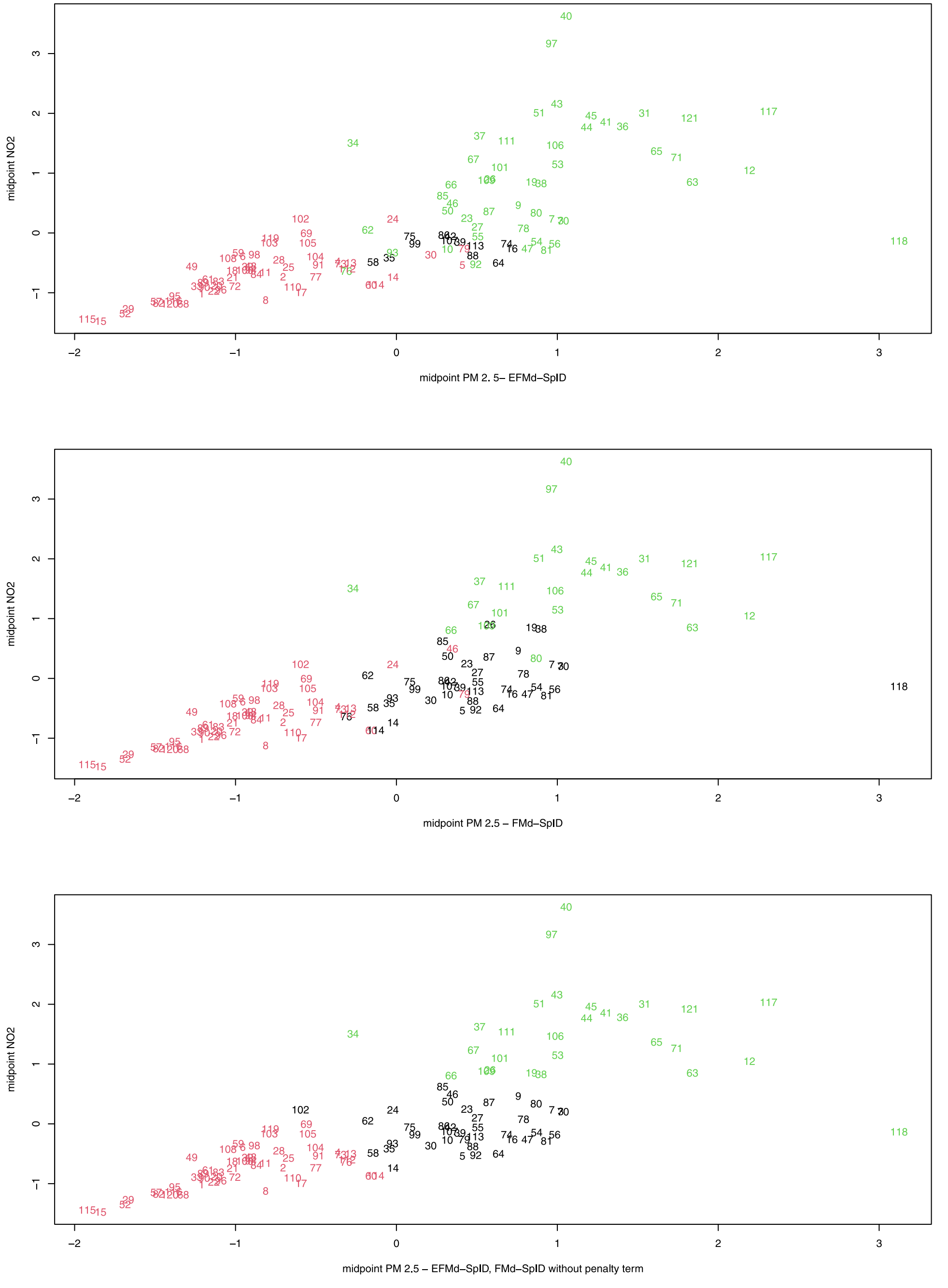

Fig. 6. Clustering EFMd-SpID, FMd-SpID.

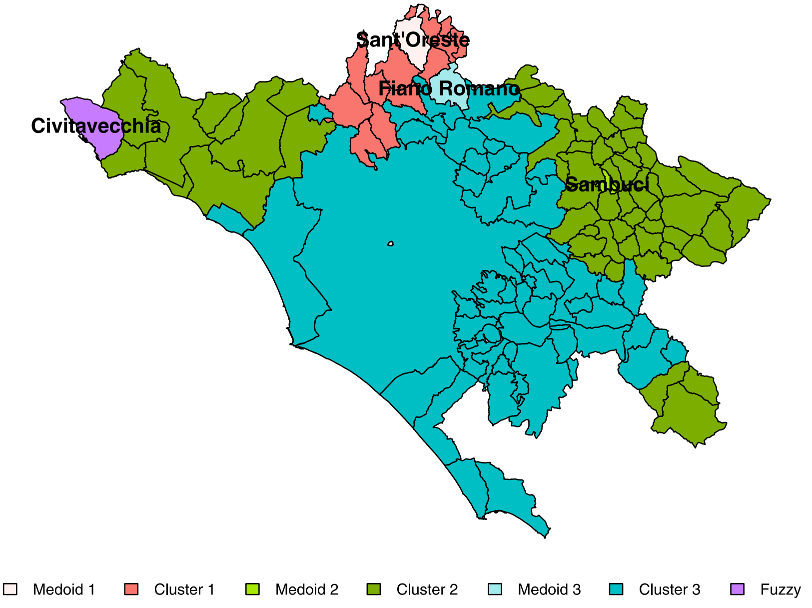

Fig. 7. Clustering EFMd-SpID with spatial penalty term.

are all contiguous to units in cluster 3. We also observe (see Fig. 9) that all the units in the largest connected component are in cluster 3 except for 11. This is because the attributes of unit 11 are so far from those of the medoid of cluster 3 that the presence of the spatial penalty is not enough to change its classification.

4.3. Application to real estate quotations

The real estate quotations identify, for each delimited homogeneous territorial zone (OMI "Osservatorio Mercato Immobiliare" zone), a minimum/maximum range – per unit of area (euros per square meter) – of market and rental values, by property type and state of maintenance and preservation. OMI quotations, available in a six-month period, are related to the municipalities surveyed in the land records. A pivotal process in the activities of the Real Estate Market Observatory is the definition of OMI zones, i.e., homogeneous market areas within the perimeter of which historical data on real estate units bought and sold and rented are recorded. Specifically, the OMI zone is defined as a continuous portion of the municipal territory that reflects a homogeneous compartment of the local real estate market, in which there is uniformity of appreciation for economic and socio-environmental conditions. The decision to establish a territorial segmentation ex ante stems from the fact that the heterogeneity of assets often tends to be expressed by territorial clusters. Particularly in urban areas, the "positional factor" is known to be the leading factor in explaining the prices of real estate units. The basic assumption of the OMI is, therefore, that this factor is the most explanatory of the differences in value between various real estate units, particularly for those with residential use. This implies that there may be a significant reduction in the variability of the investigated phenomenon within a homogeneous spatial segment, the OMI zone. There is, however, the concrete empirical difficulty in objectively drawing the boundaries of a zone in each municipality. The objective of the OMI survey is, as mentioned above, to constitute

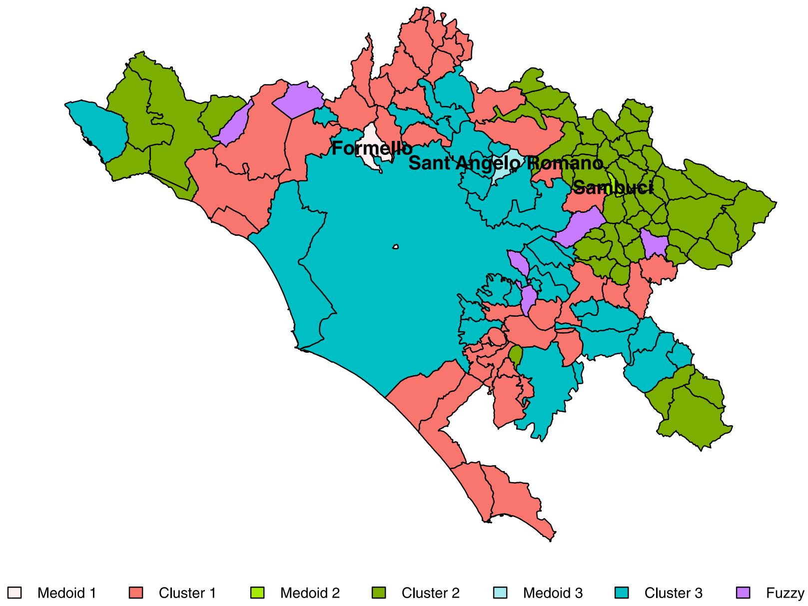

Fig. 8. Clustering EFMd-SpID without spatial penalty term.

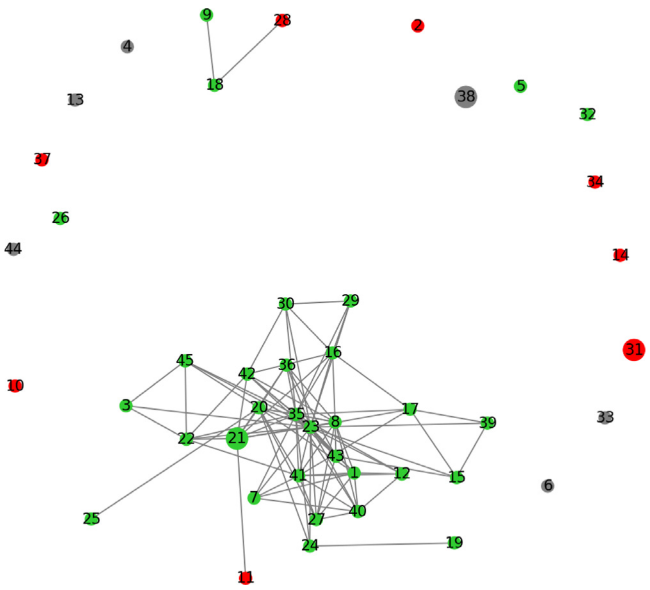

Fig. 9. Network structure of the 45 fact-checkers. The bigger dots represent the medoids.

Table 11
Environmental data of the 121 municipalities: minimum, maximum, midpoint, radius of the variables PM 2.5 and NO2.

	Municipality	*min* PM2.5	*max* PM2.5	*min* NO2	*max* NO2	*mid* PM2.5	*mid* NO2	*radius* PM2.5	*radius* NO2
1	Affile	4.93	6.65	2.57	4.52	5.79	3.55	0.86	0.98
2	Agosta	4.80	9.81	2.52	7.46	7.31	4.99	2.51	2.47
3	Albano Laziale	8.68	16.25	6.33	13.24	12.47	9.78	3.79	3.45
4	Allumiere	6.47	10.18	4.84	7.77	8.32	6.31	1.85	1.46
5	Anguillara Sabazia	7.39	13.89	3.84	8.13	10.64	5.98	3.25	2.14
6	Anticoli Corrado	5.45	7.66	3.09	10.40	6.56	6.74	1.10	3.66
7	Anzio	10.00	14.60	6.49	13.36	12.30	9.92	2.30	3.44
8	Arcinazzo Romano	4.49	9.47	2.20	3.80	6.98	3.00	2.49	0.80
9	Ardea	9.92	13.44	8.66	13.53	11.68	11.10	1.76	2.43
10	Ariccia	7.92	12.80	5.02	9.67	10.36	7.35	2.44	2.33
11	Arsoli	4.73	9.22	2.36	8.39	6.98	5.37	2.25	3.02
12	Artena	6.93	25.02	5.72	22.43	15.97	14.07	9.05	8.35
13	Bellegra	6.11	11.01	4.40	7.95	8.56	6.17	2.45	1.77
14	Bracciano	6.31	12.39	2.78	7.14	9.35	4.96	3.04	2.18
15	Camerata Nuova	3.20	4.63	0.82	1.59	3.91	1.20	0.71	0.39
16	Campagnano di Roma	8.30	14.82	5.30	9.50	11.56	7.40	3.26	2.10
17	Canale Monterano	6.20	9.08	3.14	4.18	7.64	3.66	1.44	0.52
18	Canterano	5.98	6.74	4.43	6.61	6.36	5.52	0.38	1.09
19	Capena	8.90	14.94	5.95	20.20	11.92	13.07	3.02	7.12
20	Capranica Prenestina	5.31	6.82	2.79	5.66	6.07	4.22	0.75	1.43
21	Carpineto Romano	4.13	8.61	1.15	8.74	6.37	4.95	2.24	3.79
22	Casape	5.55	6.49	2.88	4.68	6.02	3.78	0.47	0.90
23	Castel Gandolfo	8.49	12.96	6.33	13.64	10.72	9.98	2.24	3.65
24	Castel Madama	5.46	13.24	3.22	16.62	9.35	9.92	3.89	6.70
25	Castel San Pietro Romano	6.14	8.67	3.62	8.00	7.41	5.81	1.27	2.19
26	Castelnuovo di Porto	8.79	13.52	6.26	20.38	11.15	13.32	2.36	7.06
27	Cave	7.26	14.58	5.24	13.25	10.92	9.25	3.66	4.00
28	Cerreto Laziale	5.50	8.95	3.00	9.85	7.23	6.43	1.73	3.42
29	Cervara di Roma	3.40	5.45	0.94	3.57	4.43	2.25	1.03	1.32
30	Cerveteri	6.47	13.64	3.37	10.36	10.05	6.86	3.58	3.50
31	Ciampino	10.58	17.48	12.72	25.18	14.03	18.95	3.45	6.23
32	Ciciliano	5.31	7.98	2.94	8.77	6.65	5.85	1.33	2.91
33	Cineto Romano	5.25	6.15	2.74	5.62	5.70	4.18	0.45	1.44
34	Civitavecchia	6.72	10.49	4.89	27.91	8.61	16.40	1.88	11.51
35	Civitella San Paolo	8.66	9.90	5.96	7.26	9.28	6.61	0.62	0.65
36	Colleferro	6.88	20.36	7.36	28.27	13.62	17.81	6.74	10.46
37	Colonna	9.89	12.03	14.70	19.31	10.96	17.01	1.07	2.30
38	Fiano Romano	9.06	15.14	6.33	19.57	12.10	12.95	3.04	6.62
39	Filacciano	9.88	11.31	7.14	8.76	10.59	7.95	0.72	0.81
40	Fiumicino	8.74	16.39	4.86	49.59	12.57	27.23	3.82	22.37
41	Fonte Nuova	11.25	15.37	16.08	20.32	13.31	18.20	2.06	2.12
42	Formello	8.15	12.69	5.80	11.10	10.42	8.45	2.27	2.65
43	Frascati	8.85	15.94	9.25	30.23	12.39	19.74	3.55	10.49
44	Gallicano nel Lazio	8.92	16.98	9.10	26.40	12.95	17.75	4.03	8.65
45	Gavignano	10.41	15.65	15.03	22.43	13.03	18.73	2.62	3.70
46	Genazzano	7.09	13.81	6.01	16.49	10.45	11.25	3.36	5.24
47	Genzano di Roma	8.96	14.73	5.76	9.04	11.85	7.40	2.89	1.64
48	Gerano	6.15	7.26	4.52	7.25	6.71	5.89	0.56	1.37
49	Gorga	4.04	7.17	1.16	10.58	5.61	5.87	1.57	4.71
50	Grottaferrata	7.95	12.78	5.15	16.11	10.36	10.63	2.42	5.48
51	Guidonia Montecelio	8.57	15.56	10.66	27.30	12.07	18.98	3.49	8.32
52	Jenne	3.38	5.36	0.86	2.85	4.37	1.85	0.99	1.00
53	Labico	10.70	14.12	10.37	18.79	12.41	14.58	1.71	4.21
54	Ladispoli	9.20	14.84	5.51	10.47	12.02	7.99	2.82	2.48
55	Lanuvio	9.27	12.58	6.30	10.54	10.93	8.42	1.66	2.12
56	Lariano	7.30	17.41	4.54	11.07	12.35	7.81	5.06	3.26
57	Licenza	4.55	5.34	1.86	3.88	4.95	2.87	0.40	1.01
58	Magliano Romano	8.54	9.42	5.20	7.28	8.98	6.24	0.44	1.04
59	Mandela	5.33	7.61	3.39	10.67	6.47	7.03	1.14	3.64
60	Manziana	6.62	11.27	3.08	5.50	8.94	4.29	2.33	1.21
61	Marano Equo	5.26	6.56	3.09	6.47	5.91	4.78	0.65	1.69
62	Marcellina	5.76	12.00	4.16	13.81	8.88	8.99	3.12	4.83
63	Marino	8.95	20.88	7.31	18.83	14.92	13.07	5.96	5.76
64	Mazzano Romano	8.54	14.08	5.20	7.14	11.31	6.17	2.77	0.97
65	Mentana	9.99	18.51	11.11	20.30	14.25	15.71	4.26	4.60
66	Monte Compatri	7.95	12.91	4.31	21.38	10.43	12.84	2.48	8.53
67	Monte Porzio Catone	8.40	13.29	7.45	22.58	10.84	15.01	2.44	7.57
68	Monteflavio	4.47	6.43	1.77	3.60	5.45	2.68	0.98	0.91
69	Montelanico	4.89	10.59	1.74	15.66	7.74	8.70	2.85	6.96
70	Montelibretti	8.20	16.83	5.77	13.76	12.51	9.76	4.32	3.99
71	Monterotondo	10.55	18.70	11.12	19.24	14.63	15.18	4.08	4.06
72	Montorio Romano	4.62	8.20	1.96	6.42	6.41	4.19	1.79	2.23
73	Moricone	5.29	11.46	3.41	8.79	8.38	6.10	3.08	2.69
74	Morlupo	8.92	14.01	6.00	9.62	11.46	7.81	2.54	1.81
75	Nazzano	8.86	10.46	6.14	10.73	9.66	8.44	0.80	2.29
76	Nemi	7.70	9.26	4.85	6.09	8.48	5.47	0.78	0.62
77	Nerola	6.69	9.14	3.83	6.13	7.91	4.98	1.22	1.15
78	Nettuno	8.29	15.26	7.19	11.04	11.77	9.12	3.49	1.92
79	Olevano Romano	6.59	14.76	4.41	10.37	10.67	7.39	4.08	2.98
80	Palestrina	7.28	16.74	5.39	15.49	12.01	10.44	4.73	5.05
81	Palombara Sabina	4.78	19.63	2.18	12.35	12.21	7.26	7.42	5.08
82	Percile	4.64	5.34	2.10	3.39	4.99	2.74	0.35	0.64

(*continued on next page*)

Table 11 (*continued*).

83	Pisoniano	5.40	6.81	2.79	6.40	6.10	4.59	0.71	1.80
84	Poli	5.66	7.96	2.88	7.53	6.81	5.20	1.15	2.32
85	Pomezia	8.79	11.76	7.24	16.56	10.27	11.90	1.48	4.66
86	Ponzano Romano	8.79	11.81	6.14	10.94	10.30	8.54	1.51	2.40
87	Riano	8.87	13.39	7.26	13.86	11.13	10.56	2.26	3.30
88	Rignano Flaminio	8.60	13.07	5.72	7.88	10.83	6.80	2.24	1.08
89	Riofreddo	5.32	6.32	2.75	6.22	5.82	4.49	0.50	1.74
90	Rocca Canterano	5.45	6.23	3.00	5.21	5.84	4.11	0.39	1.10
91	Rocca di Cave	6.50	9.43	4.17	7.84	7.96	6.01	1.47	1.84
92	Rocca di Papa	6.59	15.19	3.37	8.73	10.89	6.05	4.30	2.68
93	Rocca Priora	6.84	11.84	3.86	10.23	9.34	7.05	2.50	3.19
94	Rocca Santo Stefano	6.11	7.30	4.40	6.78	6.70	5.59	0.59	1.19
95	Roccagiovine	4.87	5.71	2.37	4.31	5.29	3.34	0.42	0.97
96	Roiate	5.11	7.19	2.57	5.19	6.15	3.88	1.04	1.31
97	Roma	7.96	16.63	5.28	44.48	12.30	24.88	4.33	19.60
98	Roviano	5.87	7.67	4.28	9.48	6.77	6.88	0.90	2.60
99	Sacrofano	8.15	11.37	5.93	9.68	9.76	7.81	1.61	1.88
100	Sambuci	5.85	7.34	3.67	7.43	6.59	5.55	0.75	1.88
101	San Cesareo	8.01	14.66	6.74	21.90	11.34	14.32	3.32	7.58
102	San Gregorio da Sassola	5.31	9.94	2.88	16.98	7.63	9.93	2.31	7.05
103	San Polo dei Cavalieri	4.69	9.40	1.96	13.81	7.04	7.89	2.36	5.93
104	San Vito Romano	6.32	9.49	5.04	8.37	7.90	6.70	1.58	1.66
105	Santa Marinella	6.40	9.12	4.43	11.32	7.76	7.87	1.36	3.45
106	Sant'Angelo Romano	9.20	15.52	11.16	21.26	12.36	16.21	3.16	5.05
107	Sant'Oreste	8.60	12.20	5.90	10.25	10.40	8.08	1.80	2.17
108	Saracinesco	5.50	7.06	3.11	10.03	6.28	6.57	0.78	3.46
109	Segni	4.94	17.23	1.74	24.71	11.08	13.23	6.15	11.49
110	Subiaco	3.26	11.70	0.84	7.39	7.48	4.11	4.22	3.28
111	Tivoli	7.00	15.93	6.81	26.37	11.46	16.59	4.47	9.78
112	Tolfa	6.28	10.70	3.33	8.02	8.49	5.68	2.21	2.35
113	Torrita Tiberina	10.15	11.60	6.98	8.29	10.88	7.64	0.72	0.66
114	Trevignano Romano	6.66	11.39	3.22	5.40	9.03	4.31	2.37	1.09
115	Vallepietra	3.11	4.22	0.72	2.04	3.66	1.38	0.56	0.66
116	Vallinfreda	5.10	5.42	2.25	3.61	5.26	2.93	0.16	0.68
117	Valmontone	11.21	21.45	11.06	27.14	16.33	19.10	5.12	8.04
118	Velletri	7.09	30.42	4.01	12.08	18.75	8.04	11.67	4.03
119	Vicovaro	5.15	8.98	2.85	13.71	7.07	8.28	1.91	5.43
120	Vivaro Romano	5.12	5.28	2.10	3.25	5.20	2.68	0.08	0.57
121	Zagarolo	10.71	19.02	10.65	26.41	14.86	18.53	4.16	7.88

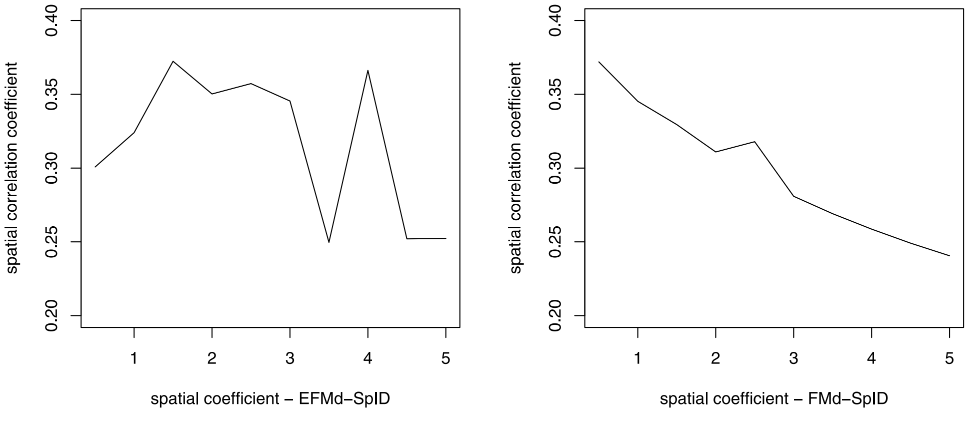

Fig. 10. Clustering EFMd-SpID, FMd-SpID.

a sample that is sufficiently representative of the population of sales/purchases/rentals and that, properly processed, can support the determination of a range of values (min–max) relative to a building type in a homogeneous area of a municipality. A statistical processing procedure is used based on estimating the confidence interval for the price. The confidence interval developed, however, represents an informative state that a special Committee may assume or modify in order to define the range of quotations, depending on any additional information, as well as on the opinion expressed by an Advisory Committee of a purely technical nature.

In the following, the *zone* is a portion of the territorial belt that reflects a homogeneous segment of the local housing market, in which there is substantial uniformity of appreciation for economic and socio-environmental conditions. The *band* is the aggregation of contiguous homogeneous zones,

Table 12
Membership degrees and highest membership cluster - $C = 3$ EFMd-SpID and FMd-SpID. In bold the EFMd-SpID medoids, in italic the EFMd-SpID fuzzy municipalities.

	Municipality	Membership EFMd-SpID			*Cluster*	Membership FMd-SpID			*Cluster*
		Cluster 1	*Cluster* 2	*Cluster* 3		*Cluster* 1	*Cluster* 2	*Cluster* 3	
1	Affile	1.000	0.000	0.000	1	0.999	0.001	0.000	1
2	Agosta	1.000	0.000	0.000	1	0.999	0.001	0.000	1
3	Albano Laziale	0.000	0.000	1.000	3	0.021	0.940	0.039	2
4	Allumiere	0.993	0.007	0.001	1	0.863	0.128	0.009	1
5	Anguillara Sabazia	0.747	0.146	0.107	1	0.099	0.862	0.039	2
6	Anticoli Corrado	1.000	0.000	0.000	1	1.000	0.000	0.000	1
7	Anzio	0.001	0.044	0.955	3	0.006	0.985	0.009	2
8	Arcinazzo Romano	1.000	0.000	0.000	1	0.997	0.003	0.000	1
9	Ardea	0.000	0.000	1.000	3	0.017	0.957	0.026	2
10	Ariccia	0.000	0.000	1.000	3	0.000	1.000	0.000	2
11	Arsoli	1.000	0.000	0.000	1	1.000	0.000	0.000	1
12	Artena	0.000	0.000	1.000	3	0.054	0.204	0.742	3
13	Bellegra	1.000	0.000	0.000	1	0.989	0.009	0.002	1
14	Bracciano	1.000	0.000	0.000	1	0.370	0.605	0.025	2
15	Camerata Nuova	1.000	0.000	0.000	1	0.971	0.025	0.004	1
16	Campagnano di Roma	0.001	0.999	0.001	2	0.011	0.980	0.008	2
17	Canale Monterano	0.995	0.005	0.000	1	0.991	0.009	0.001	1
18	Canterano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
19	Capena	0.000	0.022	0.978	3	0.029	0.881	0.091	2
20	Capranica Prenestina	1.000	0.000	0.000	1	0.995	0.004	0.001	1
21	Carpineto Romano	0.995	0.005	0.000	1	0.994	0.005	0.001	1
22	Casape	1.000	0.000	0.000	1	0.998	0.002	0.000	1
23	Castel Gandolfo	0.000	0.000	1.000	3	0.061	0.867	0.072	2
24	Castel Madama	0.991	0.001	0.008	1	0.652	0.241	0.108	1
25	Castel San Pietro Romano	0.989	0.001	0.011	1	0.880	0.103	0.017	1
26	Castelnuovo di Porto	0.000	0.106	0.894	3	0.027	0.899	0.074	2
27	Cave	0.073	0.007	0.921	3	0.291	0.504	0.206	2
28	Cerreto Laziale	1.000	0.000	0.000	1	1.000	0.000	0.000	1
29	Cervara di Roma	1.000	0.000	0.000	1	0.990	0.008	0.002	1
30	Cerveteri	0.984	0.003	0.012	1	0.240	0.705	0.055	2
31	Ciampino	0.000	0.000	1.000	3	0.004	0.017	0.979	3
32	Ciciliano	1.000	0.000	0.000	1	0.999	0.001	0.000	1
33	Cineto Romano	1.000	0.000	0.000	1	0.999	0.001	0.000	1
34	*Civitavecchia*	0.462	0.054	*0.484*	3	0.204	0.286	0.510	3
35	Civitella San Paolo	0.000	0.999	0.001	2	0.008	0.991	0.002	2
36	Colleferro	0.000	0.000	1.000	3	0.005	0.015	0.980	3
37	Colonna	0.000	0.001	0.999	3	0.043	0.157	0.800	3
38	**Fiano Romano**	0.000	0.000	**1.000**	3	0.025	0.872	0.103	2
39	Filacciano	0.001	0.999	0.000	2	0.002	0.997	0.001	2
40	Fiumicino	0.000	0.000	1.000	3	0.066	0.149	0.785	3
41	Fonte Nuova	0.000	0.000	1.000	3	0.007	0.022	0.971	3
42	Formello	0.004	0.943	0.053	2	0.000	1.000	0.000	2
43	Frascati	0.000	0.000	1.000	3	0.003	0.010	0.988	3
44	Gallicano nel Lazio	0.000	0.000	1.000	3	0.000	0.000	1.000	3
45	Gavignano	0.000	0.004	0.996	3	0.012	0.047	0.942	3
46	Genazzano	0.385	0.001	0.613	3	0.357	0.349	0.294	1
47	Genzano di Roma	0.000	0.001	0.999	3	0.009	0.986	0.005	2
48	Gerano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
49	Gorga	0.997	0.003	0.000	1	0.990	0.009	0.001	1
50	Grottaferrata	0.000	0.000	1.000	3	0.127	0.454	0.420	2

(*continued on next page*)

Table 12 (*continued*).

51	Guidonia Montecelio	0.000	0.000	1.000	3	0.020	0.049	0.931	3
52	Jenne	1.000	0.000	0.000	1	0.982	0.015	0.002	1
53	Labico	0.000	0.004	0.996	3	0.021	0.125	0.854	3
54	Ladispoli	0.098	0.211	0.692	3	0.055	0.899	0.046	2
55	Lanuvio	0.001	0.009	0.989	3	0.005	0.993	0.002	2
56	Lariano	0.000	0.000	1.000	3	0.064	0.850	0.086	2
57	Licenza	1.000	0.000	0.000	1	0.994	0.006	0.001	1
58	Magliano Romano	0.000	1.000	0.000	2	0.011	0.987	0.002	2
59	Mandela	1.000	0.000	0.000	1	1.000	0.000	0.000	1
60	Manziana	0.998	0.002	0.000	1	0.860	0.132	0.008	1
61	Marano Equo	1.000	0.000	0.000	1	1.000	0.000	0.000	1
62	Marcellina	0.134	0.041	0.825	3	0.401	0.489	0.109	2
63	Marino	0.000	0.000	1.000	3	0.044	0.276	0.680	3
64	Mazzano Romano	0.003	0.995	0.003	2	0.006	0.992	0.002	2
65	Mentana	0.000	0.000	1.000	3	0.009	0.035	0.957	3
66	Monte Compatri	0.000	0.000	1.000	3	0.087	0.293	0.621	3
67	Monte Porzio Catone	0.000	0.000	1.000	3	0.029	0.111	0.861	3
68	Monteflavio	1.000	0.000	0.000	1	0.978	0.019	0.003	1
69	Montelanico	0.773	0.013	0.214	1	0.677	0.236	0.086	1
70	Montelibretti	0.002	0.000	0.998	3	0.185	0.622	0.193	2
71	Monterotondo	0.000	0.000	1.000	3	0.049	0.359	0.592	3
72	Montorio Romano	0.999	0.001	0.000	1	0.974	0.023	0.003	1
73	Moricone	0.920	0.015	0.065	1	0.785	0.198	0.017	1
74	Morlupo	0.001	0.851	0.148	2	0.001	0.998	0.001	2
75	Nazzano	0.000	1.000	0.000	2	0.001	0.999	0.000	2
76	Nemi	0.016	0.011	0.974	3	0.096	0.896	0.009	2
77	Nerola	0.933	0.056	0.011	1	0.906	0.091	0.004	1
78	Nettuno	0.011	0.212	0.776	3	0.005	0.991	0.005	2
79	Olevano Romano	0.986	0.007	0.008	1	0.608	0.331	0.061	1
80	Palestrina	0.000	0.000	1.000	3	0.132	0.254	0.614	3
81	Palombara Sabina	0.000	0.000	1.000	3	0.249	0.388	0.363	2
82	Percile	1.000	0.000	0.000	1	0.993	0.006	0.001	1
83	Pisoniano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
84	Poli	1.000	0.000	0.000	1	1.000	0.000	0.000	1
85	Pomezia	0.002	0.018	0.980	3	0.054	0.865	0.081	2
86	Ponzano Romano	0.000	1.000	0.000	2	0.000	1.000	0.000	2
87	Riano	0.000	0.024	0.976	3	0.040	0.900	0.060	2
88	Rignano Flaminio	0.000	0.999	0.001	2	0.001	0.999	0.000	2
89	Riofreddo	1.000	0.000	0.000	1	0.999	0.001	0.000	1
90	Rocca Canterano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
91	Rocca di Cave	0.967	0.004	0.029	1	0.834	0.149	0.016	1
92	Rocca di Papa	0.000	0.000	1.000	3	0.071	0.875	0.054	2
93	Rocca Priora	0.000	0.000	1.000	3	0.191	0.729	0.080	2
94	Rocca Santo Stefano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
95	Roccagiovine	1.000	0.000	0.000	1	0.995	0.004	0.001	1
96	Roiate	1.000	0.000	0.000	1	0.997	0.003	0.000	1
97	Roma	0.000	0.000	1.000	3	0.098	0.217	0.684	3
98	Roviano	1.000	0.000	0.000	1	1.000	0.000	0.000	1
99	Sacrofano	0.001	0.904	0.095	2	0.013	0.982	0.005	2
100	**Sambuci**	**1.000**	0.000	0.000	1	0.999	0.001	0.000	1
101	San Cesareo	0.000	0.000	1.000	3	0.032	0.135	0.833	3
102	San Gregorio da Sassola	0.999	0.000	0.001	1	0.806	0.132	0.062	1
103	San Polo dei Cavalieri	0.967	0.000	0.033	1	0.817	0.135	0.049	1
104	San Vito Romano	1.000	0.000	0.000	1	0.968	0.028	0.004	1
105	Santa Marinella	0.999	0.001	0.000	1	0.764	0.216	0.020	1
106	Sant'Angelo Romano	0.000	0.000	1.000	3	0.005	0.018	0.978	3
107	**Sant'Oreste**	0.000	**1.000**	0.000	2	0.000	1.000	0.000	2
108	Saracinesco	1.000	0.000	0.000	1	1.000	0.000	0.000	1
109	Segni	0.000	0.000	1.000	3	0.045	0.113	0.842	3
110	Subiaco	1.000	0.000	0.000	1	0.994	0.004	0.001	1

(*continued on next page*)

Table 12 (*continued*).

111	Tivoli	0.007	0.000	0.993	3	0.109	0.177	0.714	3
112	Tolfa	1.000	0.000	0.000	1	0.859	0.128	0.014	1
113	Torrita Tiberina	0.003	0.994	0.002	2	0.004	0.995	0.001	2
114	Trevignano Romano	0.774	0.219	0.007	1	0.185	0.799	0.017	2
115	Vallepietra	1.000	0.000	0.000	1	0.968	0.027	0.005	1
116	Vallinfreda	1.000	0.000	0.000	1	0.995	0.004	0.001	1
117	Valmontone	0.000	0.000	1.000	3	0.013	0.043	0.944	3
118	Velletri	0.000	0.000	1.000	3	0.123	0.497	0.380	2
119	Vicovaro	1.000	0.000	0.000	1	0.951	0.038	0.012	1
120	Vivaro Romano	0.995	0.005	0.000	1	0.986	0.013	0.001	1
121	Zagarolo	0.000	0.000	1.000	3	0.003	0.010	0.988	3

Table 13
Fuzzy silhouette - EFMd-SpID.

$C\backslash p$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0.47	0.48	0.48	0.50	0.50	0.50	0.51	0.51	0.51	0.51
3	0.48	0.49	0.49	0.51	0.51	0.51	0.51	0.51	0.51	0.52
4	0.45	0.46	0.46	0.36	0.47	0.38	0.39	0.46	0.37	0.38
5	0.35	0.40	0.41	0.41	0.42	0.29	0.30	0.42	0.43	0.44
6	0.33	0.38	0.34	0.38	0.36	0.36	0.33	0.36	0.38	0.38

Table 14
Fuzzy silhouette - FMd-SpID.

$C\backslash m$	1.0	1.5	1.7
2	0.46	0.47	0.47
3	0.46	0.51	0.51
4	0.44	0.48	0.33
5	0.40	0.37	0.14
6	0.33	0.27	0.16

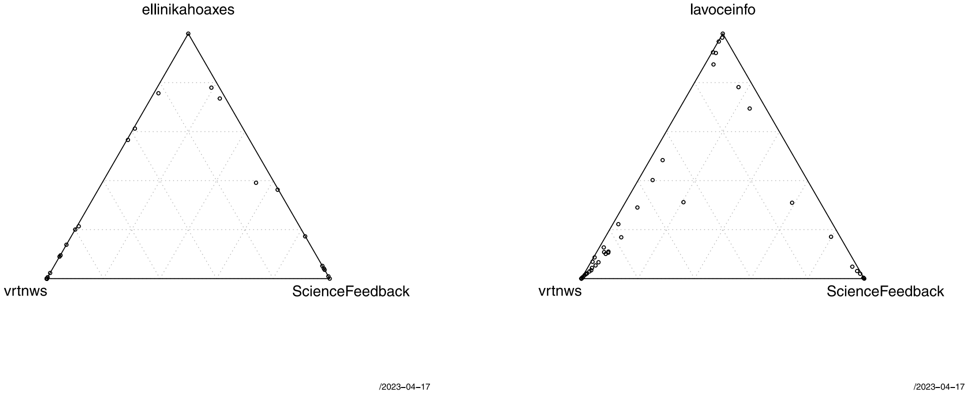

Fig. 11. Ternary plots, EFMd-SpID (left), FMd-SpID (right).

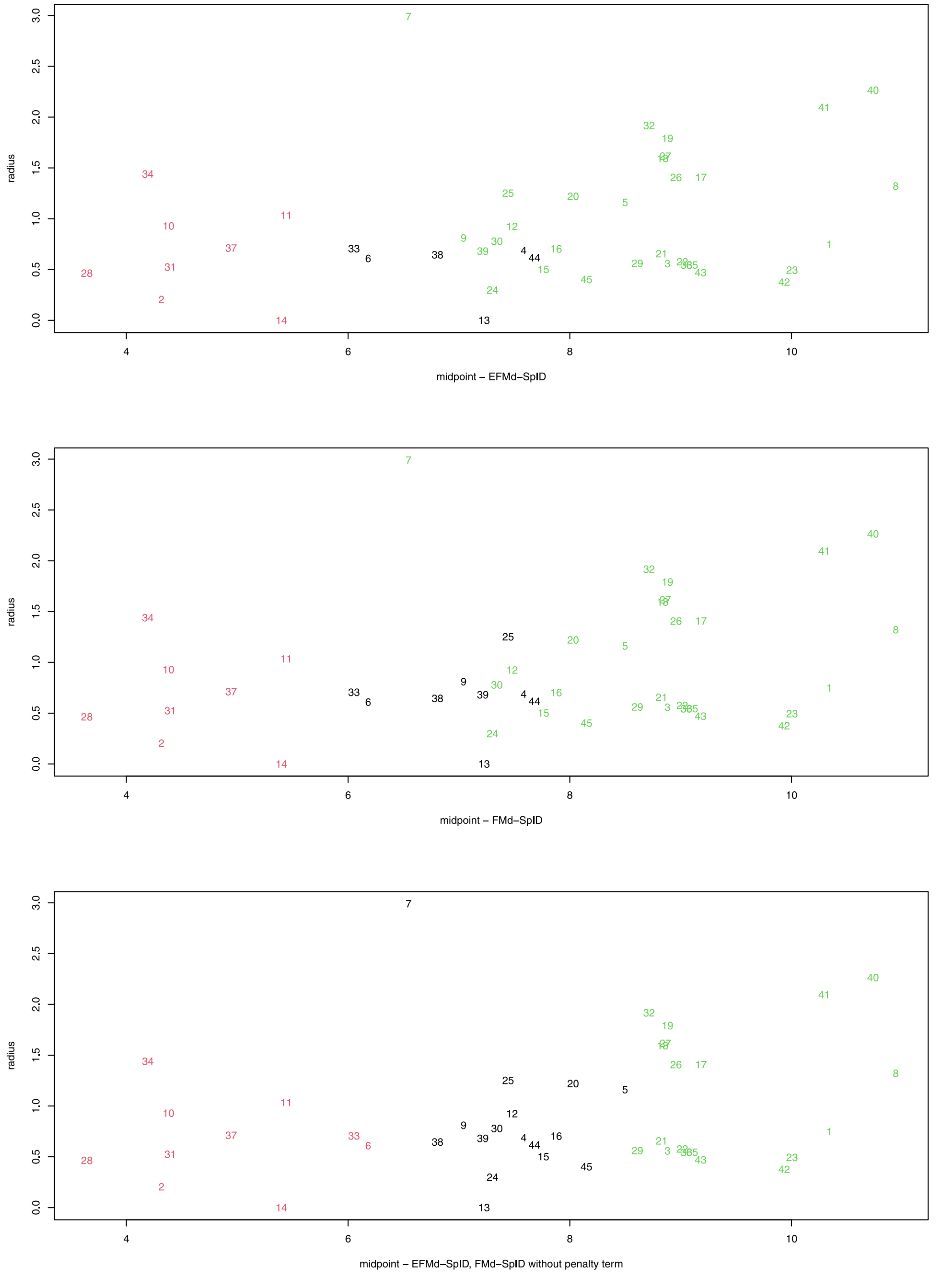

Fig. 12. Clustering EFMd-SpID, FMd-SpID.

Table 15
Interval-value average number of impressions received by the tweets of 45 fact-checkers (logarithmic scale).

	Fact-checker	Minimum	Maximum	Midpoint	Radius
1	AfpFactuel	9.59	11.09	10.34	0.75
2	AEJBulgaria	4.11	4.52	4.32	0.21
3	Correctiv_org	8.32	9.44	8.88	0.56
4	FactCheckNI	6.90	8.27	7.58	0.69
5	Faktisk_no	7.34	9.66	8.50	1.16
6	FaktografHR	5.57	6.79	6.18	0.61
7	Observateurs	3.56	9.54	6.55	2.99
8	FullFact	9.62	12.26	10.94	1.32
9	istinomer	6.23	7.85	7.04	0.81
10	Istinomjer	3.45	5.31	4.38	0.93
11	JanJagers	4.41	6.48	5.44	1.04
12	LesSurligneurs	6.55	8.40	7.48	0.92
13	mdfgeo	7.23	7.23	7.23	0.00
14	med_transparent	5.40	5.40	5.40	0.00
15	Open_gol	7.26	8.26	7.76	0.50
16	Publico	7.17	8.58	7.88	0.70
17	PagellaPolitica	7.78	10.59	9.18	1.41
18	Raskrikavanje	7.24	10.43	8.84	1.59
19	rebaltica	7.09	10.67	8.88	1.79
20	StopFakingNews	6.81	9.25	8.03	1.22
21	vrtnws	8.17	9.48	8.83	0.66
22	dpa	8.44	9.59	9.01	0.58
23	franceinfo	9.51	10.50	10.00	0.49
24	DelfiLV	7.00	7.60	7.30	0.30
25	DemagogPL	6.19	8.69	7.44	1.25
26	DemagogCZ	7.55	10.36	8.96	1.41
27	FerretScot	7.25	10.48	8.86	1.62
28	vistinomer	3.18	4.11	3.64	0.46
29	Newtral	8.05	9.17	8.61	0.56
30	observadorpt	6.56	8.12	7.34	0.78
31	ScienceFeedback	3.87	4.92	4.40	0.52
32	TjekDet	6.80	10.63	8.71	1.92
33	voxukraine	5.35	6.76	6.05	0.71
34	HibridInfo	2.75	5.63	4.19	1.44
35	20Minutes	8.56	9.65	9.10	0.54
36	dagensnyheter	8.51	9.59	9.05	0.54
37	DELFILietuva	4.23	5.66	4.94	0.71
38	ellinikahoaxes	6.16	7.45	6.81	0.65
39	lavoceinfo	6.53	7.90	7.21	0.68
40	decodeurs	8.47	13.00	10.74	2.26
41	CheckNewsfr	8.20	12.39	10.29	2.10
42	NUnl	9.56	10.31	9.93	0.38
43	thejournal_ie	8.71	9.65	9.18	0.47
44	Faktabaari	7.06	8.30	7.68	0.62
45	ZDDK_	7.75	8.55	8.15	0.40

representing a territorial area with precise geographic location in the municipality and reflecting an established urban location. The municipal territory is divided into the following bands: Central, Semi-central, Peripheral, Suburban, Extra-urban (B, C, D, E, R, respectively).

The data and the contiguity matrix among 211 zones with residential use destination were downloaded by the "Agenzia Entrate - OMI" website[2] (Fig. 13 and Table 18 in Appendix).

The Entropy-based Fuzzy C-Medoids Clustering for Spatial Interval-Value Data (EFMd-SpID) algorithm has been applied to the values of the midpoints and radius of the variable representing the interval of values of the price (thousands of euros per square meter).

[2] https://www1.agenziaentrate.gov.it/servizi/Consultazione/ricerca.htm

Table 16
Membership degrees and highest membership cluster - $C = 3$ EFMd-SpID and FMd-SpID. In bold the EFMd-SpID medoids, in italic the EFMd-SpID fuzzy fact-checkers.

	Fact-checker	Membership EFMd-SpID			*Cluster*	Membership FMd-SpID			*Cluster*
		Cluster 1	*Cluster* 2	*Cluster* 3		*Cluster* 1	*Cluster* 2	*Cluster* 3	
1	AfpFactuel	0.000	0.000	1.000	3	0.044	0.016	0.941	3
2	AEJBulgaria	0.041	0.959	0.000	2	0.002	0.998	0.001	2
3	Correctiv_org	0.000	0.000	1.000	3	0.000	0.000	1.000	3
4	FactCheckNI	0.613	0.005	0.382	1	0.968	0.002	0.030	1
5	Faktisk_no	0.200	0.000	0.800	3	0.087	0.004	0.910	3
6	FaktografHR	0.780	0.192	0.029	1	0.782	0.165	0.054	1
7	Observateurs	0.000	0.000	1.000	3	0.312	0.205	0.483	3
8	FullFact	0.000	0.000	1.000	3	0.054	0.023	0.923	3
9	istinomer	0.214	0.006	0.780	3	0.874	0.030	0.096	1
10	Istinomjer	0.052	0.948	0.000	2	0.004	0.995	0.001	2
11	JanJagers	0.391	0.544	0.065	2	0.310	0.590	0.100	2
12	LesSurligneurs	0.000	0.000	1.000	3	0.101	0.036	0.863	3
13	mdfgeo	0.757	0.016	0.227	1	0.921	0.015	0.064	1
14	med_transparent	0.363	0.635	0.003	2	0.171	0.797	0.032	2
15	Open_gol	0.000	0.000	1.000	3	0.290	0.053	0.657	3
16	Publico	0.000	0.000	1.000	3	0.021	0.009	0.969	3
17	PagellaPolitica	0.000	0.000	1.000	3	0.108	0.026	0.865	3
18	Raskrikavanje	0.023	0.000	0.977	3	0.402	0.051	0.547	3
19	rebaltica	0.006	0.000	0.994	3	0.127	0.016	0.857	3
20	StopFakingNews	0.000	0.000	1.000	3	0.030	0.014	0.957	3
21	**vrtnws**	0.000	0.000	**1.000**	3	0.000	0.000	1.000	3
22	dpa	0.000	0.000	1.000	3	0.001	0.000	0.999	3
23	franceinfo	0.000	0.000	1.000	3	0.034	0.018	0.948	3
24	DelfiLV	0.000	0.000	1.000	3	0.169	0.057	0.775	3
25	DemagogPL	0.095	0.001	0.904	3	0.484	0.046	0.471	1
26	DemagogCZ	0.090	0.000	0.910	3	0.069	0.005	0.926	3
27	FerretScot	0.000	0.000	1.000	3	0.020	0.007	0.973	3
28	vistinomer	0.009	0.991	0.000	2	0.049	0.934	0.018	2
29	Newtral	0.000	0.000	1.000	3	0.002	0.001	0.997	3
30	observadorpt	0.000	0.000	1.000	3	0.106	0.042	0.852	3
31	**ScienceFeedback**	0.000	**1.000**	0.000	2	0.000	1.000	0.000	2
32	TjekDet	0.138	0.000	0.862	3	0.222	0.020	0.758	3
33	voxukraine	0.735	0.244	0.021	1	0.694	0.248	0.058	1
34	HibridInfo	0.036	0.964	0.000	2	0.031	0.959	0.010	2
35	20Minutes	0.000	0.000	1.000	3	0.005	0.003	0.992	3
36	dagensnyheter	0.000	0.000	1.000	3	0.001	0.000	0.999	3
37	DELFILietuva	0.173	0.827	0.001	2	0.020	0.976	0.004	2
38	**ellinikahoaxes**	**1.000**	0.000	0.000	1	0.984	0.006	0.010	1
39	lavoceinfo	0.000	0.000	1.000	3	1.000	0.000	0.000	1
40	decodeurs	0.000	0.000	1.000	3	0.110	0.041	0.849	3
41	CheckNewsfr	0.000	0.000	1.000	3	0.067	0.027	0.906	3
42	NUnl	0.000	0.000	1.000	3	0.018	0.007	0.975	3
43	Thejournal_ie	0.000	0.000	1.000	3	0.006	0.002	0.992	3
44	*Faktabaari*	*0.566*	0.004	0.430	1	0.924	0.004	0.073	1
45	ZDDK_	0.000	0.000	1.000	3	0.014	0.004	0.982	3

To determine the best clustering of the municipalities, the optimal iterative solutions were obtained by solving the EFMd-SpID method with the Lagrangian multipliers method.

The results of the fuzzy silhouette in the method without penalty term are presented in Table 17. The selected values of the parameters are $p = 0.15, C = 4$. The selected value of the spatial coefficient γ by using the spatial correlation coefficient is 0.50 (Fig. 15). The value of the weight (of the radius) w is 0.50. The memberships to the 4 clusters and the highest membership cluster are presented in Table 19 in Appendix. The sizes of the clusters are 32, 128, 40, 11, respectively.

The clustering results of the EFMd-SpID are presented as a geographic map in Fig. 14. The partitions obtained, with and without penalty term, are visualized in Fig. 17 (from bottom left to top right the four clusters 4, 3, 2, 1).

Table 17
Fuzzy silhouette - EFMd-SpID.

$C \backslash p$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0.61	0.62	0.62	0.62	0.60	0.59	0.56	0.53	0.50	0.50
3	0.57	0.47	0.41	0.35	0.37	0.33	0.33	0.32	0.27	0.19
4	0.53	0.48	0.63	0.39	0.31	0.28	0.30	0.27	0.27	0.04
5	0.53	0.49	0.40	0.40	0.33	0.31	0.28	0.30	0.25	0.01
6	0.51	0.48	0.42	0.42	0.39	0.35	0.34	0.31	0.23	0.22
7	0.50	0.40	0.25	0.44	0.39	0.31	0.31	0.25	0.25	0.02
8	0.47	0.42	0.26	0.20	0.32	0.30	0.31	0.31	0.24	0.22

The medoids are SALLUSTIANO-CASTRO PRETORIO (PIAZZA INDIPENDENZA) (6), PODERE ROSA (VIA DIEGO FABBRI) (93), TORRENOVA (VIA DELLA TENUTA DI TORRENOVA) (133), ARDEATINA SELVOTTA-SPREGAMORE (VIA DEL FOSSO DELLA SOLFARATA) (174) (Fig. 14). They are in bands B, C, D, E and show – as expected – decreasing midpoints of the price.

As a general comment the ''bands' are identified by the clustering process; though the interval-valued nature of the data and the presence of the contiguity constraint place some zones in 'bands' different from those expected.

The zone ESQUILINO (PIAZZA VITTORIO) (7), in band B, is a member of cluster 1. Table 18 and Fig. 16 show that its interval of price is more similar to that of the medoid in cluster 2 (it is close to zone 157 - Fig. 17); nonetheless its adjacency to the zones SALLUSTIANO-CASTRO PRETORIO (PIAZZA INDIPENDENZA) (6), VIMINALE (VIA TORINO) (10), CELIO (VIA CLAUDIA) (11), MONTI (VIA DEI SERPENTI) (14) places ESQUILINO in cluster 1 (Fig. 16). Similarly the zone 56 APPIO METRONIO (PIAZZA TUSCOLO), in band C, is a member of cluster 1, being adjacent to the zones SAN SABA (PIAZZA G.L. BERNINI) (8) and CELIO (VIA CLAUDIA) (11) in cluster 1, despite presenting a lower interval value of the price.

The zones CAMILLUCCIA (VIA DELLA CAMILLUCCIA) (44) in cluster 2 and MARCO POLO (VIALE MARCO POLO) (35) in cluster 1, both in band C, even showing almost the same midpoint and radius are not adjacent.

The zones numbered from 26 to 31, 33, 46, 49, 51 e 56, in band C, have the same radius of the zones in cluster 1 which they belong to; slightly less similar the zones 29 and 30, showing lower radius and related lower membership to cluster 1 (Fig. 16).

For the same reasons there are some movements of zones between the clusters 3 and 4, with respect to their 'bands'.

The spatial constraint reinforces in the same cluster zones similar with respect to the interval-valued variable representing the price and contiguous, whereas splits into different clusters zones similar but not contiguous.

5. Conclusions

In this paper, two fuzzy clustering methods for spatial interval-valued data are proposed, i.e. the fuzzy C-Medoids clustering of spatial interval-valued data with and without an entropy regularization. The methods are based on the fuzzy Partitioning Around Medoids (PAM) algorithm, inheriting the great advantage of obtaining non-fictitious representative units for each cluster. A very general definition of contiguity is adopted.

The performances of the methods are proved either by simulation or applications to real data. The main evidence related to the influence of spatial contiguity on clustering results is as follows: when contiguity reflects the structure of the clusters, it leads to a membership degree matrix U with higher values than in the case of the absence of spatial information, i.e., it reinforces the membership of each unit in its natural cluster; when contiguity does not replicate natural clusters, a more complex balance between internal cohesion and spatial constraint is required and, however, in both cases, mainly affects fuzzy units.

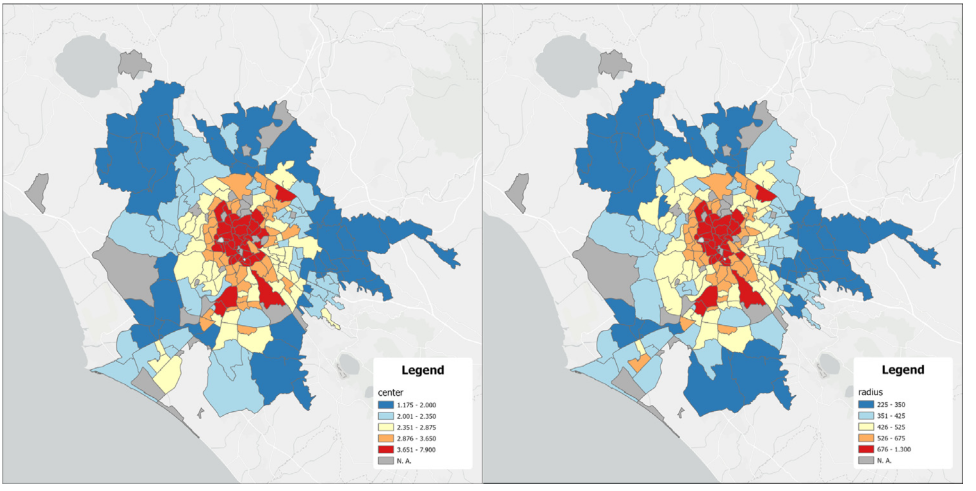

Fig. 13. Center (left) and radius (right) of the 211 OMI zones (quintiles of the distribution).

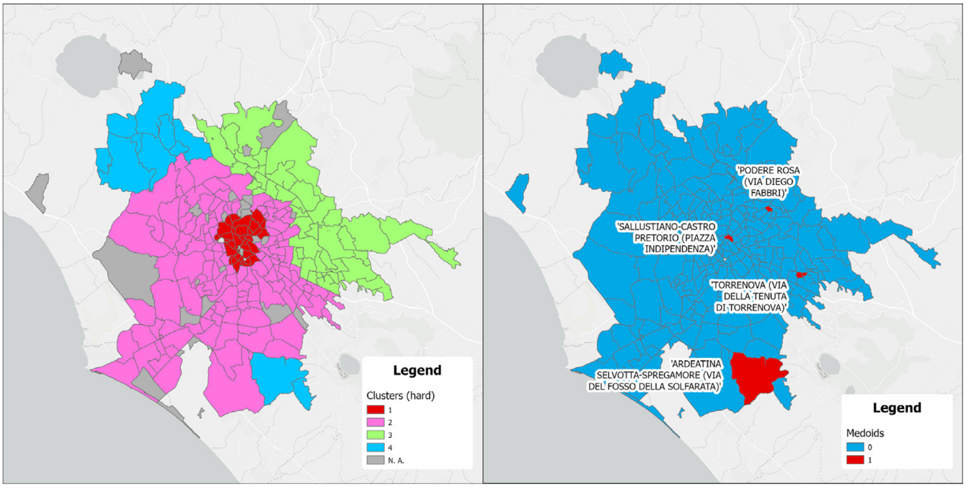

Fig. 14. Clustering EFMd-SpID with spatial penalty term (left partition; right medoids.

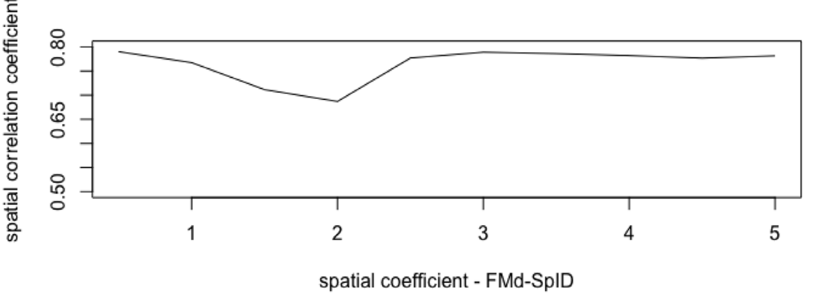

Fig. 15. Multivariate spatial autocorrelation for different values of the spatial coefficient (26).

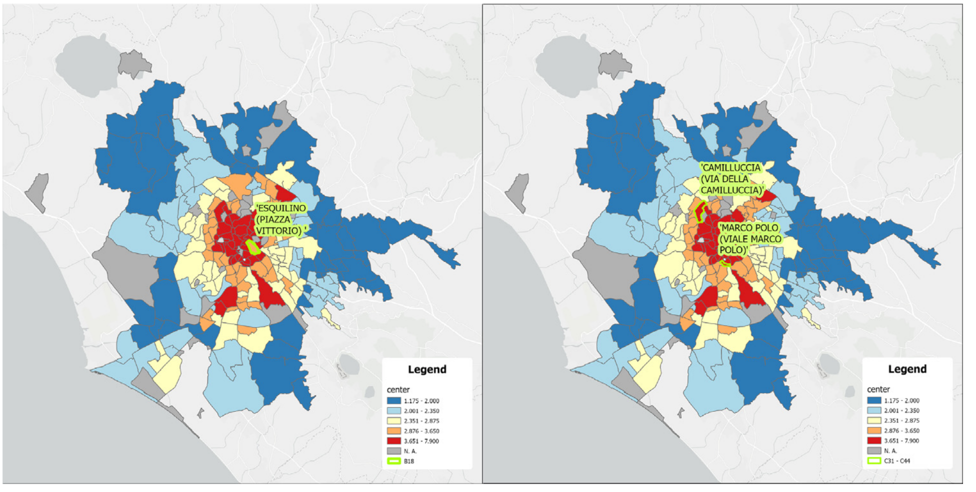

Fig. 16. Clustering - zone SALLUSTIANO (left); zones CAMILLUCCIA and MARCO POLO (right).

In this paper *undirected* contiguity relations have been considered, that is, the matrix **A** is symmetric. Extensions can consider *directed* contiguity matrices, in which this constraint is not necessarily satisfied. Moreover, it is possible to consider *weighted* contiguity matrices, in which the entries are not binary (0 for non-contiguous, 1 for contiguous elements) but are only required to be non-negative. In this case, $a_{ii'}$ represents the strength of the connection between unit i and unit i'.

Further extensions include robust versions, spatial fuzzy clustering methods for other kinds of imprecise data (e.g. histogram-valued data and interval-valued time series) and clustering methods based on other theoretical approaches for managing the uncertainty in the clustering process (e.g. possibilistic approach).

Acknowledgments

The research of Livia De Giovanni and Lorenzo Federico is supported by the Horizon 2020 Framework Programme through the grant Media Futures: Data-driven innovation hub for the media value chain (grant number 951962) and by the European Media and Information Fund through the grant Leveraging argument technology for impartial fact-checking (LATIF, Grant Agreement number 268755).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Tables 18 and 19

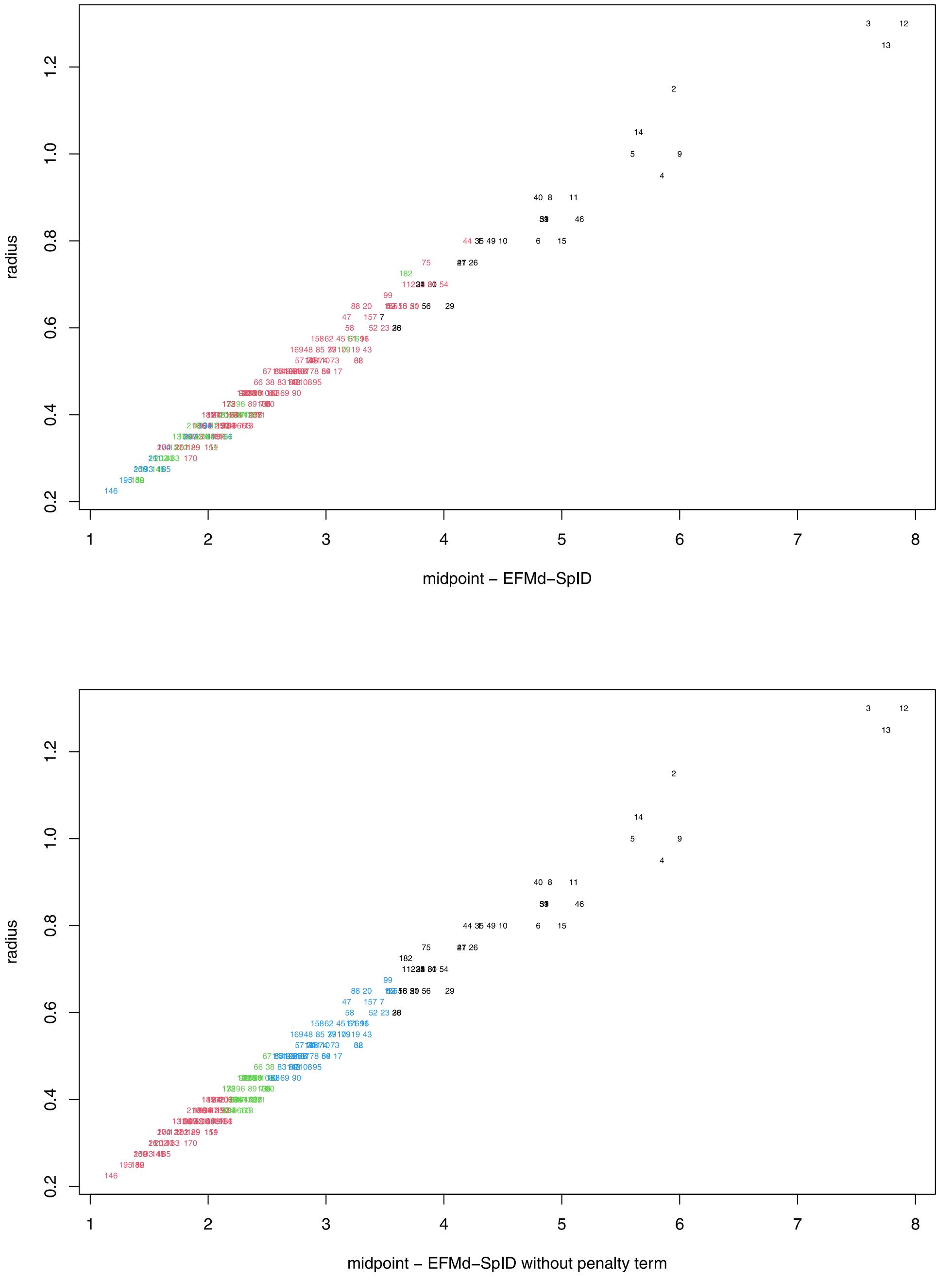

Fig. 17. Clustering EFMd-SpID.

Table 18
Values (euros per square meter) of the 211 zones of the municipality of Rome: minimum, maximum, midpoint, radius and zone.

	Zone name	Min	Max	Mid	Radius	Zone
1	TESTACCIO (PIAZZA S. MARIA LIBERATRICE)	3500	5100	4300	800	B1
2	AVENTINO (RIPA-VIA DI S. SABINA)	4800	7100	5950	1150	B12
3	SANT'ANGELO-CAMPITELLI (VIA DEL PORTICO D'OTTAVIA)	6300	8900	7600	1300	B13
4	TRASTEVERE (VIA DELLA LUNGARA)	4900	6800	5850	950	B14
5	BORGO (VIA DELLA CONCILIAZIONE)	4600	6600	5600	1000	B15
6	SALLUSTIANO-CASTRO PRETORIO (PIAZZA INDIPENDENZA)	4000	5600	4800	800	B17
7	ESQUILINO (PIAZZA VITTORIO)	2850	4100	3475	625	B18
8	SAN SABA (PIAZZA G.L. BERNINI)	4000	5800	4900	900	B2
9	LUDOVISI (VIA VENETO)	5000	7000	6000	1000	B25
10	VIMINALE (VIA TORINO)	3700	5300	4500	800	B29
11	CELIO (VIA CLAUDIA)	4200	6000	5100	900	B3
12	C. STORICO: TRIDENTE (CAMPO MARZIO, COLONNA, PIGNA, TREVI)	6600	9200	7900	1300	B31
13	C. STORICO: CORSO VITTORIO (PONTE, PARIONE, REGOLA, S. EUSTACHIO)	6500	9000	7750	1250	B32
14	MONTI (VIA DEI SERPENTI)	4600	6700	5650	1050	B4
15	PARIOLI (PIAZZA EUCLIDE)	4200	5800	5000	800	C1
16	GARBATELLA (LARGO DELLE SETTE CHIESE)	2750	3900	3325	575	C10
17	MARCONI (PIAZZA ENRICO FERMI)	2600	3600	3100	500	C11
18	MONTEVERDE VECCHIO (VIA POERIO)	3000	4300	3650	650	C12
19	MONTEVERDE NUOVO (VIA DI DONNA OLIMPIA)	2700	3800	3250	550	C13
20	AURELIO MONTE DI CRETA (PIAZZA IRNERIO)	2700	4000	3350	650	C14
21	AURELIO GREGORIO VII (VIA GREGORIO VII)	3100	4400	3750	650	C15
22	CAVALLEGGERI (VIA DELLE FORNACI)	3100	4500	3800	700	C16
23	BALDUINA GIOVENALE (VIA DELLE MEDAGLIE DORO)	2900	4100	3500	600	C17
24	CIPRO (VIA ANGELO EMO)	3100	4500	3800	700	C18
25	PONTE MILVIO-FARNESINA (VIA DELLA FARNESINA)	3100	4500	3800	700	C19
26	SALARIO (VIA NIZZA)	3500	5000	4250	750	C2
27	SALARIO AFRICANO (VIALE LIBIA)	3400	4900	4150	750	C21
28	BATTERIA NOMENTANA-LANCIANI (VIA COSTANTINO MAES)	3000	4200	3600	600	C22
29	BOLOGNA (VIA LIVORNO)	3400	4700	4050	650	C24
30	VILLAGGIO OLIMPICO (VIALE DE COUBERTIN)	3200	4600	3900	700	C26
31	PORTA PORTESE (VIA ETTORE ROLLI)	3100	4500	3800	700	C27
32	CASAL BERTONE-PORTONACCIO (VIA DI CASAL BERTONE)	2750	3800	3275	525	C29
33	PINCIANO (VIA GIOVANNI PAISIELLO)	4000	5700	4850	850	C3
34	PIGNETO (PIAZZA DEL PIGNETO)	2350	3400	2875	525	C30
35	MARCO POLO (VIALE MARCO POLO)	3500	5100	4300	800	C31
36	OSTIENSE (VIA DEL PORTO FLUVIALE)	3000	4200	3600	600	C32
37	AURELIO MADONNA DEL RIPOSO (VIA BENTIVOGLIO)	2500	3600	3050	550	C35
38	CASILINO MARRANELLA (VIA LABICO)	2050	3000	2525	475	C38
39	NOMENTANO TORLONIA (PIAZZA GALENO)	4000	5700	4850	850	C4
40	PRATI (VIA COLA DI RIENZO)	3900	5700	4800	900	C40
41	DELLA VITTORIA (PIAZZA MAZZINI)	3400	4900	4150	750	C41
42	VIGNA CLARA (VIA DI VIGNA STELLUTI)	2900	4200	3550	650	C42
43	TRIONFALE IGEA (VIA MARIO FANI)	2800	3900	3350	550	C43
44	CAMILLUCCIA (VIA DELLA CAMILLUCCIA)	3400	5000	4200	800	C44
45	NOCETTA (VIA DELLA NOCETTA)	2550	3700	3125	575	C45
46	SALARIO TRIESTE (CORSO TRIESTE)	4300	6000	5150	850	C46
47	BALDUINA BELSITO (PIAZZA MADONNA DEL CENACOLO)	2550	3800	3175	625	C47
48	TOR MARANCIA NAVIGATORI (VIA C.T. ODESCALCHI)	2300	3400	2850	550	C48
49	FLAMINIO (VIA G. RENI)	3600	5200	4400	800	C49
50	SAN LORENZO (VIA DEI SABELLI)	3100	4400	3750	650	C5
51	FLAMINIO PORTA DEL POPOLO (PIAZZA DELLA MARINA)	4000	5700	4850	850	C50
52	APPIO LATINO (VIA LATINA)	2800	4000	3400	600	C51
53	COLLINA FLEMING (VIA BEVAGNA)	2900	4200	3550	650	C52
54	APPIO VILLA FIORELLI (VIA TARANTO)	3300	4700	4000	700	C7
55	APPIO NOCERA UMBRA (PIAZZA S. MARIA AUSILIATRICE)	3000	4300	3650	650	C8

(*continued on next page*)

Table 18 (*continued*).

56	APPIO METRONIO (PIAZZA TUSCOLO)	3200	4500	3850	650	C9
57	PRENESTINO LABICANO (VIALE PARTENOPE, VIA TEANO)	2250	3300	2775	525	D1
58	MONTESACRO (VIALE ADRIATICO)	2600	3800	3200	600	D11
59	VALMELAINA-TUFELLO (VIA DELLE ISOLE CURZOLANE)	2500	3500	3000	500	D12
60	CENTOCELLE (PIAZZA DEI MIRTI)	2100	2950	2525	425	D14
61	ARDEATINO OTTAVO COLLE (VIA DEL SERAFICO)	2650	3800	3225	575	D15
62	PIETRALATA TIBURTINO (VIA FILIPPO MEDA)	2450	3600	3025	575	D16
63	COLLATINO (VIA DELLA SERENISSIMA)	1950	2700	2325	375	D17
64	ALESSANDRINO (VIALE ALESSANDRINO)	1900	2700	2300	400	D18
65	CINECITTA' DON BOSCO (PIAZZA S. GIOVANNI BOSCO)	2100	3100	2600	500	D19
66	TOR PIGNATTARA (VIA DI TOR PIGNATTARA)	1950	2900	2425	475	D2
67	CINECITTA' LAMARO (VIA RAIMONDO SCINTU)	2000	3000	2500	500	D20
68	APPIO CLAUDIO (VIALE GIULIO AGRICOLA)	2750	3800	3275	525	D21
69	STATUARIO-CAPANNELLE (VIA DEL CALICE)	2200	3100	2650	450	D22
70	QUARTO MIGLIO (VIA APPIA PIGNATELLI)	2350	3400	2875	525	D23
71	PRIMAVALLE-TORREVECCHIA (VIA DI TORREVECCHIA)	2050	2850	2450	400	D24
72	SAN BASILIO (VIA POLLENZA)	1700	2500	2100	400	D26
73	TALENTI (VIA UGO OJETTI)	2550	3600	3075	525	D27
74	GIULIANO DALMATA (VIA MATTEO BARTOLI)	2450	3500	2975	525	D28
75	EUR (VIALE EUROPA)	3100	4600	3850	750	D29
76	TOMBA DI NERONE (VIA DI GROTTAROSSA)	2050	2900	2475	425	D30
77	CASSIA DUE PONTI (VIA ORIOLO ROMANO)	2300	3300	2800	500	D31
78	FONTE MERAVIGLIOSA-ARDEATINO MILLEVOI (VIA STEFANO GRADI)	2400	3400	2900	500	D32
79	MONTAGNOLA (VIA PICO DELLA MIRANDOLA)	2500	3600	3050	550	D34
80	GROTTA PERFETTA-ROMA 70 (VIALE ERMINIO SPALLA)	2100	3000	2550	450	D36
81	APPIA ANTICA (VIA DI TOR CARBONE)	3200	4600	3900	700	D37
82	TINTORETTO (VIA BALLARIN)	2250	3200	2725	475	D38
83	PINETA SACCHETTI (VIA MATTIA BATTISTINI)	2150	3100	2625	475	D45
84	CONCA D'ORO (VIA VAL DI LANZO)	2500	3500	3000	500	D46
85	SACCO PASTORE (VIA VAL TROMPIA)	2400	3500	2950	550	D47
86	PIETRALATA (VIA DI PIETRALATA)	2300	3300	2800	500	D48
87	CASAL BRUCIATO (VIA C. FACCHINETTI)	2100	3100	2600	500	D49
88	SAN PAOLO (VIA TULLIO LEVI CIVITA)	2600	3900	3250	650	D5
89	CASILINO VILLA DE SANTIS (VIA ROMOLO BALZANI)	1950	2800	2375	425	D52
90	QUADRARO (VIA DEI QUINTILI)	2300	3200	2750	450	D53
91	ARCO DI TRAVERTINO-TOR FISCALE (VIA DEMETRIADE)	2750	3900	3325	575	D54
92	NOMENTANO KANT (VIALE KANT)	1850	2750	2300	450	D57
93	PODERE ROSA (VIA DIEGO FABBRI)	2200	3200	2700	500	D58
94	CASAL DEI PAZZI (VIA DI CASAL DEI PAZZI)	1850	2650	2250	400	D59
95	COLLI ANIENE-VERDEROCCA (VIA GROTTA DI GREGNA)	2450	3400	2925	475	D61
96	TOR TRE TESTE (VIA DAVIDE CAMPARI)	1850	2700	2275	425	D63
97	TORRE SPACCATA (VIA DEI ROMANISTI)	1750	2450	2100	350	D64
98	TORRACCIA DI SAN BASILIO (VIA DONATO MENICHELLA)	1850	2650	2250	400	D66
99	CORTINA D'AMPEZZO (VIA CORTINA D'AMPEZZO)	2850	4200	3525	675	D68
100	MONTE MARIO ALTO (VIA AUGUSTO CONTI)	2200	3200	2700	500	D69
101	PORTUENSE (VIA PROSPERO COLONNA)	2300	3300	2800	500	D7
102	PORTUENSE AFFOGALASINO (VIA AFFOGALASINO)	2200	3200	2700	500	D70
103	NUOVO SALARIO-PRATI FISCALI (VIA MONTE CERVIALTO)	2350	3400	2875	525	D71
104	CASETTA MATTEI (VIA DEGLI ADIMARI)	2050	2950	2500	450	D73
105	TRULLO (VIA MONTE DELLE CAPRE)	2050	2900	2475	425	D74
106	BOSCO DEGLI ARVALI (VIA DI GENEROSA)	1750	2550	2150	400	D75
107	PISANA-BRAVETTA (VIA DEI GONZAGA)	2200	3200	2700	500	D76
108	AURELIO VAL CANNUTA (VIA DI VAL CANNUTA)	2350	3300	2825	475	D77
109	FERRATELLA (VIALE CESARE PAVESE)	2600	3700	3150	550	D78
110	TORRINO SUD (VIA DEL FIUME GIALLO)	2450	3500	2975	525	D79
111	MAGLIANA (VIA DELL'IMPRUNETA)	1900	2800	2350	450	D8

(*continued on next page*)

Table 18 (*continued*).

112	TORRINO NORD-CITTA' D'EUROPA (VIA DELLE COSTELLAZIONI)	3000	4400	3700	700	D80
113	DECIMA (VIALE CAMILLO SABATINI)	1950	2700	2325	375	D81
114	MOSTACCIANO (VIALE BEATA VERGINE DEL CARMELO)	2100	3100	2600	500	D82
115	COLLI PORTUENSI (PIAZZALE EUGENIO MORELLI)	2350	3400	2875	525	D9
116	MORENA-CASAL MORENA (VIA CASAL MORENA)	1750	2450	2100	350	E10
117	CASAL MONASTERO (VIA BELMONTE IN SABINA)	1950	2750	2350	400	E101
118	BOCCEA QUARTACCIO (VIA ANDERSEN)	2250	3200	2725	475	E105
119	GIARDINETTI (VIA MARCANTONIO RAIMONDI)	1700	2350	2025	325	E11
120	CASALOTTI SELVA CANDIDA (VIA DI SELVA NERA)	1850	2750	2300	450	E113
121	OTTAVIA LUCCHINA (VIA DELLA LUCCHINA)	1900	2800	2350	450	E114
122	SETTECAMINI (VIA RUBELLIA)	1400	2050	1725	325	E115
123	CASALOTTI (PIAZZA ORMEA)	1800	2600	2200	400	E12
124	LA CINQUINA-BUFALOTTA (VIA FEO BELCARI)	1950	2850	2400	450	E123
125	CASALOTTI PANTAN MONASTERO (VIA DI CASAL SELCE)	1700	2450	2075	375	E125
126	CASILINO DUE TORRI-VILLAVERDE (VIA DEGAS)	1450	2150	1800	350	E127
127	VERMICINO (VIA DEL CASALE ANTONIONI)	1650	2450	2050	400	E128
128	FONTE OSTIENSE (VIALE IGNAZIO SILONE)	2200	3200	2700	500	E13
129	LA RUSTICA (VIA ACHILLE VERTUNNI)	1550	2200	1875	325	E132
130	TOR BELLA MONACA PEEP (VIA DELL'ARCHEOLOGIA)	1150	1650	1400	250	E138
131	TOR BELLA MONACA-VALLE FIORITA-DUE LEONI (VIA ACQUARONI)	1400	2100	1750	350	E139
132	TOR SAPIENZA-ZONA INDUSTRIALE VIA DELL'OMO (VIALE GIORGIO DE CHIRICO)	2000	2800	2400	400	E14
133	TORRENOVA (VIA DELLA TENUTA DI TORRENOVA)	1650	2350	2000	350	E140
134	CINECITTA' EST (VIALE ANTONIO CIAMARRA)	2050	2900	2475	425	E141
135	TOR VERGATA UNIVERSITA'-PASSO LOMBARDO (VIA DI PASSO LOMBARDO)	1750	2600	2175	425	E143
136	ROMANINA (VIA SCIMONELLI)	1650	2350	2000	350	E144
137	ANAGNINA VALLE MARCIANA-FOSSO SANT'ANDREA	2000	2800	2400	400	E145
138	MORENA GASPERINA (VIA CROPANI)	1900	2800	2350	450	E146
139	GREGNA SANT'ANDREA (VIA PIETRO CROSTAROSA)	1600	2400	2000	400	E147
140	TORRE ANGELA (VIA DEL TORRACCIO DI TORRENOVA)	1350	1950	1650	300	E15
141	PISANA-PONTE GALERIA (VIA ETTORE SCANDALE, VIA USINI)	1600	2400	2000	400	E156
142	CASTEL DI LEVA (VIA DI TOR CHIESACCIA)	2250	3200	2725	475	E157
143	DIVINO AMORE-FALCOGNANA (VIA DEI CASALI DI PORTA MEDAGLIA)	1550	2250	1900	350	E158
144	LA STORTA CASALE SAN NICOLA (VIA G.B. PARAVIA)	1650	2350	2000	350	E159
145	CASTEL GIUBILEO-BEL POGGIO (VIA CASTORANO)	1300	1850	1575	275	E16
146	TRAGLIATELLA (VIA PIOSSASCO)	950	1400	1175	225	E160
147	VALLE MURICANA-MONTE PIETRA PERTUSA (VIA DI VALLE MURICANA)	1850	2650	2250	400	E162
148	LUNGHEZZA-CASTELVERDE-FOSSO SAN GIULIANO (VIA DEL FOSSO DELL'OSA)	1300	1850	1575	275	E165
149	CORCOLLE-SAN VITTORINO (VIA SANT'ELPIDIO AL MARE)	1150	1650	1400	250	E166
150	PRENESTINO COLLE DEL SOLE-LAGO REGILLO (VIA OLLOLAI)	1150	1700	1425	275	E168
151	OTTAVIA PALMAROLA (VIA DELLA PALMAROLA)	1700	2350	2025	325	E17
152	COLLE MATTIA-FONTANA CANDIDA (VIA DEL CASALE CIMINELLI)	1150	1650	1400	250	E170
153	CASALOTTI VALLE SANTA (VIA VENDRAMINI)	1550	2250	1900	350	E171
154	ROMANINA TOR VERGATA (VIA BERNARDINO ALIMENA)	1800	2600	2200	400	E179
155	GROTTAROSSA-SAXA RUBRA (VIA CARLO EMERY)	2900	4200	3550	650	E18
156	PARCO DI VEIO PRATO DELLA CORTE (VIA FORMELLESE)	1800	2500	2150	350	E181
157	MEZZOCAMMINO (VIA DI MEZZOCAMMINO)	2750	4000	3375	625	E184
158	FONTE LAURENTINA (VIA EDOARDO AMALDI)	2350	3500	2925	575	E185
159	BORGHESIANA-FINOCCHIO (VIA DI BORGHESIANA)	1500	2200	1850	350	E19
160	SPINACETO-TOR DE' CENCI (VIALE DEGLI EROI DI CEFALONIA)	1600	2350	1975	375	E21
161	ACQUA VERGINE (PRATO FIORITO-COLLE PRENESTINO, MONFORTANI)	1250	1850	1550	300	E22
162	SETTECAMINI CASE ROSSE (VIA DELLE CASE ROSSE)	1450	2100	1775	325	E23
163	LA GIUSTINIANA (VIA ITALO PICCAGLI)	1900	2800	2350	450	E24
164	PRIMA PORTA (VIA DELLA VILLA DI LIVIA)	1450	2150	1800	350	E25
165	LA STORTA (VIA DELLA TORRE DI SPIZZICHINO)	1800	2600	2200	400	E26
166	LABARO (VIA GEMONA DEL FRIULI)	1550	2300	1925	375	E27
167	TORRE GAIA-VILLAGGIO BREDA (VIA DI GROTTE CELONI)	1650	2400	2025	375	E28
168	TORRE MAURA (VIA WALTER TOBAGI)	1650	2350	2000	350	E3
169	CASAL PALOCCO (VIALE GORGIA DI LEONTINI)	2200	3300	2750	550	E30
170	ACILIA NORD (VIA DEI MONTI DI SAN PAOLO)	1550	2150	1850	300	E31
171	OSTIA ANTICA (VIA DEL CASTELLO)	1650	2450	2050	400	E32

(*continued on next page*)

Table 18 (*continued*).

172	ACILIA SUD (VIA DI PRATO CORNELIO)	1750	2600	2175	425	E33
173	OSTIA (VIA DELLE BALENIERE)	1750	2600	2175	425	E34
174	ARDEATINA SELVOTTA-SPREGAMORE (VIA DEL FOSSO DELLA SOLFARATA)	1300	1950	1625	325	E38
175	CASTEL DI DECIMA-CASTEL ROMANO (VIA NAZZARENO STRAMPELLI)	1700	2400	2050	350	E39
176	COLLE SALARIO (VIA MONTE GIBERTO)	2650	3800	3225	575	E40
177	FIDENE-VILLA SPADA (VIA RADICOFANI)	1750	2500	2125	375	E41
178	SERPENTARA (VIALE LINA CAVALIERI)	1950	2750	2350	400	E42
179	VIGNE NUOVE-PORTA DI ROMA (VIA DELLE VIGNE NUOVE)	2600	3700	3150	550	E47
180	OSTERIA DEL CURATO-LUCREZIA ROMANA (VIA DELLE CAPANNELLE)	1950	2850	2400	450	E49
181	SETTEBAGNI (VIA S. ANTONIO DA PADOVA)	1800	2500	2150	350	E50
182	CASAL BOCCONE BUFALOTTA (VIA PAOLO MONELLI)	2950	4400	3675	725	E51
183	PONTE DI NONA (VIA LUIGI GASTINELLI)	1400	2000	1700	300	E53
184	TRIGORIA (VIA DI TRIGORIA)	1650	2450	2050	400	E62
185	SANTA PALOMBA-PIAN SAVELLI (VIA DELLA STAZIONE DI PAVONA)	1350	1900	1625	275	E64
186	FIORANELLO (VIA DI FIORANELLO)	1550	2300	1925	375	E65
187	VALLERANO (VIA DI VALLERANELLO)	2300	3300	2800	500	E66
188	TORRESINA-MONTE DEL MARMO (VIA DEL PODERE DI SAN GIUSTO)	2100	3000	2550	450	E7
189	PIANA DEL SOLE - FIERA DI ROMA (VIA CRISTOFORO SABBADINO)	1550	2200	1875	325	E71
190	CASTEL DI GUIDO-MALAGROTTA (VIA DI CASTEL DI GUIDO)	1750	2500	2125	375	E72
191	MASSIMINA-CASAL LUMBROSO (VIA MASSIMILLA)	1600	2350	1975	375	E73
192	CECCHIGNOLA-TOR PAGNOTTA (VIA DELLA CECCHIGNOLA)	1750	2500	2125	375	E75
193	CESANO (VIA DI BACCANELLO)	1200	1750	1475	275	E77
194	ISOLA FARNESE (VIA CERQUETTA)	1600	2350	1975	375	E78
195	OSTERIA NUOVA-CASACCIA (VIA ANGUILLARESE)	1050	1550	1300	250	E79
196	MONTESPACCATO (VIA CORNELIA)	1850	2600	2225	375	E8
197	VITINIA (VIA SARSINA)	1500	2200	1850	350	E81
198	ACILIA NUOVA-MADONNETTA (VIA BEPI ROMAGNONI)	1950	2850	2400	450	E83
199	AXA (VIA ARISTOFANE)	1800	2550	2175	375	E84
200	DRAGONCELLO (VIALE A. RUSPOLI)	1300	1950	1625	325	E87
201	DRAGONA (VIA DI DRAGONE)	1450	2100	1775	325	E88
202	TIBERINA-MALBORGHETTO (VIA MALBORGHETTO-TIBERINA KM 7)	1300	1900	1600	300	E89
203	TOR CERVARA-PONTE MAMMOLO (VIA VANNINA)	1750	2550	2150	400	E9
204	STAGNI DI OSTIA-LONGARINA (VIA FEDERICO BAZZINI)	1800	2550	2175	375	E92
205	INFERNETTO (VIA ERMANNO WOLF FERRARI)	2000	2800	2400	400	E94
206	SANTA CORNELIA (VIA DEL FOSSO DI MONTE OLIVIERO)	1600	2300	1950	350	E96
207	OLGIATA (LARGO DELL'OLGIATA)	1500	2200	1850	350	E97
208	MURATELLA (VIALE GAETANO ARTURO CROCCO)	1900	2800	2350	450	E98
209	SANTA MARIA DI GALERIA (VIA DI SANTA MARIA DI GALERIA)	1150	1700	1425	275	R40
210	RISERVA DELLA MARCIGLIANA (VIA DI SANTA COLOMBA)	1500	2250	1875	375	R7
211	AGRO ROMANO OVEST (VIA SANTA MARIA DI GALERIA)	1250	1850	1550	300	R9

Table 19
Membership degrees and highest membership cluster - $C = 4$ EFMd-SpID. In bold the EFMd-SpID medoids, in italic the EFMd-SpID fuzzy municipalities.

	Zone name	Membership				Cluster
		1	2	3	4	
1	TESTACCIO (PIAZZA S. MARIA LIBERATRICE)	1.00	0.00	0.00	0.00	1
2	AVENTINO (RIPA-VIA DI S.S ABINA)	1.00	0.00	0.00	0.00	1
3	SANT'ANGELO-CAMPITELLI (VIA DEL PORTICO D'OTTAVIA)	1.00	0.00	0.00	0.00	1
4	TRASTEVERE (VIA DELLA LUNGARA)	1.00	0.00	0.00	0.00	1
5	BORGO (VIA DELLA CONCILIAZIONE)	1.00	0.00	0.00	0.00	1
6	**SALLUSTIANO-CASTRO PRETORIO (PIAZZA INDIPENDENZA)**	**1.00**	0.00	0.00	0.00	1
7	ESQUILINO (PIAZZA VITTORIO)	0.89	0.11	0.00	0.00	1
8	SAN SABA (PIAZZA G.L. BERNINI)	1.00	0.00	0.00	0.00	1
9	LUDOVISI (VIA VENETO)	1.00	0.00	0.00	0.00	1
10	VIMINALE (VIA TORINO)	1.00	0.00	0.00	0.00	1
11	CELIO (VIA CLAUDIA)	1.00	0.00	0.00	0.00	1

(*continued on next page*)

Table 19 (*continued*).

12	C. STORICO: TRIDENTE (CAMPO MARZIO. COLONNA. PIGNA. TREVI)	1.00	0.00	0.00	0.00	1
13	C. STORICO: CORSO VITTORIO (PONTE. PARIONE. REGOLA. S. EUSTACHIO)	1.00	0.00	0.00	0.00	1
14	MONTI (VIA DEI SERPENTI)	1.00	0.00	0.00	0.00	1
15	PARIOLI (PIAZZA EUCLIDE)	1.00	0.00	0.00	0.00	1
16	GARBATELLA (LARGO DELLE SETTE CHIESE)	0.00	1.00	0.00	0.00	2
17	MARCONI (PIAZZA ENRICO FERMI)	0.00	1.00	0.00	0.00	2
18	MONTEVERDE VECCHIO (VIA POERIO)	0.02	0.98	0.00	0.00	2
19	MONTEVERDE NUOVO (VIA DI DONNA OLIMPIA)	0.00	1.00	0.00	0.00	2
20	AURELIO MONTE DI CRETA (PIAZZA IRNERIO)	0.00	1.00	0.00	0.00	2
21	AURELIO GREGORIO VII (VIA GREGORIO VII)	0.00	1.00	0.00	0.00	2
22	CAVALLEGGERI (VIA DELLE FORNACI)	0.00	1.00	0.00	0.00	2
23	BALDUINA GIOVENALE (VIA DELLE MEDAGLIE D'ORO)	0.00	1.00	0.00	0.00	2
24	*CIPRO (VIA ANGELO EMO)*	*0.53*	0.47	0.00	0.00	1
25	PONTE MILVIO-FARNESINA (VIA DELLA FARNESINA)	0.01	0.99	0.00	0.00	2
26	SALARIO (VIA NIZZA)	1.00	0.00	0.00	0.00	1
27	SALARIO AFRICANO (VIALE LIBIA)	0.96	0.04	0.00	0.00	1
28	BATTERIA NOMENTANA-LANCIANI (VIA COSTANTINO MAES)	0.99	0.01	0.00	0.00	1
29	BOLOGNA (VIA LIVORNO)	0.88	0.12	0.00	0.00	1
30	VILLAGGIO OLIMPICO (VIALE DE COUBERTIN)	0.76	0.24	0.00	0.00	1
31	PORTA PORTESE (VIA ETTORE ROLLI)	0.98	0.02	0.00	0.00	1
32	CASAL BERTONE-PORTONACCIO (VIA DI CASAL BERTONE)	0.00	1.00	0.00	0.00	2
33	PINCIANO (VIA GIOVANNI PAISIELLO)	1.00	0.00	0.00	0.00	1
34	PIGNETO (PIAZZA DEL PIGNETO)	0.00	1.00	0.00	0.00	2
35	MARCO POLO (VIALE MARCO POLO)	0.98	0.02	0.00	0.00	1
36	OSTIENSE (VIA DEL PORTO FLUVIALE)	1.00	0.01	0.00	0.00	1
37	AURELIO MADONNA DEL RIPOSO (VIA BENTIVOGLIO)	0.00	1.00	0.00	0.00	2
38	CASILINO MARRANELLA (VIA LABICO)	0.00	1.00	0.00	0.00	2
39	NOMENTANO TORLONIA (PIAZZA GALENO)	1.00	0.00	0.00	0.00	1
40	PRATI (VIA COLA DI RIENZO)	1.00	0.00	0.00	0.00	1
41	DELLA VITTORIA (PIAZZA MAZZINI)	0.96	0.04	0.00	0.00	1
42	VIGNA CLARA (VIA DI VIGNA STELLUTI)	0.00	1.00	0.00	0.00	2
43	TRIONFALE IGEA (VIA MARIO FANI)	0.00	1.00	0.00	0.00	2
44	CAMILLUCCIA (VIA DELLA CAMILLUCCIA)	0.00	1.00	0.00	0.00	2
45	NOCETTA (VIA DELLA NOCETTA)	0.00	1.00	0.00	0.00	2
46	SALARIO TRIESTE (CORSO TRIESTE)	1.00	0.00	0.00	0.00	1
47	BALDUINA BELSITO (PIAZZA MADONNA DEL CENACOLO)	0.00	1.00	0.00	0.00	2
48	TOR MARANCIA NAVIGATORI (VIA C.T. ODESCALCHI)	0.00	1.00	0.00	0.00	2
49	FLAMINIO (VIA G. RENI)	1.00	0.00	0.00	0.00	1
50	SAN LORENZO (VIA DEI SABELLI)	0.02	0.98	0.00	0.00	2
51	FLAMINIO PORTA DEL POPOLO (PIAZZA DELLA MARINA)	1.00	0.00	0.00	0.00	1
52	APPIO LATINO (VIA LATINA)	0.00	1.00	0.00	0.00	2
53	COLLINA FLEMING (VIA BEVAGNA)	0.00	1.00	0.00	0.00	2
54	APPIO VILLA FIORELLI (VIA TARANTO)	0.09	0.91	0.00	0.00	2
55	APPIO NOCERA UMBRA (PIAZZA S. MARIA AUSILIATRICE)	0.00	1.00	0.00	0.00	2
56	APPIO METRONIO (PIAZZA TUSCOLO)	0.99	0.01	0.00	0.00	1
57	PRENESTINO LABICANO (VIALE PARTENOPE.VIA TEANO)	0.00	1.00	0.00	0.00	2
58	MONTESACRO (VIALE ADRIATICO)	0.00	1.00	0.00	0.00	2
59	VALMELAINA-TUFELLO (VIA DELLE ISOLE CURZOLANE)	0.00	1.00	0.00	0.00	2
60	CENTOCELLE (PIAZZA DEI MIRTI)	0.00	1.00	0.00	0.00	2
61	ARDEATINO OTTAVO COLLE (VIA DEL SERAFICO)	0.00	1.00	0.00	0.00	2
62	PIETRALATA TIBURTINO (VIA FILIPPO MEDA)	0.00	1.00	0.00	0.00	2
63	COLLATINO (VIA DELLA SERENISSIMA)	0.00	1.00	0.00	0.00	2
64	ALESSANDRINO (VIALE ALESSANDRINO)	0.00	0.03	0.97	0.00	3
65	CINECITTA‘ DON BOSCO (PIAZZA S. GIOVANNI BOSCO)	0.00	1.00	0.00	0.00	2
66	TOR PIGNATTARA (VIA DI TOR PIGNATTARA)	0.00	1.00	0.00	0.00	2
67	CINECITTA‘ LAMARO (VIA RAIMONDO SCINTU)	0.00	1.00	0.00	0.00	2
68	APPIO CLAUDIO (VIALE GIULIO AGRICOLA)	0.00	1.00	0.00	0.00	2
69	STATUARIO-CAPANNELLE (VIA DEL CALICE)	0.00	1.00	0.00	0.00	2
70	QUARTO MIGLIO (VIA APPIA PIGNATELLI)	0.00	1.00	0.00	0.00	2
71	PRIMAVALLE-TORREVECCHIA (VIA DI TORREVECCHIA)	0.00	1.00	0.00	0.00	2

(*continued on next page*)

Table 19 (*continued*).

72	SAN BASILIO (VIA POLLENZA)	0.00	0.00	1.00	0.00	3
73	TALENTI (VIA UGO OJETTI)	0.00	1.00	0.00	0.00	2
74	GIULIANO DALMATA (VIA MATTEO BARTOLI)	0.00	1.00	0.00	0.00	2
75	EUR (VIALE EUROPA)	0.00	1.00	0.00	0.00	2
76	TOMBA DI NERONE (VIA DI GROTTAROSSA)	0.00	1.00	0.00	0.00	2
77	CASSIA DUE PONTI (VIA ORIOLO ROMANO)	0.00	1.00	0.00	0.00	2
78	FONTE MERAVIGLIOSA-ARDEATINO MILLEVOI (VIA STEFANO GRADI)	0.00	1.00	0.00	0.00	2
79	MONTAGNOLA (VIA PICO DELLA MIRANDOLA)	0.00	1.00	0.00	0.00	2
80	GROTTA PERFETTA-ROMA 70 (VIALE ERMINIO SPALLA)	0.00	1.00	0.00	0.00	2
81	APPIA ANTICA (VIA DI TOR CARBONE)	0.00	1.00	0.00	0.00	2
82	TINTORETTO (VIA BALLARIN)	0.00	1.00	0.00	0.00	2
83	PINETA SACCHETTI (VIA MATTIA BATTISTINI)	0.00	1.00	0.00	0.00	2
84	CONCA D'ORO (VIA VAL DI LANZO)	0.00	1.00	0.00	0.00	2
85	SACCO PASTORE (VIA VAL TROMPIA)	0.00	1.00	0.00	0.00	2
86	PIETRALATA (VIA DI PIETRALATA)	0.00	1.00	0.00	0.00	2
87	CASAL BRUCIATO (VIA C. FACCHINETTI)	0.00	1.00	0.00	0.00	2
88	SAN PAOLO (VIA TULLIO LEVI CIVITA)	0.00	1.00	0.00	0.00	2
89	CASILINO VILLA DE SANTIS (VIA ROMOLO BALZANI)	0.00	1.00	0.00	0.00	2
90	QUADRARO (VIA DEI QUINTILI)	0.00	1.00	0.00	0.00	2
91	ARCO DI TRAVERTINO-TOR FISCALE (VIA DEMETRIADE)	0.00	1.00	0.00	0.00	2
92	NOMENTANO KANT (VIALE KANT)	0.00	1.00	0.00	0.00	2
93	**PODERE ROSA (VIA DIEGO FABBRI)**	0.00	**1.00**	0.00	0.00	2
94	CASAL DEI PAZZI (VIA DI CASAL DEI PAZZI)	0.00	1.00	0.00	0.00	2
95	COLLI ANIENE-VERDEROCCA (VIA GROTTA DI GREGNA)	0.00	0.99	0.01	0.00	2
96	TOR TRE TESTE (VIA DAVIDE CAMPARI)	0.00	0.00	1.00	0.00	3
97	TORRE SPACCATA (VIA DEI ROMANISTI)	0.00	1.00	0.00	0.00	2
98	TORRACCIA DI SAN BASILIO (VIA DONATO MENICHELLA)	0.00	0.00	1.00	0.00	3
99	CORTINA D'AMPEZZO (VIA CORTINA D'AMPEZZO)	0.00	1.00	0.00	0.00	2
100	MONTE MARIO ALTO (VIA AUGUSTO CONTI)	0.00	1.00	0.00	0.00	2
101	PORTUENSE (VIA PROSPERO COLONNA)	0.00	1.00	0.00	0.00	2
102	PORTUENSE AFFOGALASINO (VIA AFFOGALASINO)	0.00	1.00	0.00	0.00	2
103	NUOVO SALARIO-PRATI FISCALI (VIA MONTE CERVIALTO)	0.00	0.90	0.10	0.00	2
104	CASETTA MATTEI (VIA DEGLI ADIMARI)	0.00	1.00	0.00	0.00	2
105	TRULLO (VIA MONTE DELLE CAPRE)	0.00	1.00	0.00	0.00	2
106	BOSCO DEGLI ARVALI (VIA DI GENEROSA)	0.00	1.00	0.00	0.00	2
107	PISANA-BRAVETTA (VIA DEI GONZAGA)	0.00	1.00	0.00	0.00	2
108	AURELIO VAL CANNUTA (VIA DI VAL CANNUTA)	0.00	1.00	0.00	0.00	2
109	FERRATELLA (VIALE CESARE PAVESE)	0.00	1.00	0.00	0.00	2
110	TORRINO SUD (VIA DEL FIUME GIALLO)	0.00	1.00	0.00	0.00	2
111	MAGLIANA (VIA DELL'MPRUNETA)	0.00	1.00	0.00	0.00	2
112	TORRINO NORD-CITTA' D'EUROPA (VIA DELLE COSTELLAZIONI)	0.00	1.00	0.00	0.00	2
113	DECIMA (VIALE CAMILLO SABATINI)	0.00	1.00	0.00	0.00	2
114	MOSTACCIANO (VIALE BEATA VERGINE DEL CARMELO)	0.00	1.00	0.00	0.00	2
115	COLLI PORTUENSI (PIAZZALE EUGENIO MORELLI)	0.00	1.00	0.00	0.00	2
116	MORENA-CASAL MORENA (VIA CASAL MORENA)	0.00	1.00	0.00	0.00	2
117	CASAL MONASTERO (VIA BELMONTE IN SABINA)	0.00	0.00	1.00	0.00	3
118	BOCCEA QUARTACCIO (VIA ANDERSEN)	0.00	1.00	0.00	0.00	2
119	GIARDINETTI (VIA MARCANTONIO RAIMONDI)	0.00	0.00	1.00	0.00	3
120	CASALOTTI SELVA CANDIDA (VIA DI SELVA NERA)	0.00	1.00	0.00	0.00	2
121	OTTAVIA LUCCHINA (VIA DELLA LUCCHINA)	0.00	1.00	0.00	0.00	2
122	SETTECAMINI (VIA RUBELLIA)	0.00	0.00	1.00	0.00	3
123	CASALOTTI (PIAZZA ORMEA)	0.00	1.00	0.00	0.00	2
124	LA CINQUINA-BUFALOTTA (VIA FEO BELCARI)	0.00	0.00	1.00	0.00	3
125	CASALOTTI PANTAN MONASTERO (VIA DI CASAL SELCE)	0.00	1.00	0.00	0.00	2
126	CASILINO DUE TORRI-VILLAVERDE (VIA DEGAS)	0.00	0.00	1.00	0.00	3
127	VERMICINO (VIA DEL CASALE ANTONIONI)	0.00	0.94	0.06	0.00	2
128	FONTE OSTIENSE (VIALE IGNAZIO SILONE)	0.00	1.00	0.00	0.00	2
129	LA RUSTICA (VIA ACHILLE VERTUNNI)	0.00	0.00	1.00	0.00	3
130	TOR BELLA MONACA PEEP (VIA DELL'ARCHEOLOGIA)	0.00	0.00	1.00	0.00	3

(*continued on next page*)

Table 19 (*continued*).

131	TOR BELLA MONACA-VALLE FIORITA-DUE LEONI (VIA ACQUARONI)	0.00	0.00	1.00	0.00	3
132	TOR SAPIENZA-ZONA INDUSTRIALE VIA DELL'OMO (VIALE GIORGIO DE CHIRICO)	0.00	0.00	1.00	0.00	3
133	**TORRENOVA (VIA DELLA TENUTA DI TORRENOVA)**	0.00	0.00	**1.00**	0.00	3
134	CINECITTA' EST (VIALE ANTONIO CIAMARRA)	0.00	1.00	0.00	0.00	2
135	TOR VERGATA UNIVERSITA'-PASSO LOMBARDO (VIA DI PASSO LOMBARDO)	0.00	0.02	0.98	0.00	3
136	ROMANINA (VIA SCIMONELLI)	0.00	1.00	0.00	0.00	2
137	ANAGNINA VALLE MARCIANA-FOSSO SANT'ANDREA	0.00	0.95	0.03	0.01	2
138	MORENA GASPERINA (VIA CROPANI)	0.00	1.00	0.00	0.00	2
139	GREGNA SANT'ANDREA (VIA PIETRO CROSTAROSA)	0.00	1.00	0.00	0.00	2
140	TORRE ANGELA (VIA DEL TORRACCIO DI TORRENOVA)	0.00	0.00	1.00	0.00	3
141	PISANA-PONTE GALERIA (VIA ETTORE SCANDALE.VIA USINI)	0.00	1.00	0.00	0.00	2
142	CASTEL DI LEVA (VIA DI TOR CHIESACCIA)	0.00	1.00	0.00	0.00	2
143	DIVINO AMORE-FALCOGNANA (VIA DEI CASALI DI PORTA MEDAGLIA)	0.00	0.91	0.00	0.09	2
144	LA STORTA CASALE SAN NICOLA (VIA G.B. PARAVIA)	0.00	0.00	0.00	1.00	4
145	CASTEL GIUBILEO-BEL POGGIO (VIA CASTORANO)	0.00	0.00	1.00	0.00	3
146	TRAGLIATELLA (VIA PIOSSASCO)	0.00	0.00	0.02	0.98	4
147	VALLE MURICANA-MONTE PIETRA PERTUSA (VIA DI VALLE MURICANA)	0.00	0.00	1.00	0.00	3
148	LUNGHEZZA-CASTELVERDE-FOSSO SAN GIULIANO (VIA DEL FOSSO DELL'OSA)	0.00	0.00	1.00	0.00	3
149	CORCOLLE-SAN VITTORINO (VIA SANT'ELPIDIO AL MARE)	0.00	0.00	0.94	0.06	3
150	PRENESTINO COLLE DEL SOLE-LAGO REGILLO (VIA OLLOLAI)	0.00	0.00	1.00	0.00	3
151	OTTAVIA PALMAROLA (VIA DELLA PALMAROLA)	0.00	1.00	0.00	0.00	2
152	COLLE MATTIA-FONTANA CANDIDA (VIA DEL CASALE CIMINELLI)	0.00	0.00	0.94	0.06	3
153	CASALOTTI VALLE SANTA (VIA VENDRAMINI)	0.00	1.00	0.00	0.00	2
154	ROMANINA TOR VERGATA (VIA BERNARDINO ALIMENA)	0.00	1.00	0.00	0.00	2
155	GROTTAROSSA-SAXA RUBRA (VIA CARLO EMERY)	0.00	1.00	0.00	0.00	2
156	PARCO DI VEIO PRATO DELLA CORTE (VIA FORMELLESE)	0.00	0.03	0.00	0.97	4
157	MEZZOCAMMINO (VIA DI MEZZOCAMMINO)	0.00	1.00	0.00	0.00	2
158	FONTE LAURENTINA (VIA EDOARDO AMALDI)	0.00	1.00	0.00	0.00	2
159	BORGHESIANA-FINOCCHIO (VIA DI BORGHESIANA)	0.00	0.00	1.00	0.00	3
160	SPINACETO-TOR DE' CENCI (VIALE DEGLI EROI DI CEFALONIA)	0.00	1.00	0.00	0.00	2
161	ACQUA VERGINE (PRATO FIORITO-COLLE PRENESTINO.MONFORTANI)	0.00	0.00	1.00	0.00	3
162	SETTECAMINI CASE ROSSE (VIA DELLE CASE ROSSE)	0.00	0.00	1.00	0.00	3
163	LA GIUSTINIANA (VIA ITALO PICCAGLI)	0.00	1.00	0.00	0.00	2
164	PRIMA PORTA (VIA DELLA VILLA DI LIVIA)	0.00	0.00	1.00	0.00	3
165	LA STORTA (VIA DELLA TORRE DI SPIZZICHINO)	0.00	0.97	0.00	0.03	2
166	LABARO (VIA GEMONA DEL FRIULI)	0.00	0.01	0.99	0.00	3
167	TORRE GAIA-VILLAGGIO BREDA (VIA DI GROTTE CELONI)	0.00	0.00	1.00	0.00	3
168	TORRE MAURA (VIA WALTER TOBAGI)	0.00	0.00	1.00	0.00	3
169	CASAL PALOCCO (VIALE GORGIA DI LEONTINI)	0.00	1.00	0.00	0.00	2
170	ACILIA NORD (VIA DEI MONTI DI SAN PAOLO)	0.00	1.00	0.00	0.00	2
171	OSTIA ANTICA (VIA DEL CASTELLO)	0.00	1.00	0.00	0.00	2
172	ACILIA SUD (VIA DI PRATO CORNELIO)	0.00	1.00	0.00	0.00	2
173	OSTIA (VIA DELLE BALENIERE)	0.00	1.00	0.00	0.00	2
174	**ARDEATINA SELVOTTA-SPREGAMORE (VIA DEL FOSSO DELLA SOLFARATA)**	0.00	0.00	0.00	**1.00**	4
175	CASTEL DI DECIMA-CASTEL ROMANO (VIA NAZZARENO STRAMPELLI)	0.00	1.00	0.00	0.00	2
176	COLLE SALARIO (VIA MONTE GIBERTO)	0.00	0.00	1.00	0.00	3
177	FIDENE-VILLA SPADA (VIA RADICOFANI)	0.00	0.00	1.00	0.00	3
178	SERPENTARA (VIALE LINA CAVALIERI)	0.00	0.00	1.00	0.00	3
179	VIGNE NUOVE-PORTA DI ROMA (VIA DELLE VIGNE NUOVE)	0.00	0.11	0.89	0.00	3
180	OSTERIA DEL CURATO-LUCREZIA ROMANA (VIA DELLE CAPANNELLE)	0.00	1.00	0.00	0.00	2
181	SETTEBAGNI (VIA S. ANTONIO DA PADOVA)	0.00	0.00	1.00	0.00	3
182	CASAL BOCCONE BUFALOTTA (VIA PAOLO MONELLI)	0.00	0.00	1.00	0.00	3
183	PONTE DI NONA (VIA LUIGI GASTINELLI)	0.00	0.00	1.00	0.00	3
184	TRIGORIA (VIA DI TRIGORIA)	0.00	1.00	0.00	0.00	2
185	SANTA PALOMBA-PIAN SAVELLI (VIA DELLA STAZIONE DI PAVONA)	0.00	0.00	0.03	0.97	4
186	FIORANELLO (VIA DI FIORANELLO)	0.00	1.00	0.00	0.00	2
187	VALLERANO (VIA DI VALLERANELLO)	0.00	1.00	0.00	0.00	2

(*continued on next page*)

Table 19 (*continued*).

188	TORRESINA-MONTE DEL MARMO (VIA DEL PODERE DI SAN GIUSTO)	0.00	1.00	0.00	0.00	2
189	PIANA DEL SOLE - FIERA DI ROMA (VIA CRISTOFORO SABBADINO)	0.00	1.00	0.00	0.00	2
190	CASTEL DI GUIDO-MALAGROTTA (VIA DI CASTEL DI GUIDO)	0.00	1.00	0.00	0.00	2
191	MASSIMINA-CASAL LUMBROSO (VIA MASSIMILLA)	0.00	1.00	0.00	0.00	2
192	CECCHIGNOLA-TOR PAGNOTTA (VIA DELLA CECCHIGNOLA)	0.00	1.00	0.00	0.00	2
193	CESANO (VIA DI BACCANELLO)	0.00	0.00	0.00	1.00	4
194	ISOLA FARNESE (VIA CERQUETTA)	0.00	0.00	0.00	1.00	4
195	OSTERIA NUOVA-CASACCIA (VIA ANGUILLARESE)	0.00	0.00	0.00	1.00	4
196	MONTESPACCATO (VIA CORNELIA)	0.00	1.00	0.00	0.00	2
197	VITINIA (VIA SARSINA)	0.00	1.00	0.00	0.00	2
198	ACILIA NUOVA-MADONNETTA (VIA BEPI ROMAGNONI)	0.00	1.00	0.00	0.00	2
199	AXA (VIA ARISTOFANE)	0.00	1.00	0.00	0.00	2
200	DRAGONCELLO (VIALE A. RUSPOLI)	0.00	1.00	0.00	0.00	2
201	DRAGONA (VIA DI DRAGONE)	0.00	1.00	0.00	0.00	2
202	TIBERINA-MALBORGHETTO (VIA MALBORGHETTO-TIBERINA KM 7)	0.00	0.00	1.00	0.00	3
203	TOR CERVARA-PONTE MAMMOLO (VIA VANNINA)	0.00	0.02	0.98	0.00	3
204	STAGNI DI OSTIA-LONGARINA (VIA FEDERICO BAZZINI)	0.00	1.00	0.00	0.00	2
205	INFERNETTO (VIA ERMANNO WOLF FERRARI)	0.00	1.00	0.00	0.00	2
206	SANTA CORNELIA (VIA DEL FOSSO DI MONTE OLIVIERO)	0.00	0.00	1.00	0.00	3
207	OLGIATA (LARGO DELL'OLGIATA)	0.00	0.00	0.00	1.00	4
208	MURATELLA (VIALE GAETANO ARTURO CROCCO)	0.00	1.00	0.00	0.00	2
209	SANTA MARIA DI GALERIA (VIA DI SANTA MARIA DI GALERIA)	0.00	0.00	0.00	1.00	4
210	RISERVA DELLA MARCIGLIANA (VIA DI SANTA COLOMBA)	0.00	0.00	1.00	0.00	3
211	AGRO ROMANO OVEST (VIA SANTA MARIA DI GALERIA)	0.00	0.00	0.00	1.00	4

References

Campello, R., 2007. A fuzzy extension of the rand index and other related indexes for clustering and classification assessment. Pattern Recognit. Lett. 28 (7), 833–841.

Campello, R., Hruschka, E., 2006. A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets and Systems 157, 2858–2875.

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

Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., Chen, T.-J., 2006. Fuzzy c-means clustering with spatial information for image segmentation. Comput. Med. Imaging Graph. 30 (1), 9–15.

Coppi, R., D'Urso, P., 2006. Fuzzy unsupervised classification of multivariate time trajectories with the Shannon entropy regularization. Comput. Statist. Data Anal. 50 (6), 1452–1477.

Coppi, R., D'Urso, P., Giordani, P., 2010. A fuzzy clustering model for multivariate spatial time series. J. Classification 27 (1), 54–88.

Coppi, R., Giordani, P., D'Urso, P., 2006. Component models for fuzzy data. Psychometrika 71 (4), 733.

De Carvalho, F.d.A.T., 2007. Fuzzy c-means clustering methods for symbolic interval data. Pattern Recognit. Lett. 28, 423–437.

De Carvalho, F.d.A., Lechevallier, Y., 2009a. Dynamic clustering of interval-valued data based on adaptive quadratic distances. IEEE Trans. Syst. Man Cybern. 39, 1295–1306.

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

De Carvalho, F.d.A., Tenório, C.P., 2010. Fuzzy K-means clustering algorithms for interval-valued data based on adaptive quadratic distances. Fuzzy Sets and Systems 161 (23), 2978–2999.

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

Disegna, M., D'Urso, P., Durante, F., 2017. Copula-based fuzzy clustering of spatial time series. Spatial Stat. 21, 209–225.

Durante, F., Sempi, C., 2016. Principles of Copula Theory. CRC/Chapman & Hall, Boca Raton, FL.

D'Urso, P., 2015. Fuzzy clustering. In: Hennig, C., Meila, M., Murtagh, F., Rocci, R. (Eds.), Handbook of Cluster Analysis. Chapman & Hall, pp. 545–574.

D'Urso, P., De Giovanni, L., 2014. Robust clustering of imprecise data. Chemometr. Intell. Lab. Syst. 136, 58–80.

D'Urso, P., De Giovanni, L., Alaimo, L., Mattera, R., Vitale, V., 2023a. Fuzzy clustering with entropy regularization for interval-valued data with an application to scientific journal citations. Ann. Oper. Res. 1–24. http://dx.doi.org/10.1007/s10479-023-05180-1.

D'Urso, P., De Giovanni, L., Disegna, M., Massari, R., 2019a. Fuzzy clustering with spatial-temporal information. Spatial Stat. 30, 71–102.

D'Urso, P., De Giovanni, L., Maharaj, E.A., Brito, P., Teles, P., 2023b. Wavelet-based fuzzy clustering of interval time series. Internat. J. Approx. Reason. 152, 136–159.

D'Urso, P., De Giovanni, L., Massari, R., 2014. Self-organizing maps for imprecise data. Fuzzy Sets and Systems 237, 63–89.
D'Urso, P., De Giovanni, L., Massari, R., 2015. Trimmed fuzzy clustering for interval-valued data. Adv. Data Anal. Classif. 9 (1), 21–40.
D'Urso, P., De Giovanni, L., Massari, R., 2016. GARCH-based robust clustering of time series. Fuzzy Sets and Systems 305, 1–28.
D'Urso, P., De Giovanni, L., Massari, R., Sica, F.G.M., 2019b. Cross sectional and longitudinal fuzzy clustering of the NUTS and positioning of the Italian regions with respect to the regional competitiveness index (RCI) indicators with contiguity constraints. Soc. Indic. Res. 146 (3), 609–650.
D'Urso, P., De Giovanni, L., Sica, F.G., Vitale, V., 2022a. Measuring competitiveness at NUTS3 level and territorial partitioning of the Italian provinces. Soc. Indic. Res. 1–43. http://dx.doi.org/10.1007/s11205-021-02836-y.
D'Urso, P., De Giovanni, L., Vitale, V., 2022b. Spatial robust fuzzy clustering of COVID 19 time series based on B-splines. Spatial Stat. 49.
D'Urso, P., Giordani, P., 2004. A least squares approach to principal component analysis for interval valued data. Chemometr. Intell. Lab. Syst. 70 (2), 179–192.
D'Urso, P., Giordani, P., 2005. A possibilistic approach to latent component analysis for symmetric fuzzy data. Fuzzy Sets and Systems 150 (2), 285–305.
D'Urso, P., Giordani, P., 2006a. A robust fuzzy k-means clustering model for interval valued data. Comput. Statist. 21 (2), 251–269.
D'Urso, P., Giordani, P., 2006b. A weighted fuzzy c-means clustering model for fuzzy data. Comput. Statist. Data Anal. 50 (6), 1496–1523.
D'Urso, P., Leski, J., 2016. Fuzzy c-ordered medoids clustering for interval-valued data. Pattern Recognit. 58, 49–67.
D'Urso, P., Massari, R., De Giovanni, L., Cappelli, C., 2017. Exponential distance-based fuzzy clustering for interval-valued data. Fuzzy Optim. Decis. Mak. 16 (1), 51–70.
D'Urso, P., Mucciardi, M., Otranto, E., Vitale, V., 2021. Community mobility in the European regions during COVID-19 pandemic: A partitioning around medoids with noise cluster based on space–time autoregressive models. Spatial Stat. 49, 100531. http://dx.doi.org/10.1016/j.spasta.2021.100531.
D'Urso, P., Vitale, V., 2020. A robust hierarchical clustering for georeferenced data. Spatial Stat. 35, 100407.
Everitt, B., Landau, S., Leese, M., 2001. Cluster Analysis, forth ed. Arnold Press, London.
Fouedjio, F., 2016. A hierarchical clustering method for multivariate geostatistical data. Spatial Stat. 18, 333–351.
Fu, K., Albus, J., 1977. Syntactic Pattern Recognition. Springer-Verlag.
Gao, Y., Wang, D., Pan, J., Wang, Z., Chen, B., 2019. A novel fuzzy c-means clustering algorithm using adaptive norm. Int. J. Fuzzy Syst. 21 (8), 2632–2649.
Giordani, P., Kiers, H.A., 2004. Principal component analysis of symmetric fuzzy data. Comput. Statist. Data Anal. 45 (3), 519–548.
Gittleman, J., Kot, M., 1990. Adaptation: Statistics and a null model for estimating phylogenetic effects. Syst. Zool. 39 (3), 227–241.
Gowda, K.C., Diday, E., 1991. Symbolic clustering using a new dissimilarity measure. Pattern Recognit. 24 (6), 567–578.
Guru, D., Kiranagi, B.B., Nagabhushan, P., 2004. Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns. Pattern Recognit. Lett. 25 (10), 1203–1213.
Handl, J., Knowles, J., Kell, D., 2005. Computational cluster validation in post-genomic data analysis. Bioinformatics 21 (15), 3201–3212.
Hwang, H., DeSarbo, W., Takane, Y., 2007. Fuzzy clusterwise generalized structured component analysis. Psychometrika 72 (2), 181–198.
Ichihashi, H., 2000. Gaussian mixture PDF approximation and fuzzy c-means clustering with entropy regularization. In: Proc. 4th Asian Fuzzy Systems Symposium, 2000. pp. 217–221.
Ichino, M., Yaguchi, H., 1994. Generalized Minkowski metrics for mixed feature-type data analysis. IEEE Trans. Syst. Man Cybern. 24 (4), 698–708.
Kabir, S., Wagner, C., Havens, T.C., Anderson, D.T., Aickelin, U., 2017. Novel similarity measure for interval-valued data based on overlapping ratio. In: 2017 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE, IEEE, pp. 1–6.
Kahali, S., Sing, J.K., Saha, P.K., 2019. A new entropy-based approach for fuzzy c-means clustering and its application to brain MR image segmentation. Soft Comput. 23 (20), 10407–10414.
Kaufman, L., Rousseeuw, P., 2005. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons.
Krishnapuram, R., Joshi, A., Nasraoui, O., Yi, L., 2001. Low-complexity fuzzy relational clustering algorithms for web mining. IEEE Trans. Fuzzy Syst. 9 (4), 595–607.
Krishnapuram, R., Joshi, A., Yi, L., 1999. A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering. In: Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE'99. 1999 IEEE International, Vol. 3. IEEE, pp. 1281–1286.
Li, R.-P., Mukaidono, M., 1995. A maximum-entropy approach to fuzzy clustering. In: Proceedings of 1995 IEEE International Conference on Fuzzy Systems, Vol. 4. IEEE, pp. 2227–2232.
Liew, A., Leung, S., Lau, W., 2000. Fuzzy image clustering incorporating spatial continuity. IEE Proc., Vis. Image Signal Process. 147 (2), 185–192.
Liew, A.-C., Leung, S.H., Lau, W.H., 2003. Segmentation of color lip images by spatial fuzzy clustering. IEEE Trans. Fuzzy Syst. 11 (4), 542–549.
López-Oriona, Á., D'Urso, P., Vilar, J.A., Lafuente-Rego, B., 2022. Spatial weighted robust clustering of multivariate time series based on quantile dependence with an application to mobility during COVID-19 pandemic. IEEE Trans. Fuzzy Syst. 30 (9), 3990–4004.

Maharaj, E., D'Urso, P., 2011. Fuzzy clustering of time series in the frequency domain. Inform. Sci. 181 (7), 1187–1211.
Maharaj, E.A., Teles, P., Brito, P., 2019. Clustering of interval time series. Stat. Comput. 29 (5), 1011–1034.
Miyagishi, K., Yasutomi, Y., Ichihashi, H., Honda, K., 2000. Fuzzy clustering with regularization by KL information. In: 16th Fuzzy System Symposium. pp. 549–550.
Newman, M.E.J., 2003. Mixing patterns in networks. Phys. Rev. E 67, 026126.
Páez, A., Scott, D., 2005. Spatial statistics for urban analysis: A review of techniques with examples. GeoJournal 61, 53–67.
Pham, D., 2001. Spatial models for fuzzy clustering. Comput. Vis. Image Underst. 84 (2), 285–297.
Pham, D.L., Prince, J.L., 1999. Adaptive fuzzy segmentation of magnetic resonance images. IEEE Trans. Med. Imaging 18 (9), 737–752.
Tolias, Y.A., Panas, S.M., 1998a. Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions. IEEE Trans. Syst. Man Cybern. A 28 (3), 359–369.
Tolias, Y.A., Panas, S.M., 1998b. On applying spatial constraints in fuzzy image clustering using a fuzzy rule-based system. IEEE Signal Process. Lett. 5 (10), 245–247.
Yao, J., Dash, M., Tan, S., Liu, H., 2000. Entropy-based fuzzy clustering and fuzzy modeling. Fuzzy Sets Syst. 113 (3), 381–388.
Zarinbal, M., Zarandi, M.F., Turksen, I., 2014. Relative entropy fuzzy c-means clustering. Inform. Sci. 260, 74–97.