Spatial Statistics 69 (2025) 100914

Contents lists available at ScienceDirect

Spatial Statistics

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

# Spatial robust fuzzy clustering of mixed data with electoral study

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

[a] Department of Social Science and Economics, Sapienza University, Piazzale Aldo Moro 5, 00185 Rome, Lazio, Italy
[b] Department of AI, Data and Decision Sciences, Luiss University, Viale Romania 32, 00197 Rome, Lazio, Italy
[c] Data Lab, Luiss University, Viale Pola 12, 00198 Rome, Lazio, Italy

ARTICLE INFO

*Keywords:*
Mixed data
Fuzzy *C*-medoids clustering
Contiguity matrix
Robustness
European elections

ABSTRACT

A robust fuzzy clustering model for data with mixed features and spatial constraints is proposed to analyze the turnout and the preferences of the voters at the provincial level in the European elections. The 2024 European elections in Italy were held in June to elect the 76 members of the European Parliament due to Italy. The clustering model accommodates various types of variables or attributes by integrating dissimilarity measures for each one through a weighting approach. This method produces a composite distance (or dissimilarity) metric that captures multiple attribute types. The weights are determined objectively during the optimization process and indicate the importance of each attribute type. The model also incorporates robustness via the introduction of a Noise cluster, and accounts for a spatial component. The application shows consistency of the results both at the level of units' attributes and at a spatial level.

## 1. Introduction

Datasets may contain information that is not embedded in numeric variables or attributes. For instance, socio-economic data often include a variety of variables, some quantitative (education, wage, labor experience, etc.), and some qualitative (gender, marital status, employment status, etc.). In the case of longitudinal socio-economic datasets, among the quantitative variables, some are time-invariant, at least over a given period (e.g., years of education, household size), while others vary over time (e.g., wage, labor experience); qualitative variables may also vary over time, especially when units are observed for long periods (e.g., marital status, employment status), yielding ordered sequences of items. Recently, the importance of processing spatial data with mixed-type attributes has become more prominent due to the wide availability of remote sensing geographical data, for example, to predict landslide susceptibility (Ado et al., 2022), or detect different types of crops (Abdali et al., 2023), or for landfill sites optimization (Kontos and Zevgolis, 2024). These datasets often include visual data, quantitative topographic data, and qualitative data on soil composition. Hence, there is a need to apply clustering algorithms to data with mixed attributes. When more than one attribute type is collected, ignoring one or more in the clustering process could compromise the final results. Most clustering algorithms are designed for a single data type. A first approach to managing mixed variables involves preprocessing to convert all variables to a uniform type — either all numeric or all categorical (Guha et al., 1999). A second approach uses a dissimilarity measure capable of handling mixed data, possibly by assigning a weighting scheme to account for the relevance of each attribute type (Gower, 1971). In this paper, the second approach is adopted within a fuzzy framework (see, e.g., Antoni et al., 2014). Tables 1 and 2 in D'Urso and Massari (2019) provide clustering methods and an admittedly non-exhaustive list of papers that address mixed data. Mixed data in fuzzy clustering models are also discussed in D'Urso et al. (2023b). Several clustering techniques for spatial

* Corresponding author.
*E-mail address:* dcangemi@luiss.it (D. Cangemi).

https://doi.org/10.1016/j.spasta.2025.100914
Received 24 April 2025; Received in revised form 16 June 2025; Accepted 18 June 2025
Available online 2 July 2025

units have been proposed in the literature. The approach followed in this paper belongs to the broad class of spatially constrained clustering techniques (Hu and Sung, 2006; Ambroise and Dang, 2009; Viroli, 2011; Torabi, 2016). These models include a spatial penalization term in the objective function. The role of this term, and the related tuning parameter, is to smooth the membership degrees of all units adjacent to the generic $i$th unit for clusters to which the $i$th unit does not belong. Spatial constraints in fuzzy clustering models have been addressed in D'Urso et al. (2019, 2022, 2023a). The main goal of this paper is, to our knowledge, to fill a gap by presenting a robust clustering model for mixed data with spatial regularization.

The characteristics of the proposed model are:

*Mixed data*: the proposed clustering model handles mixed data by combining dissimilarity measures for each attribute via a weighting scheme, resulting in an overall distance (or dissimilarity) measure for multiple attributes;

*Clustering procedure*: using the PAM (Partitioning Around Medoids) approach, cluster prototypes (i.e., medoids) are actual observed units, not "virtual" ones like the "centroids" in fuzzy c-means clustering (Bezdek, 1981). The use of real representative units makes interpretation of the resulting clusters easier (Kaufman and Rousseeuw, 2005). Additionally, the PAM method offers a "timid robustification" of c-means clustering (García-Escudero and Gordaliza, 1999; García-Escudero et al., 2010);

*Fuzziness*: Fuzzy clustering methods offer significant advantages over conventional (hard) clustering techniques, particularly by allowing observations to exhibit partial membership across multiple clusters. Unlike traditional clustering, which assigns observations exclusively to a single cluster, fuzzy clustering provides a nuanced representation of data structures. This effectively captures inherent uncertainty and overlapping characteristics frequently present in real-world datasets, enhancing flexibility while accurately managing complexity. Specifically, fuzzy clustering is more appealing than traditional methods when cluster boundaries are ambiguous (McBratney and Moore, 1985; Wedel and Kamakura, 2000). Membership degrees reveal whether a second-best cluster is nearly as good as the best one, a feature unavailable in traditional clustering (Everitt et al., 2011). Furthermore, fuzzy clustering aligns well with distribution-free methods (Hwang et al., 2007) and is computationally efficient (McBratney and Moore, 1985; Heiser and Groenen, 1997). See D'Urso (2015) for more details;

*Spatial information*: the proposed clustering model incorporates spatial information through a spatial penalty term based on the following assumption: "...when a spatial unit belongs to a cluster with a high membership degree, then the penalty term forces the neighboring spatial units to have high membership degrees in the cluster, as much as possible. In other words, it is expected that a spatial unit with high (low) membership in one cluster will have neighbors with low (high) membership in the remaining clusters. Hence, the spatial penalty term aims to spatially smooth membership degrees under the empirical evidence that neighboring units often share similar features. Nevertheless, neighboring units can also be described by rather distinct profiles. A parameter controls the strength of this spatial penalty in the clustering process" (Coppi et al., 2010);

*Robustness*: robustness against outliers is achieved by introducing a noise cluster, provided all noise points can be assigned to it. According to Davé (1991), "Noise prototype is a universal entity such that it is always at the same distance from every point in the dataset". Given the noise cluster distance, objects closer to the noise cluster than any other will be classified into it. The noise cluster is represented by a noise prototype, equidistant from all units.

An electoral application is considered. The variables considered are voter turnout, which can be measured by taking into account several factors (Geys, 2006), and voter preferences at the provincial level. Most studies focus on the ratio of voters to the voting-age population or to the number registered to vote. Others use the total votes cast or the ratio to eligible voters. Ultimately, it is difficult to identify a definitively superior turnout measure. In this paper, turnout is measured as the share of the age-eligible population that votes, available at provincial level.

In recent years, researchers have increasingly emphasized the geographic influences on voting behavior. They argue that voting is not solely explained by individual or aggregate characteristics, but is the outcome of a complex, multidimensional process influenced by the social and geographical environment in which individuals are located (Pattie and and, 2000; Eagles, 1995). These localized factors stem from shared backgrounds and values and may influence behavior via mechanisms like social networks and political discussions (Huckfeldt, 1995) or shared economic conditions (Pattie et al., 2015; McKay, 2019). The study by Fiorino et al. (2021) investigates spillover effects between regions in explaining turnout variation in European Parliamentary elections. De Giovanni et al. (2023) uses Dynamic Panel Data with Instrumental Variables to examine potential endogeneity in the relationship between the press and voting behavior at the regional level in Italy. This literature suggests that localized analyses may better explain variations in turnout and preferences than general models.

In this paper, a robust fuzzy clustering model for mixed data with spatial constraints is proposed to analyze voter turnout and preferences at the provincial level in European elections. The 2024 European elections in Italy were held in June to elect the 76 members of the European Parliament allocated to Italy. The resulting clusters are profiled using socio-economic variables.

The paper is structured as follows: Section 2 presents the model and the algorithm procedure, Section 3 shows the application of the model to the 2024 European elections and conducts a fuzzy profiling of the resulting clusters, Section 4 reports the conclusions.

## 2. The model

Let $\mathcal{X} = \{X_1, \dots, X_P\}$ be a set of $P$ variables of mixed type; let $S$ be the number of different types, and let $p_s$ be the number of variables of type $s$, for $s = 1, \dots, S$, so that $\sum_{s=1}^{S} p_s = P$. Let $\mathcal{X}_{is}$ denote the set of values observed for the $i$th unit on the $p_s$ variables of type $s$. Let ${}_s d$ be a distance (or dissimilarity measure, in general) for variables of type $s$: the distance between units $i$ and $j$ relative to type $s$ is

$$ {}_s d_{ij} = d(\mathcal{X}_{is}, \mathcal{X}_{js}) $$

The distance between units $i$ and $j$ is then computed as the overall weighted squared sum of distances relative to each type:

$$ d_{ij}^2 = \sum_{s=1}^{S} (w_s \cdot {}_s d_{ij})^2 = \sum_{s=1}^{S} (w_s \cdot d(\mathcal{X}_{is}, \mathcal{X}_{js}))^2 \tag{1} $$

where the weights $w_s$ are not exogenous, but are parameters to be determined within the clustering procedure themselves. Indeed, $w_s$ is a measure of the within-cluster similarity among variables of type $s$, since its value increases as the intra-cluster deviance decreases (cfr. Proposition 1); this weighting system thus agrees with the fact that attributes that cause a better separation into clusters should have a more significant role than the rest, other than having the advantage of determining the weights' values autonomously, as opposed to having the user input them.

As explained in Deza and Deza (2009), such defined $d$ is a valid distance over the product of all $s$ attribute spaces as long as every ${}_s d$ is.

The clustering process follows the Partitioning-Around-Medoids (PAM) algorithm, identifying a most representative unit for each one of the "good" $C-1$ clusters (Davé, 1991). The robustness of the model is attained by introducing a $C$th cluster, referred to as *Noise cluster*, which outliers end up having the highest membership to; in order to do so, a *Noise prototype* is introduced, which is assumed to have a fixed distance $\delta$ from all units. The parameter $\delta$ plays a crucial role in outlier identification: if it is set too small, normal data points may be erroneously absorbed into the noise cluster; conversely, if it is too large, true outliers may be misclassified as members of regular clusters. Therefore, $\delta$ must be chosen to reflect the underlying structure of the dataset, and it is hereby computed as a statistical average of interpoint distances, tuned by an exogenous parameter $\lambda$:

$$ \delta^2 = \frac{\lambda}{n(C-1)} \left( \sum_{i=1}^{n} \sum_{c=1}^{C-1} d_{ic}^2 \right) \tag{2} $$

Let $\mathbf{U} = (u_{ic})_{\substack{i=1,\dots,n \\ c=1,\dots,C}}$ be a $(n \times C)$ matrix with values in $[0,1]$, where $u_{ic}$ denotes the degree of membership of unit $i$ to cluster $c$; once the $C-1$ medoids are determined, so that one can compute the (overall weighted) distance $d_{ic}$ between each unit $i$ and each medoid $c$, building upon the FCMd-MD-SP model (D'Urso et al., 2025), the clustering algorithm translates into the optimization problem:

$$ \begin{aligned} \min_{U,w_s} : \quad & \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m d_{ic}^2 + \frac{\beta}{2} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m + \sum_{i=1}^{n} u_{iC}^m \delta^2 \\ & = \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m \Big( \sum_{s=1}^{S} w_s^2 \cdot {}_s d_{ic}^2 \Big) + \frac{\beta}{2} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m + \sum_{i=1}^{n} u_{iC}^m \delta^2 \\ & = \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m \Big( \sum_{s=1}^{S} w_s^2 \cdot d^2(\mathcal{X}_{is}, \mathcal{X}_{cs}) \Big) + \frac{\beta}{2} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u_{ic}^m \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ii'} u_{i'c'}^m + \sum_{i=1}^{n} u_{iC}^m \delta^2 \\ \text{s.t.} : \quad & u_{ic} \geq 0 \quad \text{and} \quad \sum_{c=1}^{C} u_{ic} = 1 \qquad \text{for} \quad i = 1, \dots, n, \quad c = 1, \dots, C \\ & w_s \geq 0 \quad \text{and} \quad \sum_{s=1}^{S} w_s = 1 \end{aligned} \tag{3} $$

where

- $u_{ic}$ denotes the membership of the $i$th unit to the $c$th cluster
- $\mathcal{X}_{cs}$ denotes the $s$th component of the $c$th medoid
- $C-1$ is the number of "actual" clusters, cluster $C$ being the noise cluster
- the parameter $m > 1$ controls the partition's fuzziness
- $d_{ic}^2$ is the weighted squared distance between unit $i$ and medoid $c$, among all $S$ types
- $C_c = \{1, \dots, C\} \setminus \{c\}$ is the set of all cluster indices except for $c$;
- the parameter $\beta > 0$ tunes the prominence of the spatial penalty term
- $\delta^2$ is the noise distance defined in (2)
- $w_s$ is the weighting parameter associated to attributes' type $s$
- $a_{ij}$ reflects the contiguity of units $i$ and $j$, i.e. it is the element in the $i$th row and $j$th column of the adjacency matrix

The spatial penalty term in the optimization process ensures that, if unit $i$ has high membership degree to cluster $c$, the membership degrees of its adjacent units to clusters different than $c$ are as lowest as possible.

### 2.1. Derivation of optimizing solutions

**Proposition 1.** *The solutions to the minimization problem (3) are*

$$u_{jc} = \frac{\left(d^2_{jc} + \beta \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'}\right)^{\frac{1}{1-m}}}{\sum_{c=1}^{C-1} \left(d^2_{jc} + \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'}\right)^{\frac{1}{1-m}} + \delta^{\frac{2}{1-m}}} \qquad \text{for} \quad c = 1, \dots, C-1 \tag{4}$$

$$u_{jc} = \frac{\delta^{\frac{2}{1-m}}}{\sum_{c=1}^{C-1} \left(d^2_{jc} + \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'}\right)^{\frac{1}{1-m}} + \delta^{\frac{2}{1-m}}} \qquad \text{for} \quad c = C \tag{5}$$

$$w_r = \Big( \sum_{s=1}^{S} \frac{\sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} \cdot {}_r d^2_{ic} + \frac{\lambda}{n(C-1)} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{iC} \cdot {}_r d^2_{ic}}{\sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} \cdot {}_s d^2_{ic} + \frac{\lambda}{n(C-1)} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{iC} \cdot {}_s d^2_{ic}} \Big)^{-1} \qquad \text{for} \quad r = 1, \dots, S \tag{6}$$

**Proof.** For $j = 1, \dots, n$, let $\mathbf{u}_j = (u_{j1}, u_{j2}, \dots, u_{jC})$ and let the Lagrangian function be:

$$L(\mathbf{u}_j, \gamma) = \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} d^2_{ic} + \frac{\beta}{2} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ii'} u^m_{i'c'}$$
$$+ \sum_{i=1}^{n} u^m_{iC} \delta^2 - \gamma \Big( \sum_{c=1}^{C} u_{jc} - 1 \Big)$$

Differentiating with respect to $\gamma$ leads to:

$$\frac{\partial L}{\partial \gamma} = \sum_{c=1}^{C} u_{jc} - 1 = 0 \Longleftrightarrow \sum_{c=1}^{C} u_{jc} = 1 \tag{7}$$

As concerns the derivative with respect to $u_{jc}$, if $1 \le c < C$ then:

$$\frac{\partial L}{\partial u_{jc}} = m u^{m-1}_{jc} d^2_{jc} + \beta m u^{m-1}_{jc} \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'} - \gamma$$

so that

$$\frac{\partial L}{\partial u_{jc}} = 0 \Longleftrightarrow u_{jc} = \Big( \frac{\gamma}{m(d^2_{jc} + \beta \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'})} \Big)^{\frac{1}{m-1}} \tag{8}$$

For $c = C$, in turn:

$$\frac{\partial L}{\partial u_{jC}} = m u^{m-1}_{jC} \delta^2 - \gamma = 0 \quad \Longleftrightarrow \quad u_{jC} = \Big( \frac{\gamma}{m\delta^2} \Big)^{\frac{1}{m-1}} \tag{9}$$

Introducing the notation

$$\theta_{jc} = \begin{cases} d^2_{jc} + \beta \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ji'} u^m_{i'c'} & \text{for} \quad 1 \le c < C \\ \delta^2 & \text{for} \quad c = C \end{cases}$$

due to (7), one can thus write

$$1 = \sum_{c=1}^{C} u_{jc} = \sum_{c=1}^{C} \Big( \frac{\gamma}{m\theta_{jc}} \Big)^{\frac{1}{m-1}} \quad \Longleftrightarrow \quad \Big( \frac{\gamma}{m} \Big)^{\frac{1}{m-1}} = \frac{1}{\sum_{c=1}^{C} \theta_{jc}^{\frac{1}{1-m}}}$$

so that (8) and (9) can be rewritten as:

$$u_{jc} = \frac{\theta_{jc}^{\frac{1}{1-m}}}{\sum_{c=1}^{C} \theta_{jc}^{\frac{1}{1-m}}} \qquad 1 \le c \le C, \quad 1 \le j \le n \tag{10}$$

Next, in order to compute the expressions of the weights $\mathbf{w} = (w_1, \dots, w_S)$, we consider the Lagrangian function

$$L(\mathbf{w}, \xi) = \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} (\sum_{s=1}^{S} w^2_s \cdot {}_s d^2_{ic}) + B + \frac{\lambda}{n(C-1)} \sum_{i=1}^{n} u^m_{iC} \sum_{s=1}^{S} \sum_{c=1}^{C-1} w^2_s \cdot {}_s d^2_{ic} - \xi \Big( \sum_{s=1}^{S} w_s - 1 \Big)$$

where $B = \frac{\beta}{2} \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} \sum_{i'=1}^{n} \sum_{c' \in C_c} a_{ii'} u^m_{i'c'}$ is the spatial penalty term. Thus the partial derivatives with respect to $w_r$ and $\xi$ are

$$\frac{\partial L}{\partial w_r} = 2 \sum_{i=1}^{n} \sum_{c=1}^{C-1} u^m_{ic} w_r \cdot {}_r d^2_{ic} + \frac{2\lambda}{n(C-1)} \sum_{i=1}^{n} u^m_{iC} \sum_{c=1}^{C-1} w_r \cdot {}_r d^2_{ic} - \xi \tag{11}$$

$$\frac{\partial L}{\partial \xi} = \sum_{s=1}^{S} w_s - 1 \tag{12}$$

Setting (11) equal to zero yields to

$$w_r\Big(\sum_{i=1}^{n}\sum_{c=1}^{C-1} u_{ic}^m \cdot {}_r d_{ic}^2 + \frac{\lambda}{n(C-1)}\sum_{i=1}^{n}\sum_{c=1}^{C-1} u_{iC}^m \cdot {}_r d_{ic}^2\Big) = \frac{\xi}{2} \tag{13}$$

which we rewrite as $w_r = \frac{\xi}{2}\frac{1}{\sigma_r}$, having set

$$\sigma_r := \sum_{i=1}^{n}\sum_{c=1}^{C-1} u_{ic}^m \cdot {}_r d_{ic}^2 + \frac{\lambda}{n(C-1)}\sum_{i=1}^{n}\sum_{c=1}^{C-1} u_{iC}^m \cdot {}_r d_{ic}^2 \,.$$

Thus, by (12):

$$1 = \sum_{s=1}^{S} w_s = \frac{\xi}{2}\sum_{s=1}^{S}\frac{1}{\sigma_s} \quad \Longleftrightarrow \quad \frac{\xi}{2} = \frac{1}{\sum_{s=1}^{S}\frac{1}{\sigma_s}}$$

so that (13) becomes

$$w_r = \frac{1}{\sum_{s=1}^{S}\frac{\sigma_r}{\sigma_s}} \quad \square$$

The clustering algorithm alternates between updating the membership matrix $\mathbf{U}$ and the weights $w_s$ according to the equations of Proposition 1 and updating the medoids via brute-force search. Its structure is hereby illustrated in Algorithm 1.

**Algorithm 1** Robust Fuzzy C-Medoids Clustering for Mixed Data and Spatial constraints (RFCMd-MDSP) algorithm

Fix $C$ and $max.iter$; fix $conv = 10^{-9}$ and initialize $\mathbf{U}$ randomly;
Set $iter = 0$;
Pick initial medoids: $\tilde{\mathbf{x}} = \{\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{C-1}\}$;
**repeat**
  Store current membership matrix: $\mathbf{U}_{\mathbf{old}} = \mathbf{U}$
  Select new medoids $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{C-1}$ as follows:
  **for** $c = 1$ to $C-1$ **do**
    $q = \underset{1\le i'\le n}{\arg\min} \sum_{i''=1}^{n} u_{i''c}^m \sum_{s=1}^{S}(w_s \cdot {}_s d_{i',i''})^2$
    **return** $\Rightarrow \tilde{\mathbf{x}}_{\mathbf{c}} = \tilde{\mathbf{x}}_{\mathbf{q}}$
  **end for**
  Update $\mathbf{w}$ by using (6);
  Update $\delta$ by using (2);
  Update $\mathbf{u}_i$ $(i = 1, \dots, n)$ by using (4) and (5);
  $iter \leftarrow iter + 1$;
**until** $\|\mathbf{U}_{\mathbf{old}} - \mathbf{U}\| < conv$ or $iter = max.iter$

### 2.2. Validity measure

We assess the quality of an output partition $U$ via the Xie-Beni index (Xie and Beni, 1991), as it is a standard benchmark in fuzzy clustering literature:

$$F(U) = \frac{\sum_{i=1}^{n}\sum_{c=1}^{C-1} u_{ic}^m d_{ic}^2}{n \cdot \min_{c,c'} d_{cc'}^2}, \qquad c, c' = 1, \dots, C-1 \tag{14}$$

This validity index measures how compact the clusters are (at the numerator) versus how well separated they are (at the denominator). This implies that a lower value of $F(U)$ translates to a better partition $U$ in terms of the above mentioned properties. We only compute compactness and separateness of the "good" clusters $c = 1, \dots, C-1$.

## 3. Application to european electoral data

The model was applied to a dataset related to the 2024 European Parliament elections, recording in particular how the $n = 107$ Italian provinces voted. The first type of variable we take into consideration is qualitative as it reports vote preferences for political parties at provincial level, and we consider, for every unit $i$, the 15-dimensional vector reporting the distribution of votes in province $i$ among the 15 overall political parties. The second type of variable we take into account is quantitative and it reports the *turnout* of province $i$, i.e. the ratio between voters and electors in province $i$. The dissimilarity between distributions is computed as the Cosine dissimilarity, whereas L1 distance is applied for the turnout variable. Both are defined as follows:


**Fig. 1.** Network of Italian provinces.

$$
{}_1d_{ij} = 1 - cos(\hat{\theta}_{ij}) = 1 - \frac{\boldsymbol{x_i} \cdot \boldsymbol{x_j}}{\|\boldsymbol{x_i}\| \cdot \|\boldsymbol{x_j}\|}
$$

$$
{}_2d_{ij} = |y_i - y_j|
$$

$\boldsymbol{x_i}$ being the vector of votes distribution for unit $i$, $\|\boldsymbol{x_i}\|$ denoting the L2 norm of $\boldsymbol{x_i}$, and $y_i$ the turnout for unit $i$. The overall weighted distance is then computed according to (1), once the weights are computed at each iteration.

The adjacency structure for the units comes from the geographical one, thus assigning a link between two nodes (provinces) if they are neighbors, and can be visualized in Fig. 1.

We implemented the computations for our algorithm with a scaled value of the turnout variable, $\tilde{y}_i = \frac{y_i - \min\{y_i\}}{\max\{y_i\} - \min\{y_i\}}$. The box plots for the turnout variable (in both versions) and the distribution of votes variable can be seen in Fig. 2.

The model was tested with varying values for the parameters $C, \beta$ and $\lambda$, each time assessing the partition via (14): the choice for the range of $\beta$ is due to the partitions' sensibility to small changes of its value, whereas, in relation to $\lambda$, it is common to consider a neighborhood of 0.5 (Davé, 1991).

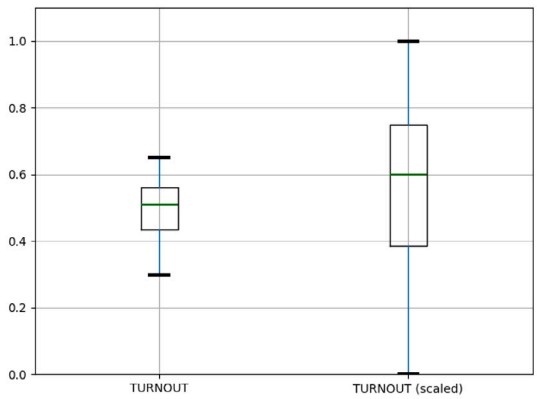
For readability reasons, the results are reported in Table 1 according to each value of $\beta$, showing the corresponding $C$ value and $\lambda$ value that minimize the Xie-Beni index. The overall minimum is attained for $(\beta, C, \lambda) = (0.0016, 4, 0.4)$. Fig. 3 shows the output partition with respect to this choice of parameters, where a unit is assigned to the cluster it has highest membership degree to; fuzzier units are still assigned to a cluster by the same criterion, but their (cluster's) color appears paler in the figure: these units are discussed in detail below. It should be pointed out that, even though the optimal $\beta$ is small, the data are highly sensitive to variations in its value. For comparison's sake, Appendix D shows the output partition for $\beta = 0.04$, $C = 3$ and $\lambda = 2$ (otherwise every unit would end up in the noise cluster): the spatial component becomes predominant, the three "good" clusters depicting three marked areas of Italy.

A sensitivity analysis of the parameter $\lambda$ was also carried out to show the role that the noise distance plays in grouping outliers: for fixed values $(\beta, C) = (0.0016, 4)$, the optimal ones, the output partitions as $\lambda$ increases can be seen in Fig. 4.

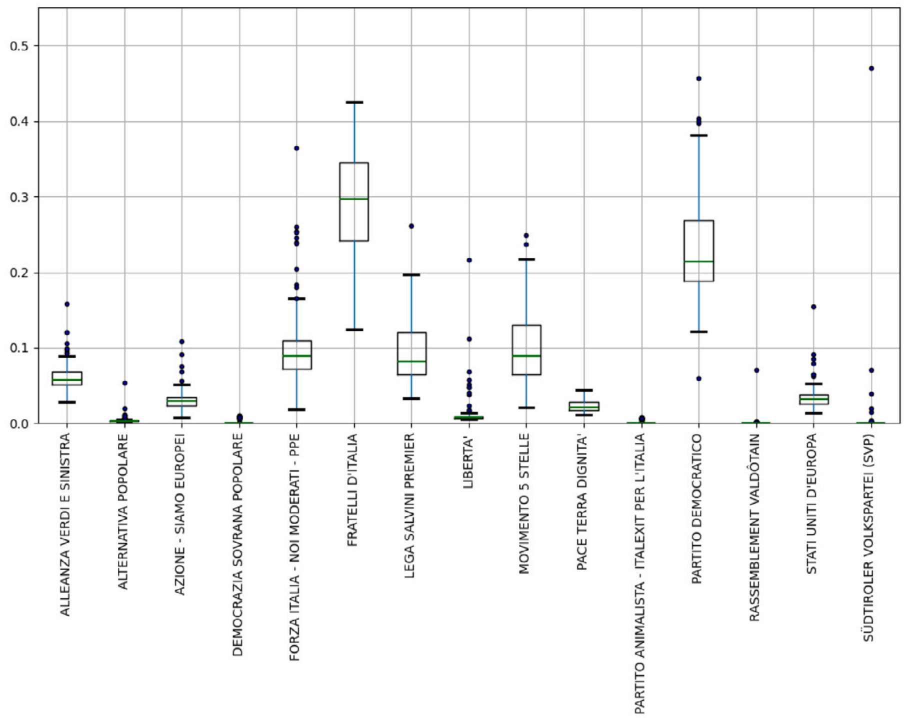
The final selection of weights for the parameters listed above is $(w_1, w_2) = (0.986, 0.014)$, indicating that the turnout's relevance to the partition is nearly null. As discussed in Section 2, this means that the dissimilarity measure of the first type determines well-separated clusters much more significantly than the second one: this is not surprising, since a distribution has superior informative power in general. This selection of weights is not modified, in fact, by changes in the value of the parameter $\beta$. It is also worth mentioning how the proposed model endogenously filters which variables are relevant for the clustering process.

We hereby give a description of each cluster (the full description of the medoids' variables is given in Appendix A):

- **Cluster 1**: covers most of the northern part of the Country and part of the center; its medoid, Alessandria, has high values for right-wing parties Fratelli D'Italia (33%), Lega (12%) and Forza Italia (10%), whereas the left-wing leading party Partito Democratico gained 21% of votes, and centrist party Movimento 5 Stelle only gained 7%.




(a) Voter turnout and scaled turnout by province.




(b) Distribution of party vote shares by province.

**Fig. 2.** Box plots of party vote distributions and electoral turnout at the provincial level.

- **Cluster 2**: covers most of the southern part of the country, and half of the region of Sardinia; results for its medoid Ragusa show above the average outcomes for the party Movimento 5 Stelle (16%) and center-right party Forza Italia (18%), whereas right-wing parties performed worse than average (Fratelli D'Italia, the overall winning party, gained 23% of votes).
- **Cluster 3**: covers the central-northern part of the Country, the region of Puglia and half of the region of Sardinia; as concerns its medoid, Pisa, the leading left-wing party Partito Democratico scored the highest result (32%), also defeating the overall winner of the elections, Fratelli D'Italia (27%); other right-wing parties also scored below the average (Forza Italia 5.5%, Lega 7.5%).

The algorithm labeled five units as belonging to the Noise cluster, as reported in Appendix A. Looking at the voting distributions for these units, some anomalies can be detected that make this outcome coherent. The province of Bolzano, for example, is one of the few units where Südtiroler VolksPartei (SVP) gained votes, and the only one where SVP scored 47%; in the province of Messina,

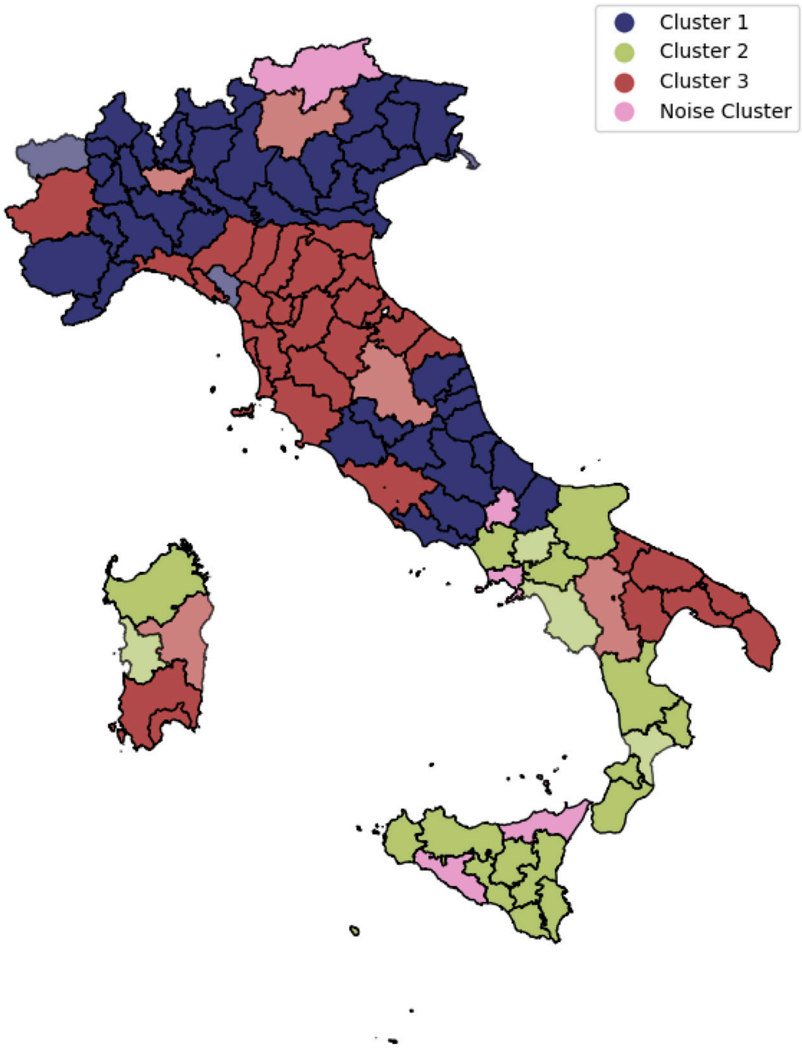



**Fig. 3.** Partition of Italian provinces into 3 clusters plus the noise clusters, with model parameters $(\beta, C, \lambda) = (0.0016, 4, 0.4)$. Fuzzier units are paler than the rest. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**
Minimum values of Xie-Beni index for each value of $\beta$ and $3 \leq C \leq 6$, $0.4 \leq \lambda \leq 0.9$. The overall minimum is stressed in bold.

| $\beta$ | $\min_{C,\lambda} F(U)$ | $C$ | $\lambda$ |
|---|---|---|---|
| 0 | 0.1911 | 4 | 0.5 |
| 0.0008 | 0.1907 | 4 | 0.5 |
| 0.0016 | **0.1741** | 4 | 0.4 |
| 0.0024 | 0.1745 | 4 | 0.4 |
| 0.0032 | 0.3860 | 4 | 0.4 |
| 0.0040 | 0.2138 | 4 | 0.4 |

the party Libertà scored 26%, while gaining 0% in the majority of the remaining units; in the province of Agrigento, the right-wing party Forza Italia scored 36%, far above its average among all units. It is worth mentioning that SVP is a regional political party based in the autonomous province of Bolzano, whose main mission is to represent the German- and Ladin-speaking minorities in the region, defending their cultural and linguistic rights; it also strongly supports regional autonomy. On the opposite, Libertà is a new political party launched in 2024 by Cateno De Luca, a Sicilian politician; Libertà positions itself on the center-right wing and advocates for greater regional autonomy, particularly in southern Italy. What is also worth noticing is that the leading right-wing party and overall winner of the elections, Fratelli D'Italia, scored below average in every one of the units in the Noise clusters.

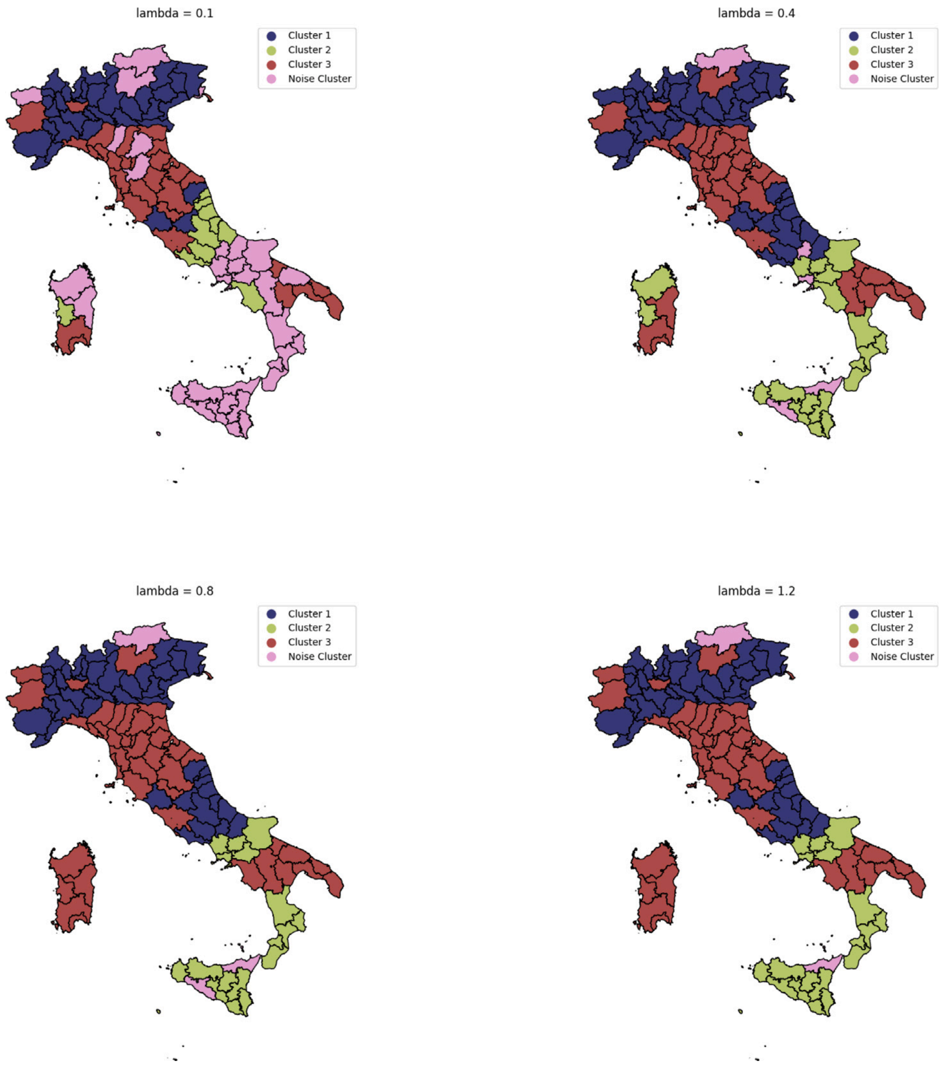



**Fig. 4.** Partition of Italian provinces into 3 clusters plus the noise clusters, with model parameters $(\beta, C) = (0.0016, 4)$ and increasing values of $\lambda$.

Appendix A reports the units' votes distributions as well as the cluster they have been grouped in (the medoids being stressed in bold); the membership values of each unit to each cluster are listed in Appendix B.

If a unit's highest membership value is below 0.6, it is considered to be fuzzy. Table 2 lists the fuzzy units together with their membership to each cluster: as one can see by looking at the distribution of votes, some units' memberships are almost equally split among more than one cluster (the province of Salerno, for example, has almost equal membership degrees to Cluster 1 and 2: this is coherent with the fact that, in Salerno, FDI scored the highest result of 27%, like the majority of Cluster 1 units, but also FI and M5S scored a high result of 10% and 15% respectively, like most units in Cluster 2).

### 3.1. Fuzzy profiling of the clusters

Results of cluster analysis can be summarized in the profiling phase where internal and external variables – i.e., variables involved and not involved in the cluster algorithm, respectively – are used to characterize and interpret the clusters (Everitt et al., 2011; Hair et al., 1998). In the case of fuzzy clustering algorithms, the $(n \times C)$ membership degrees matrix $\mathbf{U}$ can be used to properly weigh the observations on profiling variables and further describe the final clusters (D'Urso et al., 2013, 2016).

Let $X = \{x_1, \ldots, x_n\}$ be a quantitative variable observed on the sample. The weighted average of $X$ in the $c$th cluster is:

$$\mu_{X_c} = \frac{\sum_{i=1}^{n} u_{ic} x_i}{\sum_{i=1}^{n} u_{ic}}. \tag{15}$$

**Table 2**
Membership degrees of fuzzy units.

| Province | Cluster 1 | Cluster 2 | Cluster 3 | Noise cluster |
|---|---|---|---|---|
| Aosta | 0.485 | 0.018 | 0.293 | 0.205 |
| Benevento | 0.045 | 0.540 | 0.015 | 0.401 |
| Catanzaro | 0.320 | 0.563 | 0.020 | 0.098 |
| Massa–Carrara | 0.363 | 0.247 | 0.353 | 0.036 |
| Milano | 0.401 | 0.004 | 0.583 | 0.012 |
| Nuoro | 0.034 | 0.392 | 0.436 | 0.138 |
| Oristano | 0.266 | 0.456 | 0.219 | 0.059 |
| Perugia | 0.416 | 0.003 | 0.574 | 0.007 |
| Potenza | 0.275 | 0.214 | 0.315 | 0.195 |
| Salerno | 0.381 | 0.382 | 0.186 | 0.050 |
| Trento | 0.368 | 0.003 | 0.578 | 0.052 |
| Trieste | 0.552 | 0.001 | 0.444 | 0.004 |

**Table 3**
Fuzzy profiling of Clusters 1 to 3.

| | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| GDP | 32 404.820 | 21 254.860 | 33 498.504 |
| Unemployment | 6.530 | 14.291 | 7.291 |
| Graduates | 27.512 | 21.269 | 29.249 |
| Incl. Rent | 0.276 | 0.312 | 0.271 |
| Excl. Rent | 0.303 | 0.344 | 0.299 |
| Turnout | 52.5 | 39.6 | 52.6 |

As it can be seen, the greater is the membership degree of unit $i$ to cluster $c$, the greater is the contribution of observation $x_i$ to the weighted average.

Similarly, let $Y = \{y_1, \dots, y_n\}$ be a categorical variable with $L$ ($L \geq 2$) categories. Let $l$ be the generic category, and $y_{il}$ the observation in the $i$th unit, which is equal to 1 if the category is observed on the $i$th unit and 0 otherwise. The weighted proportion of the $l$th category in the $c$th cluster is:

$$w_{Y_{lc}} = \frac{\sum_{i=1}^{n} y_{il} u_{ic}}{\sum_{i=1}^{n} u_{ic}}. \tag{16}$$

The concept of weighted averages and weighted proportions can be easily extended to other attribute types. For instance, for time series we have the weighted average of the quantitative time-variant variable $X$ observed at time $t$ in the $c$th cluster:

$$\mu_{X_{tc}} = \frac{\sum_{i=1}^{n} u_{ic} x_{it}}{\sum_{i=1}^{n} u_{ic}} \tag{17}$$

where $t = 1, \dots, T$.

In order to carry out the fuzzy profiling of the identified clusters of Italian provinces, we took into account the following economic and social indicators:

- **GDP** (€): GPD at market prices per inhabitant. It indicates the level of local economic activity and allows for comparisons of wealth generation across regions. The values refer to the year 2022
- **Unemployment** (Percentage): Unemployment rate, calculated as percentage of people actively seeking work out of the total labor force (employed and unemployed). The values refer to the year 2022
- **Graduates** (Percentage): Percentage of university graduates by province, calculated as the share of people aged 25–39 with a university degree out of the total population in that age group. The values refer to the year 2021
- **Incl. Rent** and **Excl. Rent** (Gini index): Measure of economic equity or lack thereof, referring to the "imputed rent", i.e. the estimated value of the housing services that homeowners receive by living in their own property, without actually paying rent. When analyzing household income distribution, one can distinguish between income including and excluding imputed rents, the former giving a more accurate picture of economic well-being, as it accounts for the advantage homeowners have over renters. The values are at regional level, so that every province belonging to one region is assigned the same value; they refer to the year 2022
- **Turnout** (Percentage): Same metric used in the clustering process, computed as the ratio between voters and electors in each province. The values refer to the year 2024

Table 3 presents the weighted average values of the above mentioned economic and social indicators, computed according to (15). Cluster 3 emerges as the most economically and socially advanced group, displaying the highest average GDP per capita, the highest share of university graduates, and an unemployment rate close to the lowest. Cluster 1 follows closely in economic performance, with a slightly lower GDP and graduate percentage, yet exhibiting the lowest unemployment rate among the three clusters.

In contrast, Cluster 2 appears socioeconomically disadvantaged, characterized by the lowest GDP, the highest unemployment rate, and the smallest share of graduates. Moreover, it shows the highest levels of income inequality, as indicated by the greater gap between income including and excluding imputed rents. Voter turnout is also noticeably lower in Cluster 2 compared to Clusters 1 and 3, suggesting a potential link between economic conditions and civic engagement. Appendix C reports the full list of indicator values for each province.

## 4. Conclusions

In this paper, we proposed a robust version of the *Fuzzy C-medoids clustering for mixed data model with spatial constraints* (FCMd-MD-SP) model (D'Urso et al., 2025), accounting for a Noise cluster aimed at detecting the outliers. The model benefits from a weighted distance taking into account variables of different type, where the weights themselves are to be determined within the clustering procedure to optimize the partition into well-separated clusters. In addition, the model interpolates between the units' attributes vector and their spatial (adjacency) structure, due to the spatial penalty term, in determining the optimal partition.

An application to the 2024 European elections data is carried out, pertaining Italy at a provincial level; both the province's votes distribution and turnout are taken into account. Following the output partition into optimal clusters, a fuzzy profiling of the clusters is implemented by computing a score based on 6 socio-economic indicators, weighted by the membership degrees of each province to each cluster. The outcome shows significant discrepancy among the clusters' scores, hinting at a possible interconnection between the economic and social context and the electoral outcome of the provinces. Further research could build upon the descriptive findings in Section 3 by incorporating causal and motivational analyses, to enrich the interpretation and implications of the results.

## Appendix A. Table of votes per province

| Province | AVS | AP | ASE | DSP | FI | FDI | LE | LI | M5S | PTD | PAI | PD | RV | SUD | SVP | Turnout | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Agrigento | 0.031 | 0.010 | 0.009 | 0.000 | 0.364 | 0.177 | 0.074 | 0.038 | 0.138 | 0.012 | 0.000 | 0.131 | 0.000 | 0.016 | 0.000 | 0.365 | Noise |
| **Alessandria** | **0.052** | **0.003** | **0.030** | **0.000** | **0.104** | **0.333** | **0.127** | **0.009** | **0.075** | **0.018** | **0.000** | **0.214** | **0.002** | **0.032** | **0.000** | **0.550** | **1** |
| Ancona | 0.071 | 0.004 | 0.033 | 0.009 | 0.061 | 0.309 | 0.063 | 0.007 | 0.105 | 0.031 | 0.000 | 0.275 | 0.000 | 0.033 | 0.000 | 0.520 | 3 |
| Aosta | 0.121 | 0.003 | 0.076 | 0.000 | 0.079 | 0.243 | 0.089 | 0.007 | 0.046 | 0.031 | 0.000 | 0.200 | 0.071 | 0.034 | 0.000 | 0.425 | 1 |
| Arezzo | 0.057 | 0.004 | 0.029 | 0.008 | 0.072 | 0.320 | 0.067 | 0.006 | 0.078 | 0.022 | 0.000 | 0.295 | 0.000 | 0.044 | 0.000 | 0.606 | 3 |
| Ascoli Piceno | 0.046 | 0.005 | 0.029 | 0.009 | 0.072 | 0.363 | 0.072 | 0.008 | 0.113 | 0.029 | 0.000 | 0.229 | 0.000 | 0.025 | 0.000 | 0.579 | 1 |
| Asti | 0.069 | 0.004 | 0.026 | 0.000 | 0.118 | 0.323 | 0.136 | 0.009 | 0.075 | 0.020 | 0.000 | 0.188 | 0.002 | 0.031 | 0.000 | 0.569 | 1 |
| Avellino | 0.055 | 0.003 | 0.028 | 0.000 | 0.147 | 0.194 | 0.061 | 0.008 | 0.162 | 0.018 | 0.008 | 0.230 | 0.000 | 0.085 | 0.000 | 0.472 | 2 |
| Bari | 0.031 | 0.003 | 0.009 | 0.000 | 0.061 | 0.236 | 0.042 | 0.014 | 0.104 | 0.015 | 0.005 | 0.457 | 0.000 | 0.022 | 0.000 | 0.475 | 3 |
| Barletta-A.-T. | 0.039 | 0.007 | 0.029 | 0.000 | 0.076 | 0.294 | 0.051 | 0.005 | 0.124 | 0.013 | 0.003 | 0.331 | 0.000 | 0.027 | 0.000 | 0.382 | 3 |
| Belluno | 0.065 | 0.004 | 0.035 | 0.000 | 0.076 | 0.367 | 0.116 | 0.010 | 0.043 | 0.022 | 0.000 | 0.193 | 0.000 | 0.030 | 0.039 | 0.431 | 1 |
| Benevento | 0.043 | 0.003 | 0.017 | 0.000 | 0.166 | 0.206 | 0.076 | 0.007 | 0.141 | 0.019 | 0.007 | 0.161 | 0.000 | 0.154 | 0.000 | 0.473 | 2 |
| Bergamo | 0.054 | 0.003 | 0.031 | 0.000 | 0.082 | 0.352 | 0.162 | 0.006 | 0.041 | 0.015 | 0.000 | 0.225 | 0.001 | 0.026 | 0.000 | 0.597 | 1 |
| Biella | 0.060 | 0.004 | 0.031 | 0.000 | 0.100 | 0.373 | 0.111 | 0.008 | 0.061 | 0.019 | 0.000 | 0.185 | 0.003 | 0.046 | 0.000 | 0.596 | 1 |
| Bologna | 0.085 | 0.003 | 0.032 | 0.000 | 0.046 | 0.245 | 0.045 | 0.006 | 0.073 | 0.028 | 0.000 | 0.403 | 0.000 | 0.033 | 0.001 | 0.608 | 3 |
| Bolzano | 0.158 | 0.005 | 0.068 | 0.000 | 0.019 | 0.124 | 0.033 | 0.012 | 0.021 | 0.015 | 0.000 | 0.060 | 0.000 | 0.014 | 0.470 | 0.496 | Noise |
| Brescia | 0.059 | 0.003 | 0.035 | 0.000 | 0.088 | 0.362 | 0.158 | 0.006 | 0.047 | 0.015 | 0.000 | 0.194 | 0.001 | 0.031 | 0.000 | 0.605 | 1 |
| Brindisi | 0.036 | 0.002 | 0.016 | 0.000 | 0.115 | 0.283 | 0.065 | 0.009 | 0.156 | 0.015 | 0.005 | 0.271 | 0.000 | 0.027 | 0.000 | 0.395 | 3 |
| Cagliari | 0.106 | 0.003 | 0.020 | 0.000 | 0.104 | 0.236 | 0.052 | 0.006 | 0.152 | 0.042 | 0.000 | 0.253 | 0.000 | 0.027 | 0.000 | 0.427 | 3 |
| Caltanissetta | 0.032 | 0.005 | 0.011 | 0.000 | 0.254 | 0.206 | 0.066 | 0.052 | 0.218 | 0.015 | 0.000 | 0.122 | 0.000 | 0.019 | 0.000 | 0.481 | 2 |
| Campobasso | 0.052 | 0.011 | 0.018 | 0.000 | 0.083 | 0.276 | 0.140 | 0.009 | 0.160 | 0.031 | 0.008 | 0.189 | 0.000 | 0.023 | 0.000 | 0.506 | 1 |
| Caserta | 0.048 | 0.003 | 0.032 | 0.000 | 0.138 | 0.205 | 0.089 | 0.007 | 0.208 | 0.012 | 0.007 | 0.189 | 0.000 | 0.062 | 0.000 | 0.463 | 2 |
| Catania | 0.051 | 0.006 | 0.012 | 0.000 | 0.245 | 0.227 | 0.097 | 0.048 | 0.154 | 0.019 | 0.000 | 0.125 | 0.000 | 0.016 | 0.000 | 0.371 | 2 |
| Catanzaro | 0.061 | 0.003 | 0.047 | 0.000 | 0.138 | 0.224 | 0.146 | 0.012 | 0.142 | 0.018 | 0.005 | 0.170 | 0.000 | 0.035 | 0.000 | 0.368 | 2 |
| Chieti | 0.055 | 0.004 | 0.021 | 0.000 | 0.092 | 0.334 | 0.079 | 0.011 | 0.134 | 0.029 | 0.006 | 0.209 | 0.000 | 0.027 | 0.000 | 0.436 | 1 |
| Como | 0.060 | 0.003 | 0.039 | 0.000 | 0.096 | 0.327 | 0.180 | 0.006 | 0.053 | 0.017 | 0.000 | 0.180 | 0.002 | 0.035 | 0.000 | 0.536 | 1 |
| Cosenza | 0.063 | 0.003 | 0.032 | 0.000 | 0.155 | 0.191 | 0.079 | 0.012 | 0.210 | 0.015 | 0.005 | 0.142 | 0.000 | 0.092 | 0.000 | 0.449 | 2 |
| Cremona | 0.052 | 0.004 | 0.030 | 0.000 | 0.103 | 0.356 | 0.136 | 0.007 | 0.056 | 0.018 | 0.000 | 0.208 | 0.002 | 0.029 | 0.000 | 0.588 | 1 |
| Crotone | 0.048 | 0.003 | 0.034 | 0.000 | 0.184 | 0.193 | 0.090 | 0.008 | 0.209 | 0.011 | 0.003 | 0.169 | 0.000 | 0.047 | 0.000 | 0.338 | 2 |
| Cuneo | 0.057 | 0.004 | 0.033 | 0.000 | 0.114 | 0.347 | 0.133 | 0.009 | 0.053 | 0.023 | 0.000 | 0.189 | 0.002 | 0.037 | 0.000 | 0.603 | 1 |
| Enna | 0.034 | 0.003 | 0.008 | 0.000 | 0.253 | 0.194 | 0.069 | 0.048 | 0.142 | 0.012 | 0.000 | 0.221 | 0.000 | 0.016 | 0.000 | 0.352 | 2 |
| Fermo | 0.053 | 0.004 | 0.035 | 0.010 | 0.098 | 0.353 | 0.080 | 0.008 | 0.093 | 0.029 | 0.000 | 0.211 | 0.000 | 0.026 | 0.000 | 0.487 | 1 |
| Ferrara | 0.048 | 0.003 | 0.024 | 0.000 | 0.072 | 0.347 | 0.088 | 0.007 | 0.066 | 0.019 | 0.000 | 0.302 | 0.000 | 0.024 | 0.001 | 0.599 | 3 |
| Firenze | 0.094 | 0.003 | 0.030 | 0.009 | 0.049 | 0.225 | 0.046 | 0.005 | 0.076 | 0.033 | 0.000 | 0.366 | 0.000 | 0.065 | 0.000 | 0.651 | 3 |
| Foggia | 0.028 | 0.003 | 0.026 | 0.000 | 0.111 | 0.237 | 0.065 | 0.009 | 0.237 | 0.018 | 0.006 | 0.236 | 0.000 | 0.025 | 0.000 | 0.427 | 2 |
| Forli'-Cesena | 0.056 | 0.003 | 0.026 | 0.000 | 0.066 | 0.301 | 0.059 | 0.007 | 0.069 | 0.020 | 0.000 | 0.362 | 0.000 | 0.029 | 0.001 | 0.613 | 3 |
| Frosinone | 0.054 | 0.004 | 0.020 | 0.005 | 0.128 | 0.337 | 0.119 | 0.006 | 0.107 | 0.017 | 0.000 | 0.170 | 0.000 | 0.034 | 0.000 | 0.530 | 1 |
| Genova | 0.084 | 0.003 | 0.041 | 0.000 | 0.076 | 0.241 | 0.074 | 0.009 | 0.112 | 0.031 | 0.000 | 0.287 | 0.002 | 0.039 | 0.000 | 0.493 | 3 |
| Gorizia | 0.066 | 0.003 | 0.027 | 0.000 | 0.057 | 0.247 | 0.195 | 0.011 | 0.065 | 0.032 | 0.000 | 0.251 | 0.000 | 0.026 | 0.020 | 0.514 | 1 |
| Grosseto | 0.055 | 0.004 | 0.026 | 0.009 | 0.066 | 0.343 | 0.072 | 0.009 | 0.082 | 0.028 | 0.000 | 0.268 | 0.000 | 0.039 | 0.000 | 0.544 | 3 |
| Imperia | 0.064 | 0.004 | 0.027 | 0.000 | 0.106 | 0.334 | 0.123 | 0.009 | 0.081 | 0.029 | 0.000 | 0.188 | 0.002 | 0.034 | 0.000 | 0.498 | 1 |
| Isernia | 0.050 | 0.009 | 0.026 | 0.000 | 0.084 | 0.244 | 0.261 | 0.007 | 0.111 | 0.019 | 0.006 | 0.152 | 0.000 | 0.032 | 0.000 | 0.415 | Noise |

| Province | AVS | AP | ASE | DSP | FI | FDI | LE | LI | M5S | PTD | PAI | PD | RV | SUD | SVP | Turnout | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L'Aquila | 0.046 | 0.003 | 0.023 | 0.000 | 0.106 | 0.360 | 0.107 | 0.008 | 0.093 | 0.030 | 0.005 | 0.196 | 0.000 | 0.025 | 0.000 | 0.418 | 1 |
| La Spezia | 0.068 | 0.003 | 0.025 | 0.000 | 0.083 | 0.276 | 0.084 | 0.007 | 0.099 | 0.035 | 0.000 | 0.283 | 0.001 | 0.036 | 0.000 | 0.511 | 3 |
| Latina | 0.051 | 0.003 | 0.021 | 0.004 | 0.165 | 0.348 | 0.091 | 0.005 | 0.095 | 0.017 | 0.000 | 0.150 | 0.000 | 0.048 | 0.000 | 0.447 | 1 |
| Lecce | 0.054 | 0.002 | 0.016 | 0.000 | 0.069 | 0.319 | 0.092 | 0.008 | 0.120 | 0.016 | 0.008 | 0.271 | 0.000 | 0.025 | 0.000 | 0.463 | 3 |
| Lecco | 0.066 | 0.003 | 0.041 | 0.000 | 0.092 | 0.335 | 0.136 | 0.006 | 0.048 | 0.016 | 0.000 | 0.220 | 0.002 | 0.034 | 0.000 | 0.582 | 1 |
| Livorno | 0.075 | 0.008 | 0.022 | 0.010 | 0.050 | 0.252 | 0.066 | 0.006 | 0.112 | 0.036 | 0.000 | 0.331 | 0.000 | 0.033 | 0.000 | 0.586 | 3 |
| Lodi | 0.056 | 0.003 | 0.030 | 0.000 | 0.097 | 0.350 | 0.135 | 0.006 | 0.064 | 0.016 | 0.000 | 0.211 | 0.002 | 0.030 | 0.000 | 0.588 | 1 |
| Lucca | 0.076 | 0.003 | 0.032 | 0.009 | 0.068 | 0.321 | 0.078 | 0.006 | 0.082 | 0.031 | 0.000 | 0.244 | 0.000 | 0.051 | 0.000 | 0.482 | 3 |
| Macerata | 0.052 | 0.004 | 0.030 | 0.010 | 0.083 | 0.367 | 0.089 | 0.009 | 0.090 | 0.030 | 0.000 | 0.207 | 0.000 | 0.029 | 0.000 | 0.509 | 1 |
| Mantova | 0.056 | 0.005 | 0.039 | 0.000 | 0.077 | 0.339 | 0.125 | 0.006 | 0.067 | 0.018 | 0.000 | 0.237 | 0.002 | 0.030 | 0.000 | 0.523 | 1 |
| Massa–Carrara | 0.065 | 0.003 | 0.025 | 0.008 | 0.156 | 0.251 | 0.077 | 0.006 | 0.090 | 0.041 | 0.000 | 0.247 | 0.000 | 0.031 | 0.000 | 0.469 | 1 |
| Matera | 0.041 | 0.002 | 0.056 | 0.000 | 0.077 | 0.295 | 0.045 | 0.012 | 0.121 | 0.030 | 0.004 | 0.283 | 0.000 | 0.033 | 0.000 | 0.391 | 3 |
| Messina | 0.036 | 0.005 | 0.012 | 0.000 | 0.204 | 0.181 | 0.085 | 0.217 | 0.103 | 0.015 | 0.000 | 0.122 | 0.000 | 0.020 | 0.000 | 0.412 | Noise |
| Milano | 0.088 | 0.003 | 0.049 | 0.000 | 0.091 | 0.269 | 0.085 | 0.006 | 0.065 | 0.023 | 0.000 | 0.272 | 0.001 | 0.049 | 0.000 | 0.529 | 3 |
| Modena | 0.059 | 0.003 | 0.035 | 0.000 | 0.060 | 0.260 | 0.063 | 0.006 | 0.070 | 0.019 | 0.000 | 0.397 | 0.000 | 0.026 | 0.001 | 0.608 | 3 |
| Monza–Brianza | 0.071 | 0.003 | 0.043 | 0.000 | 0.110 | 0.312 | 0.106 | 0.007 | 0.060 | 0.017 | 0.000 | 0.229 | 0.001 | 0.039 | 0.000 | 0.541 | 1 |
| Napoli | 0.089 | 0.002 | 0.048 | 0.000 | 0.087 | 0.157 | 0.045 | 0.007 | 0.249 | 0.017 | 0.006 | 0.242 | 0.000 | 0.052 | 0.000 | 0.420 | Noise |
| Novara | 0.056 | 0.005 | 0.032 | 0.000 | 0.101 | 0.351 | 0.117 | 0.008 | 0.072 | 0.016 | 0.000 | 0.204 | 0.002 | 0.035 | 0.000 | 0.584 | 1 |
| Nuoro | 0.121 | 0.002 | 0.013 | 0.000 | 0.107 | 0.207 | 0.053 | 0.005 | 0.167 | 0.041 | 0.000 | 0.257 | 0.000 | 0.025 | 0.000 | 0.298 | 3 |
| Oristano | 0.099 | 0.007 | 0.015 | 0.000 | 0.098 | 0.280 | 0.060 | 0.006 | 0.161 | 0.038 | 0.000 | 0.213 | 0.000 | 0.025 | 0.000 | 0.310 | 2 |
| Padova | 0.059 | 0.003 | 0.051 | 0.000 | 0.086 | 0.372 | 0.116 | 0.008 | 0.049 | 0.021 | 0.000 | 0.204 | 0.000 | 0.030 | 0.002 | 0.561 | 1 |
| Palermo | 0.062 | 0.005 | 0.019 | 0.000 | 0.239 | 0.192 | 0.050 | 0.040 | 0.195 | 0.020 | 0.000 | 0.152 | 0.000 | 0.026 | 0.000 | 0.389 | 2 |
| Parma | 0.068 | 0.003 | 0.046 | 0.000 | 0.071 | 0.296 | 0.087 | 0.006 | 0.071 | 0.025 | 0.000 | 0.292 | 0.000 | 0.034 | 0.001 | 0.553 | 3 |
| Pavia | 0.054 | 0.003 | 0.033 | 0.000 | 0.100 | 0.313 | 0.185 | 0.007 | 0.063 | 0.018 | 0.000 | 0.191 | 0.002 | 0.031 | 0.000 | 0.562 | 1 |
| Perugia | 0.058 | 0.008 | 0.026 | 0.008 | 0.087 | 0.329 | 0.067 | 0.006 | 0.087 | 0.023 | 0.000 | 0.270 | 0.000 | 0.031 | 0.000 | 0.629 | 3 |
| Pesaro e Urbino | 0.053 | 0.003 | 0.037 | 0.007 | 0.058 | 0.294 | 0.105 | 0.008 | 0.085 | 0.022 | 0.000 | 0.302 | 0.000 | 0.025 | 0.000 | 0.625 | 3 |
| Pescara | 0.057 | 0.003 | 0.022 | 0.000 | 0.115 | 0.325 | 0.069 | 0.008 | 0.119 | 0.041 | 0.008 | 0.207 | 0.000 | 0.026 | 0.000 | 0.544 | 1 |
| Piacenza | 0.053 | 0.005 | 0.031 | 0.000 | 0.089 | 0.365 | 0.102 | 0.007 | 0.064 | 0.020 | 0.000 | 0.237 | 0.000 | 0.025 | 0.002 | 0.552 | 1 |
| **Pisa** | **0.078** | **0.004** | **0.027** | **0.008** | **0.055** | **0.277** | **0.075** | **0.006** | **0.086** | **0.029** | **0.000** | **0.320** | **0.000** | **0.036** | **0.000** | **0.610** | **3** |
| Pistoia | 0.068 | 0.004 | 0.030 | 0.008 | 0.070 | 0.306 | 0.072 | 0.006 | 0.080 | 0.027 | 0.000 | 0.284 | 0.000 | 0.045 | 0.000 | 0.552 | 3 |
| Pordenone | 0.049 | 0.003 | 0.035 | 0.000 | 0.069 | 0.417 | 0.122 | 0.010 | 0.046 | 0.022 | 0.000 | 0.194 | 0.000 | 0.032 | 0.002 | 0.466 | 1 |
| Potenza | 0.047 | 0.003 | 0.108 | 0.000 | 0.094 | 0.239 | 0.058 | 0.017 | 0.130 | 0.022 | 0.006 | 0.211 | 0.000 | 0.065 | 0.000 | 0.445 | 3 |
| Prato | 0.062 | 0.003 | 0.028 | 0.007 | 0.068 | 0.312 | 0.057 | 0.006 | 0.071 | 0.021 | 0.000 | 0.322 | 0.000 | 0.043 | 0.000 | 0.641 | 3 |
| **Ragusa** | **0.062** | **0.002** | **0.014** | **0.000** | **0.180** | **0.231** | **0.053** | **0.057** | **0.160** | **0.025** | **0.000** | **0.184** | **0.000** | **0.032** | **0.000** | **0.311** | **2** |
| Ravenna | 0.059 | 0.003 | 0.032 | 0.000 | 0.056 | 0.272 | 0.061 | 0.008 | 0.070 | 0.025 | 0.000 | 0.381 | 0.000 | 0.032 | 0.002 | 0.581 | 3 |
| Reggio Calabria | 0.058 | 0.020 | 0.039 | 0.000 | 0.260 | 0.211 | 0.082 | 0.010 | 0.097 | 0.013 | 0.005 | 0.168 | 0.000 | 0.038 | 0.000 | 0.383 | 2 |
| Reggio Emilia | 0.065 | 0.003 | 0.032 | 0.000 | 0.059 | 0.239 | 0.060 | 0.007 | 0.082 | 0.023 | 0.000 | 0.400 | 0.000 | 0.028 | 0.002 | 0.609 | 3 |
| Rieti | 0.057 | 0.011 | 0.019 | 0.008 | 0.070 | 0.367 | 0.125 | 0.024 | 0.091 | 0.020 | 0.000 | 0.182 | 0.000 | 0.027 | 0.000 | 0.584 | 1 |
| Rimini | 0.061 | 0.003 | 0.024 | 0.000 | 0.068 | 0.329 | 0.065 | 0.007 | 0.074 | 0.025 | 0.000 | 0.313 | 0.000 | 0.030 | 0.001 | 0.530 | 3 |
| Roma | 0.096 | 0.003 | 0.038 | 0.006 | 0.054 | 0.317 | 0.052 | 0.006 | 0.108 | 0.027 | 0.000 | 0.252 | 0.000 | 0.040 | 0.000 | 0.451 | 3 |
| Rovigo | 0.044 | 0.003 | 0.027 | 0.000 | 0.087 | 0.406 | 0.128 | 0.006 | 0.057 | 0.020 | 0.000 | 0.197 | 0.000 | 0.023 | 0.002 | 0.534 | 1 |
| Salerno | 0.053 | 0.003 | 0.032 | 0.000 | 0.104 | 0.274 | 0.056 | 0.009 | 0.146 | 0.021 | 0.008 | 0.214 | 0.000 | 0.079 | 0.000 | 0.451 | 2 |
| Sassari | 0.099 | 0.005 | 0.015 | 0.000 | 0.095 | 0.252 | 0.059 | 0.006 | 0.187 | 0.033 | 0.000 | 0.228 | 0.000 | 0.023 | 0.000 | 0.406 | 2 |
| Savona | 0.071 | 0.003 | 0.031 | 0.000 | 0.094 | 0.294 | 0.111 | 0.010 | 0.089 | 0.028 | 0.000 | 0.231 | 0.002 | 0.035 | 0.000 | 0.546 | 1 |
| Siena | 0.064 | 0.004 | 0.028 | 0.009 | 0.056 | 0.273 | 0.053 | 0.006 | 0.075 | 0.026 | 0.000 | 0.360 | 0.000 | 0.047 | 0.000 | 0.641 | 3 |
| Siracusa | 0.037 | 0.004 | 0.015 | 0.000 | 0.238 | 0.212 | 0.053 | 0.068 | 0.160 | 0.015 | 0.000 | 0.180 | 0.000 | 0.018 | 0.000 | 0.341 | 2 |
| Sondrio | 0.052 | 0.003 | 0.033 | 0.000 | 0.098 | 0.355 | 0.197 | 0.006 | 0.038 | 0.019 | 0.000 | 0.162 | 0.002 | 0.036 | 0.000 | 0.524 | 1 |
| Sud Sardegna | 0.084 | 0.002 | 0.014 | 0.000 | 0.089 | 0.268 | 0.055 | 0.006 | 0.171 | 0.040 | 0.000 | 0.252 | 0.000 | 0.020 | 0.000 | 0.314 | 3 |
| Taranto | 0.045 | 0.002 | 0.011 | 0.000 | 0.077 | 0.288 | 0.065 | 0.007 | 0.170 | 0.016 | 0.005 | 0.273 | 0.000 | 0.041 | 0.000 | 0.383 | 3 |
| Teramo | 0.051 | 0.003 | 0.041 | 0.000 | 0.123 | 0.318 | 0.069 | 0.008 | 0.127 | 0.030 | 0.007 | 0.200 | 0.000 | 0.024 | 0.000 | 0.495 | 1 |
| Terni | 0.053 | 0.054 | 0.024 | 0.008 | 0.074 | 0.316 | 0.072 | 0.007 | 0.094 | 0.025 | 0.000 | 0.244 | 0.000 | 0.029 | 0.000 | 0.547 | 1 |
| Torino | 0.093 | 0.004 | 0.034 | 0.000 | 0.092 | 0.265 | 0.078 | 0.010 | 0.095 | 0.026 | 0.000 | 0.261 | 0.002 | 0.042 | 0.000 | 0.556 | 3 |
| Trapani | 0.056 | 0.005 | 0.016 | 0.000 | 0.151 | 0.199 | 0.118 | 0.112 | 0.162 | 0.016 | 0.000 | 0.145 | 0.000 | 0.022 | 0.000 | 0.380 | 2 |
| Trento | 0.079 | 0.003 | 0.045 | 0.000 | 0.050 | 0.263 | 0.115 | 0.008 | 0.045 | 0.030 | 0.000 | 0.252 | 0.000 | 0.038 | 0.071 | 0.447 | 3 |
| Treviso | 0.052 | 0.003 | 0.041 | 0.000 | 0.081 | 0.367 | 0.169 | 0.009 | 0.043 | 0.020 | 0.000 | 0.179 | 0.000 | 0.034 | 0.002 | 0.504 | 1 |
| Trieste | 0.084 | 0.004 | 0.038 | 0.000 | 0.064 | 0.273 | 0.117 | 0.016 | 0.072 | 0.044 | 0.000 | 0.241 | 0.000 | 0.034 | 0.015 | 0.441 | 1 |
| Udine | 0.058 | 0.003 | 0.034 | 0.000 | 0.077 | 0.346 | 0.165 | 0.011 | 0.049 | 0.025 | 0.000 | 0.196 | 0.000 | 0.031 | 0.004 | 0.503 | 1 |
| Varese | 0.063 | 0.004 | 0.042 | 0.000 | 0.100 | 0.325 | 0.144 | 0.007 | 0.062 | 0.018 | 0.000 | 0.195 | 0.002 | 0.038 | 0.000 | 0.517 | 1 |
| Venezia | 0.068 | 0.003 | 0.038 | 0.000 | 0.078 | 0.346 | 0.110 | 0.008 | 0.060 | 0.022 | 0.000 | 0.232 | 0.000 | 0.033 | 0.002 | 0.489 | 1 |
| Verbano–Cusio–Ossola | 0.065 | 0.004 | 0.026 | 0.000 | 0.093 | 0.332 | 0.142 | 0.012 | 0.065 | 0.018 | 0.000 | 0.206 | 0.002 | 0.034 | 0.000 | 0.541 | 1 |
| Vercelli | 0.049 | 0.006 | 0.028 | 0.000 | 0.104 | 0.358 | 0.139 | 0.008 | 0.064 | 0.016 | 0.000 | 0.191 | 0.002 | 0.035 | 0.000 | 0.569 | 1 |
| Verona | 0.062 | 0.003 | 0.037 | 0.000 | 0.109 | 0.387 | 0.127 | 0.009 | 0.048 | 0.018 | 0.000 | 0.160 | 0.000 | 0.037 | 0.002 | 0.548 | 1 |
| Vibo Valentia | 0.048 | 0.010 | 0.091 | 0.000 | 0.141 | 0.231 | 0.075 | 0.018 | 0.130 | 0.015 | 0.006 | 0.181 | 0.000 | 0.053 | 0.000 | 0.395 | 2 |
| Vicenza | 0.068 | 0.003 | 0.042 | 0.000 | 0.075 | 0.396 | 0.139 | 0.009 | 0.043 | 0.020 | 0.000 | 0.172 | 0.000 | 0.029 | 0.003 | 0.548 | 1 |
| Viterbo | 0.051 | 0.004 | 0.023 | 0.008 | 0.078 | 0.425 | 0.076 | 0.008 | 0.091 | 0.022 | 0.000 | 0.189 | 0.000 | 0.025 | 0.000 | 0.557 | 1 |

## Appendix B. Membership values of units to each cluster

| Province | Cluster 1 | Cluster 2 | Cluster 3 | Noise Cluster |
|---|---|---|---|---|
| Agrigento | 0.00 | 0.17 | 0.00 | 0.83 |
| Alessandria | 1.00 | 0.00 | 0.00 | 0.00 |
| Ancona | 0.04 | 0.00 | 0.96 | 0.00 |
| Aosta | 0.49 | 0.02 | 0.29 | 0.20 |
| Arezzo | 0.01 | 0.00 | 0.99 | 0.00 |
| Ascoli Piceno | 0.98 | 0.00 | 0.02 | 0.00 |
| Asti | 1.00 | 0.00 | 0.00 | 0.00 |
| Avellino | 0.01 | 0.90 | 0.03 | 0.05 |
| Bari | 0.00 | 0.00 | 0.91 | 0.08 |
| Barletta–Andria–Trani | 0.01 | 0.00 | 0.99 | 0.00 |
| Belluno | 0.99 | 0.00 | 0.00 | 0.00 |
| Benevento | 0.04 | 0.54 | 0.01 | 0.40 |
| Bergamo | 1.00 | 0.00 | 0.00 | 0.00 |
| Biella | 1.00 | 0.00 | 0.00 | 0.00 |
| Bologna | 0.00 | 0.00 | 1.00 | 0.00 |
| Bolzano | 0.00 | 0.00 | 0.00 | 1.00 |
| Brescia | 1.00 | 0.00 | 0.00 | 0.00 |
| Brindisi | 0.19 | 0.18 | 0.61 | 0.03 |
| Cagliari | 0.05 | 0.23 | 0.67 | 0.05 |
| Caltanissetta | 0.00 | 0.88 | 0.00 | 0.11 |
| Campobasso | 0.85 | 0.06 | 0.04 | 0.05 |
| Caserta | 0.02 | 0.90 | 0.01 | 0.07 |
| Catania | 0.00 | 0.96 | 0.00 | 0.04 |
| Catanzaro | 0.32 | 0.56 | 0.02 | 0.10 |
| Chieti | 0.98 | 0.01 | 0.01 | 0.00 |
| Como | 1.00 | 0.00 | 0.00 | 0.00 |
| Cosenza | 0.01 | 0.86 | 0.00 | 0.13 |
| Cremona | 1.00 | 0.00 | 0.00 | 0.00 |
| Crotone | 0.00 | 0.97 | 0.00 | 0.02 |
| Cuneo | 1.00 | 0.00 | 0.00 | 0.00 |
| Enna | 0.00 | 0.97 | 0.00 | 0.03 |
| Fermo | 1.00 | 0.00 | 0.00 | 0.00 |
| Ferrara | 0.23 | 0.00 | 0.77 | 0.00 |
| Firenze | 0.00 | 0.00 | 1.00 | 0.00 |
| Foggia | 0.03 | 0.75 | 0.05 | 0.16 |
| Forli'-Cesena | 0.00 | 0.00 | 1.00 | 0.00 |
| Frosinone | 0.99 | 0.00 | 0.00 | 0.00 |
| Genova | 0.01 | 0.00 | 0.98 | 0.00 |
| Gorizia | 0.64 | 0.00 | 0.27 | 0.09 |
| Grosseto | 0.34 | 0.00 | 0.66 | 0.00 |
| Imperia | 1.00 | 0.00 | 0.00 | 0.00 |
| Isernia | 0.33 | 0.02 | 0.02 | 0.64 |
| L'Aquila | 1.00 | 0.00 | 0.00 | 0.00 |
| La Spezia | 0.00 | 0.00 | 0.99 | 0.00 |
| Latina | 0.94 | 0.04 | 0.00 | 0.02 |
| Lecce | 0.27 | 0.00 | 0.73 | 0.00 |
| Lecco | 1.00 | 0.00 | 0.00 | 0.00 |
| Livorno | 0.00 | 0.00 | 1.00 | 0.00 |
| Lodi | 1.00 | 0.00 | 0.00 | 0.00 |
| Lucca | 0.28 | 0.00 | 0.71 | 0.01 |

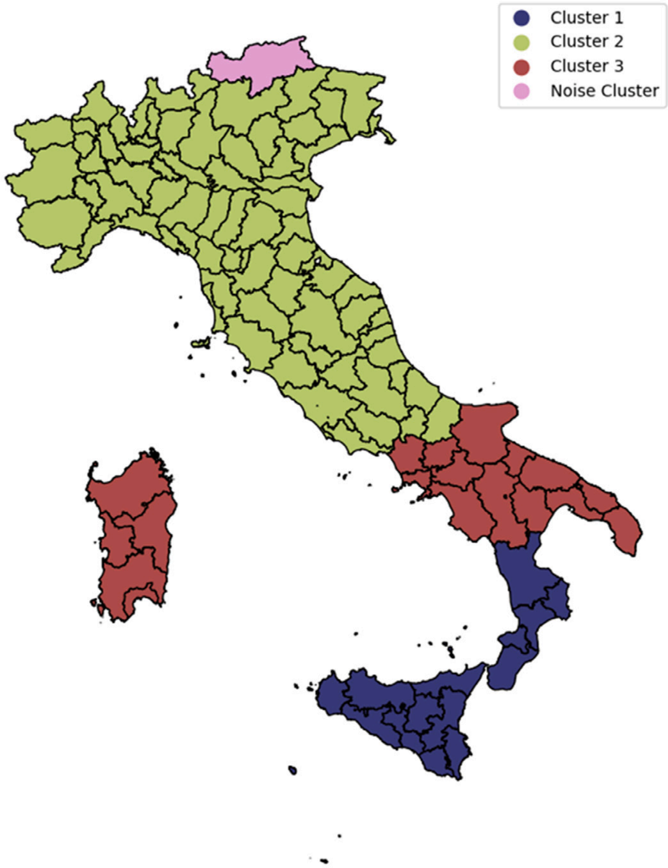
| Province | Cluster 1 | Cluster 2 | Cluster 3 | Noise Cluster |
|---|---|---|---|---|
| Macerata | 0.99 | 0.00 | 0.00 | 0.00 |
| Mantova | 0.96 | 0.00 | 0.03 | 0.00 |
| Massa–Carrara | 0.36 | 0.25 | 0.35 | 0.04 |
| Matera | 0.04 | 0.00 | 0.95 | 0.01 |
| Messina | 0.01 | 0.11 | 0.00 | 0.88 |
| Milano | 0.40 | 0.00 | 0.58 | 0.01 |
| Modena | 0.00 | 0.00 | 1.00 | 0.00 |
| Monza e della Brianza | 1.00 | 0.00 | 0.00 | 0.00 |
| Napoli | 0.01 | 0.24 | 0.06 | 0.69 |
| Novara | 1.00 | 0.00 | 0.00 | 0.00 |
| Nuoro | 0.03 | 0.39 | 0.44 | 0.14 |
| Oristano | 0.27 | 0.46 | 0.22 | 0.06 |
| Padova | 1.00 | 0.00 | 0.00 | 0.00 |
| Palermo | 0.00 | 0.98 | 0.00 | 0.02 |
| Parma | 0.09 | 0.00 | 0.91 | 0.00 |
| Pavia | 1.00 | 0.00 | 0.00 | 0.00 |
| Perugia | 0.42 | 0.00 | 0.57 | 0.01 |
| Pesaro e Urbino | 0.00 | 0.00 | 1.00 | 0.00 |
| Pescara | 0.98 | 0.01 | 0.01 | 0.00 |
| Piacenza | 0.99 | 0.00 | 0.01 | 0.00 |
| Pisa | 0.00 | 0.00 | 1.00 | 0.00 |
| Pistoia | 0.00 | 0.00 | 1.00 | 0.00 |
| Pordenone | 1.00 | 0.00 | 0.00 | 0.00 |
| Potenza | 0.28 | 0.21 | 0.32 | 0.20 |
| Prato | 0.00 | 0.00 | 1.00 | 0.00 |
| Ragusa | 0.00 | 1.00 | 0.00 | 0.00 |
| Ravenna | 0.00 | 0.00 | 1.00 | 0.00 |
| Reggio Calabria | 0.02 | 0.82 | 0.00 | 0.16 |
| Reggio nell’Emilia | 0.00 | 0.00 | 0.99 | 0.01 |
| Rieti | 1.00 | 0.00 | 0.00 | 0.00 |
| Rimini | 0.00 | 0.00 | 1.00 | 0.00 |
| Roma | 0.33 | 0.00 | 0.65 | 0.02 |
| Rovigo | 1.00 | 0.00 | 0.00 | 0.00 |
| Salerno | 0.38 | 0.38 | 0.19 | 0.05 |
| Sassari | 0.08 | 0.61 | 0.22 | 0.08 |
| Savona | 0.98 | 0.00 | 0.01 | 0.00 |
| Siena | 0.00 | 0.00 | 1.00 | 0.00 |
| Siracusa | 0.00 | 1.00 | 0.00 | 0.00 |
| Sondrio | 0.98 | 0.00 | 0.00 | 0.02 |
| Sud Sardegna | 0.10 | 0.21 | 0.64 | 0.05 |
| Taranto | 0.09 | 0.04 | 0.85 | 0.03 |
| Teramo | 0.96 | 0.02 | 0.01 | 0.00 |
| Terni | 0.71 | 0.00 | 0.28 | 0.01 |
| Torino | 0.29 | 0.01 | 0.69 | 0.01 |
| Trapani | 0.01 | 0.93 | 0.00 | 0.06 |
| Trento | 0.37 | 0.00 | 0.58 | 0.05 |
| Treviso | 1.00 | 0.00 | 0.00 | 0.00 |
| Trieste | 0.55 | 0.00 | 0.44 | 0.00 |
| Udine | 1.00 | 0.00 | 0.00 | 0.00 |
| Varese | 1.00 | 0.00 | 0.00 | 0.00 |
| Venezia | 1.00 | 0.00 | 0.00 | 0.00 |
| Verbano–Cusio–Ossola | 1.00 | 0.00 | 0.00 | 0.00 |
| Vercelli | 1.00 | 0.00 | 0.00 | 0.00 |
| Verona | 1.00 | 0.00 | 0.00 | 0.00 |
| Vibo Valentia | 0.13 | 0.80 | 0.02 | 0.05 |
| Vicenza | 0.99 | 0.00 | 0.00 | 0.00 |
| Viterbo | 0.97 | 0.00 | 0.01 | 0.02 |

## Appendix C. Socio-economic indicators per province

| Province | GDP | Unemployment | Graduates | Incl. Rent | Excl. Rent | Turnout |
|---|---|---|---|---|---|---|
| Agrigento | 17 961.399 | 19.305 | 21.400 | 0.313 | 0.347 | 36.5 |
| Alessandria | 32 295.404 | 7.178 | 17.700 | 0.275 | 0.298 | 55.0 |
| Ancona | 33 253.899 | 7.886 | 37.400 | 0.259 | 0.285 | 52.0 |
| Aosta | 43 904.221 | 5.420 | 27.700 | 0.261 | 0.302 | 42.5 |
| Arezzo | 34 286.996 | 5.322 | 32.200 | 0.270 | 0.297 | 60.6 |
| Ascoli piceno | 29 187.717 | 6.950 | 28.400 | 0.259 | 0.285 | 57.9 |
| Asti | 28 553.580 | 7.942 | 22.000 | 0.275 | 0.298 | 56.9 |
| Avellino | 21 216.892 | 14.028 | 24.600 | 0.309 | 0.336 | 47.2 |
| Bari | 25 686.083 | 9.308 | 29.900 | 0.277 | 0.304 | 47.5 |
| Barletta–Andria–Trani | 18 226.782 | 11.019 | 16.800 | 0.277 | 0.304 | 38.2 |
| Belluno | 37 399.194 | 2.881 | 28.500 | 0.266 | 0.292 | 43.1 |
| Benevento | 19 863.688 | 7.659 | 26.600 | 0.309 | 0.336 | 47.3 |
| Bergamo | 38 642.165 | 3.447 | 21.800 | 0.277 | 0.302 | 59.7 |
| Biella | 30 654.481 | 3.853 | 29.300 | 0.275 | 0.298 | 59.6 |
| Bologna | 46 647.901 | 3.600 | 44.300 | 0.257 | 0.284 | 60.8 |
| Bolzano | 56 022.342 | 2.324 | 24.900 | 0.268 | 0.299 | 49.6 |
| Brescia | 40 504.302 | 4.131 | 25.800 | 0.277 | 0.302 | 60.5 |
| Brindisi | 20 328.601 | 13.107 | 22.000 | 0.277 | 0.304 | 39.5 |
| Cagliari | 32 185.273 | 13.874 | 32.100 | 0.285 | 0.318 | 42.7 |
| Caltanissetta | 18 802.873 | 20.129 | 19.300 | 0.313 | 0.347 | 48.1 |
| Campobasso | 25 421.402 | 11.059 | 27.800 | 0.245 | 0.281 | 50.6 |
| Caserta | 19 785.777 | 14.486 | 23.500 | 0.309 | 0.336 | 46.3 |
| Catania | 21 179.368 | 16.329 | 17.900 | 0.313 | 0.347 | 37.1 |
| Catanzaro | 22 272.462 | 13.600 | 20.300 | 0.339 | 0.372 | 36.8 |
| Chieti | 29 603.324 | 11.416 | 23.200 | 0.269 | 0.297 | 43.6 |
| Como | 32 733.221 | 6.552 | 30.400 | 0.277 | 0.302 | 53.6 |
| Cosenza | 18 087.602 | 16.035 | 28.500 | 0.339 | 0.372 | 44.9 |
| Cremona | 37 541.915 | 5.391 | 26.900 | 0.277 | 0.302 | 58.8 |
| Crotone | 19 810.049 | 17.193 | 10.500 | 0.339 | 0.372 | 33.8 |
| Cuneo | 37 201.930 | 3.745 | 21.200 | 0.275 | 0.298 | 60.3 |
| Enna | 18 222.080 | 13.040 | 22.200 | 0.313 | 0.347 | 35.2 |
| Fermo | 27 424.152 | 4.397 | 24.100 | 0.259 | 0.285 | 48.7 |
| Ferrara | 28 889.216 | 8.267 | 32.900 | 0.257 | 0.284 | 59.9 |
| Firenze | 43 724.438 | 6.088 | 34.900 | 0.270 | 0.297 | 65.1 |
| Foggia | 20 421.828 | 16.882 | 16.300 | 0.277 | 0.304 | 42.7 |
| Forli'-cesena | 37 159.642 | 4.115 | 32.400 | 0.257 | 0.284 | 61.3 |
| Frosinone | 25 167.270 | 9.005 | 27.400 | 0.310 | 0.348 | 53.0 |
| Genova | 37 794.370 | 6.936 | 36.400 | 0.294 | 0.319 | 49.3 |
| Gorizia | 31 797.101 | 5.929 | 25.700 | 0.273 | 0.305 | 51.4 |
| Grosseto | 28 385.609 | 6.124 | 23.000 | 0.270 | 0.297 | 54.4 |
| Imperia | 26 899.856 | 9.500 | 19.200 | 0.294 | 0.319 | 49.8 |
| Isernia | 23 778.055 | 10.106 | 26.100 | 0.245 | 0.281 | 41.5 |
| L'aquila | 29 074.202 | 9.207 | 33.900 | 0.269 | 0.297 | 41.8 |
| La spezia | 36 046.512 | 8.192 | 23.500 | 0.294 | 0.319 | 51.1 |
| Latina | 25 650.564 | 9.598 | 18.700 | 0.310 | 0.348 | 44.7 |
| Lecce | 20 116.384 | 13.070 | 21.100 | 0.277 | 0.304 | 46.3 |
| Lecco | 36 518.340 | 2.884 | 33.600 | 0.277 | 0.302 | 58.2 |
| Livorno | 31 217.498 | 4.952 | 27.700 | 0.270 | 0.297 | 58.6 |
| Lodi | 30 570.927 | 5.181 | 25.200 | 0.277 | 0.302 | 58.8 |
| Lucca | 33 125.817 | 6.840 | 26.600 | 0.270 | 0.297 | 48.2 |

| Province | GDP | Unemployment | Graduates | Incl. Rent | Excl. Rent | Turnout |
|---|---|---|---|---|---|---|
| Macerata | 30 706.307 | 6.242 | 24.400 | 0.259 | 0.285 | 50.9 |
| Mantova | 36 508.759 | 4.391 | 24.300 | 0.277 | 0.302 | 52.3 |
| Massa–Carrara | 29 425.532 | 8.945 | 33.500 | 0.270 | 0.297 | 46.9 |
| Matera | 23 324.621 | 7.843 | 23.800 | 0.278 | 0.304 | 39.1 |
| Messina | 20 889.147 | 21.547 | 16.900 | 0.313 | 0.347 | 41.2 |
| Milano | 66 556.156 | 5.510 | 40.400 | 0.277 | 0.302 | 52.9 |
| Modena | 44 997.155 | 5.119 | 29.200 | 0.257 | 0.284 | 60.8 |
| Monza e della brianza | 35 901.376 | 4.285 | 36.100 | 0.277 | 0.302 | 54.1 |
| Napoli | 22 370.405 | 20.960 | 20.000 | 0.309 | 0.336 | 42.0 |
| Novara | 34 596.908 | 6.635 | 21.800 | 0.275 | 0.298 | 58.4 |
| Nuoro | 22 186.560 | 7.770 | 20.100 | 0.285 | 0.318 | 29.8 |
| Oristano | 20 225.166 | 13.638 | 20.700 | 0.285 | 0.318 | 31.0 |
| Padova | 39 117.995 | 4.566 | 38.100 | 0.266 | 0.292 | 56.1 |
| Palermo | 21 868.888 | 17.952 | 22.900 | 0.313 | 0.347 | 38.9 |
| Parma | 42 738.175 | 5.379 | 28.700 | 0.257 | 0.284 | 55.3 |
| Pavia | 29 413.632 | 6.079 | 20.600 | 0.277 | 0.302 | 56.2 |
| Perugia | 29 710.847 | 7.272 | 32.300 | 0.269 | 0.298 | 62.9 |
| Pesaro e urbino | 31 855.876 | 4.966 | 34.600 | 0.259 | 0.285 | 62.5 |
| Pescara | 27 377.154 | 11.304 | 32.900 | 0.269 | 0.297 | 54.4 |
| Piacenza | 37 565.187 | 6.540 | 24.900 | 0.257 | 0.284 | 55.2 |
| Pisa | 35 631.743 | 6.687 | 31.600 | 0.270 | 0.297 | 61.0 |
| Pistoia | 28 715.470 | 6.163 | 20.600 | 0.270 | 0.297 | 55.2 |
| Pordenone | 34 896.708 | 3.050 | 27.200 | 0.273 | 0.305 | 46.6 |
| Potenza | 27 302.499 | 6.964 | 27.700 | 0.278 | 0.304 | 44.5 |
| Prato | 34 650.174 | 6.289 | 19.900 | 0.270 | 0.297 | 64.1 |
| Ragusa | 20 637.224 | 9.909 | 19.500 | 0.313 | 0.347 | 31.1 |
| Ravenna | 35 829.016 | 5.476 | 33.300 | 0.257 | 0.284 | 58.1 |
| Reggio calabria | 20 124.904 | 13.983 | 20.700 | 0.339 | 0.372 | 38.3 |
| Reggio nell' emilia | 41 976.059 | 4.359 | 29.900 | 0.257 | 0.284 | 60.9 |
| Rieti | 24 336.870 | 10.056 | 29.700 | 0.310 | 0.348 | 58.4 |
| Rimini | 33 917.921 | 6.710 | 33.300 | 0.257 | 0.284 | 53.0 |
| Roma | 45 030.317 | 7.517 | 36.400 | 0.310 | 0.348 | 45.1 |
| Rovigo | 29 382.662 | 8.197 | 26.800 | 0.266 | 0.292 | 53.4 |
| Salerno | 22 185.736 | 14.469 | 24.100 | 0.309 | 0.336 | 45.1 |
| Sassari | 24 824.469 | 10.336 | 16.700 | 0.285 | 0.318 | 40.6 |
| Savona | 32 898.767 | 5.139 | 30.300 | 0.294 | 0.319 | 54.6 |
| Siena | 34 768.110 | 4.511 | 27.600 | 0.270 | 0.297 | 64.1 |
| Siracusa | 29 291.645 | 14.794 | 21.800 | 0.313 | 0.347 | 34.1 |
| Sondrio | 33 529.083 | 6.570 | 23.600 | 0.277 | 0.302 | 52.4 |
| Sud sardegna | 17 855.228 | 12.267 | 21.700 | 0.285 | 0.318 | 31.4 |
| Taranto | 21 844.886 | 13.630 | 18.300 | 0.277 | 0.304 | 38.3 |
| Teramo | 28 219.773 | 6.305 | 32.500 | 0.269 | 0.297 | 49.5 |
| Terni | 26 111.111 | 6.730 | 36.800 | 0.269 | 0.298 | 54.7 |
| Torino | 35 647.603 | 7.407 | 33.300 | 0.275 | 0.298 | 55.6 |
| Trapani | 18 738.284 | 13.387 | 18.100 | 0.313 | 0.347 | 38.0 |
| Trento | 43 901.207 | 3.849 | 31.100 | 0.268 | 0.299 | 44.7 |
| Treviso | 38 232.013 | 5.212 | 28.900 | 0.266 | 0.292 | 50.4 |
| Trieste | 39 095.280 | 6.274 | 41.300 | 0.273 | 0.305 | 44.1 |
| Udine | 35 816.918 | 6.271 | 29.900 | 0.273 | 0.305 | 50.3 |
| Varese | 34 209.628 | 4.752 | 28.700 | 0.277 | 0.302 | 51.7 |
| Venezia | 35 761.597 | 4.259 | 28.400 | 0.266 | 0.292 | 48.9 |
| Verbano–Cusio–Ossola | 27 624.919 | 5.853 | 26.900 | 0.275 | 0.298 | 54.1 |
| Vercelli | 32 385.542 | 5.478 | 25.100 | 0.275 | 0.298 | 56.9 |
| Verona | 38 930.580 | 3.170 | 34.600 | 0.266 | 0.292 | 54.8 |
| Vibo valentia | 18 385.175 | 14.421 | 24.400 | 0.339 | 0.372 | 39.5 |
| Vicenza | 40 057.478 | 3.491 | 30.400 | 0.266 | 0.292 | 54.8 |
| Viterbo | 24 312.581 | 7.460 | 28.000 | 0.310 | 0.348 | 55.7 |

## Appendix D. Output partition for higher $\beta$ value




Partition of Italian provinces into 3 clusters plus the noise clusters, with model parameters $(\beta, C, \lambda) = (0.04, 4, 2)$.

## References

Abdali, E., Valadan Zoej, M.J., Taheri Dehkordi, A., Ghaderpour, E., 2023. A parallel-cascaded ensemble of machine learning models for crop type classification in google earth engine using multi-temporal sentinel-1/2 and landsat-8/9 remote sensing data. Remote. Sens. 16 (1), 127.
Ado, M., Amitab, K., Maji, A.K., Jasińska, E., Gono, R., Leonowicz, Z., Jasiński, M., 2022. Landslide susceptibility mapping using machine learning: A literature survey. Remote. Sens. 14 (13), 3029.
Ambroise, C., Dang, M., 2009. Spatial data clustering. Data Anal. 289–318.
Antoni, L., Krajči, S., Krídlo, O., Macek, B., Pisková, L., 2014. On heterogeneous formal contexts. Fuzzy Sets and Systems 234, 22–33.
Bezdek, J., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic, Norwell, MA, USA.
Coppi, R., D'Urso, P., Giordani, P., 2010. A fuzzy clustering model for multivariate spatial time series. J. Classification 27 (1), 54–88.
Davé, R.N., 1991. Characterization and detection of noise in clustering. Pattern Recognit. 12, 657–664.
De Giovanni, L., D'Urso, P., Fiorino, N., Galli, E., Garzarelli, G., Pacifico, A., 2023. The local press as an external public governance power. Appl. Econ. 55 (53), 6327–6344.
Deza, M.M., Deza, E., 2009. Encyclopedia of distances. In: Encyclopedia of Distances. Springer, pp. 1–583.
D'Urso, P., 2015. Fuzzy clustering. In: Hennig, C., Meila, M., Murtagh, F., Rocci, R. (Eds.), Handbook of Cluster Analysis. Chapman and Hall, pp. 545–573.
D'Urso, P., De Giovanni, L., Disegna, M., Massari, R., 2013. Bagged clustering and its application to tourism market segmentation. Expert Syst. Appl. 40 (12), 4944–4956.
D'Urso, P., De Giovanni, L., Disegna, M., Massari, R., 2019. Fuzzy clustering with spatial–temporal information. Spat. Stat. 30, 71–102.
D'Urso, P., De Giovanni, L., Federico, L., Vitale, V., 2023a. Fuzzy clustering of spatial interval-valued data. Spat. Stat. 57, 100764.
D'Urso, P., De Giovanni, L., Federico, L., Vitale, V., 2025. Fuzzy clustering of mixed data with spatial regularization. Spat. Stat. 65, 100874.
D'Urso, P., De Giovanni, L., Vitale, V., 2022. Spatial robust fuzzy clustering of COVID 19 time series based on B-splines. Spat. Stat. 49, 100518.
D'Urso, P., Disegna, M., Massari, R., Osti, L., 2016. Fuzzy segmentation of postmodern tourists. Tour. Manag. 55, 297–308.
D'Urso, P., Giovanni, L., Vitale, V., 2023b. A robust method for clustering football players with mixed attributes. Ann. Oper. Res. 325 (1), 9–36.
D'Urso, P., Massari, R., 2019. Fuzzy clustering of mixed data. Inform. Sci. 505, 513–534.
Eagles, M., 1995. Spatial and contextual models of political behavior: an introduction. Political Geogr. 14 (6), 499–502.
Everitt, B., Landau, S., Leese, M., Stahl, D., 2011. Cluster Analysis, fifth ed. John Wiley & Sons, Ltd, London.

Fiorino, N., Pontarollo, N., Ricciuti, R., 2021. Spatial links in the analysis of voter turnout in European parliamentary elections. Lett. Spat. Resour. Sci. 14 (1), 65–78.
García-Escudero, L.A., Gordaliza, A., 1999. Robustness properties of k means and trimmed k means. J. Amer. Statist. Assoc. 94 (447), 956–969.
García-Escudero, L.A., Gordaliza, A., Matrán, C., Mayo-Iscar, A., 2010. A review of robust clustering methods. Adv. Data Anal. Classif. 4, 89–109.
Geys, B., 2006. Explaining voter turnout: A review of aggregate-level research. Elect. Stud. 25 (4), 637–663.
Gower, J.C., 1971. A general coefficient of similarity and some of its properties. Biometrics 857–871.
Guha, S., Rastogi, R., Shim, K., 1999. ROCK: A robust clustering algorithm for categorical attributes. In: Data Engineering, 1999. Proceedings., 15th International Conference on. IEEE, pp. 512–521.
Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1998. Multivariate Data Analysis. Upper Saddle River.
Heiser, W., Groenen, P., 1997. Cluster differences scaling with a within-clusters loss component and a fuzzy successive approximation strategy to avoid local minima. Psychometrika 62 (1), 63–83.
Hu, T., Sung, S., 2006. A hybrid EM approach to spatial clustering. Comput. Statist. Data Anal. 50 (5), 1188–1205.
Huckfeldt, R., 1995. Citizens, Politics and Social Communication: Information and Influence in an Election. Cambridge University Press.
Hwang, H., Desarbo, W., Takane, Y., 2007. Fuzzy clusterwise generalized structured component analysis. Psychometrika 72 (2), 181–198.
Kaufman, L., Rousseeuw, P., 2005. Finding Groups in Data: An Introduction to Cluster Analysis. WileyBlackwell.
Kontos, T.D., Zevgolis, Y.G., 2024. A fuzzy spatial multiple criteria analysis methodology for solid waste landfill siting. J 7 (4), 502–528.
McBratney, A., Moore, A., 1985. Application of fuzzy sets to climatic classification. Agricult. Forest. Meterol. 35 (1–4), 165–185.
McKay, L., 2019. Left behind people, or places? The role of local economies in perceived community representation. Elect. Stud. 60, 102046.
Pattie, C., and, R.J., 2000. People who talk together vote together: An exploration of contextual effects in Great Britain. Ann. Assoc. Am. Geogr. 90 (1), 41–66.
Pattie, C.J., Johnston, R.J., Schipper, M., and, L.P., 2015. Are regions important in british elections? Valence politics and local economic contexts at the 2010 general election. Reg. Stud. 49 (9), 1561–1574.
Torabi, M., 2016. Hierarchical multivariate mixture generalized linear models for the analysis of spatial data: An application to disease mapping. Biom. J. 58 (5), 1138–1150.
Viroli, C., 2011. Finite mixtures of matrix normal distributions for classifying three-way data. Stat. Comput. 21 (4), 511–522.
Wedel, M., Kamakura, W., 2000. Market Segmentation: Conceptual and Methodological Foundations, vol. 8, Springer.
Xie, X., Beni, G., 1991. A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. 13 (8), 841–847.