RESEARCH ARTICLE

Seismic Event Clustering in Mainland Portugal: DBSCAN Approach

André Martins Brito[a]

[a] Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

ARTICLE HISTORY
Compiled December 9, 2024

Abstract
The identification of seismic zones continues to be a relevant topic, mainly in regions with considerable seismic activity, as is the case of Portugal, and the application of quantitative methods has shown to be highly versatile. In this work, in a first part, the DBSCAN algorithm is applied to a catalog of seismic events for Mainland Portugal and surrounding maritime area, considering the Haversine distance between epicentres of seismic events. When compared to a defined seismic zonation, the resulting clusters fit especially well within the zones. In a second part, a novel distance index is presented. This index will calculate the distances between seismic events taking into account, not only the geographical distance between epicentres, but also the time of occurrence and size of the events. The distance values calculated with the proposed distance index will be used in conjunction with the DBSCAN algorithm with the purpose of identifying sets of seismic events originated from the same geological structures but in different time periods. With simulated data, the proposed distance index shows exactly the intended behaviour. Applying it to the seismic event catalog for Mainland Portugal, also shows interesting results, being able to identify several geographically overlapped clusters but separated in time. This approach could be used to study how often seismic activity is expected from a specific geological structure and how these structures interact with each other.

KEYWORDS
Seismic zoning; Density based algorithms; spatial-temporal clustering;

1. Introduction

The Portuguese Mainland is located relatively close to a geological complex region. Just south of the Iberian Peninsula, of which Mainland Portugal is a part of, is the boundary between the Eurasian and African Plates. This boundary extends along the Azores-Gibraltar Fracture Zone, starting at the Azores Archipelago, passing through the Gibraltar strait and continuing on until the Mediterranean Sea. Due to the complexity of the Africa/Eurasia plate boundary, this region is prone to generate powerful earthquakes that, given its proximity to Mainland Portugal, can produce major material damages in land and, thus, represents a risk to the Portuguese Population.

[6] present a geophysical and geological overview of the geological structures of the Iberian Peninsula and surrounding maritime area. The authors give particular focus to the western margin of Portugal stating that, most of the seismic activity recorded

CONTACT André Brito. Email: anm.brito@campus.fct.unl.pt

in the Portuguese Mainland originates at the Africa/Eurasia plate boundary or close to it.

[11] propose several seismic zones for Mainland Portugal and surrounding maritime area based on their classification as SCR or ACR, meaning Stable Continental Region and Active Continental Region respectively, and on seismicity criteria. In the context of seismic zonation, [8] define seven zones in the Azores Archipelago area based on seismicity criteria such as, the number of seismic events in a year and the magnitude of these events. The purpose of defining different seismic zones is to establish surface areas with homogeneous seismic characteristics, thus, allowing for similar risk evaluation.

The definition of seismic zones usually results in polygonal areas that follow straight and regular shapes. However, the complexity and interaction between geological structures, such as faults and interplate boundaries, tend to produce seismic events that, when mapped by their epicentres, create agglomerates of irregular shapes and sizes. Quantitative methods, such as clustering methods, may assist in the process of seismic zone definition since their purpose of application is to group objects by identifying similar characteristics and underlying patterns, [4]. In order to demonstrate the advantages of applying quantitative methods to the definition of seismic zones, [13] used a hierarchical clustering algorithm to group a grid of area units in Iran. The usefulness of this method is highlighted, stating that, it is able to reveal tectonic evolution trends of a region and, additionally, allows for a continuous tectonic classification and assists in determining the interaction between the own variables used for classification.

With the purpose of categorising seismic data, [7] have implemented a density based algorithm, the DBSCAN, in data sourced from a South African mine. In this work, the authors demonstrated that the algorithm is able to identify clusters of precursor events that precede bigger seismic events. The DBSCAN algortihm was initially presented by [1] and enables the identification of irregular shaped clusters, without the need of previously specifying the number of clusters to be created, thus, the resulting number of clusters is influenced only by the data. Besides, the algorithm is able to classify data points as outliers or noise. [5] applied the DBSCAN algorithm to seismic data, consisting of seismic event's geographical location and depth, to identify seismic zones in India. The authors compared the resulting clusters to another seismic zonation used by an Indian governmental entity and concluded their similarity.

By identifying the necessity of the existence of algorithms that take into account not only spatial seismic data but also the complex relationships between geological structures, namely the interaction between faults, and that allow for the identification of irregular and elongated shaped clusters, [2] present a two-phased clustering method to apply on seismic data, where the first phase is a modification of the DBSCAN algorithm based on the notion of "seismic mass" and the second phase is an agglomerating phase dropping time information. This two-phased method was referred to as the "SM-DBSCAN", where "SM" stands for "seismic mass", and was applied to data sourced from the Hellenic Arc.

The identification of seismic zones continues to be a relevant topic, mainly in regions with considerable seismic activity, as is the case of Portugal, and the application of quantitative methods has shown to be highly versatile, providing important results in a continuous fashion. Having said that, the aim of this work is to obtain seismic zones for Mainland Portugal by applying the DBSCAN algorithm to seismic data, considering the geographical distance and a proposed distance index between events, and confront the results with known geological structures in the region and other seismic zonations defined by other authors.

This work will have the following structure: firstly, the dataset used is to be presented, secondly the DBSCAN algorithm is going to be presented, as well as, the proposed distance index along with reasoning behind it. Simulated data is used to present the intended results from this index. Subsequently, the DBSCAN algorithm will be applied to a seismic catalog for Mainland Portugal, firstly considering only the geographic distance between epicentres of seismic events and secondly in conjunction with proposed distance index.

2. Dataset

The dataset used contains information regarding seismic events in Mainland Portugal and its maritime surroundings. Given the complexity of the boundary between the Eurasian and African plates and the influence this structure has on the seismic activity felt in Portugal [6], it is important to include the maximum number of events that originate in the western margin of the country. The dataset also contains records starting from 1900 to 1992. For each seismic event, one has information about its epicentre's geographical location, longitude and latitude, its magnitude on the Richter scale and the time of its occurrence in decimal years. However, there is a large number of events with missing magnitude recording, thus, were not considered. Table 1 summarises the dataset's characteristics.

Table 1. Dataset's characteristics

Dataset's characteristics	
Time period (in years)	1900 - 1992
Longitude interval considered	6°W - 16°W
Latitude interval considered	34°N - 44°N
Total number of records	4132
Total number of records with magnitude recording	1856
Total number of records with missing magnitude recording	2276

3. Methods

In this section the DBSCAN algorithm will be presented, as well as, all the definitions needed to understand it. Next, the proposed distance index is presented along with the reasoning behind it. Furthermore, seismic data was simulated to better display the intended results from applying the distance index.

3.1. DBSCAN

The method used in this work, in order to cluster seismic events, is the DBSCAN algorithm. DBSCAN stands for Density Based Spatial Clustering of Applications with Noise and it was first presented by [1]. The authors' intention was to develop an efficient clustering algorithm that required little input from the user and could find clusters of irregular shapes. The DBSCAN algorithm relies on the notion of "density" to cluster objects. Two objects are clustered together if it is possible to reach one another without leaving a defined "dense" zone. In order to understand the underlying mechanics of the algorithm, one needs to comprehend several definitions firstly presented by [1]. In this work, the definitions presented are adapted from [1] and [3].

From now on, D is the set of points to be clustered.

Definition 3.1. A ϵ-neighbourhood of a given point $p \in D$, denoted by $N_\epsilon(p)$, is defined by,

$$N_\epsilon(p) = \{q \in D : dist(p, q) \leq \epsilon\} \tag{1}$$

where *dist* is any distance function.

The cardinality of a ϵ-neighbourhood of a point $p \in D$ determines the notion of density. For a given point, one can define the minimal number of points (*minPts*) that should be in the ϵ-neighbourhood in order for it to be called dense.

Definition 3.2. A point $p \in D$ is classified, given ϵ and *minPts*, as,

- a *core point*, if $N_\epsilon(p)$ is dense, *i.e.*, $|N_\epsilon(p)| \geq minPts$,
- a *border point*, if p is not a *core point*, however it belongs to the neighbourhood of a *core point* $q \in D$, *i.e.*, $p \in N_\epsilon(q)$,
- a *noise point*, otherwise.

The application of the DBSCAN algorithm also results in the identification of the so called *noise points*, a point that doesn't belong to any dense area.

Definition 3.3. A given point $p \in D$ is *directly density-reachable* from a point $q \in D$, given ϵ and *minPts* if,

(1) $p \in N_\epsilon(q)$
(2) $|N_\epsilon(q)| \geq minPts$

Where 2 is the requirement for q to be a *core point*.

Definition 3.4. A given point $p \in D$ is *density-reachable* from a point $q \in D$, given ϵ and *minPts*, if there is a chain of points $p_1, \ldots, p_n$, where $p_1 = q$ and $p_n = p$, such that p_{i+1} is *directly density-reachable* from p_i.

Definition 3.5. A given point $p \in D$ is *density-connected* to a point $q \in D$, given ϵ and *MinPts*, if there is a point $o \in D$ such that, p and q are *density-reachable* from a point o.

Definition 3.6. A *density-based cluster* C is a non-empty subset of D that satisfies the following conditions,

(1) Maximality: If $p \in C$ and q is *density-reachable* from p, then $q \in C$.
(2) Connectivity: $\forall p, q \in C$, p is *density-connected* to q.

The application of the DBSCAN relies solely on the input of ϵ, *minPts* and a defined distance function. One can manipulate what a dense neighbourhood is by setting different values for ϵ and *minPts* or even by changing the distance function. Through the distance function, one is able to "encourage" the joint grouping of two given points by decreasing the distance between them or "discourage" their joint grouping by increasing the distance between them. This opens up a new dimension within DBSCAN, where by manipulating the distance function one can also manipulate the clustering, besides adjusting ϵ or *minPts*.

The following algorithm describes DBSCAN and was adapted from [12]. Let D be the set of points to be clustered, D_s the set of points already classified and S the set of points classified as *noise point*, given ϵ and *minPts*.

Algorithm 1: DBSCAN

Result: List $\mathcal{L}$ with the cluster assignment to each point $p \in D$ through ordered pairs (i, p)

```
{Each cluster is identified by the index i};
i = 0;
D_s = ∅;
L = {};
for each point x ∈ D do
    if x ∉ D_s then
        D_s = D_s ∪ {x};
        if |N_ϵ(x)| < minPts then
            S = S ∪ {x};
        else
            i = i + 1;
            for each point p ∈ N_ϵ(x) do
                L = L + {(i, p)};
            end
            for each point y ∈ N_ϵ(x) e y ∉ D_s do
                D_s = D_s ∪ {y};
                if |N_ϵ(y)| ≥ minPts then
                    for each point p ∈ N_ϵ(y) do
                        L = L + {(i, p)};
                        if p ∈ S then
                            S = S − {y};
                        end
                    end
                end
            end
        end
    end
end
```

In the next subsection, a distance index for the clustering of seismic data is presented, as well as, its objective and the reasoning behind it.

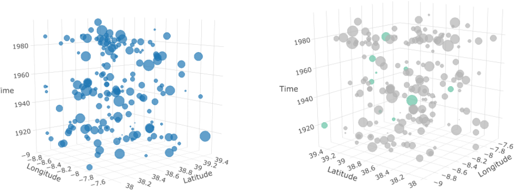
3.2. Distance Index for Seismic Data Clustering

Considering that a seismic event i is characterised by (t_i, m_i, Lo_i, La_j), one defines the distance index as:

$$\texttt{dist}(i, j) = k_t(t_i - t_j)^2 + (1 - k_s \max\{m_i, m_j\}) \times \text{Hav}((Lo_i, La_j), (Lo_j, La_j)) \quad (2)$$

where, t_i is the time of occurrence in decimal year, m_i is the magnitude and (Lo_i, La_j) are the geographical coordinates of the epicentre of a seismic event i. Also, k_t is a scalar $k_t \geq 0$, k_s is a scalar $\left(0 \leq k < \frac{1}{M}\right)$ and Hav refers to the Haversine formula. So as for the distance to remain positive, M corresponds the maximum magnitude of any given seismic event ever recorded. The Haversine Formula [10] determines the distance between two points on a sphere given their longitudes and latitudes. By manipulating the distance between points, one is able to "encourage" the joint grouping of two given objects by decreasing the distance between them or "discourage" their joint grouping by increasing the distance between them. In the context of seismic clustering, the magnitude of an event, due to the interaction of geological structures, influences the occurrence of other seismic events and this influence is directly proportional to the magnitude. The stronger the event is, the bigger its influence in surrounding structures. As stated by [2], the "importance" of a seismic event depends on its magnitude and, thus, magnitude should be taken into account when clustering. Furthermore, [9] has shown through statistical methods that, the occurrence of an earthquake depends on the location of preceding earthquakes. In order to capture this influence through the `dist` index and reflect it in the clustering, two events will be "virtually" closer, the bigger the magnitude of either events. This refers to the $(1 - k_s \max\{m_i, m_j\})$ part of the calculation, where k_s determines how much influence will magnitude have on the distance index value. On the other hand, [9] also showed that, there is a clear reduction of the dependence between events over time, therefore, the time difference between events should also be taken into account considering that, the influence of a seismic event over another lessens as time passes. To capture this effect, events will be "virtually" further apart the larger the time difference between them. This relates to the $k_t(t_i - t_j)^2$ part of the calculation, where k_t determines how much influence will the time difference have on the distance index value. Additionally, all other effects aside, the distance will also depend on the Haversine distance between the epicentres' geographical location of two seismic events. By including the time of occurrence of the seismic events, the clustering will be in a three dimensional space which, if we consider only the longitude latitude plane and project the clusters, will result in overlapping clusters. The first stage of the method presented in [2] also results in spatially overlapping cluster. The authors state that this could help identify closely located geological structures, like faults, whose seismic activity could be separated in time or even identify the seismic activity of one geological structure in different time periods.

So as to illustrate the expected results of the proposed distance index, a simulation of seismic data was made. The simulation was based on the seismic data for Mainland Portugal. Intervals for the longitude and latitude were defined and the events within that region were selected. Then, epicentres' geographical coordinates were sampled from the selected events, with longitude and latitude sampled one by one in order not to have duplicated epicentres. Magnitude for the simulated events was also sampled from the selected events' magnitude. The time of occurrence for the simulated seismic events was done by sampling from three different normal distributions with averages 1920, 1950 and 1980 respectively, and all with standard deviation of 6. There were simulated 180 seismic events. Figure 1 displays the simulated data in a three dimensional scatter plot, where the x and y axis are longitude and latitude, respectively, and z axle represents the time. The points' diameter is larger the bigger its magnitude.



(a) Simulated seismic data. (b) DBSCAN Results, only with Haversine distance.

Figure 1. Simulation Results.

Figure 1(a) clearly displays, three distinct cluster in time. By applying the BDSCAN algorithm with $\epsilon = 20$ and $minPts = 5$ to the simulated data, considering only the Haversine distance between the events' epicenters, the results, displayed in figure 1(b), show 1 cluster was obtained, which corresponds to the points displayed in grey, while the green points represent events considered noise. There were identified 10 seismic events as noise.

Figure 2 displays the results from the DBSCAN algorithm now using the proposed distance index to determine the distances between the seismic events. There were obtained three clusters clearly separated in time, after some parameter tuning. In this particular case, it was considered $k_t = 0.5$. On the other hand, these results show only 7 seismic events were identified as noise compared to the 10 obtained when only the Haversine distance was used. This is clearly an effect of the k_s scalar, which in this case was $k_s = 0.05$. The parameters for the DBSCAN algorithm used were $\epsilon = \frac{1}{1-k_t} \times 20$ and $minPts = 5$, where the unit measure for ϵ is the kilometre. The increase in the distance index generated by the time difference between two seismic events needed to be accounted for when applying the DBSCAN algorithm, especially when defining the ϵ-neighbourhood. Trough experimentation, applying the $\frac{1}{1-k_t}$ factor to the previously defined ϵ value showed the best results for clustering when using the distance index.

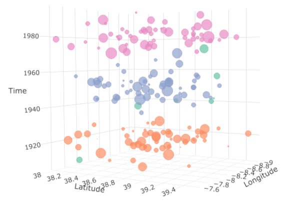


Figure 2. Results with the distance index on simulated data.

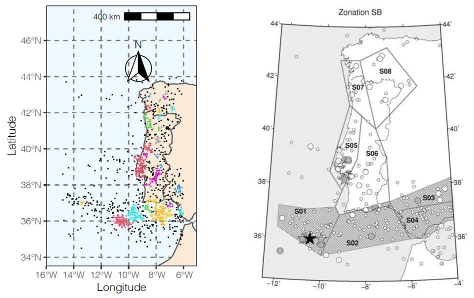
The proposed distance index, with simulated data, shows the intended behaviour was achieved, separating clusters in time and joint grouping events that would otherwise be separated through the influence of their magnitude.

4. Results & Discussion

The DBSCAN algorithm used to obtain the results presented in this and in the previous sections was implemented by [3] for the `R` software.

4.1. Seismic Zones

In this first subsection, the DBSCAN algorithm was applied considering only the spatial distance between seismic events, given by the Haversine distance between epicentres. There were obtained 29 clusters, along with the 484 events considered noise. For these results, it was used $\epsilon = 15$ and $minPts = 5$. Figure 3(a) displays the clusters plotted on Mainland Portugal's map, while figure 3 shows the zonations proposed by [11]. There is a clear match between the cluster results and the zonation presented. In figure 3(b), the Active Continental Regions (ACR) are displayed with a stripped pattern while the Stable Continental Regions (SCR) are transparent. ACR have more seismic activity then SCR, this seismic activity results in bigger denser clusters in these regions. In addition, the clusters that overlap Lisbon Metropolitan Area and the above coast fit plainly with the zone identified as S05 in figure 3(b). Furthermore, the purple cluster located close to the middle of Portugal coincides with a known geological fault, designated as the Messejana Fault. DBSCAN is clearly able to identify irregular shaped clusters despite the complex shape of geological structures like the Messejana Fault. Besides this, there's another relevant result that needs to be taken into account when using the DBSCAN algorithm, the noise points. No clusters are formed outside of the considered seismic zones meaning, all seismic events outside the identified zones are marked as noise. When a neighbourhood of a given point doesn't comply with the definition of a dense zone, that point is considered noise. Translating this in the context of seismic events, if in a given zone there is a smaller number of seismic events, these seismic events will be considered noise and the zone will considered a background zone. This shows an advantage of the way DBSCAN works and proves to be a very useful feature in the case of seismic clustering.



(a) Seismic event clusters for Mainland Portugal.

(b) Seismic zonation for Mainland Portugal, adapted from [11].

Figure 3. Clustering Results and seismic zonation proposed by [11].

The very good fit between clusters and the considered seismic zones obtained, suggest that DBSCAN algorithm is a practical and fast tool for seismic zonation. If applied continuously, considering the most recent records of seismic events, the resulting clusters could provide a clear picture of how seismic behaviour is evolving in certain ares, moreover, if this evolution happens to be geographical, the resulting clusters may include new areas indicating its growth or translation. This could be used to amend or propose changes to already defined seismic zones. For instance, with the results presented in this work, the south of Portugal is contained in a cluster that is included in the seismic zone identified as S02 in figure 3. This could suggest that more of the area of southern Portugal should be included in S02.

4.2. Clustering results with the proposed distance index

In this subsection, the DBSCAN is applied to seismic data for Mainland Portugal, where the distances between seismic events are given by the proposed distance index. A set of values for k_t and k_s was used to understand the effect of these values on the clustering. Furthermore, the parameters of the DBSCAN algorithm used were the same for each combination of k_t and k_s, those being $\epsilon = \frac{1}{1-k_t} \times 15$ and $minPts = 5$. Note that the adjustment factor $\frac{1}{1-k_t}$ was also taken into account. Table 2 summarises the results for each combination of values. Firstly, considering the number of noise points, it can be seen that increasing both k_t and k_s leads to a decrease in the number of

Table 2. Results from DBSCAN with the proposed distance index for different values of k_t and k_s.

	$k_t = 0.20$		$k_t = 0.25$	
	Number of clusters	Number of noise events	Number of clusters	Number of noise events
$k_s = 0.05$	40	641	35	628
$k_s = 0.1$	32	456	29	450

noise points, being k_s the factor that has the most influence in this aspect. Secondly, focusing now on the number of resulting clusters, these too are reduced when k_t and k_s increase, again k_s having more influence than k_t in the reduction of the number of clusters.

When $k_t = 0.25$ and $k_s = 0.1$, the number of clusters generated is equal to the number of clusters obtained when only the Haversine distance was considered. However, the number of noise events decreased by two hundreds, result of the k_s factor. Despite the number of clusters being equal, the proposed distance index does not fail to separate events in the time dimension. Figure 4 shows the clusters that resulted from the DBSCAN algorithm with the proposed distance index with parameters $k_t = 0.25$, $k_s = 0.1$, $\epsilon = \frac{1}{1-k_t} \times 15$ and $minPts = 5$, without the events considered noise. There are clear geographically overlapping clusters, meaning the seismic events originated from the same geological structures are grouped together in different time frames.

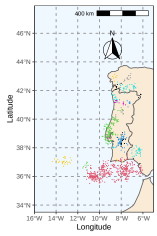
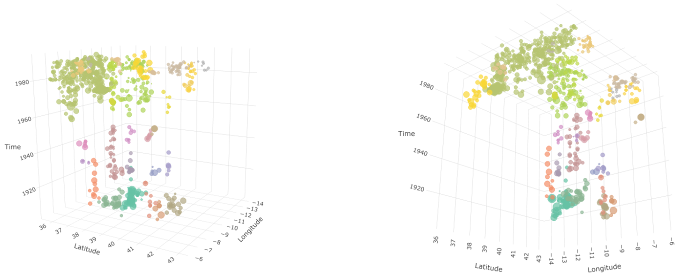


Figure 4. Seismic event clusters for Mainland Portugal, with proposed distance index ($k_t = 0.25$ and $k_s = 0.1$).

Figure 5 displays the clusters in three dimensional space, in order to better present the time separated clusters that overlap geographically. This allows for the identification of distinct sets of activity over time from the same geological structures, which could be used to study how often this activity is expected to happen. Moreover, the k_s factor aggregates seismic events to clusters that otherwise would be considered noise, hence the reduction in the number of events classified as noise. This could help relate earthquakes with epicentres in background zones to bigger events that originated elsewhere. Comparing these clusters to the previously presented seismic zones (figure 3), the earthquakes generated in the zones classified as Active Continental Regions by [11] are clustered together, fitting the defined ACR very well, aggregating almost all of the seismic events of these regions into the same cluster. This suggests that, all of the zones considered to be ACR influence each other trough larger magnitude earthquakes.

5. Conclusion

In this work, firstly, the DBSCAN algorithm is applied to a catalog of seismic events that occurred in the Portuguese Mainland and surrounding maritime area, taking only into account the geographic distance between the epicentres of seismic events. The



(a) First point of view. (b) Second point of view.

Figure 5. Results for Mainland Portugal's seismic data, DBSCAN with the proposed distance index.

resulting clusters are used to identify seismic zones for the area under consideration, which are then compared to other seismic zonation already defined in the literature. The resulting clusters reveal a very good fit with the zonation considered. Furthermore, no seismic event clusters were identified outside of the defined seismic zones, as the events outside these zones were considered noise by the DBSCAN algorithm which, in geological terms, translates to have had originated in so-called background zones. Background zones include small numbers of seismic events when compared to other areas. These results validate the DBSCAN algorithm as a useful analytical tool for identifying seismic zones. With little information, considering only the epicentre's longitude and latitude, and some parameter tuning, the algorithm provides good results when compared to seismic zones defined by geological experts. The notion of noise, inherent to the DBSCAN algorithm, shows to have a translation in a geological dimension, converting into an event from a background zones.

In a second phase, the distance index to be applied with the DBSCAN algorithm is proposed. By manipulating the distance between two given points, one can encourage the joint clustering of those points. In a seismic zoning context, given that the magnitude of an event influences the occurrence of other events and that this influences fades over time, one wants to determine how this influence can affect the clustering of events. This was the reasoning behind the proposed distance index. Simulation results show exactly what was intended, a separation in time of three geographically overlapping clusters. Subsequently, the distance index was used to cluster the seismic events for Mainland Portugal. A set of experimentations is made to determine the influence of k_t and k_t on the cluster formation. The increase of both factors determines the decrease in the number of resulting clusters and the decrease in the number of events classified as noise. The application of the distance index in this seismic catalog has clearly interesting results, as it enables the identification of distinct clusters in time that overlap geographically. Further work should focus more on the impact of both factors k_t and k_t on clustering results, and what could be done with the results achieved by combining the distance index proposed with DBSCAN, namely an analysis of how often should a cluster of seismic events be identified provided they originate from a specific geological structure.

References

[1] M. Ester, H.P. Kriegel, J. Sander, and X. Xu, *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (1996), pp. 226–231.

[2] G. Georgoulas, A. Konstantaras, E. Katsifarakis, C.D. Stylios, E. Maravelakis, and G.J. Vachtsevanos, *"Seismic-mass" density-based algorithm for spatio-temporal clustering*, Expert Systems with Applications 40 (2013), pp. 4183–4189.

[3] M. Hahsler, M. Piekenbrock, and D. Doran, *Dbscan: Fast density-based clustering with R*, Journal of Statistical Software 91 (2019).

[4] A.K. Jain, *Data clustering: 50 years beyond K-means*, Pattern Recognition Letters 31 (2010), pp. 651–666. Available at `http://dx.doi.org/10.1016/j.patrec.2009.09.011`.

[5] N.A. Karri, M. Yousuf Ansari, and A. Pathak, *Identification of seismic zones of India using DBSCAN*, 2018 International Conference on Computing, Power and Communication Technologies, GUCON 2018 (2019), pp. 65–69.

[6] L. Pinheiro, R. Wilson, R. Pena dos Reis, R. Whitmarsh, and A. Ribeiro, *The western Iberia Margin: a geophysical and geological overview*, Proceedings of the Ocean Drilling Program, 149 Scientific Results 149 (1996).

[7] D. Rebuli and S. Kohler, *Using clustering algorithms to assist short-term seismic hazard analysis in deep South African mines*, Proceedings of the Seventh International Conference on Deep and High Stress Mining (2014), pp. 699–708.

[8] M.C. Rodrigues and C.S. Oliveira, *Seismic zones for Azores based on statistical criteria*, Natural Hazards and Earth System Sciences 13 (2013), pp. 2337–2351.

[9] M.C.M. Rodrigues and C.S. Oliveira, *Considering spatial memory to estimate seismic risk: the case of the Azores Archipelago*, GEM - International Journal on Geomathematics 11 (2020), pp. 1–15. Available at `https://doi.org/10.1007/s13137-020-00152-0`.

[10] R.W. Sinnott, *Virtues of the Haversine*, Sky and Telescope 68 (1984), p. 158.

[11] S.P. Vilanova and J.F. Fonseca, *Probabilistic seismic-hazard assessment for Portugal*, Bulletin of the Seismological Society of America 97 (2007), pp. 1702–1717.

[12] P. Viswanath and R. Pinkesh, *L-DBSCAN: A fast hybrid density based clustering method*, Proceedings - International Conference on Pattern Recognition 1 (2006), pp. 912–915.

[13] A. Zamani and N. Hashemi, *Computer-based self-organized tectonic zoning: A tentative pattern recognition for Iran*, Computers and Geosciences 30 (2004), pp. 705–718.