JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 9, No. 4, October 2025, pp. 1306-1322

An ST-DBSCAN Approach to Spatio-Temporal Clustering of Earthquake Events in West Java, Indonesia

Dwi Kartika Widyawati[1], Achmad Fauzan[1*]
[1]Department of Statistics, Universitas Islam Indonesia, Indonesia
achmadfauzan@uii.ac.id

Article History:
Received : 09-07-2025
Revised : 16-08-2025
Accepted : 19-08-2025
Online : 01-10-2025

Keywords:
Earthquakes;
Clustering;
Spatio-Temporal;
ST-DBSCAN.

ABSTRACT

Earthquakes are among the most frequent and damaging natural disasters in Indonesia, particularly in West Java Province, where their unpredictable occurrence often causes casualties and severe infrastructure damage. This study aims to identify spatial and temporal patterns of earthquakes to support disaster risk mitigation efforts. A quantitative exploratory approach was applied using the Spatio-Temporal Density-Based Spatial Clustering of Applications with Noise (ST-DBSCAN) method, which groups earthquake events based on their proximity in space and time while distinguishing random noise. The analysis utilized secondary earthquake data from the Meteorology, Climatology, and Geophysics Agency (BMKG) covering the period January 2022 to December 2023. The results revealed eight distinct clusters and several high-risk zones with strong internal similarity (silhouette coefficient = 0.721), indicating stable and stationary patterns over the observed period. These findings demonstrate that ST-DBSCAN is effective in detecting consistent earthquake-prone areas. More importantly, the study provides practical implications for disaster mitigation, including the development of targeted early warning systems, prioritization of high-risk areas such as Cianjur Regency, and more efficient allocation of resources to strengthen preparedness and community safety.

Crossref
https://doi.org/10.31764/jtam.v9i4.33079

——————————— ◆ ———————————

A. INTRODUCTION

Natural disasters are events that threaten people's lives and welfare, caused by natural phenomena, either from within the earth such as earthquakes, or from weather and climate change (Chaudhary & Piracha, 2021; Prasad & Francescutti, 2017; Sealey & Logan, 2019). Indonesia is very vulnerable to earthquakes because it is located at the boundaries of three tectonic plates: Indo-Australian, Eurasian, and Pacific, their motions often triggering earthquakes due to plate collisions or shifts (Murdiaty et al., 2020; Zaccagnino & Doglioni, 2022). In addition, Indonesia is located along the Pacific Ring of Fire and the Alpide Belt, two of the world's major seismic belts, making this country one of the regions with the highest seismic activity in the world (Jufriansah et al., 2021).

West Java Province is one of the most seismically active areas in Indonesia. In 2022, the Meteorology, Climatology, and Geophysics Agency (BMKG) recorded 1,290 earthquakes in the province, most occurring on land at shallow depths. The most destructive event, a magnitude 5.6 Mw earthquake in Cianjur on November 21, 2022, caused 602 fatalities, thousands of injuries, and widespread damage to homes and public facilities (Ridwan, 2023). The province's location near the southern subduction zone and the presence of active inland faults such as the

Cimandiri, South Garut, and Lembang faults further heighten its vulnerability (Irsyam et al., 2020; Supendi et al., 2023).

Given this high seismic risk, identifying patterns in earthquake occurrences is essential for disaster risk mitigation. Clustering analysis offers a way to map high-risk areas, enabling targeted preparedness measures such as the optimal placement of shelters, evacuation routes, and emergency assembly points (Yang et al., 2025; Zhang et al., 2023). One of the clustering methods that can be used is Spatio Temporal-Density Based Spatial Clustering Applications with Noise (ST-DBSCAN). The ST-DBSCAN algorithm is included in the category of nonparametric algorithms in unsupervised learning. Unsupervised learning is a learning method that uses machine learning algorithms to analyze and group unlabeled data sets (Berry et al., 2020). In the context of ST-DBSCAN, the algorithm is able to identify patterns or groups in spatial-temporal data without requiring additional information or assumptions about data grouping. The ST-DBSCAN method considers proximity in both spatial and temporal dimensions, is able to handle noise and outliers in the data, and has parameters that are relatively easy to implement, such as spatial distance parameters (epsilon) and temporal distance parameters (epsilon time). This algorithm also has the ability to find groups based on non-spatial, spatial, and temporal values of objects (Birant & Kut, 2007). ST-DBSCAN is a very effective method in identifying clusters in large spatial databases (Gaonkar & Sawant, 2013).

Various studies have advanced spatiotemporal clustering of earthquake data using ST-DBSCAN-based approaches. Sonhaji (2023) applied ST-DBSCAN to seismic data on Java Island, yielding 13 clusters and 10 noise points with a silhouette coefficient of 0.538. Nicolis et al. (2024) developed ST-DBSCAN-EV, which adapts the epsilon radius based on point density, and demonstrated superior performance (F1 > 0.8) on major Chilean earthquakes. Meanwhile, Sharma *et.al* (2023) proposed a two-stage clustering method combining Self-Organizing Maps and density-based temporal clustering to effectively separate aftershocks from background seismicity across several global regions, including Indonesia and Chile. In addition, there is also a study using the same method for earthquake points on Sulawesi Island resulting in 60 clusters and 216 noises (Jales et al., 2021). Faraouk et al (2023) conducted an ST-DBSCAN analysis on forest fire hotspots in Riau Province from 2015 to 2020.

This study applies ST-DBSCAN to earthquake events in West Java from January 2022 to December 2023 to identify spatial and temporal clustering patterns and assess their stability. The novelty of this research lies in explicitly determining the most effective ST-DBSCAN parameter set through a structured evaluation of multiple spatial, temporal, and density thresholds using cluster quality metrics. This approach differs from previous works that often use default or heuristic parameters, ensuring that the resulting clusters are both statistically robust and contextually relevant for the study area. The findings provide practical benefits for disaster management agencies, including accurate mapping of high-risk zones, improved early warning dissemination, and better-targeted resource allocation in earthquake-prone regions.

B. METHODS

1. Data

This study is a quantitative descriptive research that uses earthquake data from West Java Province, covering the period from January 2022 to December 2023, obtained from the official BMKG repository (https://repogempa.bmkg.go.id/). The study area encompasses both land and offshore regions within the province. The variables analyzed include event date, longitude, latitude, magnitude, and depth, with detailed descriptions provided in Table 1.

Table 1. Operational Variables Definition

Variable	Definition
Date	Earthquake dates in West Java Province (Date/Month/Year)
Longitude	Latitude coordinates of the point where the earthquake occurred in West Java Province (°)
Latitude	Longitude coordinates of the point where the earthquake occurred in West Java Province (°)
Magnitude	The strength of the earthquake that occurred in West Java Province (RS)
Depth	The depth of the focus point of the earthquake that occurred in West Java Province (KM)

2. Research Methodology

This study used the (ST-DBSCAN) method to group the distribution of earthquake locations in West Java Province. The flowchart of this method is shown in Figure 1. As presented in Figure 1, the steps for analyzing the distribution of earthquakes in West Java Province from January 2022 to December 2023 include several stages as follows.

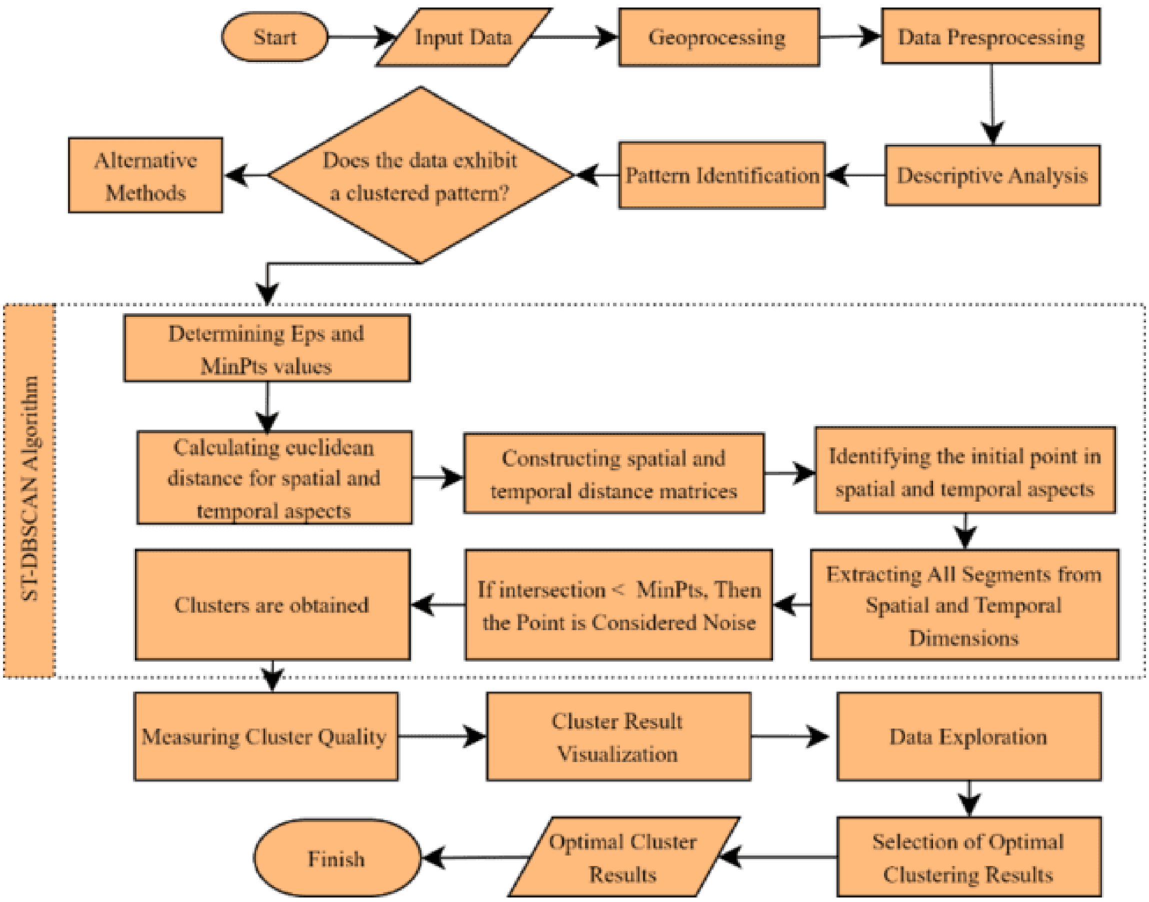

Figure 1. Research Flowchart

First, the earthquake location data in West Java Province taken from the BMKG website (https://repogempa.bmkg.go.id) is geoprocessed, then georeferencing and digitization are carried out to create regional maps, including the addition of coordinate points in shapefile format. Next, the data is processed by deleting duplicated data and data outside the West Java Province area. This step aims to maintain data accuracy and consistency, because duplication can cause inconsistencies in the analysis. Furthermore, descriptive analysis is performed to provide an overview of the distribution of earthquakes in this region during the period. Then, the data distribution pattern is analyzed using nearest-neighbor analysis (Equation 2 and Equation 4), and the results of the data distribution are examined to determine whether there is a clustered pattern (using Equation 6). If the data shows a clustered pattern, the ST-DBSCAN method is applied. The quality of the clusters is then assessed using the silhouette coefficient (Equation 7), and data exploration is carried out by changing the MinPts value so that the number of clusters formed is greater, which allows for broader identification of earthquake-prone area. Finally, visualization and interpretation of the clustering results by the ST-DBSCAN method are presented to understand the distribution of earthquake risk in West Java Province. In the data preprocessing stage, Quantum Geographic Information System (QGIS) software was employed, whereas the analysis stage was conducted using R software.

3. Nearest Neighbor Analysis and ST-DBSCAN

Nearest neighbor analysis is a method designed to see patterns in point data in two or three dimensions (Fahira & Nooraeni, 2023; Halder et al., 2024; Philo & Philo, 2022a; Soltisz et al., 2024). This method involves calculating the average distance between all points and their nearest neighbors. The nearest neighbor index is expressed as the ratio of the observed distance divided by the expected distance. The distance used is the Euclidean distance presented in Equation (1) for the spatial aspect and Equation (2) for the temporal aspect. The nearest neighbor index is expressed as R and its calculation is written in Equation (3).

$$d_s(i,j) = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \tag{1}$$

$$d_t(i,j) = \left|X_{date(i)} - X_{date(j)}\right| \tag{2}$$

$$R = \frac{\bar{d}_{obs}}{\bar{d}_{ran}} \tag{3}$$

$\bar{d}_{obs}$ is the observed mean nearest-neighbor or the average observed distance between each point and its nearest neighbor and $\bar{d}_{ran}$ is the average expected distance for the given points in a random pattern, with Equation (4).

$$\bar{d}_{obs} = \frac{\sum_{i=1}^{n} d_i}{n} \; and \; \bar{d}_{ran} = \frac{1}{2\sqrt{p}} = \frac{1}{2\sqrt{\frac{n}{A}}} \tag{4}$$

d_i is the distance between i and its nearest neighbor, n is the number of points, and A is the minimum rectangular area around all points, or the specified area value and $SE_{\bar{d}}$ is the standard error of the nearest neighbor mean and c is the test statistic, with the calculation formula given in Equations (5) and (6).

$$c = \frac{\bar{d}_{obs} - \bar{d}_{ran}}{SE_{\bar{d}}} \tag{5}$$

$$SE_{\bar{d}} = \frac{0.26136}{\sqrt{np}} = \frac{0.26136}{\sqrt{\frac{n^2}{A}}} \tag{6}$$

The nearest neighbor index has a value ranging from 0 to 2.15 (Indrawan & Adrianto, 2016). The nearest neighbor index with a value of 0 indicates a completely clustered pattern. Meanwhile, the nearest neighbor index with a value of 2.15 indicates a completely dispersion pattern. A random pattern is indicated by a nearest neighbor index value of 1 (Philo & Philo, 2022b). The sample distribution from test c is a normal distribution, therefore c is the normal standard deviation. Hypothesis testing on the c test statistic (zscore) involves the null hypothesis (H_0) stating randomly distributed data and the alternative hypothesis (H_1) stating cluster distributed data. The zscore value, which is normally distributed, is used as the test criteria. If the absolute value of Z_{score} > Z_{table}, then H_0 is rejected (Aslam, 2022, 2024).

The ST-DBSCAN algorithm requires three parameters, a parameter stating the distance between 2 objects on the earth's surface (Eps1), a parameter stating the distance between time events (Eps2), and the number of members of a cluster (MinPts). The stages of the ST-DBSCAN algorithm begin by determining the Eps parameter, which are the spatial and temporal distance parameters, and MinPts (the minimum number of cluster members) through a trial and error approach based on k-dist graph analysis (An et al., 2023; Birant & Kut, 2007; Iswari, 2022). Subsequently, we calculate the Euclidean distance between objects by considering the spatial and temporal dimensions, using Equations (1) and Equation (2). Next, a distance matrix is created that contains the distance between each pair of objects from a total of n objects by combining the spatial and temporal dimensions. From the starting point, all points are selected based on the spatial and temporal dimensions that meet the following criteria: A for spatial distance (x ≤ Eps1 in the spatial distance matrix) and B for temporal distance (x ≤ Eps2 in the temporal distance matrix). An intersection is taken between spatial and temporal dimensions such that it satisfies the condition $A \cap B$, where x is in A and B. If the number of objects in the intersection is less than MinPts, the point will be considered as noise. A cluster is formed when the selected points satisfy the criteria Eps1, Eps2, and MinPts. If point p is a border point and there are no other points in the intersection, then we proceed to the next point to form a new cluster. The steps from selecting a point to forming a cluster are repeated until all points have been processed. If there are two adjacent clusters, i.e. C1 and C2, where a point q may be in both, the algorithm will place point q in the cluster that first detects it.

4. Model Evaluation

One of the measures to assess cluster quality is the silhouette coefficient. The calculation of the silhouette coefficient is presented in Equation (7) (Belyadi & Haghighat, 2021; Subasi, 2020).

$$S(i) = \frac{b(i)-a(i)}{max\{a(i),b(i)\}} \tag{7}$$

$$a(i) = \frac{1}{|A|-1}\sum_{j\in A, j\neq 1} d(i,j) \tag{8}$$

$$b(i) = \min_{C\neq A}\ d(i,j) \tag{9}$$

$$d(i,C) = \frac{1}{|C|}\sum_{j\in C,} d(i,j) \tag{10}$$

The values |A| and |C| are the cluster sizes, which are the total number of objects in cluster A and cluster C, respectively. The value of $a(i)$ indicates the density of the cluster containing object i, while $b(i)$ describes the distance of object i from other clusters. If $a(i)$ is small and $b(i)$ is large, $S(i)$ approaches 1, indicating a dense and far-separated cluster. Conversely, if $a(i)$ is large and $b(i)$ is small, $S(i)$ approaches -1, indicating a cluster that is not dense and close to other clusters. The closer the value of $S(i)$ is to 1, the better the clustering of data. Conversely, the closer to -1, the worse the clustering (Han et al., 2012; Ikotun et al., 2025; Papakostas, 2025; Zheng & Zhao, 2018).

C. RESULT AND DISCUSSION

1. Exploratory Spatial Data Analysis

From January 2022 to December 2023, earthquakes in West Java Province occurred at 790 locations. Before grouping the data, the distribution pattern of the data is examined. As a basis for whether the earthquake points are scattered or clustered, the nearest neighbor analysis is used. Based on the calculation results using the nearest neighbor analysis, the nearest neighbor index value is 0.45 which indicates a clustered pattern. In addition, the absolute value of the z_{score} of 29.313 is greater than the z_{table} of 1.96. Based on this, the data on the location points of the distribution of earthquakes in West Java Province from January 2022 to December 2023 has a clustered data distribution pattern.

2. Clustering Using the ST-DBSCAN Algorithm

In this study, using the values of Eps1, Eps2, and MinPts, where the Eps1 interval is between 2 KM to 22 KM, then Eps2 7 days, 14 days, 21 days and 28 days, while the MinPts interval used is 3 to 9. From the combination of the three parameters, the best cluster is determined based on the silhouette coefficient. The results of the combination of the three parameters are presented in Table 2.

Table 2. Silhouette Coefficient Results

No	MinPts	Eps 1 (KM)	Eps2 (Hari)	Number of Clusters	Number of Noise Points	Silhouette Coefficient
1	3	2	7	41	103	0.245
2	3	6	7	20	13	0.543
⋮	⋮	⋮	⋮	⋮	⋮	⋮
68	5	6	28	21	37	0.556
⋮	⋮	⋮	⋮	⋮	⋮	⋮
134	8	6	21	10	132	0.560
⋮	⋮	⋮	⋮	⋮	⋮	⋮
168	9	22	28	2	0	0.676

Based on Table 2, the maximum silhouette coefficient value is 0.721, namely at MinPts = 8, Eps1 = 2, and Eps2 = 21, with 8 clusters formed and 262 noise. The silhouette coefficient value of 0.721 means that the cluster has a strong structure. The silhouette coefficient plot for each MinPts is shown in Figure 2. Figure 3 presents a 3D visualization of the silhouette coefficient at each Eps1, Eps2, and MinPts.

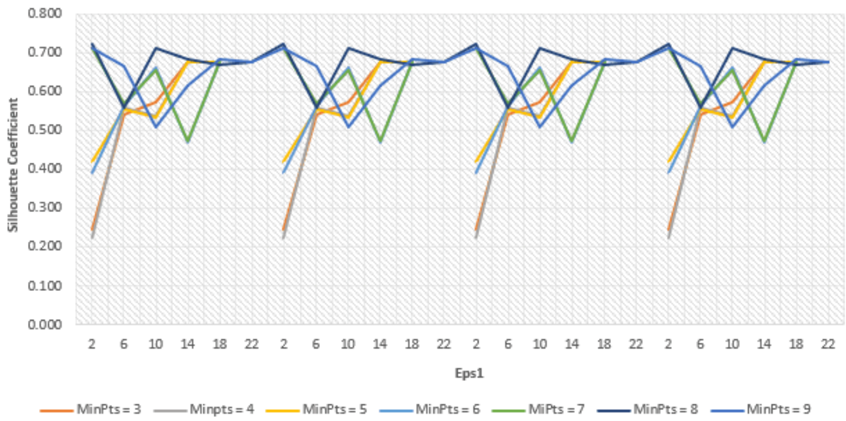

Figure 2. Silhouette Coefficient Plot against Each MinPts Value.
Source: Processed data

As presented in Figure 2, we obtained the overall average MinPts of 3 to 9, with the highest silhouette coefficient value at MinPts = 8. It is also observed that the fluctuations of the graph is the same because the time distance (days) between objects 7 days, 14 days, 21 days, and 28 days the silhouette coefficient value produced is the same or there is no change over time (days).

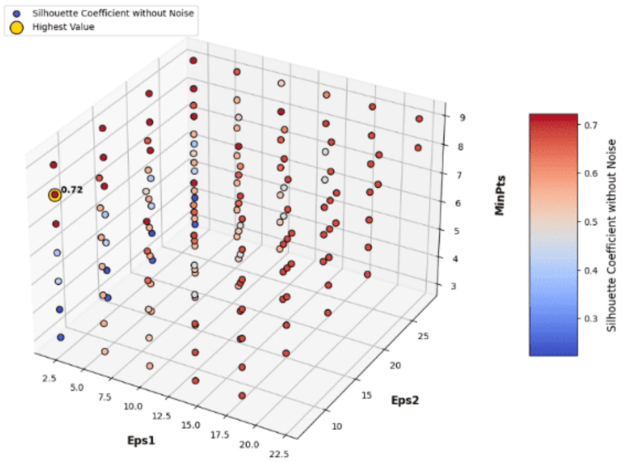

Figure 3. Visualisation of the Silhouette Coefficient (Source: Processed data)

In Figure 3, the X axis represents Eps1, the Y axis represents Eps2, and the Z axis represents MinPts. From a total of 168 pairs of Eps1, Eps2, and MinPts parameters, the highest silhouette coefficient value is 0.72, as indicated by the yellow point in the Figure, and the red points mark silhouette coefficient values that are close to 1.

3. Data Exploration

In addition to clustering based on the highest silhouette coefficient, as explained in the research methodology, data exploration for disaster mitigation was also conducted. During this stage, an earthquake incident map was created. We explored different outcomes by lowering the MinPts value, which resulted in the formation of more clusters, thereby identifying more earthquake-prone areas. The MinPts values used were 3 and 9, as a smaller MinPts value leads to more points being included in clusters. This occurs because the requirements for cluster formation become less stringent. With MinPts = 3, two clusters were formed with no noise and a silhouette coefficient of 0.676, whereas with MinPts = 9, six clusters were formed, along with 279 noise points, and a silhouette coefficient of 0.714. A visualization of the data exploration results is presented in Figure 4.

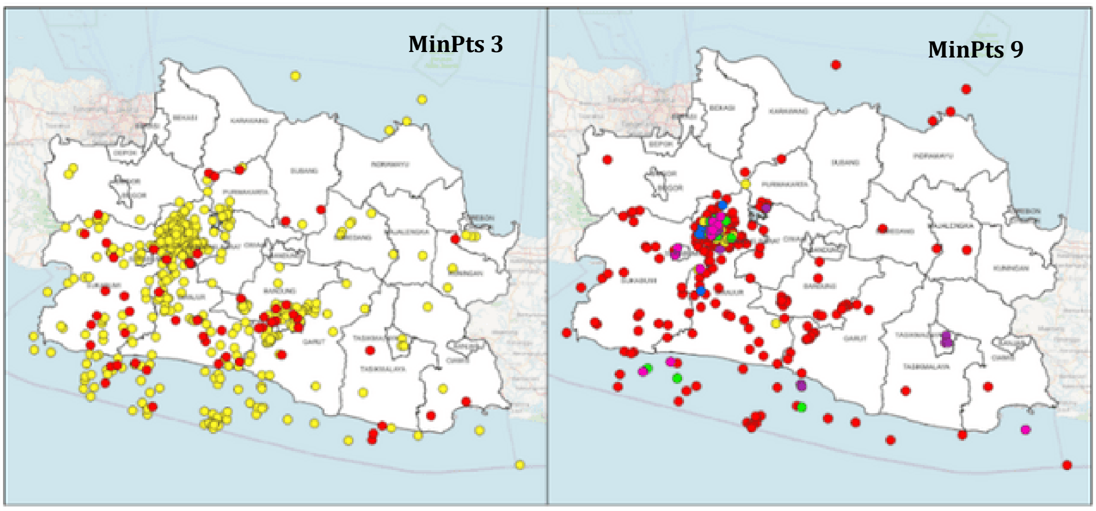

Figure 4. Visualisation of the number of clusters at MinPts=3 and MinPts=9
Source: Processed data

4. Best Clustering Selection

The best clustering is determined by comparing the clustering results and the amount of noise, based on the highest silhouette coefficient for each MinPts value. A comparison of the clustering results is presented in Table 3.

Table 3. Optimal Parameters for Each MinPts Value.

MinPts	Eps1 (Km)	Eps 2 (Days)	Cluster	Noise	Silhouette Coefficient
3	14	7	2	0	0.676
4	14	7	2	0	0.676
5	14	7	2	0	0.676
6	18	7	2	0	0.676
7	2	7	9	243	0.711
8	2	7	8	262	0.721
9	2	7	6	279	0.714

Based on Table 3, it is observed that the combination of Eps1 values from 2 to 22 and MinPts values from 3 to 9 that obtained the highest silhouette coefficient of 0.721 is MinPts = 8, Eps1 = 2, and Eps2 = 7, which means that to form a cluster, a minimum of 8 points are needed with a maximum distance between points of 2 KM and a maximum distance of 7 days. The best clustering results were obtained with the best parameters Eps1 = 2, Eps2 = 7, and MinPts = 8, involving 8 clusters and 262 noise, with a silhouette coefficient value (excluding the noise) of 0.721, suggesting that the cluster has a strong structure as the value is close to 1. The characteristics of the depth and magnitude of the earthquakes in each cluster are presented in Table 4.

Table 4. Earthquake characteristics of each cluster

Cluster	Variable	Statistic				
		Mean	Median	Min	Max	Standard Deviation
A	Magnitude (RS)	2.26	2.18	1.14	6.40	0.62
	Depth (KM)	16.32	10.00	5.00	6.40	17.00
B	Magnitude (RS)	2.40	2.23	1.64	3.75	0.56
	Depth (KM)	28.41	12.00	10.00	265.00	51.49
C	Magnitude (RS)	2.09	2.07	1.55	2.46	0.29
	Depth (KM)	30.00	13.00	10.00	123.00	38.11
D	Magnitude (RS)	2.80	2.70	1.71	4.84	0.82
	Depth (KM)	41.36	10.00	10.00	148.00	44.06
E	Magnitude (RS)	2.88	2.76	1.78	4.15	0.60
	Depth (KM)	18.86	10.00	2.29	3.84	0.43
F	Magnitude (RS)	2.86	2.72	2.38	3.84	0.43
	Depth (KM)	42.67	2.84	10.00	95.00	22.69
G	Magnitude (RS)	2.68	2.66	2.22	3.22	0.33
	Depth (KM)	23.56	10.00	10.00	105.00	31.04
H	Magnitude (RS)	3.04	3.00	2.62	3.87	0.38
	Depth (KM)	22.88	10.00	10.00	108.00	34.42

Based on Table 4, we conclude that the cluster with the highest average earthquake magnitude is cluster 8 at 3.04 RS. The following is a visualization of cluster 8 which is shown in Figure 5.

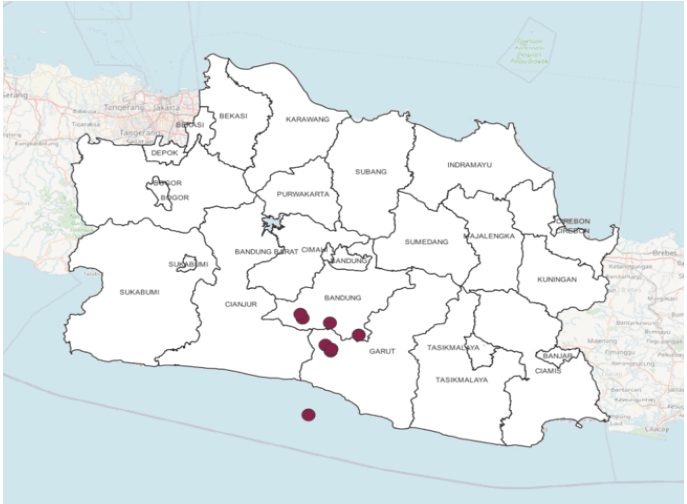

Figure 5. Visualisation of Cluster 8
Source: Processed data

In Figure 5, cluster 8 contains 8 earthquake points located in the Bandung Regency, Garut Regency and South Sea of the West Java Province, with the time span of the earthquake occurrence from September 4, 2023 to September 8, 2023. The next stage involves analyzing the Spatio-Temporal pattern, which is divided into four types: stationary, reappearing, oscillating, and track. Large clusters with at least 30 points can be analyzed for their patterns, while clusters with less than 30 points are considered small and their patterns cannot be studied (Poelitz C. & Andrienko N., 2010). Based on the research of Poelitz and Andrienko, in order to analyze the Spatio-Temporal pattern at earthquake points in West Java Province, the focus of this study is directed towards large clusters. From the best parameter results of Eps1 = 2, Eps2 = 7, and MinPts = 8, the following results of 8 clusters are obtained along with their respective number of points, as shown in Table 5.

Table 5. Number of points in each cluster

Cluster	Number of Points
A	430
B	29
C	12
D	14
E	14
F	12
G	9
H	8

Based on Table 5, there is only one cluster with more than 30 points, which is cluster A, with 430 points. The following are the characteristics of the depth and strength of earthquakes in large clusters, which is presented in Table 6.

Table 6. Earthquake Characteristics in Large Clusters

Cluster	Time Interval of Earthquake Occurences	Variable	Statistic				
			Mean	Median	Min	Max	St.Dev
A	21 November 2022 to 25 December 2022	Magnitude (RS)	2.26	2.18	1.14	6.40	0.62
		Depth (KM)	16.32	10.00	5.00	6.40	17.00

Table 6 shows that the average earthquake strength in cluster A is 2.26 RS while the average earthquake depth is 16.32 KM. In addition, in cluster A, the standard deviation value is smaller than the average earthquake magnitude value, indicating that the earthquake magnitude data in cluster A does not vary. Furthermore, cluster A is analyzed for its pattern, which is illustrated in Figure 6.

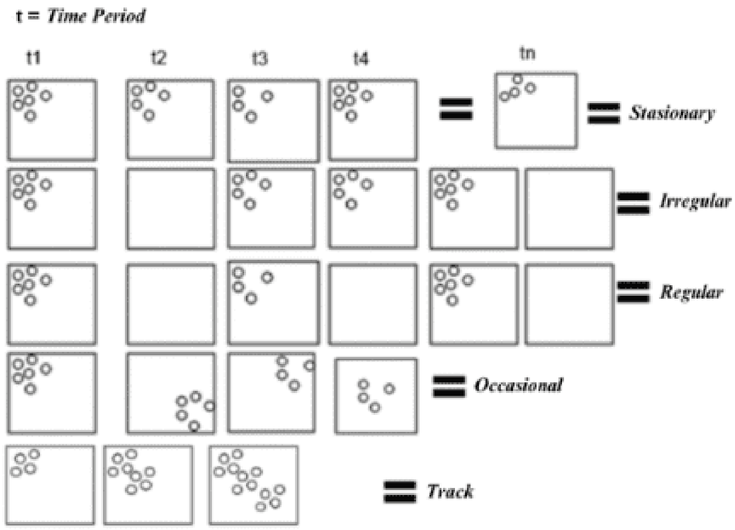

Figure 6. Spatio-Temporal Pattern Visualisation(Putri et al., 2023)

The clusters are divided into periods, where since the best parameter in this study is Eps2=7, the length of the period is seven days allowing comparison of the earthquake point distributions between periods. In cluster A, there are five periods as shown in Figure 7.

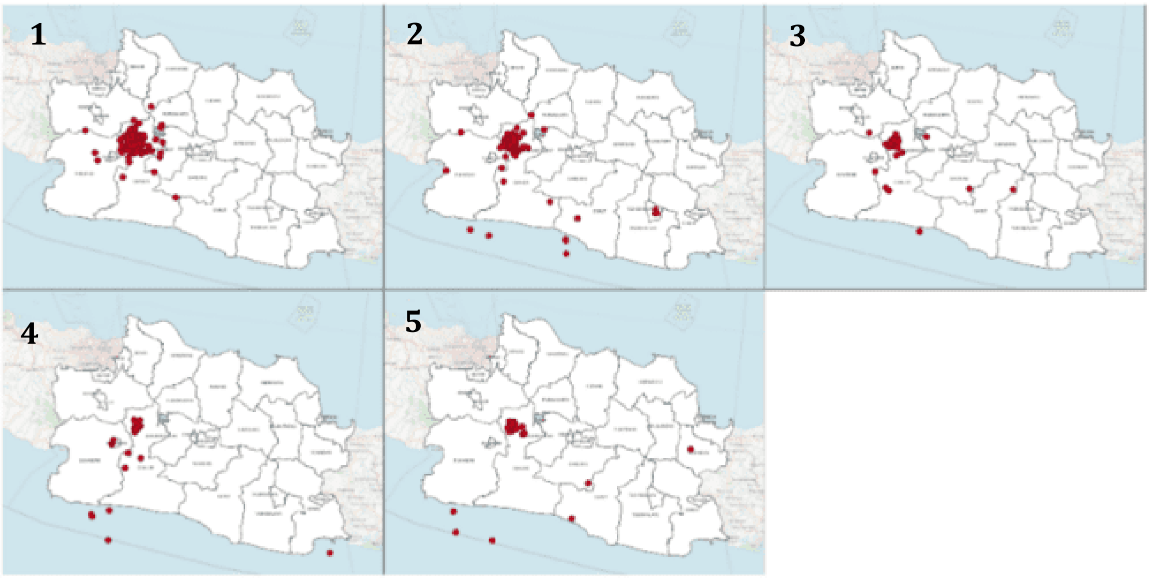

Figure 7. Spatio-Temporal Pattern Visualisation
Source: Processed Data

Figure 7 shows that in period 1 to period 5, the epicenter of the earthquake did not shift, but several points shifted not far from the epicenter. To investigate the area where the epicenters of the earthquakes are located along with the distribution of other points, we present them in Table 7.

Table 7. Areas Containing Earthquake Epicenters

Period	Description of Area
1	Centered in Cianjur Regency and other earthquake points are located in West Bandung Regency, Sukabumi Regency, Purwakarta Regency, Karawang Regency and Bogor Regency
2	Centered in Cianjur Regency and other earthquake points are located in West Bandung Regency, Sukabumi Regency, Purwakarta Regency, Garut Regency, Tasikmalaya City, and the South Sea of West Java Province
3	Centered in Cianjur Regency and other earthquake points are located in West Bandung Regency, Sukabumi Regency, Bogor Regency, Bandung Regency, Garut Regency, and the South Sea of West Java Province
4	Centered in Cianjur Regency and other earthquake points are located in Sukabumi Regency and the South Sea of West Java Province
5	Centered in Cianjur Regency and other earthquake points are located in Bandung Regency, Garut Regency, Kuningan Regency, and the South Sea of West Java Province

Based on Table 7 in period 1 to period 5 the earthquake epicenter is in Cianjur Regency. It is observed that the epicenter did not move from period 1 to period 5, which leads to the conclusion that cluster A has a stationary pattern, which refers to a collection of points or areas that tend to remain the same or do not change over time.

5. Data Exploration of the Best Cluster

This data exploration is used to examine the distribution of earthquake data in West Java Province based on the best cluster. Clusters with the number of points between 10 and 30 consist of five clusters: Cluster B with 29 points, Cluster C with 12 points, Cluster D with 14 points, Cluster E with 14 points, and Cluster F with 12 points. The distribution of earthquake data from these five clusters can be seen in Figure 8.

Based on Figure 8, in Cluster B, the distribution of earthquake points is centered in Cianjur Regency, with other points scattered across Bandung Regency, Sukabumi Regency, and the southern sea of West Java Province. In Cluster C, the earthquake points are centered in Cianjur Regency, with other points scattered in Sukabumi Regency. In Cluster D, the distribution is centered in Cianjur Regency, with other points scattered across Bandung Regency, Sukabumi Regency, Bogor Regency, and the southern sea of West Java Province. In Cluster E, the points are centered in Bandung Regency, with others scattered across Cianjur Regency, Purwakarta Regency, Sukabumi Regency, Garut Regency, and the southern sea of West Java Province. In Cluster F, the earthquake points are not centered in any specific location, but are scattered across Cianjur Regency, Bandung Regency, Garut Regency, Sumedang Regency, Subang Regency, Ciamis Regency, and the southern sea of West Java Province.

Cianjur Regency is the area that is most often the center of earthquake cluster points. Therefore, with this analysis, it is expected that the local government can take appropriate steps to reduce losses, both casualties and material losses, in earthquake-prone areas. For example, Cianjur Regency can strengthen the construction of public facilities and important buildings in accordance with earthquake-resistant building quality standards. Steps such as selecting the right materials, using shock-resistance systems (such as base isolators, sliding pendulums, or shock absorbers to reduce the force transmitted to the main structure), selecting a location that is safe from earthquake risks, and strengthening columns and beams using steel placed in concrete can increase the strength and stiffness of the structure. In addition, disaster mitigation can be carried out by conducting socialization, increasing awareness among the community, and holding basic training on disaster management for officers and residents is also an effective step.

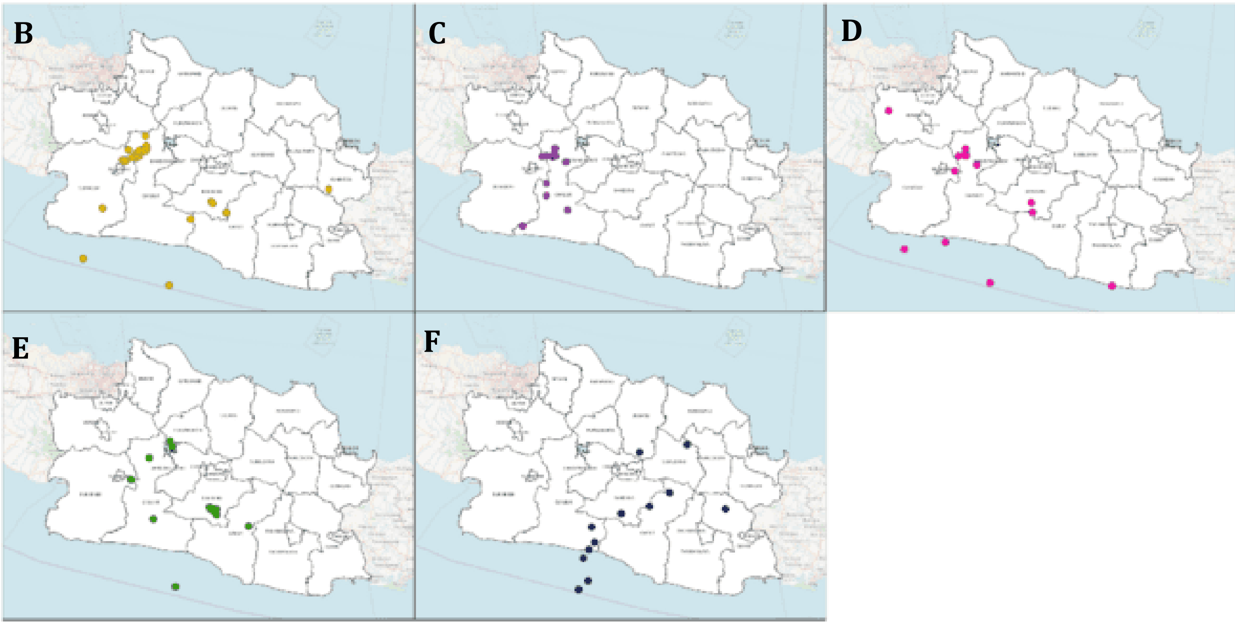

Figure 8. Distribution of Earthquakes Based on the Exploration Results (Source:Processed Data)

The results of the earthquake clustering show that earthquakes in Cianjur tend to be concentrated or have a higher density than other areas in West Java Province. This may indicate that Cianjur has the potential for greater damage due to earthquakes. The Cianjur local government needs to pay attention to these results and prioritize appropriate mitigation measures to protect its citizens from potential earthquake hazards. When an earthquake occurs in Cianjur, people in the area must immediately take steps for their safety. One step that can be taken is to move to areas located north of Cianjur. These areas such as Bekasi, Karawang, Purwakarta, and Bogor, can be considered safer alternatives because they are quite far from the epicenter and have a lower risk of damage. In addition, these areas have stronger infrastructure and can provide better access to health facilities and emergency assistance in emergency situations such as earthquakes. Therefore, it is important for the people of Cianjur to have a good evacuation plan and follow the directions of local authorities to ensure their safety and that of their families in the face of the threat of an earthquake.

The findings of this study, which identified eight earthquake clusters with a strong internal similarity (silhouette coefficient = 0.721), are consistent with previous research demonstrating the effectiveness of ST-DBSCAN in detecting spatio-temporal patterns within seismic data. Similar to the results reported by Birant & Kut (2007); Gaonkar & Sawant (2013), this study confirms the algorithm's robustness in handling noise and identifying meaningful clusters in large spatial databases. However, unlike many previous studies that typically relied on a limited set of parameters, the present research systematically combined variations of spatial distance (Eps1), temporal distance (Eps2), and minimum cluster size (MinPts), resulting in 168 parameter configurations tested. This comprehensive approach not only reinforces earlier findings on the stability of spatio-temporal clustering but also extends them by providing evidence that optimal parameter combinations enhance the precision of cluster detection. The emergence of stationary patterns in high-risk zones further supports earlier studies that emphasized the persistence of seismic activity in areas with active fault systems (Irsyam et al., 2020; Supendi et al., 2023), while at the same time offering localized, actionable insights for disaster mitigation planning in Indonesia.

D. CONCLUSION AND SUGGESTIONS

This study successfully applied the ST-DBSCAN method to identify spatio-temporal clustering patterns of earthquakes in West Java, thereby achieving the main research objective of mapping high-risk zones and understanding their stability. The analysis of 168 parameter combinations revealed the best configuration (Eps1 = 2 km, Eps2 = 7 days, MinPts = 8), producing eight clusters with a silhouette coefficient of 0.721, which indicates robust clustering quality. One major cluster (Cluster A) with 430 points was centered in Cianjur Regency and exhibited a stationary pattern, highlighting a persistent concentration of seismic activity over time. The relevance of these results lies in providing reliable insights into earthquake-prone areas that are directly aligned with disaster risk mitigation objectives. Practically, the findings can support local authorities in prioritizing Cianjur and surrounding regions for disaster preparedness measures, including strengthening earthquake-resistant infrastructure, optimizing the placement of evacuation routes and shelters, and enhancing community awareness and emergency response planning. By focusing on the significance of the clustering outcomes rather than technical details, this study demonstrates the practical value of ST-DBSCAN in supporting adaptive disaster management strategies in seismically active regions.

ACKNOWLEDGEMENT

The author would like to thank the Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia, Indonesia for assisting in the research process.

REFERENCES

An, X., Wang, Z., Wang, D., Liu, S., Jin, C., Xu, X., & Cao, J. (2023). STRP-DBSCAN: A Parallel DBSCAN Algorithm Based on Spatial-Temporal Random Partitioning for Clustering Trajectory Data. *Applied Sciences*, *13*(20), 11122. https://doi.org/10.3390/app132011122

Aslam, M. (2022). Design of a new Z-test for the uncertainty of Covid-19 events under Neutrosophic statistics. *BMC Medical Research Methodology*, *22*(1), 99. https://doi.org/10.1186/s12874-022-01593-x

Aslam, M. (2024). Analysis of imprecise measurement data utilizing z-test for correlation. *Journal of Big Data*, *11*(1), 4. https://doi.org/10.1186/s40537-023-00873-7

Belyadi, H., & Haghighat, A. (2021). Unsupervised machine learning: clustering algorithms. In *Machine Learning Guide for Oil and Gas Using Python* (pp. 125–168). Elsevier. https://doi.org/10.1016/B978-0-12-821929-4.00002-0

Berry, M. W., Mohamed, A., & Yap, B. W. (2020). *Supervised and Unsupervised Learning for Data Science* (M. W. Berry, A. Mohamed, & B. W. Yap, Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-030-22475-2

Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering*, *60*(1), 208–221. https://doi.org/10.1016/j.datak.2006.01.013

Chaudhary, M. T., & Piracha, A. (2021). Natural Disasters—Origins, Impacts, Management. *Encyclopedia*, *1*(4), 1101–1131. https://doi.org/10.3390/encyclopedia1040084

Fahira, A. N., & Nooraeni, R. (2023). Optimasi Parameter ST-DBSCAN dengan KNN dan Algoritma Genetika Studi Kasus: Data Bencana Alam di Pulau Jawa 2021. *Jurnal Komputasi*, *11*(1), 24–33. https://doi.org/10.23960/komputasi.v11i1.3175

Faraouk, K. El, Witriyono, H., Deslianti, D., & Veronika, N. D. M. (2023). ST-DBSCAN Algorithm Implementation At Riau Province Forest Fire Points (2015-2022). *Jurnal Komputer, Informasi Dan Teknologi*, *3*(1). https://doi.org/10.53697/jkomitek.v3i1.1191

Gaonkar, & Sawant. (2013). AutoEpsDBSCAN : DBSCAN with Eps Automatic for Large Dataset. *ICEISTCON*, *2*, 2319–2526.

Halder, R. K., Uddin, M. N., Uddin, Md. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*, *11*(1), 113. https://doi.org/10.1186/s40537-024-00973-y

Han, J., Kamber, M., & Pei, J. (2012). Cluster Analysis. In *Data Mining* (pp. 443–495). Elsevier. https://doi.org/10.1016/B978-0-12-381479-1.00010-1

Ikotun, A. M., Habyarimana, F., & Ezugwu, A. E. (2025). Cluster validity indices for automatic clustering: A comprehensive review. *Heliyon*, *11*(2), e41953. https://doi.org/10.1016/j.heliyon.2025.e41953

Indrawan, & Adrianto. (2016). Spatio-Temporal Clustering Hotspot di Sumatera Selatan Tahun 2002-2003 Menggunakan Algoritme ST-DBSCAN dan Bahasa Pemrograman R. *Jurnal Ilmu Komputer Dan Agri-Informatika*, *3*(2), 112. https://doi.org/10.29244/jika.3.2.112-121

Irsyam, M., Cummins, P. R., Asrurifak, M., Faizal, L., Natawidjaja, D. H., Widiyantoro, S., Meilano, I., Triyoso, W., Rudiyanto, A., Hidayati, S., Ridwan, M., Hanifa, N. R., & Syahbana, A. J. (2020). Development of the 2017 national seismic hazard maps of Indonesia. *Earthquake Spectra*, *36*(1_suppl), 112–136. https://doi.org/10.1177/8755293020951206

Iswari, L. (2022). Profiling the Spatial and Temporal Properties of Earthquake Occurrences Using ST-DBSCAN Algorithm. *2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA)*, 1–8. https://doi.org/10.1109/ICITDA55840.2022.9971295

Jales, Rahmawati, & Widiharih. (2021). *Pengelompokan Titik Gempa di Pulau Sulawesi Menggunakan Algoritma ST-DBSCAN (Spatio Temporal-Density Based Spatial Clustering Application With Noise)*. *10*, 554–561. https://doi.org/https://doi.org/10.14710/j.gauss.10.4.554-561

Jufriansah, A., Pramudya, Y., Khusnani, A., & Saputra, S. (2021). Analysis of Earthquake Activity in Indonesia by Clustering Method. *Journal of Physics: Theories and Applications*, *5*(2), 92. https://doi.org/10.20961/jphystheor-appl.v5i2.59133

Murdiaty, M., Angela, A., & Sylvia, C. (2020). Pengelompokkan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas Dengan Algoritma K-Means. *Jurnal Media Informatika Budidarma*, *4*(3), 744. https://doi.org/10.30865/mib.v4i3.2213

Nicolis, O., Delgado, L., Peralta, B., Díaz, M., & Chiodi, M. (2024). Space-time clustering of seismic events in Chile using ST-DBSCAN-EV algorithm. *Environmental and Ecological Statistics*, *31*(2), 509–536. https://doi.org/10.1007/s10651-023-00594-3

Papakostas, I. (2025). *A Silhouette-Based Deep Clustering Method*.

Philo, C., & Philo, P. (2022a). 2.15 or Not 2.15? An Historical-Analytical Inquiry into the Nearest-Neighbor Statistic. *Geographical Analysis*, *54*(2), 333–356. https://doi.org/10.1111/gean.12284

Philo, C., & Philo, P. (2022b). 2.15 or Not 2.15? An Historical-Analytical Inquiry into the Nearest-Neighbor Statistic. *Geographical Analysis*, *54*(2), 333–356. https://doi.org/10.1111/gean.12284

Poelitz C., & Andrienko N. (2010). *Finding Arbitrary Shaped Cluster with Related Extents in Space and Time*. IEEE VGTC Simposium on Visualization.

Prasad, A. S., & Francescutti, L. H. (2017). Natural Disasters. In *International Encyclopedia of Public Health* (pp. 215–222). Elsevier. https://doi.org/10.1016/B978-0-12-803678-5.00519-1

Putri, N. A., Utami, T. W., & Wasono, R. (2023). Spatial Temporal Density Based Spatial Clustering Applications With Noise (ST-DBSCAN) For Classification of Forest and Land Fire Points in Riau Province in 2021. *Prosiding Seminar Nasional UNIMUS*, 436–445.

Ridwan. (2023). *1.290 Kali Gempa Bumi Terjadi di Jabar Sepanjang Tahun 2022*. Https://Www.Detik.Com/Jabar/Berita/d-6493847/1-290-Gempa-Bumi-Terjadi-Sepanjang-2022-Di-Jawa-Barat.

Sealey, K. S., & Logan, A. (2019). The Commonwealth of the Bahamas. In *World Seas: an Environmental Evaluation* (pp. 591–615). Elsevier. https://doi.org/10.1016/B978-0-12-805068-2.00030-9

Sharma, A., Vijay, R. K., & Nanda, S. J. (2023). Identification and spatio-temporal analysis of earthquake clusters using SOM–DBSCAN model. *Neural Computing and Applications*, *35*(11), 8081–8108. https://doi.org/10.1007/s00521-022-08085-5

Soltisz, A. M., Craigmile, P. F., & Veeraraghavan, R. (2024). Spatial Pattern Analysis using Closest Events (SPACE)—A Nearest Neighbor Point Pattern Analysis Framework for Assessing Spatial Relationships from Digital Images. *Microscopy and Microanalysis*, *30*(2), 306–317. https://doi.org/10.1093/mam/ozae022

Sonhaji, N. R. (2023). *Clustering Titik Gempa di Pulau Jawa Menggunakan Algoritma St-Dbscan*. Universitas Islam Negeri Maulana Malik Ibrahim.

Subasi, A. (2020). Clustering examples. In *Practical Machine Learning for Data Analysis Using Python* (pp. 465–511). Elsevier. https://doi.org/10.1016/B978-0-12-821379-7.00007-2

Supendi, P., Winder, T., Rawlinson, N., Bacon, C. A., Palgunadi, K. H., Simanjuntak, A., Kurniawan, A., Widiyantoro, S., Nugraha, A. D., Shiddiqi, H. A., Ardianto, Daryono, Adi, S. P., Karnawati, D., Priyobudi, Marliyani, G. I., Imran, I., & Jatnika, J. (2023). A conjugate fault revealed by the destructive Mw 5.6 (November 21, 2022) Cianjur earthquake, West Java, Indonesia. *Journal of Asian Earth Sciences*, *257*, 105830. https://doi.org/10.1016/j.jseaes.2023.105830

Yang, Q., Liu, Y., Duan, Z., & Liu, X. (2025). An Accessibility Analysis of Emergency Shelters in Shenzhen Using the Gaussian-Based Two-Step Floating Catchment Area Method and Clustering. *Sustainability (Switzerland)*, *17*(12). https://doi.org/10.3390/su17125250

Zaccagnino, D., & Doglioni, C. (2022). Earth's gradients as the engine of plate tectonics and earthquakes. *La Rivista Del Nuovo Cimento*, *45*(12), 801–881. https://doi.org/10.1007/s40766-022-00038-x

Zhang, Z., Hu, Y., Lu, W., Cao, W., & Gao, X. (2023). Spatial accessibility analysis and location optimization of emergency shelters in Deyang. *Geomatics, Natural Hazards and Risk*, *14*(1). https://doi.org/10.1080/19475705.2023.2213809

Zheng, S., & Zhao, J. (2018). *States Identification of Complex Chemical Process Based on Unsupervised Learning* (pp. 2239–2244). https://doi.org/10.1016/B978-0-444-64241-7.50368-2