ASEAN Engineering Journal

Full Paper

# EARTHQUAKE CLUSTER MAPPING IN ACEH PROVINCE: AN ORDERING POINTS TO IDENTIFY THE CLUSTERING STRUCTURE (OPTICS) CLUSTERING APPROACH

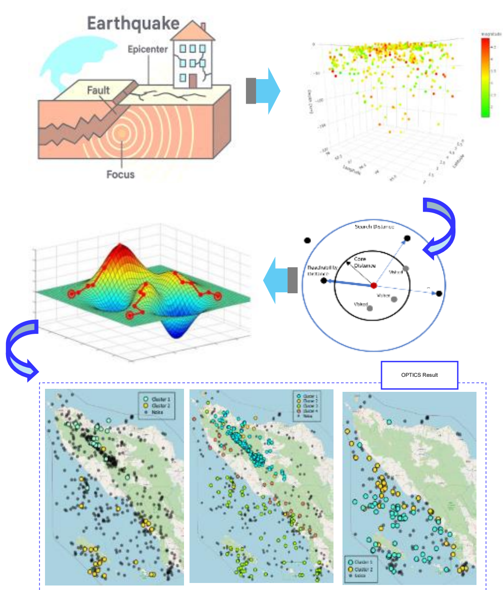
Lasta Syakila Nisa, Achmad Fauzan*

Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia, Indonesia.

**Article history**
Received
09 July 2025
Received in revised form
25 January 2026
Accepted
13 March 2026
Published online
31 August 2026

*Corresponding author
achmadfauzan@uii.ac.id

**Graphical abstract**




**Abstract**

Indonesia is known as a disaster-prone region. Located on the Pacific Ring of Fire, it is one of the countries most susceptible to various types of natural disasters. The province of Aceh is one of the regions in Indonesia that frequently experiences earthquakes due to its geographical location in a highly seismically active area. Therefore, an analysis of earthquake data clustering in the province of Aceh is necessary. The aim of this research is to determine earthquake clusters, which can be used for disaster mitigation and prevention measures. This research begins with examining the distribution pattern of data through a scatterplot. If the data distribution pattern shows varying density levels, then a clustering analysis using Ordering Points to Identify the Clustering Structure (OPTICS) is conducted. In the OPTICS algorithm, two parameters are required before clustering: minimum points (*MinPts*) and $x_i$. The clustering results are evaluated using the silhouette coefficient (SC). Further data exploration is carried out by: (1) reducing the *MinPts* value, (2) clustering based on the supremum of the SC, and (3) identifying earthquake events with potential tsunami risks. The purpose of this data exploration is to increase the number of clusters formed while still considering the SC value limit, thus expanding the regions identified as earthquake-prone. The best clustering results are determined by comparing the obtained clusters. The best clustering result was found at *MinPts* = 20 and $x_i$ = 0.05, forming 4 clusters with a silhouette coefficient value of 0.72, indicating a strong cluster structure. This information is expected to benefit the public by raising awareness of earthquake-prone areas.

*Keywords*: Disaster Mitigation, Earthquake, Spatial Clustering, OPTICS.



## 1.0 INTRODUCTION

Indonesia is known as a disaster-prone region because it is located on the Pacific Ring of Fire. In addition, geographical, climate, and topographic factors also cause Indonesia to often experience various natural disasters, including earthquakes [1] An earthquake is a natural phenomenon that occurs when there is a sudden release or shift of energy from within the earth's outer layer. This shift generally occurs along cracks called faults [2]. When two parts of the earth's crust move or approach each other, energy is released in the form of seismic waves. These waves travel through the earth and cause vibrations or shocks on the earth's surface. The speed and magnitude of these shocks can vary, and their impact also depends on the depth of the focus and the distance from the epicenter. Earthquakes can cause significant damage to buildings, infrastructure, and the environment. In addition, if an earthquake occurs under the sea, it can also trigger a tsunami [3].

Aceh, located at the western tip of Sumatra Island, is one of the regions in Indonesia that often witnesses devastating earthquakes, as its geographical location is in a very seismically active area. This is largely due to its position at the meeting point of three major tectonic plates [4]. One of the most devastating earthquakes occurred in 2004, which also triggered a massive tsunami that hit Aceh and much of the surrounding area [5]. According to the National Disaster Management Agency [6], the earthquake hazard class in this province varies from low, moderate, to high. In several regions such as South Aceh Regency, Southeast Aceh, Central Aceh, West Aceh, Greater Aceh, Pidie, Simeulue, Aceh Singkil, Southwest Aceh, Gayo Lues, Aceh Jaya, Nagan Raya, Banda Aceh City, Sabang City, and Subulussalam City, the earthquake hazard class is classified as high. Whereas, in Bener

Meriah and Pidie Jaya Regencies, the earthquake risk is at a moderate level. In East Aceh Regency, North Aceh, Bireuen, Aceh Tamiang, Lhokseumawe City, and Langsa City, the earthquake hazard class tends to be lower. this data indicates that most regions in Aceh Province have a high potential risk of earthquakes. Therefore, an analysis of the grouping of earthquake data in Aceh Province is needed as part of disaster prevention measures to reduce the potential losses that may occur due to earthquakes.

Clustering is a data analysis technique used to group objects based on their similarities, with the main goal of identifying patterns in the data [8,9]. When dealing with objects that have irregular distribution patterns, density-based clustering is an effective approach. This method groups data by finding areas with high density, making it better at detecting non-linear patterns and handling outliers compared to distance- or partition-based methods. As a result, density-based clustering works well for datasets with complex shapes and uneven distributions [10,11]. One well-known method in this category is the Ordering Points To Identify the Clustering Structure (OPTICS) algorithm [11]. The basic principle is similar to the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm which is used to identify clusters and noise in specific data. The OPTICS algorithm describes clusters based on data density [12]. In this algorithm, a point "p" is considered a cluster if it has at least a certain number of points (*MinPts*) that are not located further than a predetermined distance. The OPTICS algorithm analyzes points that belong to clusters that have higher density, so that each point is assigned a core distance value that reflects its distance to the point to its *MinPts*. If a point is not included in a cluster and is also not part of another cluster, then the point is classified as noise [13]. In addition, the OPTICS method produces a reachability plot that helps identify cluster patterns by assessing the relationship between points based on distance and number of neighbors. The reachability plot describes the characteristics of the cluster structure by showing how points in the data are connected to each other based on density [14]

Several studies have explored the use of the OPTICS clustering algorithm. Zhu [15] applied OPTICS to remove photon noise in LiDAR data from the ICESat-2 satellite. Ahmed (2020) compared K-means, DBSCAN, and OPTICS for clustering translated verses of Surah Al-Baqarah, finding that K-means was the fastest but had the lowest clustering quality, DBSCAN had the highest quality but introduced noise, while OPTICS provided a balance between the two with longer execution time [13]. Febriana (2017) used OPTICS to detect outliers in forest fire hotspot data, proving effective for non-uniform cluster distributions [16]. Fan (2023) applied OPTICS to improve clustering in finite-state automata (FSA) construction [17]. Rodriguez (2019) compared multiple clustering algorithms on synthetic data, showing that no single method is always the best. However, OPTICS was more flexible than DBSCAN, especially for clusters with varying densities [18].

Despite its potential, the OPTICS clustering has rarely been applied in disaster analysis, particularly in earthquake studies. Seismic phenomena often exhibit non-uniform distribution patterns, where certain regions experience a higher density of earthquake events than others [20,21]. Conventional clustering methods such as K-means are often less effective in grouping data with varying densities [13]. In contrast, OPTICS, with its ability to detect clusters of different densities, provides a more adaptive approach for earthquake data analysis. The limited adoption of OPTICS in seismic studies highlights a gap in spatial modeling research for earthquake hazard assessment. Given the increasing frequency and impact of earthquakes in various regions, methods capable of revealing hidden spatial patterns are crucial. The application of OPTICS in mapping earthquake-prone areas can offer more accurate insights for researchers and policymakers, supporting more effective disaster mitigation strategies. Based on these considerations, this study aims to analyze the clustering of earthquake epicenters using the OPTICS algorithm.

## 2.0 METHODOLOGY

### 2.1 Research Population and Sample

The population in this study includes all earthquake events that occurred in Aceh Province. The sample consists of earthquake epicenter location data from Aceh Province for the period 2022–2023, obtained from the BMKG earthquake repository website (repogempa.bmkg.go.id). This study employs purposive sampling, as the selection of earthquake data is based on specific criteria, namely, recorded seismic events within a defined geographical region (Aceh Province) and time frame (2022–2023). This approach ensures that the collected data is relevant for analyzing earthquake distribution patterns in the study area. This study utilizes earthquake epicenter coordinates, represented by longitude and latitude. Additionally, earthquake magnitude and depth are analyzed through further data filtering.

### 2.2 OPTICS Algorithm

OPTICS is a density-based clustering algorithm. It considers a cluster as a region in the data space where many objects (or data points) are close to each other. Conversely, regions where objects are scarce (i.e., low-density regions) are considered as noise or boundaries between clusters. It also produces what is called an 'ordering' of the data, which provides a multi-resolution representation of the cluster structure in the data. From this 'ordering', different clusters with different levels of density can be identified [21]. OPTICS computes an extended cluster ordering of objects in a database. The main advantage of this approach, compared to clustering algorithms proposed in the literature, is that it does not limit itself to a single global parameter setting. Instead, the extended cluster ordering contains information equivalent to density-based clustering corresponding to a wide range of parameter settings, thus providing a versatile basis for automated and interactive clustering analysis [14]. OPTICS introduces the concept of 'reachability distance' and 'core distance', two measures that allow to measure the local density of each point in the data [22]. The parameters used are minimum points (*MinPts*) and $x_i$ as elements of cluster formation.

OPTICS will consider the distance between adjacent points to create a reachability distance which is then used to separate clusters with varying densities from noise [23] . *MinPts* is the minimum number of points required to form a cluster. $x_i$ is a parameter that determines the extent to which two points differ from each other while being considered part of the same cluster. The unit of the $x_i$ parameter is a proportion or percentage, usually expressed as a decimal between 0 and 1, such as 0.05 which means 5%. Thus, a smaller value of $x_i$ makes the algorithm more sensitive to small variations in data density, so that small differences can cause two points to be grouped in different clusters. While a larger value of $x_i$ makes the algorithm more tolerant of such differences, so that two points with larger density differences (up to 10%) can still be considered in the same cluster. The illustration of core distance and reachability distance is presented in Figure 1. OPTICS uses the concept of maximum reachability distance, which is the distance from a point to its nearest neighbor that has not been visited by the search (points on boundary 2). OPTICS will search for all neighbor distances within the specified epsilon range, then compare each of these points with the core distance. If there is a distance smaller than the core distance (all points that enter boundary 1), then the point sets the core distance as its reachability distance. If all distances are greater than the core distance (all points on boundary 2), then the smallest distance is set as its reachability distance. These distances are then used to create a reachable plot, which is an ordered plot of the reachability distance used to detect clusters [22].

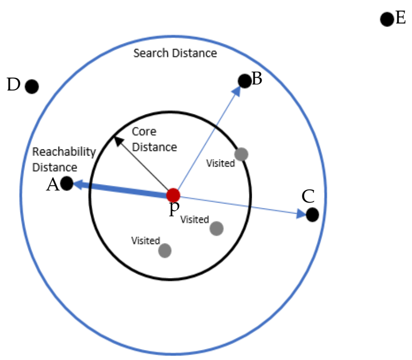



**Figure 1** Illustration of Core Distance and Reachability Distance (source: ESRI [24])

For all points in the dataset, the reachability distance is calculated because this distance will be used and mapped on the reachability plot. The valleys in the reachability plot represent different clusters for each pattern of points. The denser a cluster, the smaller its reachability distance, and the lower the valleys on the plot. Less dense clusters have larger reachability distances and higher valleys on the plot. The peaks represent the distances required from one cluster to another, or from a cluster to a noise point to another cluster, depending on the configuration of the set of points in a data set [24]. Points with large reachability distances that do not belong to any cluster are marked as noise. Finally, all points are sorted based on their reachability distances to construct a reachability plot, which visually represents the clustering structure and aids in extracting meaningful patterns from the data.

To evaluate the clustering results, cluster validation can be performed. Cluster validation is useful for measuring the level of goodness of the clusters formed, one of the internal validation indexes is the silhouette coefficient (SC) [25]. The results of the SC calculation ranges between -1 and 1, if the SC value approaches the number 1, then the stronger the clusters grouping are. Conversely, if it approaches the number -1, then the grouping of the clusters are weaker [26]. The equation of SC is presented with Equation (1) [27].

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \tag{1}$$

The SC, denoted as *s(i)* quantifies the clustering quality by evaluating the cohesion and separation of a given data point within its assigned cluster. Specifically, *a(i)* represents the mean intra-cluster distance, which measures the average dissimilarity between sample *i* and all other samples within the same cluster. Conversely, *b(i)* denotes the minimum inter-cluster distance, defined as the lowest average dissimilarity between sample *i* and any other cluster to which it does not belong. This metric effectively assesses the compactness of clusters and the degree of separation between them, thereby providing a robust evaluation of clustering performance.

Specifically, a SC within the range of 0.71 to 1.00 signifies strong clustering results, indicating well-separated and compact clusters. A value between 0.51 and 0.70 reflects reasonable clustering performance, where clusters are distinguishable but may exhibit some overlap. When the SC falls between 0.26 and 0.50, the clustering results are considered weak, suggesting that some data points may be ambiguously assigned to clusters. Finally, a SC of 0.25 or lower indicates that no meaningful clusters have been formed, as the separation between clusters is poor, and data points may be arbitrarily assigned. This metric provides valuable insights into the effectiveness of a clustering algorithm and helps in assessing the overall structure of the clustered data [28].

The OPTICS algorithm was selected due to its ability to identify clusters with varying density levels and to effectively separate noise, which are common characteristics of earthquake event data [14]. The spatial distribution of earthquakes in the Aceh region is not homogeneous; instead, it forms spatial patterns influenced by subduction zones and active fault systems. Consequently, density-based clustering methods relying on a single global threshold, such as DBSCAN, are considered less adequate for capturing these complex structures [29].

The *MinPts* parameter was defined to represent the minimum number of earthquake events required for an area to be classified as a significant seismic activity zone, ensuring that the resulting clusters reflect persistent seismic activity rather than isolated events. Meanwhile, the $x_i$ parameter was employed to control the algorithm's sensitivity to density changes during the cluster extraction process, allowing for a more effective separation between major seismic zones and noise [30]. This approach is deemed suitable for seismic clustering, as it is capable of capturing complex variations in earthquake event density driven by tectonic structures such as subduction zones and active faults [31].

### 2.3 Research Flow Chart

This study starts by downloading earthquake location data for Aceh Province from 2022 to 2023. The next steps involve geoprocessing to define the relevant study area and improve the accuracy and consistency of the spatial data. The research flowchart is presented in Figure 2. After the geoprocessing stage, a descriptive analysis is performed to provide a comprehensive overview of the earthquake data. This analysis includes the computation of key statistical measures, such as the minimum, maximum, mean, median, and standard deviation. Data visualization techniques, including plots and boxplots, are utilized to illustrate the distribution and characteristics of the dataset, facilitating the identification of patterns, trends, and potential noise. Subsequently, a spatial pattern detection analysis is conducted. If the data exhibit density variations without a well-defined geometric structure, the OPTICS algorithm is employed. Conversely, if the data demonstrate a more structured pattern, alternative clustering methods are considered to ensure optimal result.

The clustering analysis using the OPTICS algorithm begins with the selection of appropriate parameters, including *MinPts* and the value of $x_i$. Subsequently, distance and neighborhood relationships are computed by determining the core distance and reachability distance for each point. The points are then sorted in ascending order based on their reachability distance, and clusters are expanded by visiting neighboring points. In the OPTICS algorithm, distance measurements are performed using the Euclidean distance method. Following this, a reachability plot is constructed to visualize neighborhood relationships in a graphical format, facilitating the identification of cluster structures. All points are ordered according to their reachability distance, forming the reachability graph, which serves as the basis for cluster identification. The clustering process is then conducted to classify and group data points with similar characteristics, ensuring a robust and meaningful segmentation of the dataset.

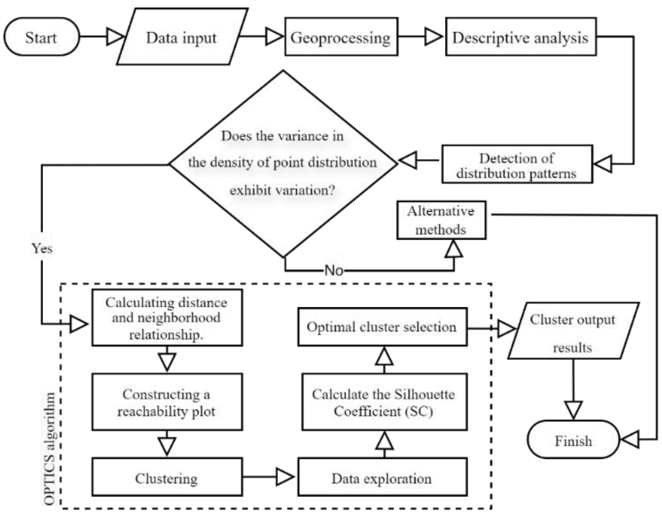



**Figure 2** Research flowchart (source: research report)

After the cluster formation, further data exploration is conducted through various scenarios: (1) reducing the *MinPts* value, (2) clustering based on the smallest upper bound (supremum) of the SC, and (3) identifying earthquake events with potential tsunami risks. The clustering results are evaluated by calculating the SC for each scenario. Subsequently, the optimal clustering outcome is determined by comparing the clusters derived from the optimal SC with those obtained from the data exploration process. Ultimately, the best clustering result is identified as the primary output of this study, marking the final stage of the research.

## 3.0 RESULTS AND DISCUSSION

This chapter discusses analysis steps and results obtained from the clustering of earthquake locations in Aceh Province from the year 2022 to 2023. The initial step to gain insight is a descriptive analysis, the explanation is as follows.

### 3.1 Descriptive Analysis

Descriptive analysis is a statistical method used to summarize and interpret the characteristics of a dataset, providing an overview of its distribution, central tendencies, and variability [32]. In this study, descriptive analysis is applied through a 3D plot representation to assess the distribution of earthquake location data in Aceh Province from the year 2022 to 2023.

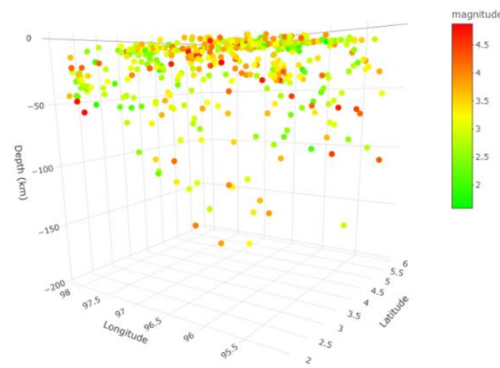



**Figure 3** The 3D Visualization of Earthquake Data (source: research report)

Figure 3 is a visualization of earthquakes in Aceh Province in 2022 to 2023 in 3D to provide a clearer perspective on the distribution and characteristics of earthquakes. The x-axis shows the longitude coordinates of each epicenter, the y-axis shows the latitude coordinates, and the z-axis shows the depth in kilometers (km) with a depth range from 0 km to -200 km. This depth is measured from sea level down, which means that earthquakes with depths approaching 0 km occur closer to the earth's surface, while earthquakes with greater depths occur deeper below the sea surface. The closer to the earth's surface, the stronger the vibrations will be and cause damage to buildings and infrastructure. In addition, shallow earthquakes that occur under the sea can cause massive movements on the seabed that can generate tsunamis. The dots on the plot indicate the location of the earthquake, with colors representing the magnitude of the earthquake. The color scale ranges from green, which indicates a low magnitude earthquake, to red, which indicates a high magnitude earthquake. Based on the image above, it can be seen that the majority of points are near the surface, with a depth of less than 50 km, which means that many earthquakes occur at relatively shallow depths.

**Table 1** Characteristics of the Earthquakes Based on the Research Variables.

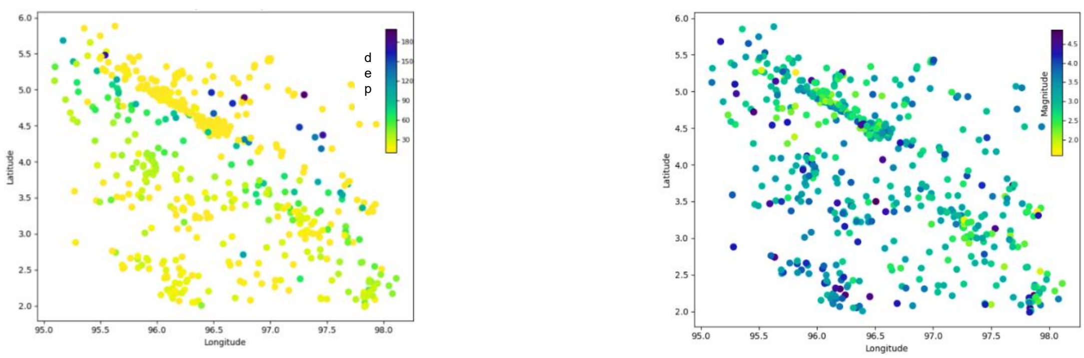
| Variable | Descriptive Statistics | | | | |
|---|---|---|---|---|---|
| | Mean | Median | Standard Deviation | Minimum | Maximum |
| Depth (Km) | 25.44 | 12 | 28.24 | 1 | 197 |
| Magnitude (RS) | 3.16 | 3.1 | 0.61 | 1.59 | 4.87 |

Table 1 displays the summary statistics of the depth and magnitude variables. The depth variable has values ranging from 1 to 197 km, with an average value of 25.44 km. This shows that overall, earthquakes in Aceh Province from 2022 to 2023 tend to be at a depth of 25.44 km (shallow earthquakes). The median or middle value of the depth variable is obtained at 12 km, which means that half of the earthquakes occur at depths of less than 12 km, indicating that most earthquakes occur at relatively shallow depths.

Meanwhile, the magnitude variable is in the range of 1.59 to 4.87 on the Richter Scale (RS), with an average value of 3.16 RS. This shows that overall, earthquakes in Aceh Province from 2022 to 2023 tend to be at a magnitude or earthquake strength of 3.16 RS (small-scale earthquakes). The median or middle value of the magnitude variable is obtained at 3.10 RS, which means that half of the earthquakes have a magnitude of less than 3.10 RS. The analysis shows that the data on the depth variable has a higher variation, indicated by the standard deviation value exceeding the average value. This indicates that there are a number of earthquakes with very varying depths, from shallow to deep. Meanwhile, the data on the magnitude variable shows a lower variation, with a standard deviation value smaller than the average value. This indicates that the strength of earthquakes in Aceh Province tends to be more consistent and does not vary as much as its depth.

### 3.2 Analyzing Data Distribution Patterns

Before proceeding with the clustering process using the OPTICS method, first look at the data distribution pattern through a scatterplot. Where the scatterplot can assess how data points are spread visually. The purpose of looking at the data distribution pattern is to determine whether the data tends to form different clusters, detect data points that may be outliers, and understand how the density of data points is distributed. The results of the scatterplot formed can be seen in Figure 4.




(a) (b)

**Figure 4** Scatterplot result, (a) based on depth variable, (b) based on magnitude variable (source: research report)

Figure 4(a) presents a scatterplot depicting the distribution of earthquake epicenters in Aceh Province from 2022 to 2023. The x-axis represents longitude, while the y-axis represents latitude. Data points are color-coded by depth, with brighter colors indicating shallower earthquakes and darker colors representing deeper ones. The depth scale ranges from 30 to 180 km. The scatterplot reveals variations in earthquake density, with some regions showing a higher concentration of events, while others have fewer or none. Additionally, most recorded earthquakes occurred at depths below 30 km. Figure 4(b) illustrates the same earthquake distribution, but with data points color-coded by magnitude instead of depth. Brighter colors represent lower magnitudes, while darker colors indicate higher magnitudes, ranging from 2.0 to 4.5 RS. The visualization suggests that most earthquakes in Aceh during 2022–2023 had magnitudes between 3.0 and 3.5 RS.

### 3.3 Clustering Using the OPTICS

The clustering process begins with model initialization, where parameters are determined through iterative simulations and error evaluations to achieve an optimal model. The results of multiple experimental repetitions are presented in Table 2. Based on Table 2, the results indicate that the differences are due to variations in the *MinPts* and $x_i$ values used. Through several trial-and-error iterations, the best SC of 0.9631 was achieved using a *MinPts* value of 20 and an $x_i$ value of 0.1. This suggests that cluster formation requires a minimum of 20 points.

**Table 2** Results of Optimum Silhouette Coefficient Experiments.

| ID. | *MinPts* | $x_i$ | SC | Number of Clusters | Number of Noise Points | Number of Points in the Cluster |
|---|---|---|---|---|---|---|
| 1 | 5 | 0.01 | 0.54 | 32 | 332 | 252 |
| 2 | 5 | 0.02 | 0.55 | 31 | 340 | 244 |
| 3 | 5 | 0.05 | 0.58 | 28 | 361 | 223 |
| 4 | 5 | 0.1 | 0.62 | 22 | 411 | 173 |
| 5 | 10 | 0.01 | 0.49 | 13 | 339 | 245 |
| 6 | 10 | 0.02 | 0.51 | 12 | 336 | 248 |
| 7 | 10 | 0.05 | 0.58 | 10 | 396 | 188 |
| 8 | 10 | 0.1 | 0.88 | 3 | 541 | 43 |
| 9 | 15 | 0.01 | 0.55 | 9 | 367 | 217 |
| 10 | 15 | 0.02 | 0.59 | 9 | 392 | 192 |
| 11 | 15 | 0.05 | 0.59 | 6 | 291 | 293 |
| 12 | 15 | 0.1 | 0.78 | 2 | 540 | 44 |

| ID. | *MinPts* | $x_i$ | SC | Number of Clusters | Number of Noise Points | Number of Points in the Cluster |
|---|---|---|---|---|---|---|
| 13 | 20 | 0.01 | 0.60 | 7 | 357 | 227 |
| 14 | 20 | 0.02 | 0.58 | 6 | 329 | 255 |
| 15 | 20 | 0.05 | 0.72 | 4 | 308 | 276 |
| 16 | 20 | 0.1 | 0.96 | 2 | 277 | 307 |

After initializing the model and obtaining optimal results, the next step is to train the model and perform data analysis to visualize the results. The purpose of this visualization is to provide a clear and easy-to-understand picture of the data structure formed after the clustering process. The results of the visualization after the clustering process are shown in Figure 5.

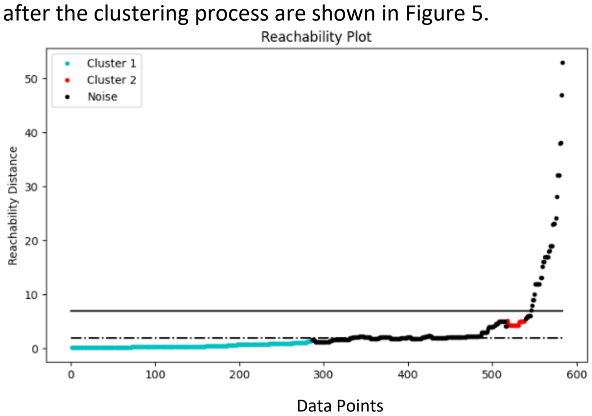
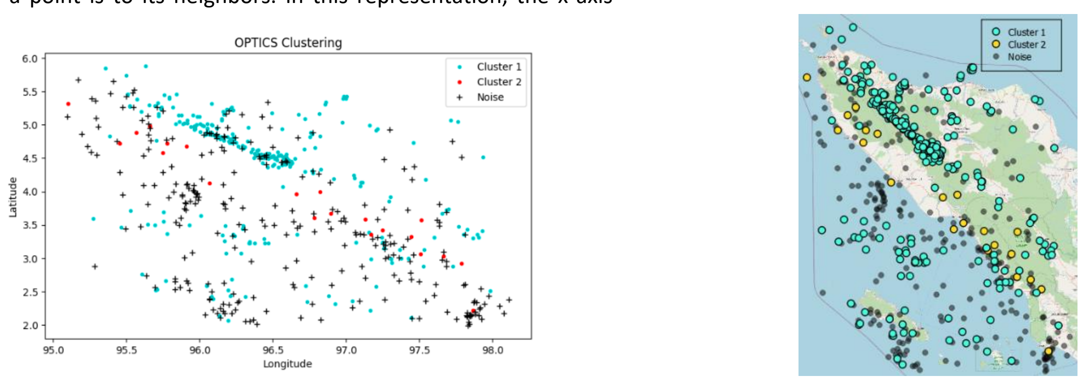



**Figure 5.** Plot of Earthquake Reachability Distances (source: research report)

Figure 5 shows the reachability distance plot using the OPTICS clustering method. Reachability distance measures how close or far a point is to its neighbors. In this representation, the x-axis shows the sequence of data generated by the OPTICS algorithm, while the y-axis shows the reachability distance value. A horizontal line is added as a reference line, with the number 7 as the upper limit and the number 2 as the lower limit. The dashed line indicates the reference line at number 2 on the y-axis, while the solid line indicates the reference line at number 7 on the y-axis. The results show variations in the data points formed, identified by light blue, red, and black colors. These variations are caused by similar or interconnected reachability distance values between data points in a cluster. Data points that are cut off by a dotted line in one color show lower values than the dotted line, indicating the proximity of the data points to other points in the same cluster. Conversely, if the data is cut off by a solid line, it indicates that the points are further away from the points within the same cluster.

Figure 6(a) and 6(b) shows the scatterplot generated using the OPTICS Clustering method, which shows the distribution of earthquake events in Aceh Province from the year 2022 to 2023. By using the optimal SC (*MinPts* parameter of 20 and the *xi* value of 0.1), two different clusters are formed, marked with light blue and red colors. The light blue cluster represents Cluster 1, while the red cluster represents Cluster 2. In addition, there are 277 points that are considered noise, marked with a black + symbol. Furthermore, the scatterplot shows a SC of 0.96309, which indicates that the cluster has a strong or good structure.




(a) (b)

**Figure 6**. (a) Visualization of OPTICS clustering, (b) Visualization of OPTICS clustering results with basemap (source: research report).

Cluster 1 is widespread almost throughout the land and waters (Andaman Sea) of Aceh Province, especially in Pidie Regency, West Aceh Regency, and Nagan Raya Regency. However, the distribution of Cluster 2 only includes certain regions of the Aceh Province such as Aceh Jaya Regency, Southwest Aceh Regency, Gayo Lues Regency, South Aceh Regency, and Southeast Aceh Regency.

### 3.4 Data Exploration and OPTICS Parameter Adjustment

In addition to clustering based on the highest SC from multiple repetitions (default), this study also explores the parameter settings of the OPTICS algorithm itself. Data exploration has several purposes, such as understanding data structures, detecting unusual patterns, identifying trends, preparing data, and selecting appropriate analysis methods. Researchers conduct data exploration in hopes that it can help reduce the impact of disasters. The exploration is conducted by conducting through various scenarios: (1) reducing the *MinPts* value (adjusting the *MinPts* parameter to modify cluster formation), (2) clustering based on the smallest upper bound/ supremum of the SC value (to refine clustering accuracy), and (3) identifying tsunami-prone earthquake events (detecting earthquakes with potential tsunami risks based on clustering outcomes). The first exploration was conducted by lowering the *MinPts* value to 15 based on the optimum SC (initially, *MinPts* had been 20). The purpose of lowering *MinPts* is to increase the number of clusters formed, thus allowing for more detailed identification of data clusters. With more clusters, the distribution of earthquake data can be analyzed even further. The visualization of the first exploration is presented in Figure 7(a).

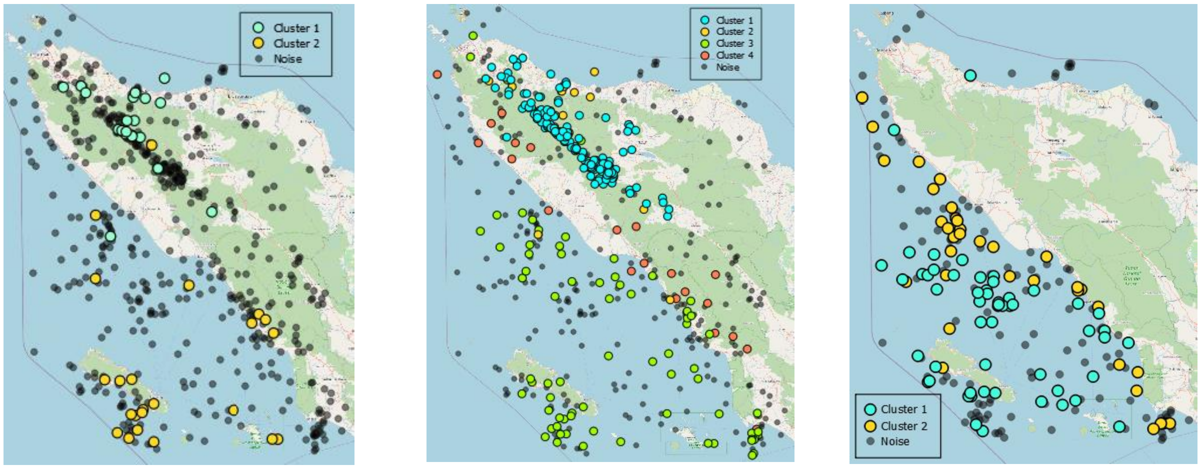



**Figure 7** Visualization of Clusters, (a) Reduce *MinPts* and optimal SC value, (b) the supremum of SC value, (c) tsunami-prone events (source: research report)

Based on the first exploration, two clusters were identified. In one of the clusters the districts are mostly located on land, meanwhile, the other cluster has regions that are mostly located in water (Andaman Sea). This is due to the existence of active faults on the island of Sumatra, such as the Great Sumatran Fault, and due to the island's proximity to the boundary between the Indo-Australian and Eurasian plates which is marked by the red line on the Indian Ocean.

The second exploration is performed by observing the number of clusters that has a silhouette value approaching or more than 0.71 as the number of clusters are increased, and 0.71 is the lowest threshold for a cluster to be considered strong. Based on Figure 7(b), it can be observed that the SC value at the supremum of 0.71 (with *MinPts* = 20, $x_i$ = 0.05, and a SC of 0.7215) formed 4 clusters. Three earthquake clusters are localized primarily on the land of Aceh Province, particularly in Aceh Besar Regency, Pidie Regency, Pidie Jaya Regency, West Aceh Regency, Nagan Raya Regency, Aceh Jaya Regency, Southwest Aceh Regency, and South Aceh Regency. Meanwhile, one other cluster is localized predominantly in the waters, especially around the Andaman Sea.

The third exploration is based on the optimum SC value. In earthquake research, analysis of the tsunami potential of earthquakes is important. Therefore, in this case, the main focus is to explore earthquake points that have the potential to trigger a tsunami in Aceh Province using the OPTICS method. Earthquakes with tsunami potential generally occur at depths of less than 70 km below sea level [33]. By expanding the scope of research to investigate earthquakes that occur at depths of less than 70 km, it is hoped to provide a better understanding of the potential for tsunami hazards in Aceh Province. The visualization of grouping results using the OPTICS method, with parameters *MinPts* = 25 and $x_i$ = 0.05, on the location data of the distribution of earthquakes with tsunami potential in Aceh Province from 2022 to 2023 is shown in Figure 7(c). The clustering results divide the tsunami-prone earthquake data in Aceh Province from 2022 to 2023 into two different clusters. Cluster 1 is marked in light blue while Cluster 2 is marked in yellow. The black dots indicate noise. The size of the dots on the noise is smaller, emphasizing the dots that belong in the clusters. This demonstrates that the OPTICS method can be used to map areas vulnerable to earthquakes with tsunami potential. Based on the analysis results in Figure 7(c), the cluster consists of 89 points in it, while 137 other points are identified as noise. The number of points considered as noise is greater than those in the clusters. The clusters are spread around the Andaman Sea. A summary of the results from all scenarios, including the default scenario, is presented in Table 3.

**Table 3** Summary of the results from all scenarios

| ID | Explorations | *MinPts* | $x_i$ | SC | Number of Clusters | Number of Points in Cluster | Number of Noise Points | Region |
|---|---|---|---|---|---|---|---|---|
| 1 | Optimal SC value (*default*) | 20 | 0.1 | 0.9631 | 2 | 307 | 277 | Sabang City, Aceh Besar Regency, Aceh Jaya, Pidie, Pidie Jaya, West Aceh, Bireuen, Central Aceh, Nagan Raya, Bener Meriah, North Aceh, East Aceh, Langsa City, Gayo Lues, South Aceh, Southeast Aceh, Aceh Singkil, Simeulue, Andaman Sea, Southwest Aceh |
| 2 | Reduce *MinPts* and optimal SC value | 15 | 0.1 | 0.7806 | 2 | 44 | 540 | Andaman Sea, Aceh Besar, Pidie, Pidie Jaya, West Aceh, Gayo Lues, Bireuen, Central Aceh, South Aceh, Simeulue |
| 3 | Supremum of SC value | 20 | 0.05 | 0.7215 | 4 | 276 | 308 | Andaman Sea, Aceh Besar Regency, Pidie, Pidie Jaya, West Aceh, Nagan Raya, Gayo Lues, Central Aceh, Bireuen, Bener Meriah, |

| ID | Explorations | *MinPts* | $x_i$ | SC | Number of Clusters | Number of Points in Cluster | Number of Noise Points | Region |
|---|---|---|---|---|---|---|---|---|
| 4 | Tsunami-prone events | 25 | 0.05 | 0.9191 | 2 | 89 | 137 | Aceh Jaya, South Aceh, Aceh Singkil, Simeulue, Southwest Aceh, Southeast Aceh Andaman Sea |

The characteristics of each explorations are presented in Table 4.

**Table 4** Characteristics of each scenario

| Explorations | Cluster | Variable | Statistics | | | | |
|---|---|---|---|---|---|---|---|
| | | | Average | Median | Standard Deviation | Minimum | Maximum |
| Optimal SC value (default) | 1 | Depth (km) | 10.07 | 10 | 0.44 | 8 | 12 |
| | | Magnitude (RS) | 3.11 | 3.06 | 0.58 | 1.59 | 4.59 |
| | 2 | Depth (km) | 57.86 | 57 | 3.09 | 53 | 64 |
| | | Magnitude (RS) | 3.21 | 3.14 | 0.73 | 2.15 | 4.85 |
| Reduce *MinPts* and optimal SC value | 1 | Depth (km) | 13.48 | 13 | 0.59 | 13 | 15 |
| | | Magnitude (RS) | 2.68 | 2.62 | 0.49 | 1.91 | 3.89 |
| | 2 | Depth (km) | 20.30 | 20 | 0.91 | 19 | 22 |
| | | Magnitude (RS) | 3.36 | 3.39 | 0.64 | 2.08 | 4.75 |
| Supremum of SC value | 1 | Depth (km) | 10 | 10 | 0 | 10 | 10 |
| | | Magnitude (RS) | 2.99 | 2.97 | 0.46 | 1.85 | 4.1 |
| | 2 | Depth (km) | 13.38 | 13 | 0.49 | 13 | 14 |
| | | Magnitude (RS) | 2.7 | 2.76 | 0.48 | 1.91 | 3.89 |
| | 3 | Depth (km) | 24.23 | 24 | 3.02 | 19 | 29 |
| | | Magnitude (RS) | 3.35 | 3.34 | 0.61 | 1.79 | 4.87 |
| | 4 | Depth (km) | 57.86 | 57 | 3.09 | 53 | 64 |
| | | Magnitude (RS) | 3.21 | 3.14 | 0.73 | 2.15 | 4.85 |
| Tsunami-prone events | 1 | Depth (km) | 10.2 | 10 | 0.59 | 9 | 12 |
| | | Magnitude (RS) | 3.31 | 3.25 | 0.52 | 2.17 | 4.65 |
| | 2 | Depth (km) | 34.14 | 34 | 1.57 | 31 | 37 |
| | | Magnitude (RS) | 3.19 | 3.18 | 0.58 | 1.99 | 4.3 |

Based on Table 4, in the default scenario, Cluster 1 has the lowest average earthquake depth of 10.07 km, while Cluster 2 has the highest average earthquake magnitude of 3.21 RS. Additionally, the standard deviation for each cluster is lower than the respective average, indicating minimal variation in earthquake depth and magnitude within each cluster. In the first exploration, the data on Table 4 indicates that the shallowest earthquake depth is in Cluster 1 (mostly located in the land area, namely Aceh Besar Regency, Pidie Regency, Pidie Jaya Regency, Bireuen Regency, West Aceh Regency, and Gayo Lues Regency), while the highest earthquake magnitude occurs in Cluster 2 (in the water area, especially around the Andaman Sea). In addition, the standard deviation in each cluster is lower than the average of each, indicating that the earthquake depth and magnitude data in each cluster have small variation.

Subsequently, in the second exploration, the shallowest earthquake depth is observed in Cluster 1 whereas the highest earthquake magnitude occurs in Cluster 3. Additionally, the standard deviation in each cluster is lower than the mean, indicating that the variation in earthquake depth and magnitude within each cluster is limited. From the results of data exploration, it is found that earthquakes in the Aceh Province region generally display a shallow depth pattern, where the majority of earthquakes have an average depth of less than 70 kilometers, and a relatively small magnitude. The shallow depth of the earthquake indicates that these earthquakes often occur in the earth's crust, closer to the surface. Therefore, even though the magnitude is small, these earthquakes can cause significant damage on the surface. However, earthquake events in water areas tend to have a much higher magnitude and depth in comparison to those that occur on land.

The results of data exploration demonstrate that reducing the *MinPts* value or $x_i$ value results in an increase in the number of clusters formed. This is because the requirements for the formation of a cluster are more easily met. As the number of clusters increases, the identification of areas vulnerable to earthquakes also expands. Therefore, it is important to carefully consider determining the *MinPts* value or *xi* value to ensure the formation of high-quality clusters. Finally, in the third scenario, the highest average earthquake depth value is in Cluster 2, which is 34.14 km, while the highest average earthquake magnitude value is in Cluster 1, which is 3.31 RS. In addition, the standard deviation in each cluster is lower than the average of each, indicating that the earthquake depth and magnitude data in each cluster have small variation. In addition to the cluster characteristics presented in Table 4, the characteristics of data points that are not assigned to any cluster (noise) for each clustering scenario are presented in Table 5.

**Table 5** Characteristics of Earthquakes Considered as Noise

| Scenario | Variable | Statistics | | | | |
|---|---|---|---|---|---|---|
| | | Average | Median | Standard Deviation | Minimum | Maximum |
| Optimal SC value (default) | Depth (km) | 38.85 | 28 | 34.31 | 1 | 197 |
| | Magnitude (RS) | 3.19 | 3.15 | 0.63 | 1.79 | 4.87 |
| Reduce *MinPts* and optimal SC value | Depth (km) | 26.12 | 10 | 29.25 | 1 | 197 |
| | Magnitude (RS) | 3.17 | 3.11 | 0.61 | 1.59 | 4.87 |
| Supremum of SC value | Depth (km) | 32.36 | 15 | 35.44 | 1 | 197 |
| | Magnitude (RS) | 3.22 | 3.17 | 0.64 | 1.59 | 4.79 |
| Tsunami-prone events | Depth (km) | 27.06 | 25 | 13.3 | 10 | 68 |
| | Magnitude (RS) | 3.37 | 3.26 | 0.63 | 2.08 | 4.87 |

Based on Tabel 5, in the default exploration, the noise data found in the optimal clustering has no significant variation in earthquake depth and magnitude, which is indicated by a standard deviation that is smaller than the mean value. In the first exploration, the noise data at the optimum *MinPts*=15 have significant variation in the earthquake depth, which is indicated by a standard deviation that exceeds the average value. In contrast, the magnitude data shows a relatively lower variation, with a standard deviation smaller than the average value. Meanwhile, in the second exploration, there is a significant variation in the earthquake depth, which is indicated by a standard deviation that exceeds the average value. This indicates that the earthquake depth in the noise data varies greatly.

The clustering results show that the majority of earthquakes occur around the Andaman Sea, especially at depths of less than 70 km below sea level. This is because there is seismic activity on the seabed, including the subduction process where the oceanic tectonic plate sinks under the continental plate, and the movement of other tectonic plates. This interaction creates pressure and tension that can cause earthquakes. In addition, shifts along fault lines on the seabed can also create seismic events in the region. Based on the previous Table 3, the four cluster results (all explorations) are widely distributed in Aceh Besar Regency, Pidie Regency, Pidie Jaya Regency, West Aceh Regency, Bireuen Regency, Central Aceh Regency, Gayo Lues Regency, South Aceh Regency, Simeulue Regency, and the Andaman Sea. However, differences appear in the results of the optimum silhouette coefficient value clustering, which covers the areas of Sabang City, Aceh Jaya Regency, Nagan Raya Regency, Bener Meriah Regency, North Aceh Regency, East Aceh Regency, Langsa City, Southeast Aceh Regency, Aceh Singkil Regency, and Southwest Aceh Regency. In addition, in the results of the second data exploration (cluster with a supremum SC of 0.71), which covers the areas of Nagan Raya Regency, Bener Meriah Regency, Aceh Jaya Regency, Aceh Singkil Regency, Southwest Aceh Regency, and Southeast Aceh Regency.

Based on the results, the best clustering results is determined by considering the number of clusters formed, especially those with the largest number. This is based on the belief that the more clusters formed, the more areas can be identified as earthquake-prone areas, which ultimately provides information for the public to increase awareness of areas prone to earthquakes. Therefore, it is concluded that the best clustering results is found in the second data exploration (*MinPts* = 20, *xi* = 0.05, and SC = 0.721), which consists of a total of 4 clusters, with 276 points in the clusters and 308 points classified as noise.

Although the best clustering result (*MinPts* = 20, *xi* = 0.05, SC = 0.72) has been identified, this result is important because it allows a broader identification of earthquake-prone areas in Aceh Province. The formation of four clusters provides a more detailed description of the spatial distribution of seismic activity compared to results with higher silhouette coefficients but fewer clusters. This makes it possible to identify additional areas that consistently experience earthquakes, both on land and offshore, which may not be clearly detected when stricter clustering settings are applied.

From a practical perspective, the wider identification of earthquake-prone zones is relevant for disaster mitigation and community preparedness. The clustering results can be used by local governments and disaster management agencies to better prioritize mitigation efforts, strengthen risk communication, and increase public awareness in vulnerable areas. Thus, this result is not only statistically acceptable but also meaningful for supporting disaster risk reduction and spatial planning.

Among the four clustering results, this is the clustering with the lowest silhouette coefficient, which is 0.721. However, this clustering still falls within the category of a strong or good clustering. Based on the results of the optimal clustering analysis, the findings indicate a strong similarity between the identified high-risk areas and the 2020 earthquake risk map of Aceh Province from Regional Disaster Management Agency (BPBD). Both sources highlight regions vulnerable to earthquakes, including Aceh Besar Regency, Pidie Regency, Pidie Jaya Regency, Aceh Barat Regency, Nagan Raya Regency, Gayo Lues Regency, Aceh Tengah Regency, Aceh Jaya Regency, Aceh Selatan Regency, Aceh Singkil Regency, Simeulue Regency, Aceh Barat Daya Regency, Aceh Tenggara Regency, as well as several areas in western Aceh, which the BPBD has classified as having a relatively high earthquake risk index.

From a policy perspective, the identified clustering results provide actionable information for disaster mitigation planning in Aceh Province. The spatial concentration of high-risk clusters enables local governments and disaster management agencies to prioritize mitigation efforts, such as strengthening early warning systems, improving seismic-resistant infrastructure, and allocating emergency response resources more efficiently in the most vulnerable regions. Moreover, the consistency between the clustering outcomes and the official earthquake risk map from the BPBD indicates that the proposed clustering framework can serve as a complementary, data-driven tool to support evidence-based disaster mitigation policies and regional risk management strategies. In practical terms, these clustering results can be used as a spatial decision-support tool by government agencies and disaster mitigation institutions to determine priority areas for mitigation actions, differentiate intervention strategies across regions, and support planning of preparedness and response activities based on the dominant seismic characteristics of each cluster.

Additionally, earthquake-prone areas in Bireuen Regency and Bener Meriah Regency have been classified by BPBD Aceh Province as regions with a moderate earthquake risk index. This vulnerability is attributed to several significant geological and tectonic factors. Most of these regions are in proximity to the Sumatra Fault, an active fault system known for frequent seismic activity. Furthermore, the Andaman Sea and the western coast of Sumatra are part of the Sunda Trench subduction zone, where the Indo-Australian Plate subducts beneath the Eurasian Plate, generating immense stress that often triggers earthquakes. Additionally, Aceh is located within the Pacific Ring of Fire, a region characterized by high tectonic activity. These factors make Aceh highly susceptible to both shallow and deep earthquakes of varying magnitudes, with the potential to generate tsunamis. Hence, disaster mitigation should be a routine and sustainable task (sustainable disaster mitigation).

Mitigation activities should be performed long before any disaster, which often comes faster than expected times, and even has a greater intensity than originally estimated. In addition, the government should also actively provide various appropriate and continuous directions in dealing with disaster events, and promptly adapt to the potential risks of natural disasters that exist. In the context of disaster risk reduction, disaster mitigation is also understood as an effort to increase the capacity of communities in disaster-prone areas to eliminate or reduce the effects of threats and levels of disaster [34]. In this regard, the clustering results obtained in this study provide a practical basis for implementing proactive mitigation by identifying and prioritizing earthquake- and tsunami-prone areas. These results can support government agencies and disaster mitigation institutions in designing location-specific mitigation strategies, optimizing resource allocation, and enhancing community preparedness in the most vulnerable regions.

## 4.0 CONCLUSION

During the period of 2022 to 2023, a total of 584 earthquake events were recorded in Aceh Province. In general, earthquakes in this region tend to have a depth of less than 70 km (shallow) and relatively low magnitudes. This study employs the OPTICS clustering method, with parameter adjustments to optimize the clustering process. Additionally, three main exploratory scenarios were conducted: (1) reducing the *MinPts* value to refine cluster formation, (2) clustering based on the supremum of the SC value to enhance clustering accuracy, and (3) identifying earthquake events with potential tsunami risks based on clustering results. Among these explorations, the second approach—clustering based on the smallest upper bound—yielded the most optimal results, as it produced a more suitable number of clusters and had direct implications for disaster mitigation. Based on the optimal clustering results, earthquakes occurring in Aceh Province during 2022–2023 were classified into four clusters. The most earthquake-prone areas, defined as regions with the highest earthquake frequency, include Aceh Besar, Pidie, Pidie Jaya, Aceh Barat, Nagan Raya, Gayo Lues, Aceh Tengah, Bireuen, Bener Meriah, Aceh Jaya, Aceh Singkil, Simeulue, Aceh Barat Daya, and Aceh Tenggara. These regions are predominantly located along the Sumatra Fault, an active fault system. Furthermore, the southern part of Aceh, including Aceh Selatan Regency, as well as offshore areas along the western and southern coasts of Aceh, particularly the Andaman Sea, exhibit high seismic activity associated with the Sunda Trench subduction zone. Future research could be expanded by integrating the OPTICS method with deep learning techniques to improve accuracy and scalability in handling large and complex datasets.

## Acknowledgement

We would like to express our gratitude to the Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia, for providing research facilities and continuous support throughout this study.

## Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper

## References

[1] Djalante, R. 2018. Review article: A Systematic Literature Review of Research Trends and Authorships on Natural Hazards, Disasters, Risk Reduction and Climate Change in Indonesia. *Natural Hazards and Earth System Sciences*. 18(6): 1785–1810. DOI: https://doi.org/10.5194/nhess-18-1785-2018.

[2] Ismail-Zadeh, A.T. 2013. Earthquake Prediction and Forecasting. In: Bobrowsky, P.T. (eds) Encyclopedia of Natural Hazards. *Encyclopedia of Earth Sciences Series*. Springer, Dordrecht. DOI: https://doi.org/10.1007/978-1-4020-4399-4_106

[3] Cheng, S., and H. Liu. 2023. General Review of the Worldwide Tsunami Research. *Journal of Marine Science and Application.* 22 (1): 14–24. DOI: https://doi.org/10.1007/s11804-023-00315-z.

[4] Putriana, S., E. Ernawati, and D. Andreswari. 2021. Clustering Data Titik Gempa dengan Metode Fuzzy Possibilistic C-Means (Studi Kasus: Titik Gempa Pulau Sumatera Tahun 2013- 2018). *Rekursif: Jurnal Informatika*. 9(1): 99-119. DOI: https://doi.org/10.33369/rekursif.v9i1.15898.

[5] Benazir, and R. S. Oktari. 2024. Assessing Tsunami Risk Along the Aceh Coast, Indonesia: A Quantitative Analysis of Fault Rupture Potential and Early Warning System Efficacy for Predicting Arrival Time and Flood Extent. *Natural Hazards*. 120(5): 4875–4900. DOI: https://doi.org/10.1007/s11069-024-06401-x.

[6] Badan Penanggulangan Bencana Aceh. 2021. *Dokumen Kajian Risiko Bencana Nasional Provinsi Aceh 2021-2025*. Indonesia: Kartsaceh.

[7] King, R.S.,. 2015. *Cluster Analysis and Data Mining: an Introduction*. United States of America: Mercury Learning and Information.

[8] Aggarwal, C.C. 2018. *An Introduction to Cluster Analysis. in Data Clustering: Algorithm and Application*. New York: Chapman and Hall/CRC.

[9] Saavedra-Nieves, P., and M. Fernández-Pérez. 2024 Directional Density-Based Clustering. *Advances in Data Analysis and Classification.* 20(1): 289-314. DOI: https://doi.org/10.1007/s11634-024-00616-3.

[10] Menardi, M., and D. De Stefano. 2022. Density-based clustering of Social Networks. *Journal of the Royal Statistical Society Series A: Statistics in Society.* 185(3): 1004–1029. DOI: https://doi.org/10.48550/arXiv.2101.08334.

[11] Liu. L., Z. Hu, C. Zhou, and G. Xu. 2018. Research on the Clustering Algorithm of the Bicycle Stations Based on OPTICS. *Concurrency and Computation: Practice and Experience*. 31(10): 1-7. DOI: https://doi.org/10.1002/cpe.4876.

[12] Wagner, T., R. Feger, and A. Stelzer. 2018. Modifications of the OPTICS Clustering Algorithm for Short-Range Radar Tracking Applications. *15th European Radar Conference (EuRAD).* DOI: https://doi.org/10.23919/EuRAD.2018.8546579.

[13] Ahmed, M.A., H. Baharin, and P. N. E. Nohuddin. 2020. Analysis of K-means, DBSCAN and OPTICS Cluster Algorithms on Al-Quran Verses. *International Journal of Advanced Computer Science and Applications.* 11(8): 248-254. DOI https://doi.org/10.14569/IJACSA.2020.0110832.

[14] Ankerst M., M. M. Breunig, H. P. Kriegel, and J. Sander. 1999. OPTICS: Ordering Points to Identify the Clustering Structure. *SIGMOD Record (ACM Special Interest Group on Management of Data).* 28(2): 49–60. DOI: https://doi.org/10.1145/304182.304187.

[15] Zhu, X., S. Nie, C. Wang, X. Xi, J. Wang, D.Li, and H. Zhou. 2021. A Noise Removal Algorithm Based on OPTICS for Photon-Counting LiDAR Data. *IEEE Geoscience and Remote Sensing Letters.* 18(8): 1471–1475. DOI: https://doi.org/10.1109/LGRS.2020.3003191.

[16] Febriana, N.L., and I. S. Sitanggang . 2017. Outlier Detection on Hotspot Data in Riau Province using OPTICS Algorithm. *IOP Conf. Series: Earth and Environmental Science*. 58(2017). DOI: https://doi.org/10.1109/LGRS.2020.3003191.

[17] Fan Y., and M. Wang,. 2024. Specification Mining Based on the Ordering Points to Identify the Clustering Structure Clustering Algorithm and Model Checking. *Algorithms.* 17(1): 28. DOI: https://doi.org/10.3390/a17010028.

[18] Rodriguez, M.Z., C.H. Comin, D. Casanova, O.M. Bruno, D.R. Amancio, L.F. Costa, and F.A. Rodrigues. 2019. Clustering Algorithms: A comparative approach. *PLoS One*. 14 (1): e0210236. DOI: https://doi.org/10.1371/journal.pone.0210236.

[19] Yin, L., X.Li, W.Zheng, Z.Lin, L.Song, L.Ge, and Q.Zeng. 2019. Fractal Dimension Analysis For Seismicity Spatial And Temporal Distribution In The Circum-Pacific Seismic Belt. *Journal of Earth System Science*. 128(1). DOI: https://doi.org/10.1007/s12040-018-1040-2.

[20] Moustafa, S.S.R., M.H.Yassien, M.Metwaly, A.M.Faried, and B.Elsaka. 2024. Applying Geostatistics to Understand Seismic Activity Patterns in the Northern Red Sea Boundary Zone. *Applied Sciences.* 14(4):1455. DOI: https://doi.org/10.3390/app14041455.

[21] Datasans. 2023. Machine Learning Cheatsheet : Model Klasifikasi. *Datasans*.

[22] Campello. R. J. G. B., P. Kröger, J. Sander, and A. Zimek. 2020. Density-based Clustering. *WIREs Data Mining and Knowledge Discovery.* 10(2): 1-15. DOI: https://doi.org/10.1002/widm.1343.

[23] Rajabi, F., and J. J. McArthur. 2024. Applying OPTICS with and without PCA for Fault Detection of Fan Coil Units Using Building Automation System Data. *Energy and Buildings.* 317: 114368. DOI: https://doi.org/10.1016/j.enbuild.2024.114368

[24] Esri. (n.d.). How density-based clustering works. ArcGIS Pro documentation. Retrieved August, 2025, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-density-based-clustering-works.htm

[25] Sowan, B., T.-P. Hong, A. Al-Qerem, M. Alauthman, and N. Matar. 2023. Ensembling validation indices to estimate the optimal number of clusters. *Applied Intelligence*. 53(9): 9933–9957. DOI: https://doi.org/10.1007/s10489-022-03939-w.

[26] Pavlopoulos, J., G. Vardakas, and A. Likas. 2025. Revisiting Silhouette Aggregation *Discovery Science*. DOI: https://doi.org/10.1007/978-3-031-78977-9_23.

[27] Chicco, D., A. Campagner, A. Spagnolo, D. Ciucci, and G. Jurman. 2025. The Silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and Gap statistic for unsupervised clustering internal evaluation of two convex clusters. *PeerJ Computer Science.* 11. DOI: https://peerj.com/articles/cs-3309/.

[28] Struyf, a., M. Hubert, and P. Rousseeuw. 1997. Clustering in an Object-Oriented Environment. *Journal of Statistical Software*. 1(4). DOI: https://doi.org/10.18637/jss.v001.i04.

[29] Ester, M., H.-P. Kriegel, J. Sander, and X. Xu,. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96)*. 226–231.

[30] Schubert, E., J. Sander, M. Ester, H. P. Kriegel, and X. Xu. 2017. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Transactions on Database Systems*. 42(3). DOI: https://doi.org/10.1145/3068335.

[31] Kagan, Y. Y. 2010. Earthquake Size Distribution: Power-Law with Exponent β≡1/2?. *Tectonophysics*. 490(1-2): 103–114. DOI: https://doi.org/10.1016/j.tecto.2010.04.034.

[32] Chattamvelli, R., and R. Shanmugam. 2023. Descriptive Statistics for Scientists and Engineers: Applications in R. *Springer Nature Switzerland*. DOI: https://doi.org/10.1007/978-3-031-32330-0.

[33] Madlazim. 2012. Toward Tsunami Early Warning System in Indonesia by Using Rapid Rupture Durations Estimation. *AIP Conference Proceedings.* 1454(1):142–145. DOI: https://doi.org/10.1063/1.4730707.

[34] Desfandi, M. 2014. Urgensi Kurikulum Pendidikan Kebencanaan Berbasis Kearifan Lokal di Indonesia. *Sosio-Didaktika: Social Science Education Journal*. 1(2): 191–198. DOI: https://doi.org/10.15408/sd.v1i2.1261