



Optimizing DBSCAN Parameters for Depth-Based Earthquake Clustering Using Grid Search

Rushendra^{1*}, Ody Octora Wijaya², Mohamad Yusuf³, Andri Setiyaji⁴, Djoko Prabowo⁵

¹Department of Informatics, Universitas Mercu Buana, Jakarta, Indonesia

²Department of System Engineer, PT. Mindotama Avia Teknik, Jakarta, Indonesia

³Department of Informatics, Universitas Mercu Buana, Jakarta, Indonesia

^{4,5}Department of Instrumentation, Sekolah Tinggi Meteorologi dan Geofisika, Jakarta, Indonesia

¹rushendra@mercubuana.ac.id, ²odyoctorawijaya@gmail.com, ³mhd.yusuf@mercubuana.ac.id, ⁴andri.setiyaji@stmkg.ac.id, ⁵djoko.prabowo@stmkg.ac.id

Abstract

This study addresses the challenge of accurately clustering earthquake events based on depth to better understand seismic activity patterns in Sulawesi from 2019 to 2023. Traditional clustering algorithms often fail to capture the complex spatial and depth-based structures of earthquake data. To overcome this, we employed the DBSCAN algorithm, which is well-suited for identifying irregularly shaped clusters and handling noise in spatial datasets. A key focus of this research is the systematic optimization of DBSCAN's parameters—epsilon (ϵ) and minimum samples ($min_samples$)—using a grid search approach. Epsilon values varied from 0.1 to 0.5, and $min_samples$ ranged from 6 to 60. The optimal parameters, determined using the Calinski-Harabasz (CH) index, were $\epsilon = 0.4$ and $min_samples = 54$. Compared with previous heuristic settings, the optimized configuration produced better separated and more interpretable clusters. Using the optimized parameters, nine distinct clusters were identified, capturing meaningful patterns in both depth and magnitude. The results revealed that shallow earthquakes (0–20 km) tend to exhibit greater magnitude variation, with some clusters averaging magnitudes up to 3.7. This suggests a higher seismic hazard potential associated with brittle crustal activity. The findings contribute to seismic hazard analysis by providing a more robust understanding of three-dimensional earthquake distribution, aiding regional risk assessment and disaster preparedness efforts. These insights can support agencies such as BMKG and BPBD in hazard mapping, sensor deployment, and contingency planning for high-risk zones.

Keywords: clustering; DBSCAN; earthquake data; grid search; Sulawesi; seismic gap

How to Cite: R. Rushendra, O. O. Wijaya, M. Yusuf, A. Setiyaji, and D. Prabowo, "Optimizing DBSCAN Parameters for Depth-Based Earthquake Clustering Using Grid Search", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 4, pp. 850 - 864, Aug. 2025.
Permalink/DOI: <https://doi.org/10.29207/resti.v9i4.6521>

Received: April 7, 02025

Accepted: August 11, 2025

Available Online: August 19, 2025

This is an open-access article under the [CC BY 4.0 License](#)

Published by [Ikatan Ahli Informatika Indonesia](#)

1. Introduction

Sulawesi Island is one of Indonesia's most seismically active regions due to its position along the Pacific Ring of Fire, a zone characterized by frequent tectonic activity. The island features complex geological structures, including multiple active faults that contribute significantly to earthquake occurrences and related hazards such as tsunamis [1]. Major fault systems in Sulawesi include the Palu-Koro and Walanae faults in the west, as well as the Matano, Lawanopo, and Gorontalo faults in the east and north [2]. Understanding the spatial distribution of earthquakes across these fault zones is essential for effective disaster risk mitigation.

In seismological research, clustering algorithms such as k-means, hierarchical clustering, and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) have been widely used to identify seismic activity patterns and delineate seismic gaps. These studies often rely on evaluation metrics like the Silhouette Score, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index to assess cluster quality. However, many reported difficulties in achieving well-separated and meaningful clusters, with common issues such as overlapping boundaries between seismic regions. For example, as demonstrated in the study by Ody Octora Wijaya (2024), a silhouette score of 0.06097 and a DBI of 2.4648 are indicative of poor separation and high internal cluster variation [3].

Traditional clustering methods such as k-means and hierarchical clustering often assume spherical cluster shapes and require prior knowledge of the number of clusters. These assumptions are problematic in seismological contexts, where earthquake distributions are irregular, sparse, and influenced by complex fault geometries. As a result, these methods frequently produce overlapping or poorly separated clusters, limiting their utility in seismic zone delineation. DBSCAN, by contrast, offers a density-based approach that can detect clusters of arbitrary shape and handle noise effectively—making it particularly suitable for earthquake data analysis in tectonically complex regions like Sulawesi.

Our previous research applied the DBSCAN algorithm to examine the spatial distribution of earthquake epicenters in Sulawesi. However, that study relied on fixed, trial-and-error parameter values without systematic tuning and did not incorporate earthquake depth—an essential factor in understanding the vertical structure of seismic activity. This limited the clarity and interpretability of the clustering outcomes [3].

Similar limitations are also found in other seismological clustering studies, where the selection of DBSCAN's parameters (epsilon and min_samples) is often heuristic and lacks reproducibility [4], [5]. Although a few studies have attempted to tune DBSCAN parameters using methods such as visual estimation or rule-of-thumb approaches, these lack systematic evaluation and scalability [6]. Some recent research has begun exploring automated tuning techniques for clustering, including the use of metaheuristic algorithms or statistical validation, but such efforts remain rare in earthquake data applications [7].

While DBSCAN has been applied in various global seismological studies, its systematic use in the Indonesian context—especially with parameter optimization and depth integration—remains limited. To our knowledge, no prior research has comprehensively applied DBSCAN with grid-based tuning and three-dimensional clustering to earthquake data in Sulawesi. This study therefore represents a novel contribution to regional seismological analysis and offers a replicable framework for other tectonically active zones in Indonesia.

In this study, we address these gaps by introducing two methodological improvements: (1) systematic optimization of DBSCAN parameters using Grid Search, and (2) inclusion of earthquake depth as a feature for three-dimensional clustering. Grid Search is selected for its exhaustive and deterministic search space exploration, offering consistent and reproducible results across combinations of epsilon and min_samples. While more computationally intensive, it provides a transparent framework for evaluating clustering performance across different parameter settings. Several studies in other domains (e.g., image segmentation, anomaly detection) have successfully

employed Grid Search for DBSCAN tuning [8], yet its application in seismic data clustering remains limited.

The novelty of this research lies in integrating parameter optimization and vertical depth information into DBSCAN-based earthquake cluster analysis—an approach not yet systematically applied in previous studies of the Sulawesi region. The contributions are threefold: (1) improving DBSCAN clustering through parameter tuning tailored to Sulawesi seismic characteristics, (2) incorporating earthquake depth to enhance the discrimination of deep vs. shallow seismic zones, and (3) delivering refined spatial-depth clustering results to support more accurate seismic risk mapping and disaster preparedness in a highly tectonic region of Indonesia. This novel, data-driven approach seeks to advance earthquake cluster analysis in Indonesia by improving the reliability and granularity of clustering results in one of its most tectonically active regions.

This data-driven approach advances existing clustering methodologies in earthquake analysis by improving both the reliability and granularity of cluster identification, ultimately contributing to a deeper understanding of seismic patterns in Southeast Asia.

2. Methods

This study builds upon our previous research on earthquake clustering in Sulawesi (2019–2023), utilizing the same dataset [3]. While the earlier analysis employed DBSCAN with heuristic parameter settings and focused solely on spatial epicenter distribution, the present study introduces an enhanced methodology. Specifically, it incorporates earthquake depth as an additional clustering dimension and applies systematic parameter optimization to more accurately represent the three-dimensional nature of seismicity in the region.

The methodological framework of this study is illustrated in Figure 1, outlining the complete process from data collection and preprocessing to feature selection, hyperparameter tuning, and clustering. This structured workflow ensures reproducibility and enhances the reliability and interpretability of the clustering outcomes.

As shown in Figure 1, the methodology follows a structured flow to support effective earthquake data analysis. It begins with data collection, followed by preprocessing steps such as integration and cleaning to ensure data quality. Feature selection is then conducted to identify key variables for clustering. Grid Search is used for hyperparameter tuning, systematically exploring combinations of epsilon and min_samples to optimize DBSCAN performance. Once the optimal parameters are determined, DBSCAN is applied to identify clusters in the earthquake dataset. The final stage involves visualization and interpretation of the clustering results to uncover meaningful seismic patterns and assess regional seismic risks.

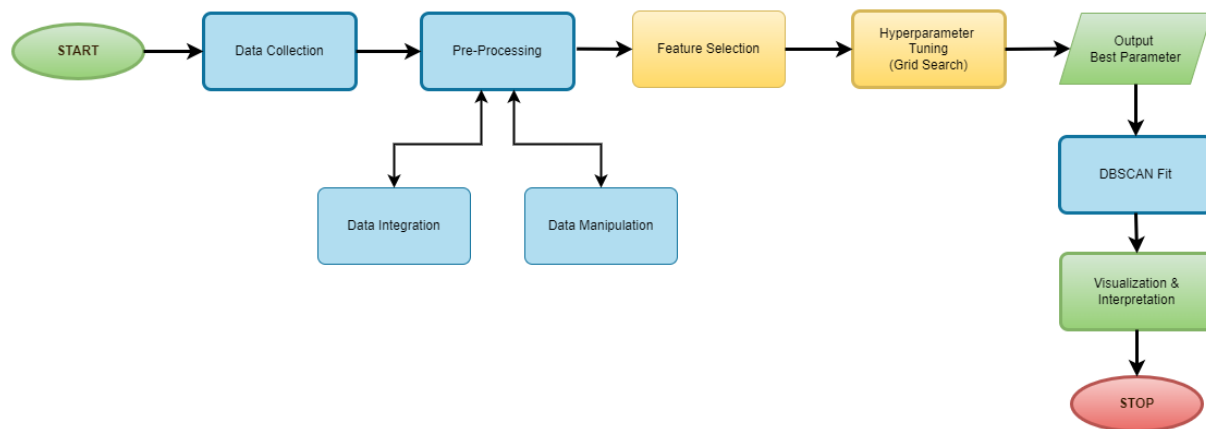


Figure 1. Research Flow

2.1 Data Collection

The earthquake dataset used in this study was sourced from the official repository of Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), Indonesia's national agency for meteorological and geophysical monitoring. It comprises earthquake events recorded in the Sulawesi region between January 1, 2019, and December 31, 2023. The dataset was selected for its comprehensiveness and the credibility of BMKG as a national authority, ensuring high standards of reliability and data verification [3].

A total of 10,255 earthquake records were retrieved, each containing attributes such as event ID, origin time, latitude, longitude, magnitude, and depth. The data were accessed in CSV format and processed using the Pandas library within a Python-based analytical environment. This structure enabled efficient data manipulation and seamless integration with subsequent clustering procedures. The dataset provides a recent and detailed foundation for analyzing seismic activity patterns in Sulawesi.

2.2 Data Preprocessing

Prior to analysis, the dataset underwent a series of preprocessing steps to ensure data quality and consistency. First, duplicate and invalid entries were removed by filtering based on the unique "Event ID" field, reducing the dataset to 10,238 valid earthquake events. This step was critical for maintaining the integrity of the clustering process and avoiding bias introduced by repeated entries [3].

An important enhancement in this study was the incorporation of earthquake depth as a clustering feature. While many previous studies focused solely on horizontal spatial attributes (latitude and longitude), depth provides essential insights into the vertical distribution of seismicity, which is often linked to tectonic processes such as subduction and faulting. To ensure uniform contribution from all features, the dataset was standardized using the z-score method, which centers and scales the variables based on their mean and standard deviation [9].

Depth plays a significant role in understanding earthquake behavior, as shallow earthquakes (depth < 70 km) are typically more damaging due to proximity to the surface, while deeper events (depth > 70 km) may signal different tectonic origins [10].

Sulawesi's complex tectonic environment characterized by the convergence of multiple microplates and subduction zones produces a diverse range of earthquake depths [11]. By emphasizing depth in the clustering process, this study aims to reveal hidden seismic patterns below the surface that are crucial for accurate hazard assessment and disaster preparedness strategies [12].

Unlike studies that apply strict filtering or outlier removal, this research intentionally retains all recorded earthquake events from 2019 to 2023 to preserve the full spectrum of seismic activity in Sulawesi. The dataset was compiled by merging annual records, followed by duplicate removal based on the unique "Event ID" field. This inclusive approach ensures that rare or extreme events—often critical for hazard interpretation—are not excluded prematurely. While no manual noise filtering was applied, data integrity was maintained through consistency checks and standardization using z-score normalization.

2.3 Parameter Tuning for DBSCAN

In this study, the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm was utilized for clustering earthquake data, with a particular focus on depth as a significant feature for cluster formation. DBSCAN is a density-based clustering method that groups data points closely packed together while treating isolated points as noise. However, the effectiveness of DBSCAN depends heavily on the careful tuning of two key parameters: epsilon (eps) and minimum points (min_samples) [13].

The eps parameter defines the maximum distance within which points are considered part of the same cluster. Optimizing eps is crucial, as too small a value may result in excessive noise (unclustered points), while too large a value may lead to overly general

clusters that lack meaningful distinctions [14]. The `min_samples` parameter specifies the minimum number of neighboring points required for a core point to initiate a cluster. Proper calibration of this value ensures that clusters represent genuine seismic groupings rather than random aggregations [15].

To optimize these parameters, we employed a systematic tuning process. First, a range of `eps` and `min_samples` value were tested to evaluate cluster quality. Several clustering performance metrics were used to assess different parameter configurations, including the Silhouette Score, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CH Index). The Silhouette Score helps assess cohesion and separation among clusters, while the DBI and CH Index provide insights into cluster compactness and distinctness. Depth was specifically used as a key feature to ensure that the resulting clusters reflected depth-related earthquake patterns [12], [16].

Through this tuning process, we identified the optimal `eps` and `min_samples` values that resulted in the best clustering performance for our dataset, particularly for analyzing depth-based earthquake distributions in Sulawesi. The selected parameters allow DBSCAN to effectively differentiate clusters based on earthquake depth, providing more granular insights into depth-related seismic patterns across the region. The parameter tuning enhances the accuracy and interpretability of the clustering outcomes, facilitating a deeper understanding of the seismic activity patterns related to the geological features in Sulawesi [13], [17].

2.4. Grid Search Optimization

The grid search method systematically evaluates each combination of parameters within a predefined range to identify the optimal set for DBSCAN. Given a set of values for `epsilon` (ϵ) and `min_samples`, the grid search optimizes the parameters by finding those that yield the highest average score across multiple cluster evaluation metrics [18].

To fine-tune the DBSCAN parameters, we used grid search with varying `epsilon` and `min_samples` over a wider range than in previous studies. We set the `epsilon` range from 0.1 to 0.5 in increments of 0.1 and `min_samples` from 6 to 60 in increments of 6.

The selection of `epsilon` (0.1–0.5) and `minPts` (6–60) ranges is based on three considerations. First, since the input features (latitude, longitude, depth) are standardized, the `epsilon` values are chosen within a small range appropriate for z-score scaled distances [19]. Second, the `minPts` values are derived from the dimensionality rule ($\text{minPts} \geq D + 1$) and expanded to accommodate both dense and sparse seismic regions [13]. Third, the extended parameter range improves upon the previous version of this study by increasing the sensitivity of DBSCAN to capture vertical (depth-based) clustering structures in Sulawesi's complex tectonic setting. This broader exploration enables more

robust identification of distinct seismic patterns, potentially valuable for disaster mitigation [20], [21].

While some studies rely on k-distance plots for parameter estimation, this research adopts an empirical approach through repeated Grid Search experiments. Each configuration was evaluated using clustering performance metrics to identify parameter sets that produced coherent, interpretable clusters aligned with depth-related seismic behavior.

2.5 Cluster Quality Evaluation

Cluster quality evaluation is essential for determining the reliability and interpretability of clustering results. In this study, Silhouette Score, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index are used to evaluate clustering performance from different perspectives [22]. During the grid search, we evaluate each parameter combination using the following three clustering metrics:

Silhouette Score (SS): This score measures the cohesion within clusters and separation between clusters. A higher Silhouette score indicates well-defined clusters, as points are closer to their own cluster than to others [23]. Defined as Equation 1.

$$SS = \frac{b-a}{\max(a,b)} \quad (1)$$

a is the mean intra-cluster distance and b is the mean nearest'-cluster distance.

Davies-Bouldin Index (DBI): Assesses cluster separation by measuring the average similarity between each cluster and the one most similar to it. Lower DBI values indicate better clustering because clusters are more compact and distinct from one another [24]. This is calculated as Equation 2.

$$DBI = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \frac{s_i + s_j}{d_{i,j}} \quad (2)$$

s_i and s_j are the intra-cluster distances and $d_{i,j}$ is the distance between the cluster centroids.

Calinski-Harabasz Index (CH): Also called the Variance Ratio Criterion, this index evaluates the ratio of cluster dispersion between and within clusters. Higher Calinski-Harabasz scores suggest a better clustering structure because cluster are well-separated [25], expressed as Equation 3.

$$CH = \frac{\text{tr}(B_k)/(k-1)}{\text{tr}(W_k)/(n-k)} \quad (3)$$

B_k is the between-cluster dispersion matrix and W_k is the within-cluster dispersion matrix.

These metrics, which are particularly used in recent clustering research, provide a robust framework for evaluating clustering structures without requiring labelled data, thereby making them widely applicable in unsupervised learning contexts. For further insights and current methodologies, refer to recent clustering evaluation metrics [26].

2.6 DBSCAN Algorithm

DBSCAN is used due to its ability to form clusters of arbitrary shapes and handle noise in data. Given the nature of earthquake data, where noise points are common, DBSCAN's density-based approach is particularly useful.

The algorithm works by setting a distance threshold (epsilon) within which points are considered part of the same cluster if they also meet the minimum number of points (min_samples) in that neighbourhood [27].

2.7 Visualization and Interpretation

The resulting clusters were visualized using color gradients representing earthquake depth, allowing clearer differentiation among clusters based on depth variations across Sulawesi. To gain a clearer understanding of the earthquake distribution in Sulawesi, the clustering results were visualized through a series of map plots, overlaid with fault lines specific to the region. The visualization scenario involved mapping each identified cluster by its geographic coordinates and color-coding clusters to reflect variations in depth and magnitude [28]. This approach enabled the identification of spatial patterns in earthquake occurrences and their proximity to known fault lines and tectonic boundaries [29].

The visual outcomes reveal distinct cluster formations across different depths and magnitudes, showing correlations between cluster locations and fault lines. For example, clusters with shallower depths tended to align closely with the Palu-Koro fault system, indicating frequent seismic activity near this fault. However, deeper clusters appeared farther from major fault lines, indicating that tectonic activity is potentially related to subduction processes [30]. Thus, scenario-driven visualization offers a comprehensive perspective on how different depth clusters are distributed relative to tectonic structures, enhancing the understanding of seismic patterns and potential earthquake risks in Sulawesi. By Visualizing these relationships, this study provides a deeper insight into regional seismic phenomena and their underlying geological influences [31].

3. Results and Discussions

The grid search results indicate the best values for epsilon and min_samples based on the three-evaluation metrics. According to the Silhouette Score, the optimal parameters were set to epsilon = 0.4 and min_samples = 50. For the Davies-Bouldin Index (DBI), the best parameters were epsilon = 0.2 and min_samples = 7. Meanwhile, the Calinski-Harabasz Index (CH Index) indicates that optimal values are epsilon = 0.2 and min_samples = 54. Table 1 presents the results obtained by the grid search method.

The grid search method identified the optimal epsilon and min_samples parameters based on the three-evaluation metrics. We then visualized the results of

these parameters within the model developed to observe clustering outcomes.

Table 1. Matrix of epsilon and min point search results

Metric	SS Score	DBI	CH Index
eps	0.4	0.2	0.2
Min samples	50	7	54
Silhouette S.	0.414768	-0.057093	0.034479
DBI	2.269853	1.187298	3.244152
CH Index	738.402986	108.612398	956.627764
Best for	Silhouette	DBI	CH Index
	Score		

Table 1 presents a matrix comparing DBSCAN clustering results across various epsilon ('eps') and minimum points ('min_samples') values, evaluated using three key metrics: the Silhouette Score, Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CH Index). Each metric led to different optimal values, highlighting distinct clustering characteristics based on the evaluation criteria used.

For the Silhouette Score, which measures the cohesion and separation of clusters, the best result was achieved with an epsilon of 0.4 and min points of 50, yielding a score of 0.414768. Although moderate, this score reflects the highest level of cohesion among the tested configurations. The DBI, a metric where lower values signify better separation, identified its optimal configuration at an epsilon of 0.2 and min points of 7, with a DBI score of 1.187298. This low DBI indicates clusters with minimal overlap, although the Silhouette Score associated with this configuration was negative (-0.057093), indicating potential limitations in separation quality.

In terms of the Calinski-Harabasz Index, which favors compact and distinct clusters, the highest score of 956.627764 was achieved with an epsilon of 0.2 and min points of 54. This indicates dense clustering with strong boundaries between clusters, although the relatively low Silhouette Score (0.034479) for this configuration indicates some limitations in achieving high-quality separation.

In summary, the optimal values for epsilon and min points vary depending on the evaluation metric being prioritized. For the Silhouette Score, the optimal parameters were epsilon 0.4 and min points 50; for DBI, epsilon 0.2 and min points 7; and for CH Index, epsilon 0.2 and min points 54. The selection of the best parameters depends on the specific clustering goals and evaluation metrics that are most relevant to the analysis.

3.1 Results

The optimal parameters for epsilon and min_samples according to the Silhouette Score from the grid search method resulted in a score of 0.414768. Figure 2 shows clustering visualization using the best epsilon and min point parameters based on the Silhouette Score evaluation.

Although this score is considered reasonably good, the visualization results indicate that objects within a

cluster are generally closer to each other than to objects in other clusters. However, there are still some objects that may be located near the cluster boundaries.

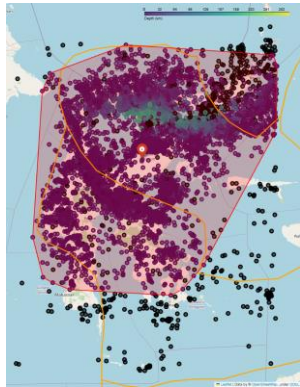


Figure 2. Visualization of clusters with optimal silhouette score parameters

Using DBSCAN with the optimal parameters obtained from Grid Search ($\epsilon = 0.4$, $\text{min_samples} = 50$), the clustering process yielded a single dominant cluster comprising 9,617 earthquake events. The evaluation

metrics indicate moderately good clustering quality, with a Silhouette Score of 0.4148, a Davies-Bouldin Index of 2.27, and a Calinski-Harabasz Index of 738.40. These values suggest that although the cluster is relatively well-defined, there may still be a degree of internal variation.

Descriptive statistics for this configuration, as shown in Table 2, reveal significant variability in depth and magnitude. Earthquake depths ranged from 1 km to 224 km, with a mean of 33.09 km and a standard deviation of 44.24 km, indicating a broad vertical spread of seismicity. Magnitudes ranged from 0.81 to 6.78, averaging 3.20 with a standard deviation of 0.72. The large number of points and wide variability suggest that the cluster may encompass multiple tectonic regimes that are not separable using a single parameter setting.

The optimal values for epsilon and min_samples according to the Davies-Bouldin Index obtained by the grid search method resulted in 16 clusters in a score of 1.187298. Figure 3 shows clustering visualization using the best epsilon and min_samples parameters based on DBI score evaluation.

Table 2. Statistic per cluster for eps=0.4 & min point=50

Cluster	Depth Mean	Depth Std	Depth Min	Depth Max	Mag Mean	Mag Std	Mag Min	Mag Max	Count
0	33.091297	44.2422	1	224	3.195402	0.715765	0.813944	6.778839	9617

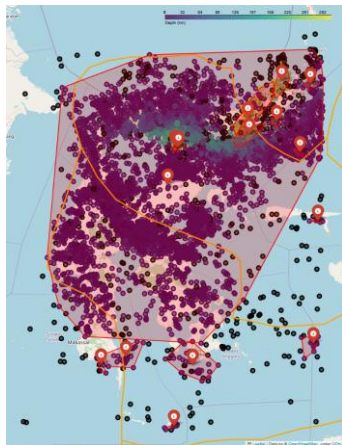


Figure 3. Visualization of clusters with optimal DBI score parameters

There are indications of overlap between clusters, which could be due to data noise, suboptimal parameter settings, or intrinsic data complexity. In addition, some clusters with low member density were identified, indicating the need to further adjust the minimum sample parameter.

In the second experiment, DBSCAN was configured with $\epsilon = 0.2$ and $\text{min_samples} = 7$, yielding the lowest Davies-Bouldin Index (1.187), suggesting relatively well-separated clusters compared to other configurations. However, the Silhouette Score was -0.057, indicating potential overlap among clusters or noise, and the Calinski-Harabasz Index was 108.61, lower than the previous trial. Nevertheless, this configuration produced 16 distinct clusters, offering a more granular perspective of the seismic distribution.

Table 3. Statistic per cluster for eps=0.2 & min point=7

Cluster	Depth Mean	Depth Std	Depth Min	Depth Max	Mag Mean	Mag Std	Mag Min	Mag Max	Count
0	32.598580	43.996759	1	216	3.190401	0.708851	0.813944	6.778839	9439
1	12.555556	3.967327	10	24	3.670204	0.822018	2.313670	5.498359	36
2	12.354839	3.411240	10	20	2.872166	0.876516	1.756392	4.948916	31
3	193.266667	8.055554	177	212	3.659167	0.648644	2.580365	5.111457	30
4	246.194030	11.595766	221	269	3.741953	0.524294	2.541016	5.337821	67
5	18.476190	9.222901	10	34	4.582934	0.598541	3.938729	6.514208	21
6	214.714286	8.527267	199	233	3.911726	0.375758	3.384588	4.949262	21
7	251.076923	10.657247	236	268	3.150695	0.607238	2.199972	4.160950	13
8	50.333333	5.830952	41	58	5.131105	0.422329	4.641149	5.900264	9
9	117.000000	5.792716	107	122	3.780582	0.595301	2.933036	4.660106	10
10	88.666667	6.461424	78	98	3.769941	0.498144	3.108414	4.709666	9
11	226.500000	5.916080	221	233	3.089935	0.574405	2.417673	3.584711	4
12	17.714286	3.683942	11	23	4.368910	0.360639	3.822556	4.923060	7
13	276.250000	4.112988	271	280	3.482434	0.541024	2.765833	4.017639	4
14	20.416667	6.501165	11	35	3.526485	0.609476	2.812387	4.949384	12
15	10.833333	2.041241	10	15	2.156563	0.447696	1.587983	2.814163	6

As detailed in Table 3, depth and magnitude characteristics vary significantly across clusters. Several clusters represent shallow events (e.g., clusters 1 and 2 with mean depths ~12 km), while others reflect deeper subduction-related activity (e.g., clusters 4, 6, 7, and 13 with mean depths over 200 km). Cluster 8 stands out with the highest average magnitude (5.13), while cluster 15 contains the weakest events (mean magnitude 2.16). These variations illustrate how smaller min_samples values allow DBSCAN to detect more localized seismic patterns that may otherwise be overlooked.

The analysis results show that the DBSCAN algorithm with parameters $\epsilon=0.2$ and min_samples=54 can produce high-quality clusters, as indicated by the very high Calinski-Harabasz Index value (956.627764). Figure 4 displays clustering visualization using optimal epsilon and mi_point parameters based on the Calinski-Harabasz Index evaluation.

Although some clusters still appear to overlap in the visualization, using the optimal epsilon and min_sample values from the grid search method can provide insights into the clusters formed and their proximity to existing faults or tectonic plates in the Sulawesi region [32].

In the third experiment, DBSCAN was applied using $\epsilon = 0.2$ and min_samples = 54, resulting in the highest Calinski-Harabasz Index (956.63) among all tested configurations. This suggests that the clusters formed are both compact and well-separated. Although the Silhouette Score remained low (0.034) and the Davies-Bouldin Index was relatively high (3.24), the visual inspection of the clustering map revealed clearer spatial groupings than in the previous two trials.

Table 4. Statistic per cluster for eps=0.2 & min point=54

Cluster	Depth Mean	Depth Std	Depth Min	Depth Max	Mag Mean	Mag Std	Mag Min	Mag Max	Count
0	10.628979	2.352131	2	28	2.877901	0.642563	1.331115	6.463147	911
1	60.466667	51.427719	4	164	3.346258	0.658888	1.883383	6.496014	1740
2	10.786726	3.098127	1	36	3.057960	0.650652	1.010857	6.778839	3601
3	15.251894	7.375211	5	39	3.724402	0.571625	2.034929	5.562747	528
4	11.458824	3.247404	5	25	3.192138	0.574624	1.747259	4.594870	170
5	164.013624	17.249925	125	200	3.028274	0.609738	1.587741	4.912976	367
6	10.168539	0.979749	10	17	3.238843	0.604896	1.952641	5.668295	89
7	10.460000	1.607992	10	21	3.433679	0.595444	2.208485	5.469223	150
8	89.432927	12.348782	67	113	3.032333	0.768144	1.661972	5.479502	164

3.2 Discussions

Based on the results obtained, we identified parameters that are optimal for processing and analysis through visualization in the model, as indicated by the CH-Index evaluation using the grid search method. The visualization results showed that nine clusters were formed, providing insights into the depth of earthquake epicenters in the Sulawesi region. The average depth and magnitude of each cluster are presented in Table 2.

The Calinski-Harabasz (CH) Index was selected as the primary evaluation metric due to its sensitivity to both intra-cluster compactness and inter-cluster separation.

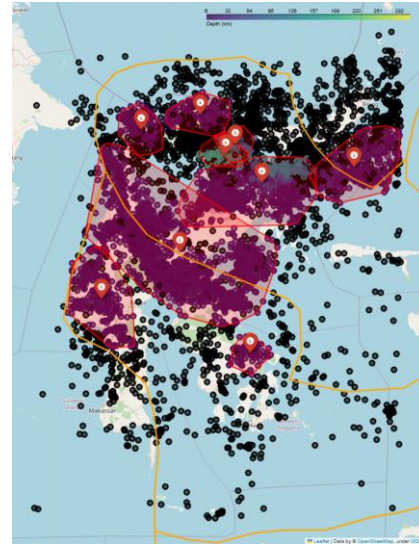


Figure 4. Visualization of clusters with optimal Calinski-Harabasz parameters

According to Table 4, this configuration produced 9 clusters, with most events concentrated in shallow zones (e.g., clusters 0, 2, 4, and 7 with mean depths around 10–15 km). Clusters 5 and 8 represented deeper seismic activity (mean depths of 164.01 km and 89.43 km, respectively), while cluster 1 exhibited a wide spread of depth values. Magnitude variation was also distinct, with cluster 3 having the highest average magnitude (3.72) and cluster 2 hosting the largest number of events (3,601) with a moderate average magnitude of 3.06. Overall, this configuration achieved a balance between statistical strength and visual interpretability, making it a compelling choice for identifying tectonic structures in Sulawesi.

Compared to the Silhouette Score and Davies-Bouldin Index, CH Index provides more consistent assessments in high-dimensional and noisy datasets, which are typical in seismic applications. Preliminary tests showed that CH Index better captured the structural coherence of clusters formed in Sulawesi's earthquake data.

To further analyze the internal characteristics of each cluster formed by the DBSCAN algorithm, a boxplot of earthquake magnitudes per cluster was generated. As shown in Figure 5, the distribution of magnitudes varies significantly across clusters.

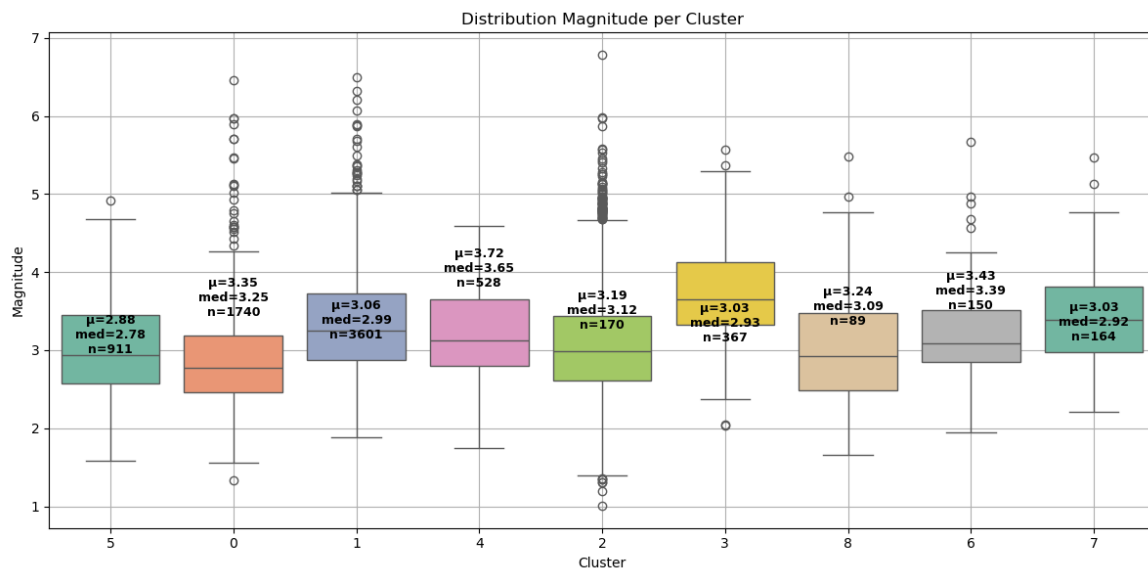


Figure 5. Magnitude distribution by cluster

Cluster 3 exhibits the highest median magnitude and a relatively wide interquartile range, indicating a concentration of moderate to strong seismic events. Clusters 1 and 7 also show higher median values compared to other clusters, suggesting zones with generally stronger earthquakes. In contrast, clusters such as 0, 2, and 4 display lower median magnitudes but include several outliers with high magnitude, which may represent sporadic yet impactful seismic activity. The presence of these outliers emphasizes the heterogeneous nature of seismic energy release within

certain clusters. These observations provide additional insight into the seismological patterns of each cluster and may reflect underlying tectonic variations in the Sulawesi region.

Table 5 provides details of the clusters identified using DBSCAN with parameters optimized according to the Calinski-Harabasz (CH) Index. The table summarizes each cluster's number of points, average depth (in kilometers), and average magnitude, offering insight into the distinct characteristics of each cluster.

Table 5. Cluster information from optimal parameters

Cluster No	Number of Points	Average Depth (km)	Average Magnitude
0	911	10.628979	2.877901
1	1740	60.466667	3.346258
2	3601	10.786726	3.057960
3	528	15.251894	3.724402
4	170	11.458824	3.192138
5	367	164.013624	3.028274
6	89	10.168539	3.238843
7	150	10.460000	3.433679
8	164	89.432927	3.032333

Cluster 0, comprising 911 points, has an average depth of 10.63 km and an average magnitude of 2.88. Cluster 1, the largest group with 1,740 points, had a significantly greater average depth of 60.47 km and a slightly higher average magnitude of 3.35. Cluster 2, with 3,601 points, is the most populated cluster, with an average depth of 10.79 km and an average magnitude of 3.06. Cluster 3, smaller with 528 points, has an average depth of 15.25 km and the highest average magnitude of 3.72. Cluster 4, comprising 170 points, has an average depth of 11.46 km and a moderate average magnitude of 3.19. Cluster 5, with 367 points, is the deepest, with an average depth of 164.01 km and an average magnitude of 3.03. Cluster 6, with 89 points, has an average depth of 10.17 km and an average magnitude of 3.24. Cluster 7, which contains 150 points, has an average depth of 10.46 km and a

relatively high average magnitude of 3.43. Finally, Cluster 8 includes 164 points with a depth average of 89.43 km and an average magnitude of 3.03.

Figure 6 presents the spatial distribution of earthquake clusters across Sulawesi as identified by the DBSCAN algorithm, using epsilon = 0.2 and minPts = 54. Each cluster is represented by a distinct color, highlighting spatially coherent groupings of seismic events based on their geographic coordinates and depth. The clustering results reveal meaningful patterns aligned with Sulawesi's complex tectonic framework. Dense clusters are observable in several areas, such as the Palu-Koro fault zone, the eastern arm of Sulawesi (e.g., Luwu Timur), and the northern region (e.g., Bolaang Mongondow), indicating concentrated seismicity likely associated with active faults or subduction interfaces

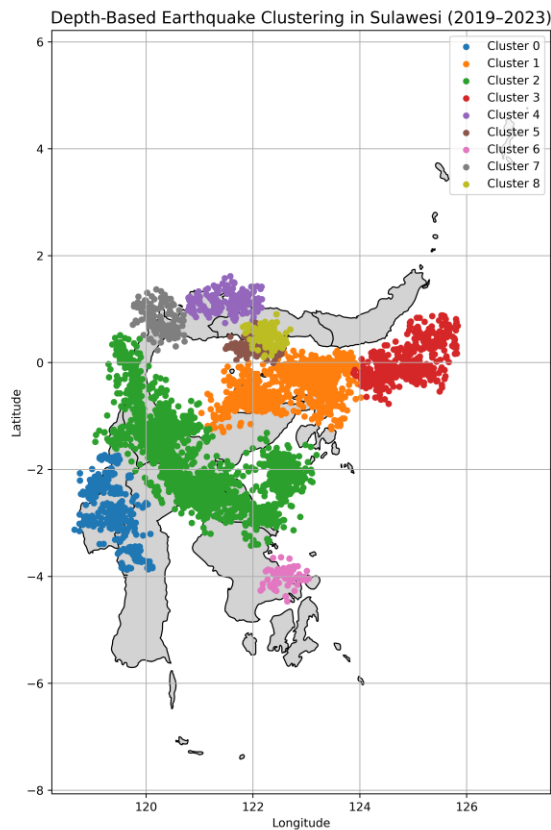


Figure 6. Distribution of Earthquake Clusters by Depth in Sulawesi Island (2019–2023)

These results demonstrate that DBSCAN effectively captures spatial groupings of seismic events while accounting for vertical distribution (depth), which is crucial for understanding tectonic segmentation and potential hazard zones. The visual evidence in Figure 6 reinforces the significance of using depth as a clustering feature and supports the regional interpretation of earthquake behavior for mitigation purposes.

In this study, a Hexbin plot was used to visualize the relationship between average depth and average magnitude across different earthquake clusters. Hexbin plots are a powerful data visualization tool that aggregates points into hexagonal bins, providing an effective means of identifying density patterns and relationships in large datasets. Unlike traditional scatter plots, Hexbin plots are particularly useful when dealing with large or overlapping data points because they reduce overplotting and reveal the underlying data distributions more clearly [33]. The use of a heatmap overlay in the Hexbin plot allows for additional analysis such as identifying areas with higher average magnitudes, and enhancing the interpretability of the data by combining spatial and intensity information into a single plot. This visualization approach is beneficial for exploring the correlation between variables like earthquake depth and magnitude, especially in the context of cluster analysis, where the density and characteristics of each cluster can provide insights into seismic activity trends.

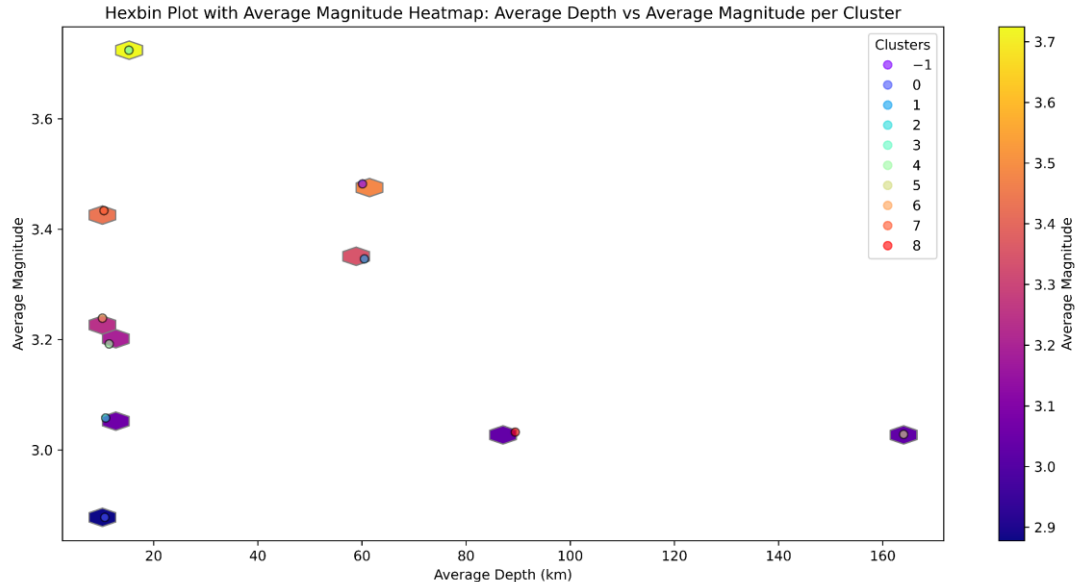


Figure 7. Hexbin plot average depth vs average magnitude

Based on the results of the Hexbin plot with average depth vs. average magnitude (Figure 7). The analysis indicates no strong linear correlation between earthquake depth and average magnitude. However, shallow earthquakes (0-20 km) tended to show higher magnitude variations, with some clusters (e.g., clusters 5 and 7) exhibiting the highest average magnitudes, reaching around 3.6-3.7. In contrast, deeper earthquakes (>100 km) tend to have lower average

magnitudes (~3.0-3.2). The clusters with shallow depths and high average magnitudes may indicate higher seismic hazard potential, possibly related to the brittle nature of the Earth's crust at these depths. Cluster -1, identified as noise, has a wide depth range but relatively lower average magnitudes, which may indicate weaker seismic events or poorly clustered data points. This analysis suggests the need for further studies to examine the characteristics of shallow

earthquakes and the source mechanisms within each cluster, especially those with higher magnitudes at shallow depths, to better understand the potential seismic risks in the Sulawesi area [34].

Cluster 0 was identified in the western region of Sulawesi, with 911 earthquake events and an average depth of 10.628979 km. Earthquakes in the study area are classified as shallow [35]. Figure 8 shows the location and shape of Cluster 0.

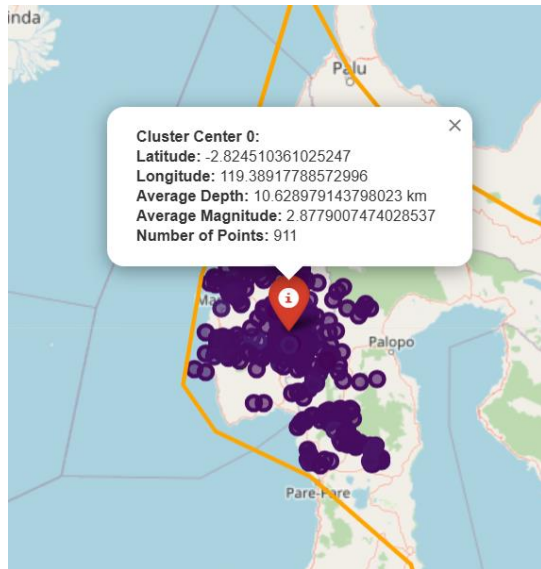


Figure 8. Visualization of cluster 0

The visualization shows that the center of cluster 0 is located in the Mamasa district. Earthquakes in this cluster tend to occur at shallow depths and have relatively low magnitudes. This indicates that although these earthquakes may not cause significant damage, they can still be felt by people in the vicinity of the epicenter. Shallow depths can result in stronger surface shaking, even when the shaking magnitude is small [6].

Cluster 1 include 1,740 earthquake events with an average depth of 60.47 km and an average magnitude of 3.35. Earthquakes in this cluster occurred at intermediate depths and had slightly higher magnitudes than those in Cluster 0. This intermediate depth may indicate that these earthquakes were occurring in a subduction zone or at depths below the Earth's crust, potentially causing different impacts than shallow earthquakes. Figure 9 shows the location and visualization of Cluster 1.

Figure 9 represents the clustering analysis of the earthquake data, focusing on the seismic characteristics within a defined region. The visualization highlights cluster center 1, at latitude: 2.037155 and longitude: 122.800466. This cluster has an average depth of approximately 86.34 km, indicating that most seismic events in this area occur at moderate depths. The average magnitude was 3.34, reflecting a predominance of moderate earthquakes. With 1740 events, this cluster offers valuable insights into the frequency and distribution of seismic activity in the vicinity of cluster

1. The color gradient represents depth variations, providing visual cues for identifying the depth patterns of seismic events across the cluster. This analysis facilitates an understanding of earthquake characteristics by depth and magnitude, which is essential for regional seismic risk assessment. The visualization of Cluster 1 shows that it was centered near Tomini Bay, with a higher concentration of earthquake events occurring in the ocean.

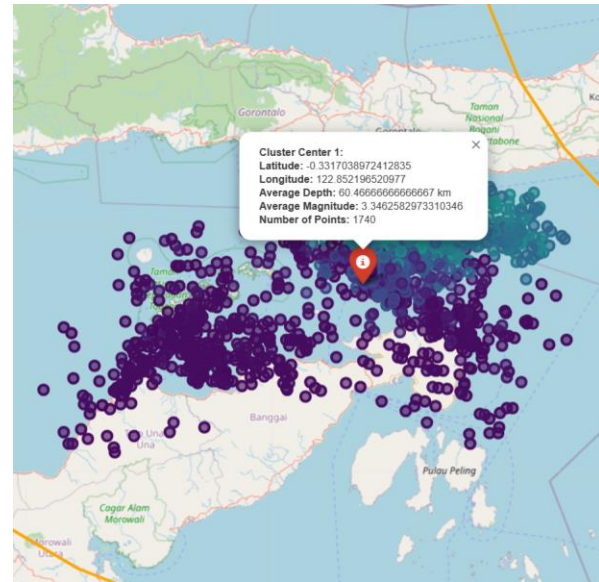


Figure 9. Visualization of cluster 1

Cluster 2 contained 3,601 earthquake events with an average depth of 10.79 km and an average magnitude of 3.06. This is the largest cluster in terms of the number of events, indicating that these earthquakes are relatively common at shallow depths with low to moderate magnitudes. The shallow depth is similar to that of Cluster 0, but the slightly higher magnitude indicates a potential for stronger shaking. Figure 10 shows the shape of the largest cluster, Cluster 2.

Figure 10 presents the spatial distribution of the Central Sulawesi during the 2019-2023 period. Clustering analysis identified a significant concentration of seismic activity along the central segment of the Palu Koro Fault, with an average depth of 10.7867 km and a mean magnitude of 3.05. The spatial distribution and characteristics of this cluster indicate that this segment of the Palu Koro Fault remains active and is under deformation. The results of this study contribute to a better understanding of the deformation mechanisms of the Palu Koro Fault and can be used to map zones with higher earthquake hazard potential.

Cluster 3 consisted of 528 earthquake events with an average depth of 15.25 km and an average magnitude of 3.72. Earthquakes in this cluster occurred at slightly greater depths and had higher magnitudes than those in the other clusters. These deeper earthquakes with higher magnitudes can cause more significant damage and be felt over a larger area [36]. The location and visualization of Cluster 3 are shown in Figure 11.

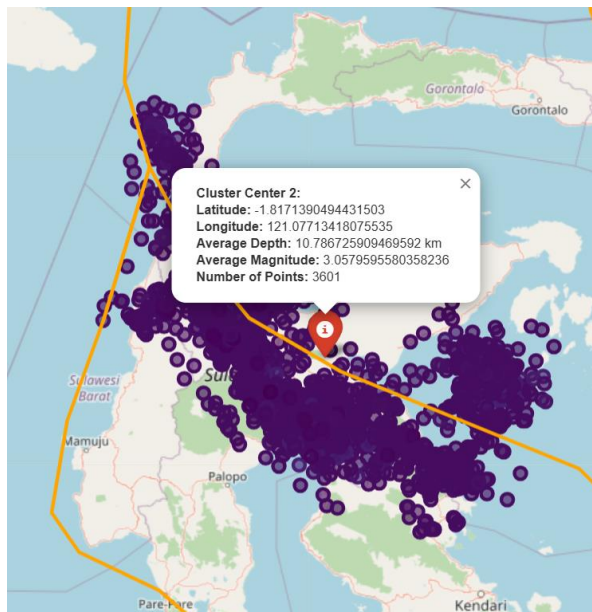


Figure 10. Visualization of cluster 2

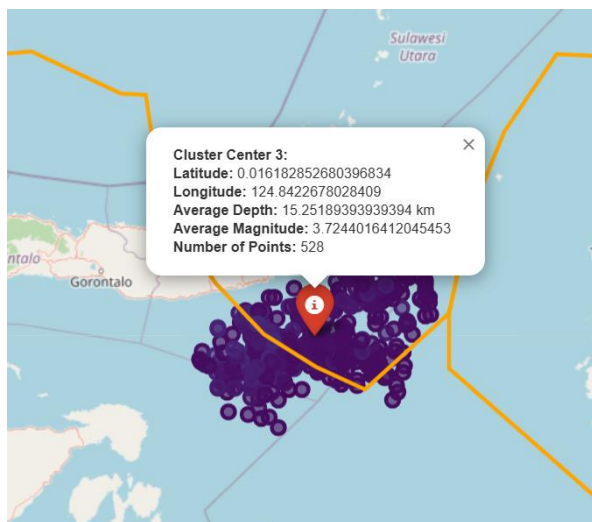


Figure 11. Visualization of cluster 3

Figure 11 illustrates the spatial distribution of earthquake epicenters in the southeastern waters near Kotamobagu. Clustering analysis revealed a significant concentration of seismic activity at 526 data points. The cluster exhibits an average depth of 15.25 km and a mean magnitude of 3.72. The spatial distribution and characteristics of this cluster indicate a high level of seismic activity in the region, likely associated with underlying active faults or subduction zones. Further research is needed to specifically identify the geological structures and mechanisms generating these earthquakes.

Cluster 4 included 170 earthquake events with an average depth of 11.46 km and an average magnitude of 3.19. Earthquakes in this cluster occurred at shallow to intermediate depths with relatively moderate magnitudes. This indicate that these earthquakes may have moderate damage potential, especially if they

occur near populated areas. The visualization of Cluster 4 is presented in Figure 12.

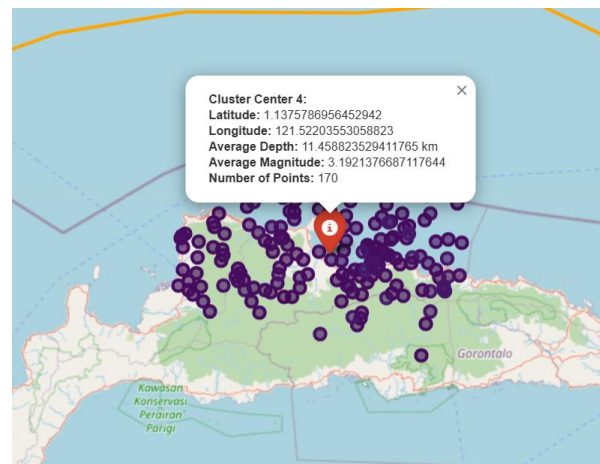


Figure 12. Visualization of cluster 4

The spatial distribution of the earthquake epicenters around Buol is presented in Figure 12. A cluster analysis identified a significant concentration of seismic activity with 170 data points, characterized by an average depth of 11.45 km and a mean magnitude of 3.19. The relatively shallow depth and moderate magnitude of these earthquakes indicate a higher likelihood of ground shaking and potential damage in the event of a larger earthquake. This highlights the importance of seismic hazard assessment and disaster preparedness in the Buol region.

Cluster 5 contained 367 earthquake events with an average depth of 164.01 km and an average magnitude of 3.03. This cluster is unique because earthquakes occur at very deep depths. Earthquakes at this depth usually occur in deep subduction zones, and although their magnitudes are not very high, their significant depth may cause them to be less perceptible at the surface. The visualization of Cluster 5 is presented in Figure 13.

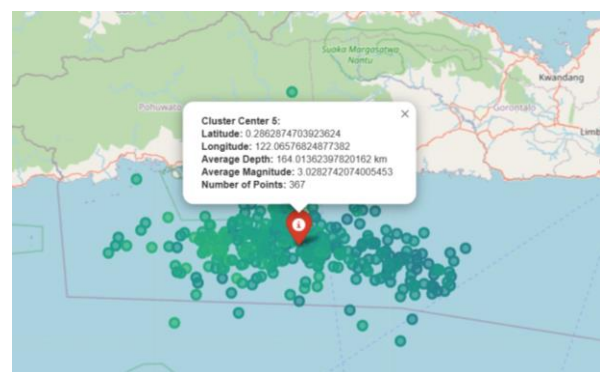


Figure 13. Visualization of cluster 5

Figure 13 presents a cluster of 367 earthquake epicenters located south of the Marisa region, with an average depth of 164 km and a mean magnitude of 3.02. Although the earthquakes in this cluster occurred at relatively deep depths, the potential for aftershocks with smaller magnitudes closer to the surface cannot be

ignored. The high seismic activity concentration in this region indicates a significant seismic hazard. These findings underscore the importance of implementing disaster mitigation strategies, particularly in coastal areas near the epicenter.

Cluster 6 consisted of 89 earthquake events with an average depth of 10.17 km and an average magnitude of 3.24. This is the smallest cluster in terms of the number of events, occurring at shallow depths with moderate magnitudes. Despite its smaller size, its shallow depth and moderate magnitude indicate that earthquakes in this cluster could be felt quite strongly on the surface. Figure 14 shows the visualization and location of Cluster 6.

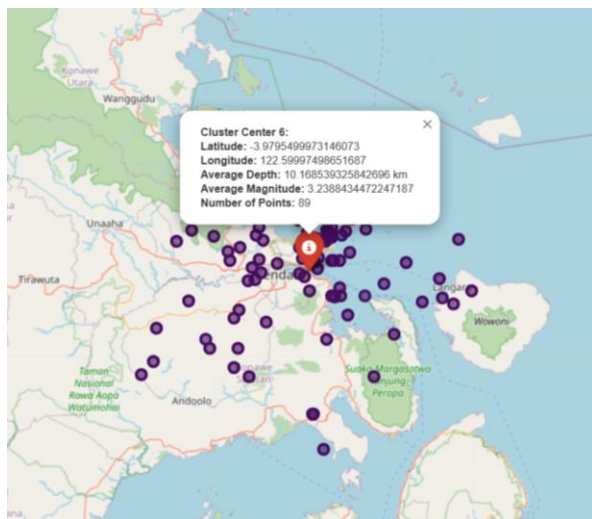


Figure 14. Visualization of cluster 6

Figure 14 shows the concentration of seismic activity around Kendari, with 89 earthquake epicenters. These earthquakes, with an average depth of 10.17 km and a mean magnitude of 3.24, indicate the activity of a local fault system, possibly a segment of the Palu-Koro fault. The relatively shallow depth increases the potential for ground shaking to affect populated areas. This analysis highlights the need for further studies on the characteristics of the Palu-Koro fault in the Kendari region to support disaster mitigation efforts, including the development of early warning systems and public education.

Cluster 7 included 150 earthquake points with an average depth of 10.46 km and an average magnitude of 3.43. Earthquakes in this cluster occur at shallow depths with relatively high magnitudes. This combination indicates the potential for significant damage, especially near the earthquake's epicenter [37]. Figure 15 illustrates the location and shape of Cluster 7.

Figure 15 illustrates the significant concentration of seismic activity in the Tolitoli region, with the epicenters clustered around the Sojol Mountain Nature Reserve. Cluster 7, exhibiting an average magnitude of 3.43, indicates a notable level of tectonic activity in the area. The average depth of these earthquakes, approximately 10.46 km, indicates that the seismic

source was located within the upper crust. The proximity of the earthquake cluster to the Sojol Mountain Nature Reserve indicate a possible correlation between seismic activity and the geological structures or volcanic activity associated with the mountain. The moderate average magnitude of 3.43 indicate the potential for minor to moderate damage to vulnerable structures in areas near the epicenters. To gain a deeper understanding of this phenomenon, further analyses are required, including detailed mapping of active faults, focal mechanism analysis, and a historical review of the earthquakes in the region. The results of these analyses can inform the development of comprehensive disaster mitigation plans that consider the seismic hazard potential of the Tolitoli area.

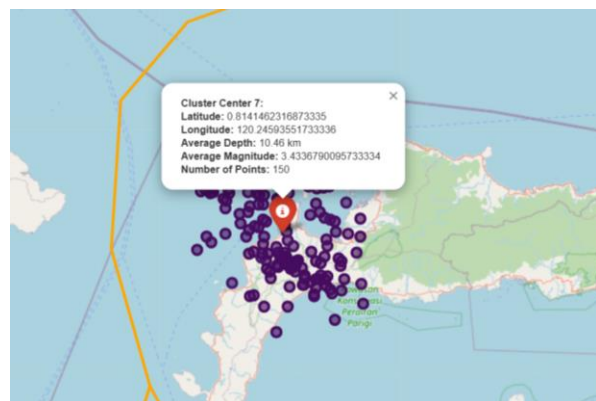


Figure 15. Visualization of cluster 7

Cluster 8 consisted of 164 earthquake events with an average depth of 89.43 km and an average magnitude of 3.03. Earthquakes in this cluster occur at considerable depths with moderate magnitudes. The significant depth indicate that these earthquakes may have occurred in subduction zones and were less likely to have been felt on the surface. Figure 16 shows cluster 8.

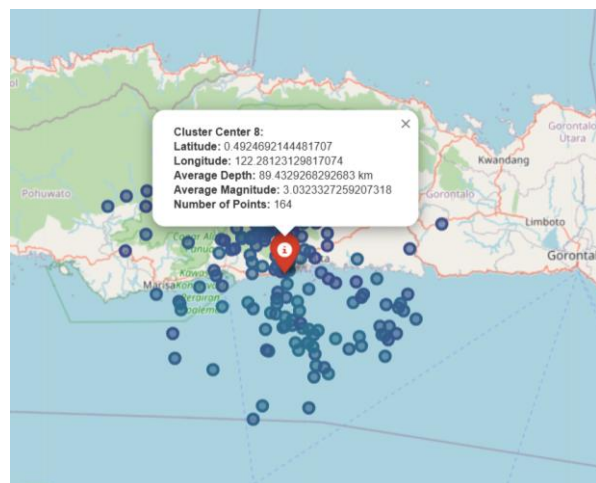


Figure 16. Visualization of Cluster 8

Figure 16 shows a high concentration of seismic activity in the Gorontalo region. The clustering of earthquakes with significant depth indicates the subduction of tectonic plates beneath the study area. This will make Gorontalo susceptible to larger

earthquakes in the future. Additionally, the potential for tsunamis should not be overlooked, considering Gorontalo's location within the Pacific Ring of Fire. Therefore, disaster mitigation efforts, such as the

construction of earthquake-resistant infrastructure, early warning systems, and public education, must be continuously improved.

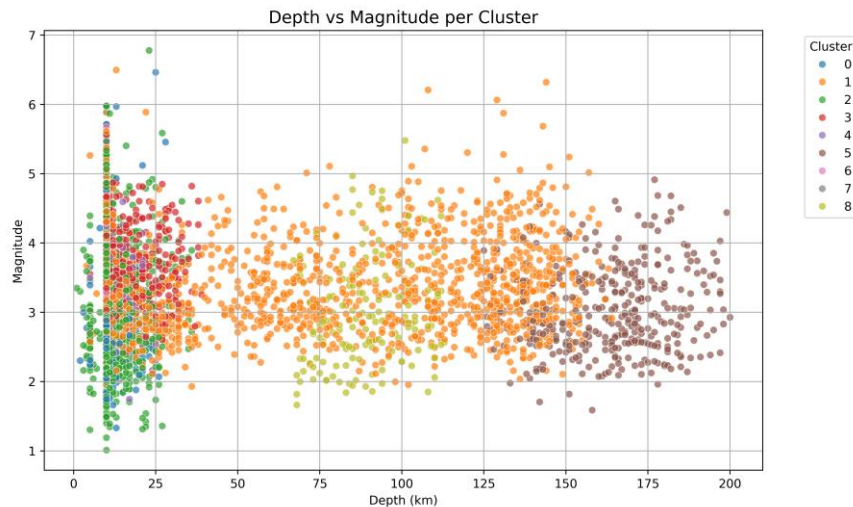


Figure 17. Depth vs Magnitude per Cluster ($\epsilon = 0.2$, $\text{min_samples} = 54$)

Figure 17 presents the depth–magnitude distribution of seismic events clustered using DBSCAN with optimized parameters ($\epsilon = 0.2$, $\text{min_samples} = 54$). Each color represents a distinct cluster. The scatter plot reveals that cluster separation primarily occurs along the depth axis, while magnitude values are more uniformly distributed across clusters. Shallow clusters (e.g., Cluster 0–4) dominate depths below 40 km, whereas deeper seismic activity is captured in Clusters 5–8, with depths reaching over 200 km. This visualization highlights the importance of depth as a clustering feature, supporting the decision to include it in the DBSCAN parameter optimization process.

From the formed clusters, we attempted to identify potential seismic gap areas. Areas with noise points, which are represented in black in the DBSCAN clustering results, are indicated in green. Several clusters surrounded these areas, indicating that they may have been seismic gap regions [38]. Figure 18 shows the potential seismic gap areas.

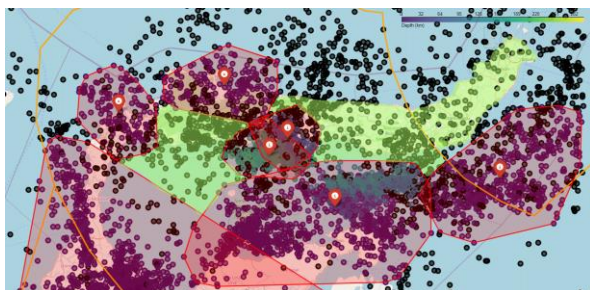


Figure 18. Potential seismic gap area

The green areas were situated between the formed earthquake clusters. The noise identified indicates that earthquakes frequently occur in these regions but do not fall into any cluster [39]. Overall, this visualization appears to be used to analyze earthquake patterns and

distributions across different regions, aiding in the interpretation of areas with higher seismic risk or specific patterns in earthquake data based on depth and earthquake density [40]. This output is highly valuable for disaster risk analysis and tectonic activity research in the area. Each cluster is represented by different colors and defines zones with similar seismic characteristics, while the points and depth scale indicate individual earthquake occurrences and their respective depths. This information is essential for understanding seismic behavior and assessing regional vulnerability to seismic events [41].

The clustering analysis has revealed diverse seismic characteristics across multiple regions in Sulawesi, each with distinct patterns in depth and magnitude. These differences are not only useful for categorizing seismic activity but also offer crucial insights for risk assessment [7]. Clusters with shallow depths and relatively high average magnitudes, such as those near Kotamobagu and Tolitoli, highlight regions that may pose greater seismic hazards due to their potential for strong ground shaking. Meanwhile, clusters located in deeper zones, like those south of Marisa and near Gorontalo, point to subduction-related seismicity that, although occurring deeper, could still have significant surface impact due to potential tsunami generation or energy transfer toward the crust [36], [42].

Moreover, the identification of noise points, especially those scattered around active clusters, sheds light on areas that may represent seismic gaps or zones of irregular seismic activity. These green-highlighted regions (as shown in Figure 18) appear to be surrounded by active clusters, suggesting that tectonic stress may be accumulating. This could indicate future rupture zones that are currently quiet but geologically active [36]. Integrating these clustering outcomes with geological

and tectonic data may enable more precise hazard zoning and early warning strategies [43]. Ultimately, this study not only maps the spatial and depth-related dynamics of seismicity in Sulawesi but also emphasizes the need for targeted monitoring in high-risk areas, contributing to a more comprehensive approach to earthquake preparedness and mitigation across the region [44].

The clustering results presented in this study offer practical implications for disaster mitigation planning by agencies such as BMKG, BPBD, and local governments. Shallow clusters with higher average magnitudes can be prioritized in hazard mapping, infrastructure risk assessment, and evacuation route planning. Meanwhile, noise points—interpreted as potential seismic gap areas—may guide BMKG in enhancing seismic monitoring networks, such as deploying additional sensors or refining early warning systems. BPBD and local authorities can utilize these insights to develop region-specific contingency plans, conduct public education campaigns, and allocate resources more effectively in high-risk zones. Integrating clustering outcomes with geological maps and spatial planning data will strengthen evidence-based mitigation strategies tailored to Sulawesi's tectonic complexity.

4. Conclusions

This study optimized the DBSCAN parameters for clustering Sulawesi earthquake data (2019-2023) by depth using a grid search method. The optimal parameters with an epsilon of 0.4 and min_samples of 54 achieved a Calinski-Harabasz (CH) index score of 956.63. Visualization revealed 9 distinct clusters, with shallow earthquakes (0-20 km) showing greater magnitude variation, and some clusters (e.g., clusters 5 and 7) exhibiting higher average magnitudes (~3.6-3.7). In contrast, deeper earthquakes (>100 km) had lower average magnitudes (~3.0-3.2). Shallow, high-magnitude clusters indicate a potential seismic hazard, likely tied to the brittle crust, while cluster -1, identified as noise, displayed a wide depth range but lower average magnitudes, hinting at weaker seismic events or dispersed data. This analysis emphasizes the significance of parameter optimization in DBSCAN for generating meaningful clusters and highlights the need for further studies on shallow, high-magnitude clusters to better assess seismic risks in Sulawesi. These clustering insights can support agencies such as BMKG, BPBD, and local governments in disaster mitigation planning. Shallow, high-magnitude clusters should be prioritized in hazard mapping and evacuation planning, while noise-identified seismic gap areas may guide sensor deployment and monitoring enhancement. Integrating these findings into regional contingency plans and public education efforts can improve preparedness and resilience in high-risk zones.

Acknowledgements

The authors gratefully acknowledge the support of the Meteorological, Climatological, and Geophysical Agency (BMKG), the Meteorology, Climatology, and Geophysics Polytechnic (STMKG), and Mercu Buana University for their essential contributions to this research. The technical support provided by PT. Mindotama Avia Teknik is also greatly appreciated. This research is dedicated to enhancing scientific understanding and earthquake risk mitigation efforts in Indonesia.

References

- [1] S. J. Hutchings and W. D. Mooney, "The Seismicity of Indonesia and Tectonic Implications," *Geochemistry, Geophysics, Geosystems*, vol. 22, no. 9, Sep. 2021, doi: 10.1029/2021GC009812.
- [2] A. Wahyu and Rushendra, "Klasterisasi Dampak Bencana Gempa Bumi Menggunakan Algoritma K-Means di Pulau Jawa," *JEPIN*, vol. 8, no. 1, 2022.
- [3] O. O. Wijaya and Rushendra, "Analysis of Sulawesi Earthquake Data from 2019 to 2023 using DBSCAN Clustering," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 4, pp. 454–465, 2024, doi: 10.29207/resti.v8i4.5819.
- [4] D. H. Natawidjaja et al., "The 2018 Mw7.5 Palu 'supershear' earthquake ruptures geological fault's multisegment separated by large bends: Results from integrating field measurements, LiDAR, swath bathymetry and seismic-reflection data," *Geophys J Int*, vol. 224, no. 2, pp. 985–1002, Feb. 2021, doi: 10.1093/gji/ggaa498.
- [5] M. Bariklana and A. Fauzan, "Implementation of the DBSCAN method for cluster mapping of earthquake spread location," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 2, pp. 867–878, Jun. 2023, doi: 10.30598/barekengvol17iss2pp0867-0878.
- [6] H. F. Yang and S. Yao, "Shallow destructive earthquakes," 2021, *Earthquake Science*. doi: 10.29382/EQS-2020-0072.
- [7] M. Ibrahim and B. Al-Bander, "An integrated approach for understanding global earthquake patterns and enhancing seismic risk assessment," *International Journal of Information Technology (Singapore)*, vol. 16, no. 4, pp. 2001–2014, Apr. 2024, doi: 10.1007/s41870-024-01778-1.
- [8] A. Jufriansah, Y. Pramudya, A. Khusnani, and S. Saputra, "Analysis of Earthquake Activity in Indonesia by Clustering Method," *Journal of Physics: Theories and Applications*, vol. 5, no. 2, p. 92, Sep. 2021, doi: 10.20961/jphys-theor-appl.v5i2.59133.
- [9] S. Sabermahani and A. W. Frederiksen, "Improved Earthquake Clustering Using a Density-Adaptive DBSCAN Algorithm: An Example from Iran," *Seismological Research Letters*, vol. 95, no. 2A, pp. 942–951, Dec. 2023, doi: 10.1785/0220220305.
- [10] C. Frohlich, *Deep Earthquakes*. Cambridge: Cambridge University Press, 2006. doi: DOI: 10.1017/CBO9781107297562.
- [11] J. A. Katili, "A review of the geotectonic theories and tectonic maps of Indonesia," *Earth Sci Rev*, vol. 7, no. 3, pp. 143–163, 1971, doi: https://doi.org/10.1016/0012-8252(71)90006-7.
- [12] T. Zheng, Q. Qiu, J. Lin, and X. Yang, "Raised Potential Earthquake and Tsunami Hazards at the North Sulawesi Subduction Zone after a Flurry of Major Seismicity," *SSRN Electronic Journal*, 2022, [Online]. Available: https://api.semanticscholar.org/CorpusID:251169199
- [13] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *ACM Transactions on Database Systems*, vol. 42, no. 3, Jul. 2017, doi: 10.1145/3068335.
- [14] S. S. Li, "An Improved DBSCAN Algorithm Based on the Neighbor Similarity and Fast Nearest Neighbor Query,"

- IEEE Access, vol. 8, pp. 47468–47476, 2020, doi: 10.1109/ACCESS.2020.2972034.
- [15] L. Yin, H. Hu, K. Li, G. Zheng, Y. Qu, and H. Chen, “Improvement of DBSCAN Algorithm Based on K-Dist Graph for Adaptive Determining Parameters,” *Electronics (Switzerland)*, vol. 12, no. 15, Aug. 2023, doi: 10.3390/electronics12153213.
- [16] B. Ma, C. Yang, A. Li, Y. Chi, and L. Chen, “A Faster DBSCAN Algorithm Based on Self-Adaptive Determination of Parameters,” *Procedia Computer Science*, vol. 221, pp. 113–120, 2023, doi: 10.1016/j.procs.2023.07.017.
- [17] R. R. A. Rahman and A. W. Wijayanto, “Pengelompokan data gempa bumi menggunakan algoritma DBSCAN,” *Jurnal Meteorologi dan Geofisika*, vol. 22, no. 1, p. 31, Oct. 2021, doi: 10.31172/jmg.v22i1.738.
- [18] A. Starczewski and A. Cader, “Grid-Based Approach to Determining Parameters of the DBSCAN Algorithm,” in *Artificial Intelligence and Soft Computing*, L. Rutkowski, R. Scherer, M. Korytkowski, W. Pedrycz, R. Tadeusiewicz, and J. M. Zurada, Eds., Cham: Springer International Publishing, 2020, pp. 555–565.
- [19] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” 1996. [Online]. Available: www.aaii.org
- [20] L. Wang, L. Mingyang, H. Xuming, and K. and Zheng, “An improved density-based spatial clustering of application with noise,” *International Journal of Computers and Applications*, vol. 40, no. 3, pp. 1–7, Jul. 2018, doi: 10.1080/1206212X.2018.1424103.
- [21] S. Scitovski, “A density-based clustering algorithm for earthquake zoning,” *Comput Geosci*, vol. 110, pp. 90–95, 2018, doi: <https://doi.org/10.1016/j.cageo.2017.08.014>.
- [22] M. Ulu and Y. S. Türkan, “Cluster Analysis and Comparative Study of Different Clustering Performance and Validity Indices,” in *Industrial Engineering in the Industry 4.0 Era*, N. M. Durakbasa and M. G. Gençyılmaz, Eds., Cham: Springer Nature Switzerland, 2024, pp. 33–45.
- [23] A. Dudek, “Silhouette Index as Clustering Evaluation Tool,” in *Classification and Data Analysis*, K. Jajuga, J. Batóg, and M. Walesiak, Eds., Cham: Springer International Publishing, 2020, pp. 19–33.
- [24] D. L. Davies and D. W. Bouldin, “A Cluster Separation Measure,” *IEEE Trans Pattern Anal Mach Intell*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979.
- [25] T. Caliński and J. Harabasz, “A dendrite method for cluster analysis,” *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, Jan. 1974, doi: 10.1080/03610927408827101.
- [26] G. Liu, “A New Index for Clustering Evaluation Based on Density Estimation,” Jul. 2022, [Online]. Available: <http://arxiv.org/abs/2207.01294>
- [27] L. Gong, T. Yamamoto, and T. Morikawa, “Identification of activity stop locations in GPS trajectories by DBSCAN-TE method combined with support vector machines,” in *Transportation Research Procedia*, Elsevier B.V., 2018, pp. 146–154. doi: 10.1016/j.trpro.2018.10.028.
- [28] Z. Zhan, “Mechanisms and Implications of Deep Earthquakes,” 2019, doi: 10.1146/annurev-earth-053018.
- [29] K. Vahldiek and F. Klawonn, “Cluster-Centered Visualization Techniques for Fuzzy Clustering Results to Judge Single Clusters,” *Applied Sciences (Switzerland)*, vol. 14, no. 3, Feb. 2024, doi: 10.3390/app14031102.
- [30] T. Kiyota, H. Furuichi, R. F. Hidayat, N. Tada, and H. Nawir, “Overview of long-distance flow-slide caused by the 2018 Sulawesi earthquake, Indonesia,” *Soils and Foundations*, vol. 60, no. 3, pp. 722–735, Jun. 2020, doi: 10.1016/j.sandf.2020.03.015.
- [31] S. Cesca, “Seiscloud, a tool for density-based seismicity clustering and visualization,” *J Seismol*, vol. 24, no. 3, pp. 443–457, Jun. 2020, doi: 10.1007/s10950-020-09921-8.
- [32] S. Pasari, A. V. H. Simanjuntak, Neha, and Y. Sharma, “Nowcasting earthquakes in Sulawesi Island, Indonesia,” Dec. 01, 2021, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1186/s40562-021-00197-5.
- [33] D. B. Carr, R. J. Littlefield, W. L. Nicholson, and J. S. Littlefield, “Scatterplot Matrix Techniques for Large N,” *J Am Stat Assoc*, vol. 82, no. 398, pp. 424–436, Jun. 1987, doi: 10.1080/01621459.1987.10478445.
- [34] B. Maruli Siahaan and A. Roma Rio, “Agglomerative Clustering of 2022 Earthquakes in North Sulawesi, Indonesia,” *Buana Information Technology and Computer Sciences (BIT and CS)*, vol. 4, no. 2, p. 77, 2023, [Online]. Available: <https://repogempa.bmkg.go.id/>
- [35] A. Kusmiran, Minarti, M. F. I. Massinai, A. Zarkasi, A. A. Maharani, and R. Desiani, “Klasifikasi Kedalaman Kejadian Gempa Menggunakan Algoritma K-Means Clustering: Studi Kasus Kejadian Gempa Di Sulawesi,” *JFT: Jurnal Fisika dan Terapannya*, vol. 9, no. 2, pp. 79–88, Dec. 2022, doi: 10.24252/jft.v9i2.29198.
- [36] T. Zheng, Q. Qiu, J. Lin, and X. Yang, “Raised potential earthquake and tsunami hazards at the North Sulawesi subduction zone after a flurry of major seismicity,” *Mar Pet Geol*, vol. 148, Feb. 2023, doi: 10.1016/j.marpetgeo.2022.106024.
- [37] A. Kusmiran, “Clustering and Risk Analysis of The Earthquake in Sulawesi Using Mini Batch K-Means, K-Medoids, and Maximum Likelihood Method,” *Journal of Islamic Science and Technology*, vol. 9, no. 1, 2023, doi: 10.22373/ekw.v9i1.13027.
- [38] Y. Rong, D. D. Jackson, and Y. Y. Kagan, “Seismic gaps and earthquakes,” *J Geophys Res Solid Earth*, vol. 108, no. B10, Oct. 2003, doi: 10.1029/2002jb002334.
- [39] W. R. McCann, S. P. Nishenko, L. R. Sykes, and J. Krause, “Seismic Gaps and Plate Tectonics: Seismic Potential for Major Boundaries,” in *Earthquake Prediction and Seismicity Patterns*, M. Wyss, Ed., Basel: Birkhäuser Basel, 1979, pp. 1082–1147. doi: 10.1007/978-3-0348-6430-5_2.
- [40] R. Guo, Y. Zheng, and J. Xu, “Stress modulation of the seismic gap between the 2008 Ms 8.0 Wenchuan earthquake and the 2013 Ms 7.0 Lushan earthquake and implications for seismic hazard,” *Geophys J Int*, vol. 221, no. 3, pp. 2113–2125, 2020, doi: 10.1093/GJI/GGAA143.
- [41] W. Thatcher, “Earthquake recurrence and risk assessment in circum-Pacific seismic gaps,” *Nature*, vol. 341, no. 6241, pp. 432–434, 1989, doi: 10.1038/341432a0.
- [42] D. L. A. Primananda and M. Muhajir, “Modelling earthquake disaster damage due depth of epicenter and magnitude using spatial regression,” *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 3, pp. 1221–1234, Sep. 2023, doi: 10.30598/barekengvol17iss3pp1221-1234.
- [43] T. Rahayu, Z. Nasution, Roesyanto, and D. Karnawati, “Regional zonation based on seismic vulnerability using local site effect analysis and potential damage to the city of Medan (North Sumatra, Indonesia) due to earthquake,” *Geoenvironmental Disasters*, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40677-022-00227-0.
- [44] M. Yuliyanti et al., “Earthquake Disaster Mitigation Strategy: Lessons Learned from Indonesia,” Jan. 2024.