Available online: https://doi.org/10.32665/james.v8i2.5170

Journal of Mathematics Education and Science

p-ISSN: 2621-1203 VOL. 8 & NO. 2 & (2025) : 162-174 e-ISSN: 2621-1211

CLUSTERING OF NATURAL DISASTER-PRONE AREAS IN EAST JAVA PROVINCE USING FUZZY C-MEANS METHOD

Deni Setiawan(1*), **Kresna Oktafianto**(2)
[1]Department of Mathematics, Universitas PGRI Ronggolawe, Indonesia
[2]Department of Mathematics, Universitas PGRI Ronggolawe, Indonesia
*Corresponding Author. E-mail: denys3009@gmail.com

ARTICLE INFO

Article History:
Received: 17-Jul. 2025
Revised: 08-Sep. 2020
Accepted: 16-Sep. 2020

Keywords:
Clustering, East Java Province, Fuzzy C-Means, Natural Disasters, Prone Areas

ABSTRACT

Bencana alam merupakan salah satu permasalahan yang sering terjadi di Indonesia. Menurut Data Informasi Bencana Indonesia (DIBI) tahun 2020-2024, Provinsi Jawa Timur menduduki peringkat ketiga dengan frekuensi kejadian bencana alam tertinggi di Pulau Jawa. Pemetaan daerah rawan bencana secara sistematis sangat diperlukan untuk mendukung upaya mitigasi dan penanganan yang lebih efektif. Penelitian ini bertujuan untuk mengelompokkan 38 kabupaten/kota di Provinsi Jawa Timur berdasarkan kerawanan terhadap bencana alam. Metode yang digunakan adalah *Fuzzy C-Means* untuk proses pengelompokan, serta *Silhouette Coefficient* sebagai alat evaluasi kualitas *cluster*. Data yang digunakan berupa data sekunder dengan indikator jumlah kejadian, jumlah korban, dan jumlah kerusakan akibat bencana banjir, tanah longsor, cuaca ekstrem, kekeringan, gempa bumi, erupsi gunung api, serta kebakaran hutan dan lahan. Hasil pengelompokkan menghasilkan tiga *cluster*, yaitu daerah dengan tingkat kerawanan tinggi, sedang, dan rendah. Evaluasi hasil *clustering* dilakukan menggunakan *Silhouette Coefficient* dengan nilai sebesar 0,2807, yang menunjukkan bahwa hasil *clustering* berada pada kategori cukup baik serta mengindikasikan adanya keterbatasan dalam pemisahan *cluster* akibat penggunaan indikator yang masih terbatas dan adanya tumpang tindih karakteristik antar wilayah. Meskipun demikian, hasil penelitian ini dapat memberikan kontribusi sebagai dasar pertimbangan dalam perumusan kebijakan mitigasi bencana, terutama untuk menentukan prioritas intervensi dan memperkuat kesiapsiagaan di wilayah dengan tingkat kerawanan bencana tinggi.

Natural disasters are one of the most common problems in Indonesia. According to Data Informasi Bencana Indonesia (DIBI) from 2020 to 2024, East Java Province ranks third in terms of the frequency of natural disasters on the island of Java. Systematic mapping of disaster-prone areas is essential to support more effective mitigation and management efforts. This study aims to group 38 districts/cities in East Java Province based on their vulnerability to natural disasters. The methods used are Fuzzy C-Means for the grouping process and Silhouette Coefficient as a tool for evaluating cluster quality. The data used is secondary data with indicators of the number of incidents, number of victims, and amount of damage caused by floods, landslides, extreme weather, drought, earthquakes, volcanic eruptions, and forest and land fires. The clustering results produced three clusters, namely areas with high, medium, and low vulnerability levels. The clustering results were evaluated using the Silhouette Coefficient with a value of 0.2807, which indicates that the clustering results are in the fairly good

SCAN ME

category and indicate limitations in cluster separation due to the use of limited indicators and overlapping characteristics between regions. Nevertheless, the results of this study can contribute as a basis for consideration in the formulation of disaster mitigation policies, especially in determining intervention priorities and strengthening preparedness in areas with high disaster vulnerability.

How to Cite: Setiawan, D., & Oktafianto, K. (2025). Clustering Of Natural Disaster-Prone Areas In East Java Province Using Fuzzy C-Means Method. *Journal of Mathematics Education and Science, 8(2),* 162-174. https://doi.org/10.32665/james.v8i2.5170

INTRODUCTION

According to Law No. 24 of 2007 on Disaster Management, a disaster is defined as a series of events that disrupts and threatens the lives and livelihoods of surrounding communities. These events are caused by natural, non-natural, or human factors and result in loss of human life, environmental damage, loss of property, and psychological impact (DPR RI, 2007). Indonesia is a country prone to natural disasters (Rahmi, 2024). This is due to geographical, geological, and climatological factors (Pascapurnama et al., 2018). Geologically, it is located in the Pacific Ring of Fire, home to 127 active volcanoes, and where three major tectonic plates converge, frequently causing earthquakes and tsunamis (Prasetyo et al., 2024).

The varied topography, ranging from mountains to lowlands, increases the risk of landslides and flooding (Septian et al., 2020). Climate change also exacerbates the frequency of extreme weather events such as heavy rains and droughts (Faruf et al., 2023). Anthropogenic factors such as land use change and urbanization further worsen disaster vulnerability (Marto et al., 2024). These conditions cause Indonesia to be hit by natural disasters that occur repeatedly every year.

According to data from the Disaster Information Database (DIBI) managed by the National Disaster Management Agency (BNPB), between 2020 and 2024, there were 20.475 natural disasters in Indonesia, with the highest number of incidents occurring on Java Island, totaling 9.131 incidents (Pusdatinkom BNPB, 2025).

East Java Province was selected as the study area because it has a high vulnerability to various types of natural disasters. Based on DIBI data from 2020 to 2024, East Java Province ranks third among the provinces with the highest frequency of natural disasters on Java Island, with 1.312 incidents. Second is Central Java Province with a total of 2.898 incidents, and first place is West Java Province with 4.171 disaster incidents (Pusdatinkom BNPB, 2025). In addition, East Java has diverse demographic and environmental characteristics, ranging from densely populated areas to agricultural and coastal areas that are vulnerable to climate change (Anggraeni, 2016). This diversity causes variations in the level of disaster vulnerability in each district/city, requiring mapping of disaster-prone areas that can handle the complexity of data that can be used as a first step in disaster risk mitigation and management.

The identification and classification of disaster-prone areas in Indonesia has been carried out using various methods, such as manual analysis using historical data or geographic-based mapping (Aulia Putri, 2025). However, these approaches are often inefficient in providing fast and accurate results, especially in classifying areas based on complex vulnerability levels.

The selection of the Fuzzy C-Means method is based on several fundamental considerations that make it superior to conventional clustering methods such as K-Means. First, the characteristics of natural disasters in East Java are multidimensional and overlapping. For example, a region like Pacitan Regency may have a high risk of both flooding and landslides due to its hilly topography and intense river flow (BNPB, 2021). The Fuzzy C-Means method can handle this complexity using fuzzy membership degrees, where an area can be assigned to multiple clusters with different membership values (Firdaus et al., 2021). This is not possible with non-fuzzy methods like K-Means, which force each data point into a single cluster. Second, the Fuzzy C-Means method offers flexibility in accommodating data uncertainty (Wijayanti et al., 2021).

Natural disaster data is often incomplete or contains noise due to variations in reporting across regions. The Fuzzy C-Means method can address this issue through an objective function that minimizes intra-cluster variance by considering fuzzy membership weights (Bezdek, 1981) in (Irabawati et al., 2016). Thus, the clustering results are more stable despite data imperfections. Third, from a policy perspective, the Fuzzy C-Means method's output in the form of membership values enables the government to prioritize mitigation based on disaster vulnerability levels. For example, areas with a membership value of 0,8 for the "high disaster vulnerability" cluster can be prioritized for disaster mitigation over areas with a membership value of 0,5.

A number of previous studies have applied clustering methods in the context of disasters. Yulianto et al. (2023) clustered provinces in Indonesia based on the number of disaster events, Nabilla Audy et al. (2024) studied disaster-prone areas in West Java using the Fuzzy C-Means method, while Ifadah et al. (2022) clustered flood-prone areas in East Java Province. Although these studies demonstrate the relevance of clustering methods, most still use limited indicators or focus on only one type of disaster, resulting in a less comprehensive representation of regional vulnerability levels.

To address these limitations, this study analyzes disaster vulnerability in 38 districts/cities in East Java Province for seven types of disasters, namely floods, landslides, extreme weather, drought, earthquakes, volcanic eruptions, and forest and land fires, covering indicators such as the number of incidents, number of victims, and amount of damage. The Fuzzy C-Means method was used in the clustering process, while the Silhouette Coefficient was applied to evaluate the quality and validity of the results. Through this multidimensional approach, the study is expected to provide a more holistic classification of disaster-prone areas, while contributing to academic development in the field of data-based spatial analysis and serving as a practical reference for policymakers in determining mitigation priorities and strengthening disaster preparedness.

METHOD

This research is quantitative research focused on the process of clustering regions based on their vulnerability to natural disasters. This process is carried out through several stages of data analysis that are arranged systematically. The steps in this research are explained as follows:

1. Data Source and Preparation
   This study uses secondary data obtained from the Indonesian Disaster Information Database (DIBI) sourced from the National Disaster Management Agency (BNPB) database for the period 2020–2024. The five-year span was chosen because it represents the most recent and complete disaster records, which are sufficient to describe the latest disaster trends while avoiding bias from older data that may not be relevant to current conditions. The data consists of 21 variables covering seven types of disasters (floods, landslides, extreme weather, droughts, earthquakes, volcanic eruptions, and forest and land fires), with three main indicators: number of incidents, number of victims, and amount of damage. Before analysis, the data was cleaned and normalized using the Min-Max method to ensure consistency in scale between variables.
2. Determination of Number of Clusters
   The number of clusters was set at three, referring to the classification of disaster risk levels (high, medium, low) as regulated in Head of BNPB Regulation No. 02/2012 on General Guidelines for Disaster Risk Assessment. This determination is also in line with disaster

management practices in Indonesia. The validity of the three-cluster structure was then tested using the Silhouette Coefficient.

3. Clustering using Fuzzy C-Means
   a. The data to be clustered is entered into matrix X, which is an n × m matrix ($\boldsymbol{s}$ = number of data samples, $\boldsymbol{p}$ = attributes of each data). $\boldsymbol{x_{ij}}$ = data sample i (i = 1, 2, ..., s), attribute $\boldsymbol{j}$ (j = 1, 2, ..., p).
   b. The next step is to determine several inputs needed in the Fuzzy C-Means calculation, namely:
      - The number of clusters (c) is the number of clusters that will be formed.
      - The exponent/weight (w) is the exponent value.
      - The maximum iteration (MaxIter) is the limit of repetition or looping. Looping will stop when the maximum iteration value is reached.
      - The minimum error (ξ) is the threshold value that causes the looping to end once the desired error value is obtained.
      - The initial objective function $(J_0 = 0)$.
      - Initial iteration $(t = 1)$, iteration is a specific property of an algorithm or computer program where a sequence or more of algorithmic steps are performed repeatedly. The initial iteration is the iteration number at which the program will begin.
   c. Generate random numbers or the initial partition matrix according to Equation (1).
   $$\sum_{k=1}^{c} \mu_{ik} = 1 \quad (1)$$
   Where $\boldsymbol{c}$ is the number of clusters, $\boldsymbol{k}$ is the cluster index from 1 to c, $\mu_{\boldsymbol{ik}}$ is the membership degree in row i and column k, and the sum of all membership degrees of data i to all clusters must be equal to 1.
   d. Calculate the cluster center or centroid according to Equation (2).
   $$V_{kj} = \frac{\sum_{k=1}^{N} (\mu_{ik})^{w} * x_k}{\sum_{k=1}^{N} (\mu_{ik})^{w}} \quad (2)$$
   Where $\boldsymbol{v_{kj}}$ is the cluster center, $\boldsymbol{w}$ is the weight, $\boldsymbol{N}$ is the number of data points in the dataset, and $\boldsymbol{x_k}$ is the kth input data point.
   e. Calculate the objective function in iteration t according to Equation (3).
   $$J_{(t)} = \sum_{i=1}^{N} \sum_{k=1}^{c} \left(\left[\sum_{j=1}^{m} \left(x_{ij} - v_{kj} \right)^2 \right] (\mu_{ik})^{w} \right) \quad (3)$$
   Where $\boldsymbol{m}$ is the number of attributes for each data point. Meanwhile, the other notations refer to the previous step.
   f. Calculate the change in the partition matrix according to Equation (4).
   $$\mu_{ik} = \frac{\left[\sum_{j=1}^{m} \left(x_{ij} - v_{kj} \right)^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c} \left[\sum_{j=1}^{m} \left(x_{ij} - v_{kj} \right)^2 \right]^{\frac{-1}{w-1}}} \quad (4)$$
   All notation used at this stage is the same as the notation explained in the previous step.
   g. Check the stopping condition:
   The iteration will continue to repeat if certain values or conditions have not been met. These conditions are as follows:
   If $\left(\left| J_{(t)} - J_{(t-1)} \right| < \xi \right)$or (t > MaxIter), then stop;
   If not: $t = t + 1$, repeat step d or calculate the cluster center.
4. Evaluate clusters using the Silhouette Coefficient method

a. Calculate the average distance between a data point and other data points in the same cluster (intra-cluster distance) according to Equation (5).

$$a(i) = \frac{1}{|A| - 1} \sum\nolimits_{j \in A, j \neq i} d(i,j) \tag{5}$$

Where $\boldsymbol{a(i)}$ is the average difference between object (i) and all other objects in cluster A, $\boldsymbol{d(i,j)}$ is the distance between data i and data j, and **A** is the cluster.

b. Calculate the average minimum distance between a data point and other clusters according to Equation (6).

$$d(i,C) = \frac{1}{|C|} \sum\nolimits_{j \in C} d(i,j) \tag{6}$$

Where $\boldsymbol{d(i,C)}$ is the average difference between object i and all other objects in C and $\boldsymbol{C}$ is another cluster other than cluster A.

c. Calculate the distance to the nearest neighboring cluster of the object that achieves the minimum value according to Equation (7).

$$b(i) = min_{C \neq A} d(i,j) \tag{7}$$

Where $\boldsymbol{b(i)}$ is the neighboring cluster of object (i) that reaches the minimum value.

d. Calculate the Silhouette Coefficient for each data point according to Equation (8).

$$s(i) = \frac{(bi - ai)}{\max\ (ai, bi)} \tag{8}$$

Where $\boldsymbol{s(i)}$ is the Silhouette Coefficient Value. Meanwhile, the other notations refer to the previous step.

To evaluate a Silhouette Coefficient value, refer to Table 1 created by Kaufman and Rousseeuw (1987) in (Rahmawati et al., 2024) as follows:

Table 1. Interpretation of Silhouette Coefficient Values

Silhouette Coefficient Value	Cluster Interpretation
0,71 – 1,00	Very Good or Strong
0,51 – 0,70	Good
0,26 – 0,50	Fairly Good or Weak
≤ 0,25	Bad or cannot be considered a cluster

5. Visualization of Clustering Results

The visualization results are displayed in the form of a clustering map using Microsoft Excel, by creating a map from the Map Chart feature. Each district/city is colored according to the cluster with the highest membership value. The resulting map displays regional clustering based on disaster vulnerability levels classified using three different colors, as follows:

a. Cluster 1 is green, indicating districts/cities with low disaster vulnerability;
b. Cluster 2 is orange, indicating districts/cities with moderate disaster vulnerability;
c. Cluster 3 is red, indicating districts/cities with high disaster vulnerability.

RESULTS

This section presents the results of data analysis to group natural disaster-prone areas in East Java Province. The data used includes 21 variables consisting of the number of incidents, number of victims, and amount of damage from seven types of natural disasters in 38 districts/cities in East Java Province. Table 2 presents some of the data as an example of the dataset structure used in the clustering process.

Table 2. Natural Disaster Data in East Java Province (DIBI BNPB)

No	Kabupaten/Kota	Number of Incidents					⋯	Number of Damages			
		V_1	V_2	V_3	V_4	V_5	⋯	V_{18}	V_{19}	V_{20}	V_{21}
1	Bangkalan	12	0	5	0	0	⋯	0	0	0	0
2	Banyuwangi	23	2	18	1	0	⋯	90	0	0	0
3	Blitar	8	11	28	1	1	⋯	29	2109	0	0
4	Bojonegoro	35	4	14	5	0	⋯	91	0	0	0
5	Bondowoso	8	0	28	3	0	⋯	57	2	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮
38	Tulungagung	5	3	9	1	0	⋯	18	81	0	0

Based on Table 2, variables V_1 to V_7 describe the number of incidents, V_8 to V_{14} indicate the number of victims, and V_{15} to V_{21} indicate the number of damages from each of the seven types of natural disasters, namely floods, landslides, extreme weather, drought, earthquakes, volcanic eruptions, and forest and land fires.

Natural Disaster Data Normalization

This stage is carried out to equalize the scale between variables so that no attribute is dominant due to differences in value ranges. This study uses Min-Max normalization, which converts each value into the range [0, 1]. The results of natural disaster data normalization can be seen in Table 3.

Table 3. Results of Natural Disaster Data Normalization

No	Kabupaten/Kota	x'									
		V_1	V_2	V_3	V_4	V_5	⋯	V_{18}	V_{19}	V_{20}	V_{21}
1	Bangkalan	0,2553	0	0,1163	0	0	⋯	0	0	0	0
2	Banyuwangi	0,4894	0,0606	0,4186	0,1429	0	⋯	0,1064	0	0	0
3	Blitar	0,1702	0,3333	0,6512	0,1429	0,5	⋯	0,0343	0,1721	0	0
4	Bojonegoro	0,7447	0,1212	0,3256	0,7143	0	⋯	0,1076	0	0	0
5	Bondowoso	0,1702	0	0,6512	0,4286	0	⋯	0,0674	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮
38	Tulungagung	0,1064	0,0909	0,2093	0,1429	0	⋯	0,0213	0,0066	0	0

Clustering Using the Fuzzy C-Means Method

1. Insert the data to be clustered into matrix X, which contains the normalized values of all variables.

$$X = \begin{bmatrix} 0{,}2553 & 0 & 0{,}1163 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0{,}4894 & 0{,}0606 & 0{,}4186 & 0{,}1429 & 0 & 0 & 0{,}0588 & \dots & 0 \\ 0{,}1702 & 0{,}3333 & 0{,}6512 & 0{,}1429 & 0{,}5000 & 0 & 0 & \dots & 0 \\ 0{,}7447 & 0{,}1212 & 0{,}3256 & 0{,}7143 & 0 & 0 & 0{,}0882 & \dots & 0 \\ 0{,}1702 & 0 & 0{,}6512 & 0{,}4286 & 0 & 0 & 0{,}2353 & \dots & 1 \\ 0{,}9787 & 0 & 0{,}1628 & 0{,}1429 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0{,}3636 & 0{,}8372 & 0{,}4286 & 0{,}5000 & 0 & 0{,}0588 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0{,}1064 & 0{,}0909 & 0{,}2093 & 0{,}1429 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

2. Determine the initial parameters required in the clustering process. These parameters consist of:
   a. Number of clusters (c) = 3
   b. Exponent/weight (w) = 2
   c. Maximum iteration (MaxIter) = 100
   d. Minimum error (ξ) = 10^{-5}

e. Initial objective function (J) = 0
f. Initial iteration (t) = 1

3. Generating random numbers as elements of the initial partition matrix U_{38x3}

$$U = \begin{bmatrix} 0,3668 & 0,3364 & 0,2967 \\ 0,4574 & 0,3014 & 0,2411 \\ 0,4916 & 0,0389 & 0,4695 \\ 0,2861 & 0,3254 & 0,3885 \\ 0,1228 & 0,4199 & 0,4572 \\ \vdots & \vdots & \vdots \\ 0,3148 & 0,6843 & 0,0009 \end{bmatrix}$$

4. Calculating the cluster center
Using Equation (2), the center of cluster V_{3X21} formed in the first iteration is obtained as:

$$V_{kj} = \begin{bmatrix} 0,2034 & 0,1135 & 0,2794 & 0,0982 & 0,0582 & 0,0049 & 0,0573 & \cdots & 0,0033 \\ 0,3288 & 0,1473 & 0,3360 & 0,1820 & 0,0991 & 0,0758 & 0,0898 & \cdots & 0,0321 \\ 0,3398 & 0,1103 & 0,2437 & 0,1424 & 0,0378 & 0,0067 & 0,0787 & \cdots & 0,0331 \end{bmatrix}$$

5. Calculate the First Iteration Objective Function according to Equation (3)

Table 4. First Iteration Objective Function Value

Data	1	2	3	4	5	6	⋯	38	∑
$\sum_{k=1}^{3} C_k \cdot \mu_{ik}^{w}$	0,0686	0,0484	0,2924	0,1991	0,7774	1,1664	⋯	0,1011	13,2314

6. Calculating the change in the partition matrix U according to Equation (4)

Table 5. First Iteration of U Partition Matrix Values

Data	1	2	3
1	0,4135	0,1458	0,4407
2	0,3455	0,2282	0,4264
3	0,3155	0,4159	0,2686
4	0,2757	0,3742	0,3501
5	0,3232	0,3573	0,3195
⋮	⋮	⋮	⋮
38	0,4819	0,2551	0,2630

7. Checking stop conditions
Because $|J_1 - J_0| = |13{,}2314 - 0| = 13{,}2314 > \xi$ in this case 10^{-5}, and iteration $= 1 < MaxIter(100)$, the process continues to iteration 2 $(t = 2)$.

The iteration is continued until $\left(\left|J_{(t)} - J_{(t-1)}\right| < \xi\right)$or (t > MaxIter), using MATLAB software, the results of the calculation are the cluster center or center, the membership degree or matrix U, and the objective function value or objective function. The first result is the calculation of the objective function value, and it is found that 30 iterations are required to obtain the optimal solution for the objective function value of 9,388695.

The second result is the calculation of the cluster center values from the last iteration as follows:

$$V_{kj} = \begin{bmatrix} 0,1517 & 0,0760 & 0,1314 & 0,0467 & 0,0117 & 0,0072 & 0,0358 & \cdots & 0,0041 \\ 0,4407 & 0,1729 & 0,4810 & 0,2839 & 0,1464 & 0,0447 & 0,1333 & \cdots & 0,0621 \\ 0,4512 & 0,1525 & 0,4082 & 0,2585 & 0,1016 & 0,0355 & 0,1439 & \cdots & 0,0386 \end{bmatrix}$$

The third result is the calculation of the Uik values. Based on the obtained μ_{ik} values, the clustering results can be presented in Table 6.

Table 6. Results of Clustering Disaster-Prone Areas in East Java Province

Cluster	Category	Number of Regencies or Cities	Regencies/ Cities
1	Low Disaster Risk	22	Bangkalan, Kediri, Kota Batu, Kota Blitar, Kota Kediri, Kota Madiun, Kota Malang, Kota Mojokerto, Kota Pasuruan, Kota Probolinggo, Kota Surabaya, Magetan, Nganjuk, Ngawi, Pacitan, Ponorogo, Probolinggo, Sampang, Sumenep, Trenggalek, Tuban, Tulungagung
2	Moderate Disaster Risk	8	Blitar, Bondowoso, Jember, Lumajang, Madiun, Malang, Pamekasan, Sidoarjo
3	High Disaster Risk	8	Banyuwangi, Bojonegoro, Gresik, Jombang, Lamongan, Mojokerto, Pasuruan, Situbondo

Evaluation of Clusters Using the Silhouette Coefficient

The clustering results were evaluated using the Silhouette Coefficient to assess the quality of separation and compactness between clusters. This coefficient value indicates the extent to which objects fit into their clusters compared to other clusters. A value close to 1 indicates good clustering results, while a value close to 0 or negative indicates overlap between clusters. The evaluation results of the clustering for each regency/city can be seen in Table 7.

Table 7. Silhouette Coefficient Value for Each Regency/City

No.	Regencies/ Cities	Cluster	s(i)
1.	Bangkalan	1	0,4966
2.	Banyuwangi	3	-0,2939
3.	Blitar	2	-0,2147
4.	Bojonegoro	3	0,0399
5.	Bondowoso	2	-0,0085
6.	Gresik	3	0,0732
7.	Jember	2	-0,2170
8.	Jombang	3	-0,1381
9.	Kediri	1	0,6560
10.	Kota Batu	1	0,5219
11.	Kota Blitar	1	0,6618
12.	Kota Kediri	1	0,6578
13.	Kota Madiun	1	0,6608
14.	Kota Malang	1	0,5359
15.	Kota Mojokerto	1	0,6525
16.	Kota Pasuruan	1	0,5468
17.	Kota Probolinggo	1	0,6585
18.	Kota Surabaya	1	0,6591
19.	Lamongan	3	-0,1618
20.	Lumajang	2	-0,0795
21.	Madiun	2	-0,0432
22.	Magetan	1	0,5487
23.	Malang	2	-0,0181
24.	Mojokerto	3	-0,1734
25.	Nganjuk	1	0,5307
26.	Ngawi	1	0,6069
27.	Pacitan	1	0,5784
28.	Pamekasan	2	-0,1769

29.	Pasuruan	3	0,1568
30.	Ponorogo	1	0,4570
31.	Probolinggo	1	0,4462
32.	Sampang	1	0,5129
33.	Sidoarjo	2	0,0634
34.	Situbondo	3	0,0311
35.	Sumenep	1	0,5933
36.	Trenggalek	1	0,2519
37.	Tuban	1	0,1722
38.	Tulungagung	1	0,4229

Based on the evaluation results using the Silhouette Coefficient method, the average Silhouette Coefficient value obtained was 0,2807. This value falls into the category of fairly good cluster interpretation, indicating that most of the data has been grouped into the appropriate clusters, although there is still some data that is close to the boundary between two clusters.

Visualization of Clustering Results

Spatial representation is used to visualize the results of disaster vulnerability clustering in East Java. Visualization is performed using the Map Chart feature in Microsoft Excel so that the clustering results are easier to understand geographically. The Fuzzy C-Means algorithm divides 38 districts/cities into three clusters displayed in different colors, so that the distribution pattern of disaster vulnerability can be observed more clearly. The following presents a visualization in the form of a map of East Java Province representing the clustering results of areas based on disaster vulnerability levels.

1. Map of Natural Disaster-Prone Areas in East Java Province

The map in Figure 1 shows the distribution of areas based on clustering results, which are divided into three clusters in red, orange, and green. The colors on the map indicate the maximum membership of an area to a cluster. Red indicates areas with a high level of vulnerability to natural disasters, orange indicates moderate vulnerability, and green indicates areas with low vulnerability.

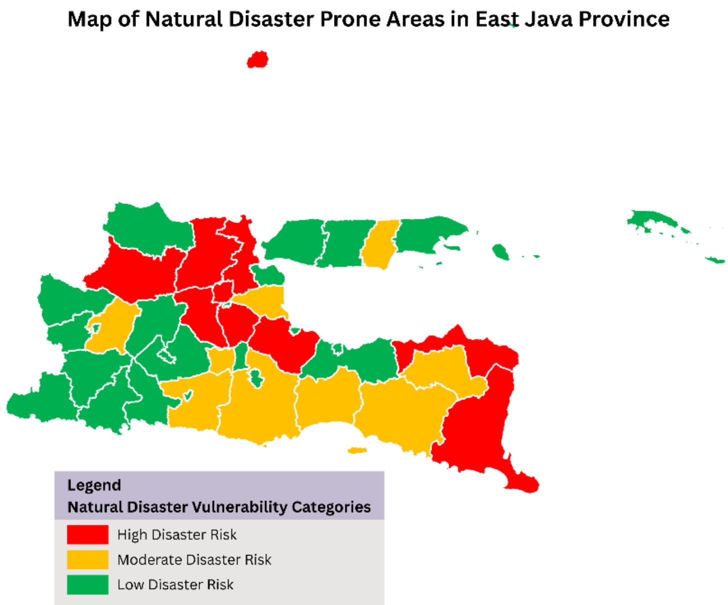

Figure 1. Map of Natural Disaster-Prone Areas in East Java Province
Source: Researcher 2025

2. Map of Low Disaster-Prone Areas in East Java Province

The map in Figure 2 shows the distribution of areas with low disaster vulnerability. Green is used as the main color gradient, where the darker the green, the higher the area's membership value in the low disaster vulnerability cluster.

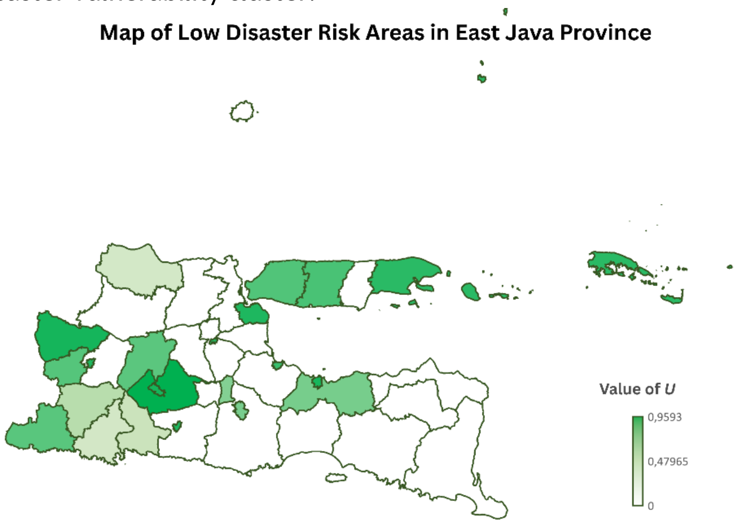

Figure 2. Map of Low Natural Disaster-Prone Areas in East Java Province
Source: Researcher 2025

3. Map of Moderate Disaster-Prone Areas in East Java Province

The map in Figure 3 shows the distribution of areas with moderate disaster vulnerability, indicated by shades of orange. The darker the orange color, the higher the area's membership value in the moderate cluster.

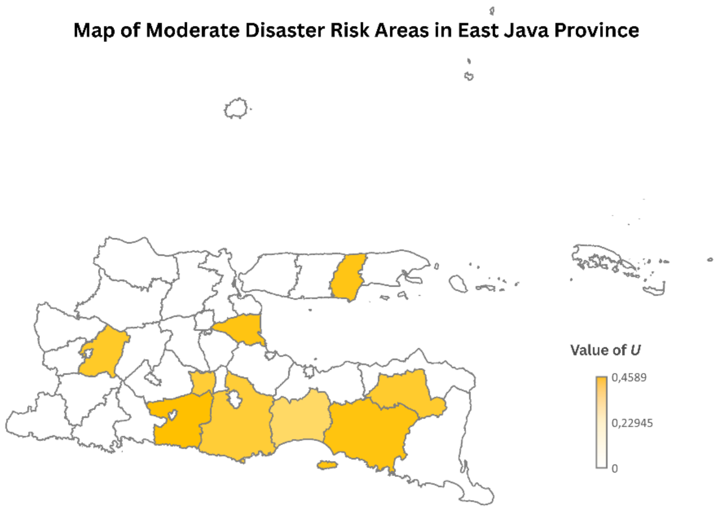

Figure 3. Map of Moderate Natural Disaster-Prone Areas in East Java Province
Source: Researcher 2025

4. Map of Moderate Disaster-Prone Areas in East Java Province

The map in Figure 4 shows the distribution of areas with high disaster vulnerability, marked in red. The darker the red color, the higher the level of vulnerability.

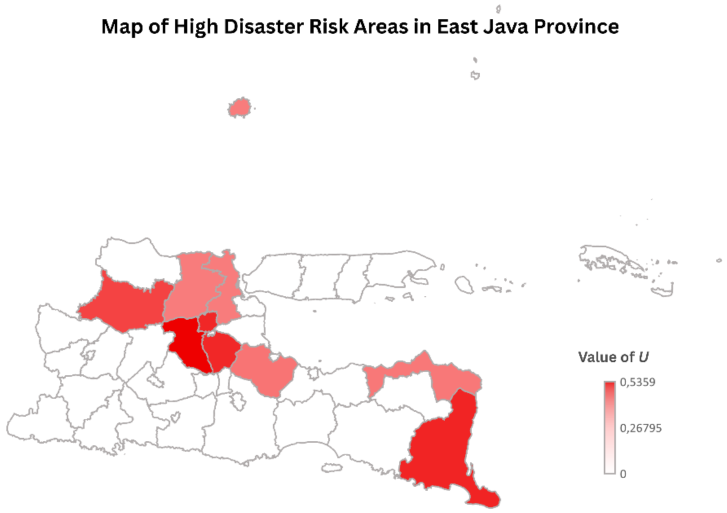

Figure 4. Map of High Natural Disaster-Prone Areas in East Java Province
Source: Researcher 2025

The four maps displayed not only show the clustering results visually, but also illustrate the distribution patterns of disaster vulnerability in East Java Province. By presenting the information in map form, the clustering results become easier to understand for various parties, including local governments, disaster management agencies, and the general public.

Each map provides a more specific overview of the level of vulnerability, ranging from low, moderate, to high, thereby serving as a foundation for formulating more targeted and effective mitigation strategies. Therefore, this visualization not only functions as a complement to data analysis but also as a tool to assist in spatial-based decision-making processes.

DISCUSSION

Based on the results of the study, the average Silhouette Coefficient value obtained was 0,2807, which indicates that in general the clustering results were in the fairly good category. However, there were a number of districts/cities that had negative values, indicating ambiguity in cluster membership. These regencies/cities include Banyuwangi, Blitar, Bondowoso, Jember, Jombang, Lamongan, Lumajang, Madiun, Malang, Mojokerto, and Pamekasan. These negative values arise due to inconsistencies between the three main indicators in the cluster, namely the number of incidents, the number of victims, and the amount of damage. For example, a region may have a high number of incidents but relatively low casualties and damage, or vice versa. These differences in characteristics make it difficult to place these regions definitively in one cluster, and they tend to overlap between clusters. From a practical standpoint, areas with negative silhouette values require special attention, as disaster mitigation strategies cannot be generalized but must be designed according to the complexity of each region's risk profile.

CONCLUSION

This study successfully grouped 38 districts/cities in East Java into three disaster vulnerability clusters using the Fuzzy C-Means method with an average Silhouette Coefficient value of 0,2807, which is classified as fairly good, although some areas showed negative values due to discrepancies between the frequency of events, number of victims, and level of damage. The limitations of this study lie in the data period, which only covers 2020–2024, and the variables, which are limited to events, casualties, and damage without involving social, economic, and disaster management capacity factors. In the future, further research is recommended to use data with a longer time span and add variables of social vulnerability and regional capacity so that the classification results are more comprehensive and relevant for the formulation of disaster mitigation policies.

AI ACKNOWLEDMENT

The authors declare that generative AI or AI-assisted technologies were not used in any way to prepare, write, or complete this manuscript. The authors confirm that they are the sole authors of this article and take full responsibility for the content therein, as outlined in COPE recommendations.

INFORMED CONSENT

The authors have obtained informed consent from all participants.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCE

Anggraeni, D. (2016). Analisis Dampak Perubahan Iklim Berdasarkan Kenaikan Muka Air Laut terhadap Wilayah Kota Surabaya. In *Tesis*. Institut Teknologi Sepuluh Nopember.

Aulia Putri, S. (2025). Pengelompokkan Daerah Rawan Bencana di Kabupaten Jember Menggunakan Metode K-Means Clustering. *JUPITER: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, *3*(1), 9–16. https://doi.org/10.19184/isj.v8i1.29542

BNPB. (2021). Dokumen Kajian Risiko Bencana Nasional Provinsi Riau 2022 - 2026. In *Kedeputian Bidang Sistem dan Strategi Direktorat Pemetaan dan Evaluasi Risiko Bencana*.

Faruf, A., Ferijal, & Yasar, M. (2023). Pengaruh Perubahan Iklim Terhadap Frekuensi dan Durasi Deret Hari Kering di Kabupaten Aceh Besar. *JURNAL ILMIAH MAHASISWA PERTANIAN*, *8*, 622–630.

Firdaus, H. S., Nugraha, A. L., Sasmito, B., & Awaluddin, M. (2021). Perbandingan Metode Fuzzy C-Means Dan K-Means Untuk Pemetaan Daerah Rawan Kriminalitas Di Kota Semarang. *Elipsoida : Jurnal Geodesi Dan Geomatika*, *4*(01), 58–64. https://doi.org/10.14710/elipsoida.2021.9219

Ifadah, C., Ratnasari, C. D., & Novitasari, D. C. R. (2022). Clustering Daerah Banjir Di Jawa Timur Dengan Algoritma Fuzzy C-Means. *Jurnal Dinamika Informatika*, *14*(2), 99–104. https://doi.org/10.35315/informatika.v14i2.8885

Irabawati, N., Wahyuningsih, S., & Syoer, R. R. (2016). Perbandingan Metode C-Means dan Fuzzy C-Means Dalam PengelompokkanWilayah Desa/Kelurahan di Kabupaten Kutai Kartanegara. *Jurnal Eksponensial*, *7*(1), 67–76. http://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/view/47%0Ahttp://jurnal.fmipa.unmul.ac.id/index.php/exponensial/article/download/47/24

Marto, F. A., Ardi, I. A., & Priyoga, I. (2024). *Analisis Kebencanaan Wilayah Berdasarkan Karakteristik Bentang Lahan di Kabupaten Kulon Progo*. *5*(1), 42–62.

Nabilla Audy, I., Nur Padilah, T., & Nurina Sari, B. (2024). Pengelompokan Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma Fuzzy C-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(4), 2799–2803. https://doi.org/10.36040/jati.v7i4.7205

Pascapurnama, D. N., Murakami, A., Chagan-Yasutan, H., Hattori, T., Sasaki, H., & Egawa, S. (2018). Integrated Health Education in Disaster Risk Reduction: Lesson Learned From Disease Outbreak Following Natural Disasters In Indonesia. *International Journal of Disaster Risk Reduction*, *29*, 94–102. https://doi.org/10.1016/j.ijdrr.2017.07.013

Prasetyo, A., Sumarno, S., Jayaputra, A., Benedictus, M., Murni, R., Nainggolan, T., Purwasantana, D., Miftah, M., Wahab, Taruna, M. M., & Wibowo, A. M. (2024). Critical Communication Of Disaster Preparedness Areas For Informational Strategies In Disaster Management In Indonesia. *Progress in Disaster Science*, *24*. https://doi.org/10.1016/j.pdisas.2024.100368

Pusdatinkom BNPB. (2025). *Data Informasi Bencana Indonesia (DIBI)*. https://dibi.bnpb.go.id/

Rahmawati, T., Wilandari, Y., & Kartikasari, P. (2024). Analisis Perbandingan Silhouette Coefficient dan Metode Elbow pada Pengelompokkan Provinsi di Indonesia Berdasarkan Indikator IPM Dengan K-Medoids. *Jurnal Gaussian*, *13*, 13–24. https://doi.org/10.14710/j.gauss.13.1.13-24

Rahmi, E. R. (2024). *Ring Of Fire Indonesia*. https://unsia.ac.id/ring-of-fire-indonesia/

Septian, A., Elvarani, A. Y., Putri, A. S., Maulia, I., Damayanti, L., Pahlevi, M. Z., & Aswad, F. H. (2020). Identifikasi Zona Potensi Banjir Berbasis Sistem Informasi Geografis Menggunakan Metode Overlay dengan Scoring di Kabupaten Agam, Sumatera Barat. *Jurnal Geosains Dan Remote Sensing*, *1*(1), 11–22. https://doi.org/10.23960/jgrs.2020.v1i1.25

Undang-Undang Republik Indonesia Nomor 24 Tahun 2007 Tentang Penanggulangan Bencana (2007).

Wijayanti, W., Rahmi HG, I., & Yanuar, F. (2021). Penggunaan Metode Fuzzy C-Means Untuk Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Kesehatan Lingkungan. *Jurnal Matematika UNAND*, *10*(1), 129. https://doi.org/10.25077/jmu.10.1.129-136.2021

Yulianto, T., Rahmah, A. F., Faisol, F., & Amalia, R. (2023). Clustering Daerah Bencana Alam di Indonesia Menggunakan Metode Fuzzy C-Means. *Unisda Journal of Mathematics and Computer Science (UJMC)*, *9*(2), 29–39. https://doi.org/10.52166/ujmc.v9i2.4776