PAPER • OPEN ACCESS

The earthquake events and impacts mapping in Bali and Nusa Tenggara using a clustering method

View the article online for updates and enhancements.

The earthquake events and impacts mapping in Bali and Nusa Tenggara using a clustering method

S Harini[1], H Fahmi[1,*], A D Mulyanto[1] and M Khudzaifah[1]

[1]Mathematics Department, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Jl Gajayana 50 Malang 65144, Indonesia

*E-mail: *hisyam.fahmi@uin-malang.ac.id*

Abstract. This article proposes the mapping of earthquake potential zones in Regional Center 3 (Bali and Nusa Tenggara, Indonesia) using the clustering method. A suitable clustering method to group spatial and non-convex data, such as earthquake data, is the Density-Based Spatial Clustering of Application with Noise (DBSCAN). We modify the epsilon parameter with the magnitude because it reflects the impact of the earthquake happens so that the epsilon value will be different for each point. Before the clustering process, magnitude is converted to kilometers (km) to determine the earthquake impact region, which will be used as epsilon. The earthquake impact region is obtained by computing the radius of the earthquake using the McCue Earthquake Perception Radius equation. This approach gives a good result for mapping the potential earthquake zones indicated by the cluster validity index that gives an average higher than 0 using the Silhouette index.

1. Introduction

The earthquake in Bali and Nusa Tenggara (Regional Center 3), Indonesia, has increased in the last three years. The most destructive earthquake in the previous 3 years was the Lombok earthquake on August 5, 2018, with a magnitude of 7, with the epicenter at sea 28 km northwest of East Lombok (8.34 S, 116.46 E) with a depth of 15 km. This earthquake is a series of earthquakes that occurred from July 28th, August 5th, 9th, and 19th, 2018 which resulted in 564 people died, 1584 injured and also 167940 housing units damaged. The last earthquake that occurred at Regional Center 3 was the earthquake on July 16th, 2019, in Nusa Dua, Bali.

From some of these earthquake events, there needs to be a mapping of earthquake zones and affected areas. This mapping aims to be able to analyze the potential earthquake zones from the history of previous earthquakes. The more frequent an area of earthquake tremors, it can be said that the area is a potential earthquake zone, but it is uncertain when and how significant the effect is because earthquake prediction is a difficult job.

Data collection on earthquake-affected areas manually is done by visiting the earthquake location and recapitulating earthquake data that has occurred. It has become the work of the Meteorology Climatology and Geophysics Council (BMKG) in general, which is to record and map the potential earthquake disasters. For example, it has been documented in the form of earthquake disaster reports in Indonesia in the range of 1821-2018 [1] and the Jembrana Earthquake Impact Survey Report on 16 July 2019 [2]. Thus, to make this mapping process effectively, an automatic method is needed to group these earthquake data so that the distribution patterns can be analyzed. Automatic earthquake mapping with the help of data mining can be done using the clustering algorithm approach.

Wang et al. [3] identify the long-aftershock candidate zone by combining information about the magnitude, several events, and the distance between the local grid node and the location of the nearest centers. The results show that areas with a high seismic density index can be used in principle to indicate the historical site of paleoseismic events. Scitovski [4] applies the Rough-DBSCAN density-based clustering algorithm for earthquake zoning. Determination of the appropriate value of the parameter ε (epsilon) needs to consider because it significantly influences the number and configuration of earthquake zones.

The clustering algorithm that is suitable to use earthquake mapping is Density-Based Spatial Clustering of Application with Noise (DBSCAN) because this algorithm can group data based on its density, not only on convex data. Therefore, this research proposes the mapping of earthquake potential zones in RC 3 (Bali and Nusa Tenggara, Indonesia) using the DBSCAN clustering method.

2. Literature Review

Scitovski [4] has implemented a density-based clustering algorithm, Rough-DBSCAN, for earthquake zoning. The application of this algorithm aims to recognize non-convex forms in earthquake zonation to provide more realistic results. The size of the epsilon parameter significantly influences the number and configuration of earthquake zones. The method is applied to the problem of earthquake zoning in the broader territory of the Republic of Croatia.

Deyasi et al. [5] have classified the earthquake zone hierarchically from the distance matrix and linked it to its proximity to the tectonic plate. The nature of the directed link indicates that each earthquake network is highly connected in that study calculated conditional probabilities of future earthquake events in each region. The conditional probabilities of each event have been compared with their stationary distribution.

Al-Ahmadi et al. [6] detecting clusters and local and global spatial patterns in the event of earthquakes in the Red Sea region using geographic information systems. In this study, several spatial pattern analysis techniques were applied, namely quadrant analysis, nearest neighbor average, global Moran's I, Getis-Ord general G, Anselin Local Moran's I, Getis-Ord Gi*, kernel density estimation, and geographical distributions. Both local and global spatial statistics indicate that earthquakes in the Red Sea region are collected into several clusters.

Hashemi and Karimi [7] tried to generate seismic vibration source areas automatically using hierarchical clustering based on earthquake history data, then made earthquake magnitude prediction models from earthquake features/characteristics using classification methods such as SVM, Decision Tree, and KNN. In his research, three combinations of features were tested, namely non-spatial features, spatial features, and a combination of both, the results of predictive accuracy were not good when only spatial features were used.

3. Methods

Among various types of clustering algorithms, density-based clustering is more efficient for determining clusters in data with different densities [8]. DBSCAN is one example of the development of clustering based on density or commonly known as density-based clustering [9]. This method is useful for clustering non-convex data.

DBSCAN does not require several clusters as a parameter, but it requires two parameters to decide whether the two closest points should be collected into one cluster or not. These two parameters are ε (epsilon) as distance threshold, and *MinPts* (minimum number of points).

The grouping process with DBSCAN can be described as a tree. It starts with any point that has at least the closest *MinPts* point in the radius ε. Then do the first search along each of these closest points. For each nearest point, check how many points there are in the radius. If it has a spot in the radius ε less than *MinPts*, this point becomes the border point. However, if it has at least *MinPts*, then that point becomes a branch, and all its neighbors are added to the FIFO queue from this first search.

3.1. Data for Earthquake Mapping

The data processed in this study is the earthquake data obtained from BMKG Sanglah in Denpasar, Bali, which is the earthquake data in RC 3 Indonesia from July 1st, 2018 to July 31st, 2019. The map of RC 3 is shown in Figure 1. The data are of the form: date, time (UTC+8), latitude (lat), longitude (long), magnitude (mi), depth (d). We try to map earthquake data and identify potential earthquake zones based on the cluster distribution.

In the process of grouping the earthquake zones, we used the Density-Based Spatial Clustering of Application with Noise (DBSCAN) method for clustering spatial and temporal earthquake data based on earthquake characteristics. The epsilon (ε) parameter modification will be used here.

Figure 1. Map of regional center 3 (rc 3) indonesia.

3.2. Data Preprocessing

Before it is processed in the DBSCAN clustering, the magnitude is converted to kilometers (km). It is to determine the earthquake impact region, which will be used as epsilon (ε). The earthquake impact region is obtained by computing the radius of the earthquake using McCue Earthquake Perception Radius Equation [10] as in Earthquake Radius (km) = (Magnitude-0.13)/1.01(1).

$$\text{Earthquake Radius (km) = (Magnitude-0.13)/1.01} \tag{1}$$

3.3. Dynamic DBSCAN

We modify the epsilon (ε) parameter with the magnitude value, so its value for each point will be different based on the magnitude. Therefore, we call it dynamic DBSCAN. The magnitude value is used as epsilon because its value affects the area damaged by the earthquake. For the *MinPts* parameter, it should be $ln(m)$ [11], where m is na umber of points. But, to reduce cthe omputational complexity for the data with only two features, *MinPts* is usually set to 4 [12].

3.4. Cluster Validation

In this experiment, three cluster validation index were used to measure the correctness of the clustering result. There are Silhouette index, Davies-Bouldin (DB) Index, and Calinski-Harabasz (CH) index. However, these three methods are proper to measure the non-convex cluster. The execution time was computed for each point of data.

3.5. Implementation

The proposed clustering method was implemented using Python. We are using the basic DBSCAN algorithm, which is implemented by Chris McCormic and we modify the ε parameter, in which the value is based on the radius of perception. The modification code for the dynamic DBSCAN is shown in Figure 2. This function computes the points which are in the region of point P based on the magnitude of point P.

4. Discussion

In this experiment, the earthquake data used on RC 3 occurred in August 2018 to July 2019. From this data, we do a clustering trial per month using Dynamic DBSCAN, and we evaluate the results of the clustering using the Silhouette cluster validity index, DB index, CH index, and execution time for each data point. The ε parameter takes the radius of percepts value of magnitude, and the *MinPts* value is set to 4.

The result shows that dynamic DBSCAN give a good clustering earthquake data for each month, indicated by the average of Silhouette index is not below 0. Also, this algorithm has an average of execution time around 1.25 seconds for each data point. The summary of the experiment result gives in Table 1.

```python
def regionQuery(D, P, eps):
    """
    This function calculates the distance between a point P and every other
point in the dataset, and then returns only those points which are within a
threshold distance `eps`.
    """
    neighbors = []
    # For each point in the dataset...
    for Pn in range(0, len(D)):
        # If the distance is below the threshold,
        # add it to the neighbors list.
        if      degrees2kilometers(numpy.linalg.norm(D.iloc[P,0:2]        -
D.iloc[Pn,0:2])) < eps.iloc[P]:
            neighbors.append(Pn)
    return neighbors
```

Figure 2. Modification code for the dynamic DBSCAN.

Table 1. Clustering result for each month.

Month	Year	Number of Clusters	Sil[a]	DB[b]	CH[c]	Average Time (s)
September	2018	7	**0.561**	**0.457**	195.988	0.906
October	2018	9	0.469	0.581	**399.217**	1.108
November	2018	16	0.392	0.654	170.178	0.841
December	2018	8	0.328	0.749	148.121	1.076
January	2019	3	0.434	0.571	151.703	**0.413**
Febebruary	2019	21	0.226	0.598	358.506	1.916
March	2019	8	0.418	0.504	145.428	1.962
April	2019	10	-0.072	0.652	50.354	1.544
May	2019	18	0.219	0.552	175.842	1.360
June	2019	2	0.361	0.908	151.811	1.155
July	2019	12	0.109	0.524	97.264	1.415
	Average		**0.313**	**0.614**	**185.857**	**1.245**

[a] Silhouette index, higher is better
[b] Davies-Bouldin index, smaller is better
[c] Calinski-Harabasz index, greater is better

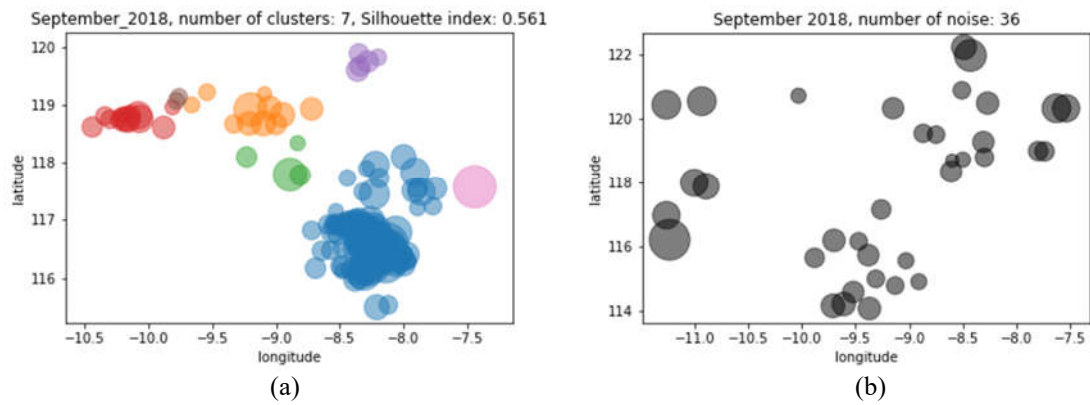(a)   (b)

Figure 3. Clustering result of earthquake and noise points in September 2018.
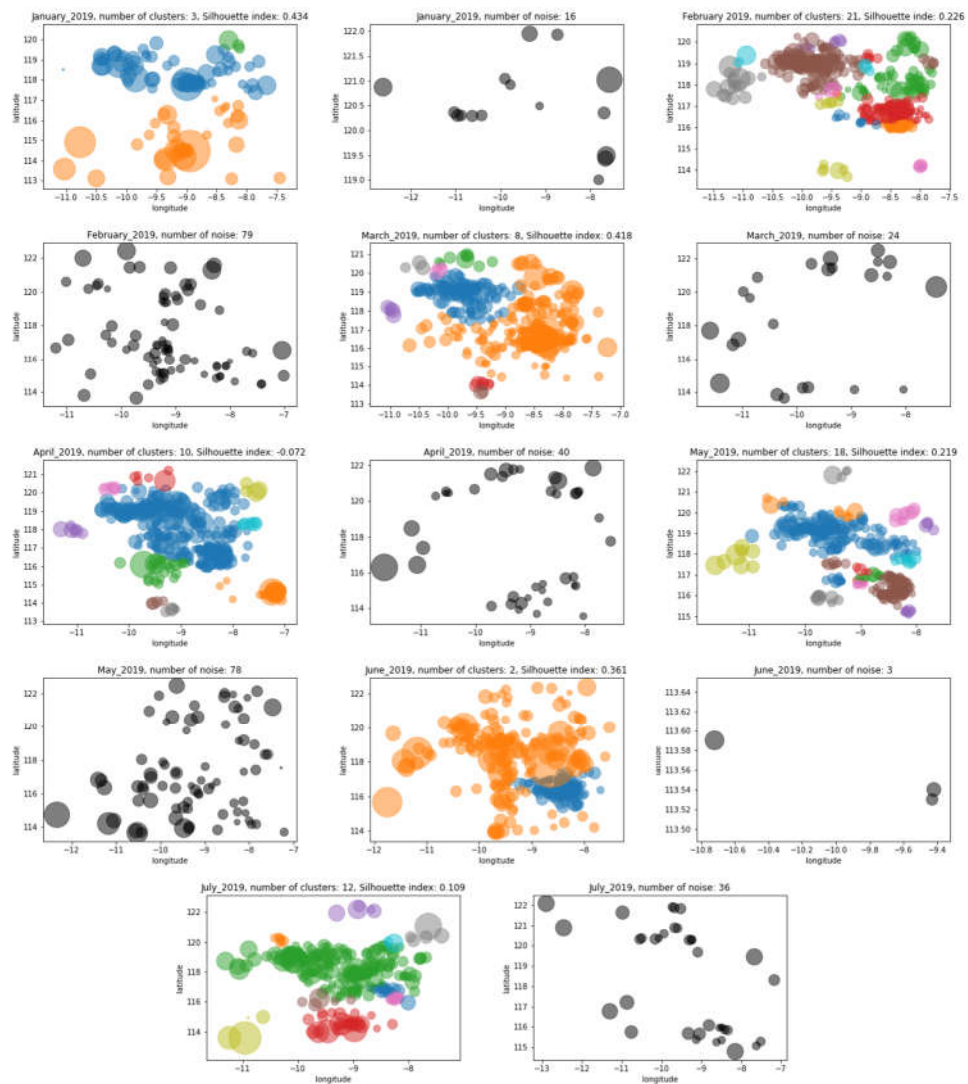

Figure 4. Mapping of potential earthquake zones in RC3 and the noise points for each month.

From the results of this experiment, it can be seen that earthquakes grouped in several areas. The results of this clustering can be a map that illustrates the grouping of earthquake points in certain areas based on its density. The denser the number of earthquake points in an area, it can be said that the area is a potential earthquake zone. Conversely, if the epicenter point of the earthquake does not have adjacent neighbors more than *MinPts* in the region of perception of its magnitude, then that point is considered as noise. Figure 3 (a) is a map of the clustering results in September 2018, which is resulted in 7 number of clusters and give the best Silhouette and DB index. Figure 3 (b) is a plot of points which are noises. Figure 4 is the mapping of potential earthquake zones in RC3 from January until July 2019 and it is noise points.

5. Conclusion

In this study, mapping of earthquake potential zones on RC 3 Indonesia was obtained. This mapping uses the DBSCAN clustering algorithm approach. Modifications were made to determine epsilon parameters based on the region of perception from the magnitude of each earthquake. The experimental results show that the grouping is done effectively in terms of execution time and with good cluster validity results, indicated by 3 cluster validity measures.

References

[1]   Sadly M 2019 *Katalog Gempabumi Signifikan dan Merusak* p 1821-2018 (Jakarta)
[2]   BMKG Wilayah III Denpasar 2019 *Laporan Survey Dampak Gempabumi Jembrana*
[3]   Wang J, Main I G and Musson R M W 2017 *Geophys. J. Int*. **211** 1005–1018
[4]   Scitovski S 2018 Comput. Geosci. **110** 90–95
[5]   Deyasi K, Chakraborty A and Banerjee A 2017 *Phys. A Stat. Mech. its Appl.* **481** 224-234
[6]   Al-Ahmadi K, Al-Amri A and See L M 2014 Artic. Arab. J. Geosci. 7 2893–2904
[7]   Hashemi M and Karimi H A 2016 *SmartCity* **360** 468-478
[8]   Matheus C J, Chan P K and Piatetsky-Shapiro G 1993 *IEEE Trans. Knowl. Data Eng.* **5** 903–913
[9]   Mumtaz K and Duraiswamy K 2010 *Indian J. Comput. Sci. Eng.* **1** 8–12
[10]   K. McCue 2019 *Earthquake Perception Radius Calculator* Available at https:// www.cqsrg.org/tools/perceptionradius
[11]   Birant D and Kut A 2006 *Data Knowl. Eng.* **60** 208-221
[12]   Jiang H, Li J, Yi S, Wang X and Hu X 2011 *Expert Syst. Appl.* **38** 9373–938