Earthquake Density Measurement Using Automatic Clustering

Amin Endah Sulistiawati[1], Ali Ridho Barakbah[2], Tri Harsono[3], Yuliana Setyowati[4]
Information and Computer Engineering, Electronic Engineering Polytechnic Institute of Surabaya
[1]aminendah@student.eepis-its.edu, [2]ridho@eepis-its.edu, [3]trison@eepis-its.edu, [4]yuliana@eepis-its.edu

Abstract

The government and earthquake associations have recorded the seismic data in spatial-temporal usually used to measure the earthquake intensity, but such information has not been processed to obtain the earthquake density. This condition makes it difficult to map all the risk, so it creates the lack of participatory development to the earthquake areas that have a high density. This research proposes a new approach for measuring the density of earthquake and performing automatic clustering with valley tracing method to detect the number of group spatial regions automatically from analysis of cluster moving variances. The density is obtained with a new approach that involves the area and the amount of data on clusters. This system follows the steps: (1) display the spatial-temporal earthquake dataset into a map, (2) create vector space data consist of temporal, spatial, and magnitude, (3) detect number of cluster with valley tracing method on automatic clustering, (4) calculate the earthquake density, and (5) display special temporal visualization with the cluster density measurement result. To perform the proposed idea, this system is examined with an experimental study for a series of quake about Indonesia and Japan during the last 50 years from ANSS catalog.

Keywords: earthquake, automatic clustering, density measurement

1. Introduction

Earthquakes are natural events that happen as tectonic stress or energy released from faults, fracture on Earth's crust [1]. It was caused by the sudden breaking and movement of tectonic plates of the earth's crust. A large earthquake can cause the great damage, such as landslides, snow avalanches, tsunamis (giant sea waves) and volcanic eruptions. If the earthquake occurs in a populated area, it may cause collapse of building or other man-made infrastructures, injuries, and deaths.

In a current research, the government and earthquake associations have recorded the seismic data in temporal and spatial. These data are very useful for measuring the frequency of earthquakes or called as intensity. However, such information has not been processed to obtain the earthquake density. This condition makes it difficult to map all the risk, so it creates the lack of participatory development to the earthquake areas that have a high density.

Knowing the earthquake-prone areas is essential. Especially in countries that often occur earthquake. There have been many researchers, who conduct research on earthquake, including: R. Sadeghian and G. R. Jalali-Naini found a new probability density function (PDF) for forecasting the time of earthquake occurrence [1], Rusnardi R.P, Junji Kiyono, and friends constructed area earthquake source model and estimated the frequency magnitude relationship by using the catalogs compiled [2], S. Das and C. Henry examined where aftershocks occur using data from several recent large earthquake [3], Masryur Irsyam, Donny T Dangkua, and friends presented the development of spectral hazard maps for Sumatra and Java islands, Indonesia [4], Hiroyuki Fujiwara, Shinichi Kawai, and friends developed an open web system, includes the hazard map results and data on seismic activity, source models, and underground structure [5]. These studies provide many benefits to human life. However, no one has ever examined the density of earthquakes. As well as information on the population density of an area, the density can provides a lot of information. By knowing which areas have the highest density, attention can be directed to the right target to analyze the likelihood of earthquakes, earthquake prediction, or early warning

2. Proposed Idea

This research proposes a new approach for measuring the density of earthquake. Earthquake density is different with earthquake intensity that only involves the number of earthquake even on an area. The density is measured after performing automatic clustering in areas affected by the earthquake. Automatic clustering is an initial process to obtain the number of cluster automatically before performing the clustering to group the earthquake dataset. This idea implements valley tracing method to detect the number of group spatial regions automatically from analysis of cluster moving variances. The valley tracing method determines the global optimum by variance differentiation and ensured the credibility ratio of the global optimum. Hierarchical clustering carried out after

ISBN : 978-602-9494-97-6

getting the ideal number of clusters to classify seismic data. From the information of each cluster, we measure the cluster density. The uniqueness point of this research is the measurement of earthquake density. A methodology how to find the density of cluster is found out because there has never been a study to calculate the density of earthquakes by performing clustering and spatial-temporal dataset. For experimental study, this research applies the earthquake dataset in Indonesia and Japan during the last 50 years from ANSS catalog [6].

We choose Indonesia and Japan because they are both located on tectonic plate boundaries. These countries have suffered from large earthquakes and volcanic eruption. Two largest earthquakes in the world during the 50 years are the 2004Sumatra-Andaman earthquake (M9.1) caused fatalities in the hundreds of thousands in Indonesia and the 2011 Tohoku Earthquake (M9.0) in the tens of thousands in Japan [7]. It means that large populations at risk from these frequent disasters.

This research is aimed to the government and associations earthquake to provide information about the earthquake density in areas that have been grouped. The result is about which areas have the highest earthquake potential based on the earthquake density measurement, so that the risk mapping and participatory development can be applied to the right areas.

3. Basic Concept of Clustering

Clustering is an exploratory data analysis tool that deals with the task of grouping objects that are similar to each other [8,9]. The clustering structure is represented as a set of subsets *C=C1,…,Ck* of *S*, such that : $S = U_{i=1}^{k} C_i$ and $Ci \cap Cj = \theta$; for $i \neq j$. Consequently, any instance in S belongs to exactly one and only one subset.

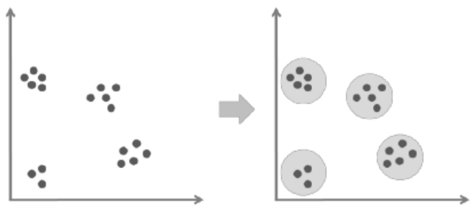


Figure 1. Clustering Data

On the figure above, the data objects have been grouped into 4 clusters based on the similarity distance.

3.1 Cluster Variance

Variance is used to measure the value of the spread of the data cluster results for unsupervised learning. There are two types of variance, variance within cluster and variance between clusters. A good cluster is when the members of a cluster have a high degree similarity to each other (external homogeneity) and are not like members of other cluster (external homogeneity).

Internal homogeneity called as variance within cluster (v_w) and external homogeneity is variance between clusters (v_b). The ideal cluster has minimum of :

$$v = \frac{v_w}{v_b} \tag{1}$$

Here are some formulas to get the value of variance within cluster (v_w) and variance between clusters (v_b) :

1) The cluster variance can be calculated with the formula bellow :

$$v_c^2 = \frac{1}{n_c - 1} \sum_{i=1}^{n_c} (d_i - d_i)^2 \tag{2}$$

where :
v_c^2 = variance on cluster c
c = 1 … k, where k = number of cluster
n_c = amount of data on cluster c
d_i = data to i on a cluster
d_i = mean of data on a cluster

2) Variance within cluster can be calculated by :

$$v_w = \frac{1}{N-k} \sum_{i=1}^{k} (n_i - 1).v_i^2 \tag{3}$$

where :
v_w = variance within cluster
N = amount of all data

3) Variance between cluster :

$$v_b = \frac{1}{k-1} \sum_{i=1}^{k} n_i (d_i - d)^2 \tag{4}$$

where :
d = mean of d_i

3.1 Hierarchical Clustering

Hierarchical clustering procedures are characterized by the tree-like structure established in the course of the analysis [10]. The hierarchical techniques in this research use a category called as agglomerative clustering. In this category, clusters are consecutively formed from objects.

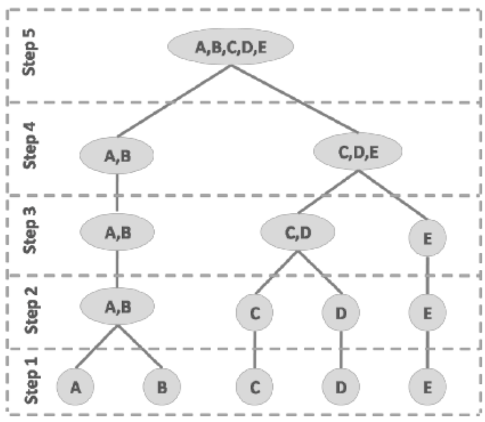


Figure 2. Agglomerative clustering

Amin Endah Sulistiawati, Ali Ridho Barakbah, Tri Harsono, Yuliana Setyowati, "Earthquake Density Measurement using Automatic Clustering",
The Third Indonesian-Japanese Conference on Knowledge Creation and Intelligent Computing (KCIC) 2014, March 25-26, 2014, Malang, Indonesia.

Initially, this type of procedure starts with each object representing an individual cluster. These clusters are then sequentially merged according to their similarity, the distance of a data to other. The two most similar clusters are merged to form a new cluster at the bottom of the hierarchy at first. In the next step, another pair of clusters is merged and linked to a higher level of the hierarchy, and so on. This allows a hierarchy of clusters to be established from the bottom up. Figure 2 shows how the methodology assigns additional objects to clusters as the cluster size increases.

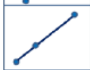
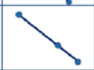
This study performs agglomerative clustering procedures include the following: single linkage, complete linkage, average linkage and centroid linkage. We use them for experiment to found out which one has the highest accuracy.

3.2. Automatic Clustering

One of the most famous methods in clustering is the classified method is k-means algorithm. The first step of K-means is initialization the number of cluster.

To find the global optimum of cluster number, people usually use the minimum value of cluster variance. From the cluster variance, finding the ideal cluster is very difficult because we can not apply directly min (V) to find the global optimum as the ideal cluster. For finding the global optimum of cluster construction and avoid the local optima, we use valley tracing method [7]. First of all we try to describe all patterns of the moving variance, then analyze the possibility of the global optimum that resides in the valley of patterns. Table 1 performs the possibility of the patterns to get the global optimum :

Table 1. Possibility of patterns to be a global optimum

Pattern	Possible ?	Pattern	Possible ?
	✓		X
	✓		X
	✓		X
	X		✓
	X		X
	X		X
	X		X
	X		

From analyzing the pattern in the Table 1, we can describe that the possibility to find the global optimum resides in stage which fulfilled:

$$V_{i-1} >= V_i \text{ and } V_{i+1} > V_i \quad (5)$$

where V_i is variance to i, for i=1..n, and n is latest stages of cluster construction.

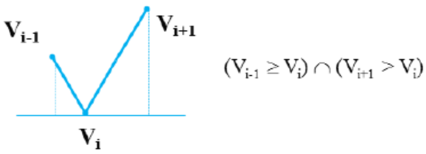


Figure 2. Different value of altitude

Then, we identify the different value of altitude ∂ for each stage.

$$\partial = (V_{i+1} - V_i) + (V_{i-1} - V_i)$$
$$= (V_{i+1} + V_{i-1}) - (2 \text{ x } V_i) \quad (6)$$

The global optimum can be obtained from maximum value of ∂. Accuracy of the valley tracing method can be acquired by defining :

$$\varphi = \frac{\max(\partial)}{\text{closer value to } \max(\partial)} \quad (7)$$

To obtain a reliable clustering process is having a minimum accuracy of 2.

So, automatic clustering is different with clustering. It concerns to find the number of cluster automatically. After getting the number of ideal cluster, the dataset will be grouped with a method of clustering, hierarchical clustering as a model performed on this case

4. System Architecture

The system architecture of the proposed research can be viewed on Figure 3. It explains the flow of finding earthquake density by collecting the earthquake dataset, finding the number of cluster with automatic clustering then measuring the density.

4.1 Spatial-Temporal Earthquake Dataset

The Earthquake datasets were collected in Indonesia and Japan from 1960 to 2012 with at least magnitude 6 on the Richter. We obtain the datasets from ANSS (Advanced National Seismic System). The datasets are displayed as a collecting of points, each have the value of longitude variable determining the position on the horizontal axis and the value of latitude variable determining the position of vertical axis.

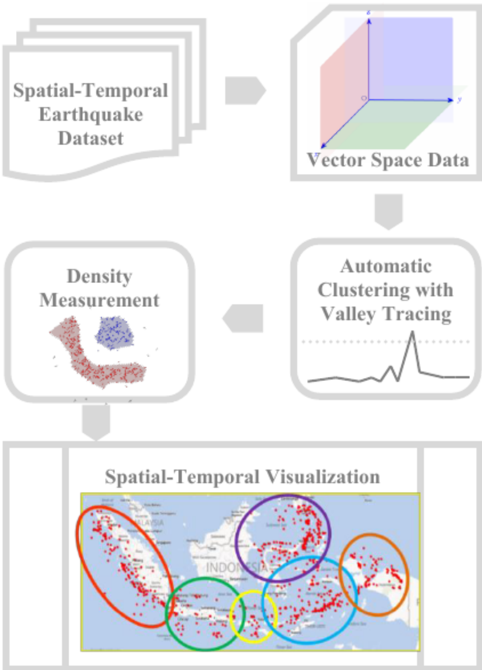


Figure 3. System Architecture

4.2 Vector Space Data

The datasets consist of location coordinates (longitude and latitude) and magnitude of the earthquake. Table 1 display example of some earthquake dataset in Indonesia.

Table 2. Vector space data

Latitude	Longitude	Magnitude
•-7.000000	•117.400002	•6.20
•-9.000000	•121.100000	•6.00
•-9.200000	•112.400000	•6.10
•2.500000	•128.800000	•6.30
•-8.200000	•115.100000	•6.10
•...	•...	•...

4.3 Valley Tracing

The first step to detect number of cluster automatically is using hierarchical clustering for grouping the earthquake datasets with number of cluster 1 to 31. With the result of clustering, we calculate the variance of each cluster.

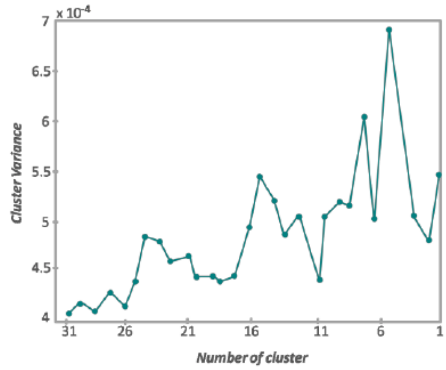


Figure 4. Cluster variance of earthquake in

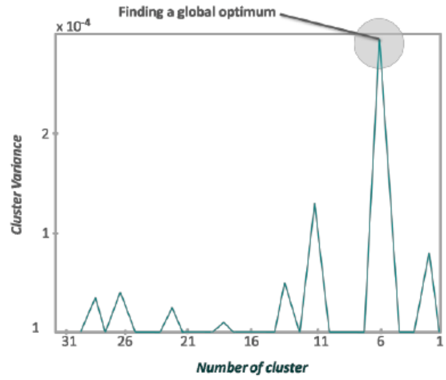


Figure 5. Variance Differentiation of earthquake in Indonesia from 1960-2012 with Mag ≥ 6

The variance results are visualized on Figure 4. The moving variance above has several similarity pattern with possibility of patterns to be a global optimum on table 3.

Based on Equation 6, we figure out the value of ∂. The global optimum of number cluster can be acquired from the maximum of ∂.

Finally, we have already gotten 6 as a cluster number for the earthquake datasets in Indonesia.

4.4 Density Measurement

This research measures the earthquake density with utilizing the wide of earthquake area and earthquake intensity. As we know that clusters usually have different shape of their spread points. But most of them have shape disposed like a circle.

The first step for measuring the density is determining the wide of earthquake clusters. Because of finding the global shape of cluster is difficult, so the counting is based on picture of regional area.

Amin Endah Sulistiawati, Ali Ridho Barakbah, Tri Harsono, Yuliana Setyowati, "Earthquake Density Measurement using Automatic Clustering",
The Third Indonesian-Japanese Conference on Knowledge Creation and Intelligent Computing (KCIC) 2014, March 25-26, 2014, Malang, Indonesia.

ISBN : 978-602-9494-97-6

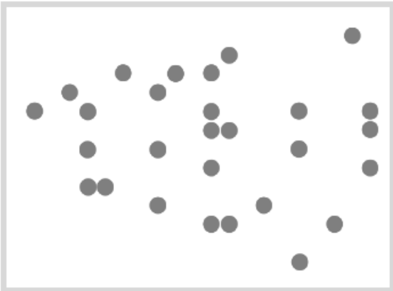


Figure 6. Cluster spread points on grid visualization

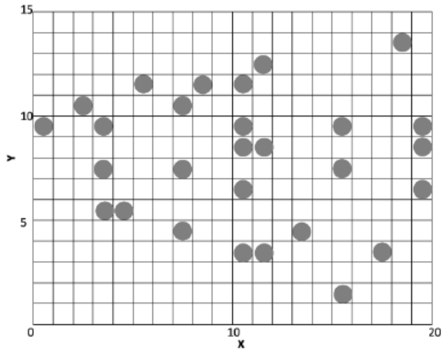


Figure 7. The dataset on grid

The spread points have random shape, so we make a polygon first from the outermost points. To create a polygon, here are some steps:

1. Define the range of grid.

   In this case we use 1x1 of latitude and longitude coordinate on each grid's cell.
2. Detect the grid length by decreasing the maximum and minimum value of earthquake dataset longitude on each cluster.
3. Round the dataset into integer to simplify the process of finding the cluster area.
4. Insert all dataset of each cluster on grid.
5. Check the grid status

   It will be difficult if the grid does not have any point or only has a point. There are 3 categories of grid status. First is status OK, if the grid has minimum 1 point on the grid minimum or the grid maximum and minimum has 2 point on other grid. The grid minimum and the grid maximum as the limit the cluster shape. The second is status when the grid only has a point besides the grid minimum and the grid maximum. The third is status when the grid does not have any point.
6. Detect the minimum and maximum limit of grid column.

   The minimum and maximum limit of grid column was detected by decreasing maximum and minimum value of earthquake dataset latitude on each grid. If on the mid of grid only has a point, it has to find out the point is the minimum or maximum values by comparing with the maximum and minimum value of its neighbored which has status grid OK. If it closes to the maximum value or the opposite if it is not.
7. Create new points

   If there is only a point or no one, we have to create other point by calculating the average point of the neighbors. Figure 8 gives the explanation clearly. The ▲ point are marked as a new data when the grid column only has a point, the ♦ point when the grid column doesn't have any data and the red ones is the old data.

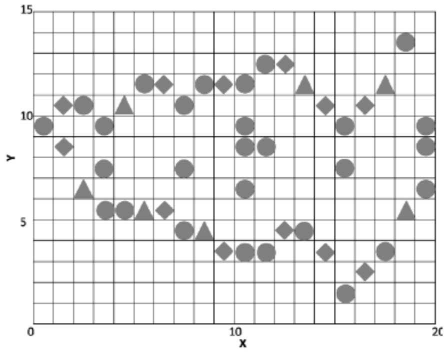


Figure 8. Create new data when the grid column only has a point or no point

After finishing the step 7, we back to step 5. The system has to recheck the status of each grid column after the new points were created. Repeat the step 6 and 7 until we get all of the status of grid column are OK. Then we can get the shape of the cluster. Finally the polygon was created. As example, it is on the Figure 9.

When it finished of creating a polygon, we measure the wide using Reimann method.

$$L = h \sum_{i=0}^{n} y_i \quad (10)$$

L = wide
h = range of grid
n = grid length
y_i = number of data on each grid

Because we use 1^o as the length of grid cell, so we have to convert the degrees into kilometers.

The second step is measuring the density is find the number of points as a number of populations on each cluster.

The last step is measuring the cluster density. It is alike the formula of human population density, the cluster density value can be found out by dividing wide of cluster and number of cluster. So the density is based on points per km^2.

Amin Endah Sulistiawati, Ali Ridho Barakbah, Tri Harsono, Yuliana Setyowati, "Earthquake Density Measurement using Automatic Clustering",
The Third Indonesian-Japanese Conference on Knowledge Creation and Intelligent Computing (KCIC) 2014, March 25-26, 2014, Malang, Indonesia.

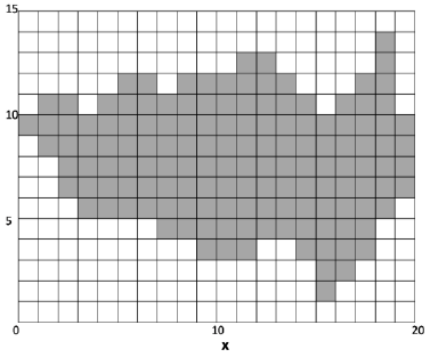


Figure 9. The polygon shape of cluster

4.5. Spatial-Temporal Visualization

This research uses Google maps for visualizing the earthquake datasets. The data needed to display the earthquake points are longitude for the x-axis and latitude for the y-axis on the map.

Figure 10. Earthquake visualization in Indonesia from ANSS catalog during 1960 to 2012

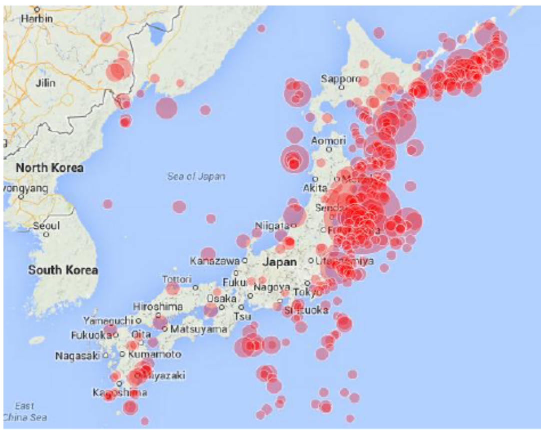


Figure 11. Earthquake visualization in Japan from ANSS catalog during 1960 to 2012

The points are depicted as colored circles with scaling size based on the magnitude. Moreover, the data must be prepared is the boundary coordinates of all provinces in Indonesia, the prefecture in Japan. The number of provinces in Indonesia is 34 and 47 for Japan prefectures. These data are needed to determine which provinces and prefecture involved in each cluster.

To connect the points of earthquake coordinate, it was conducted by finding the right position of each points compared to the minimum to maximum latitude and longitude of each provinces and prefectures.

In addition to spatial data, temporal data is also used. That is about the years when the earthquake occurred. This temporal data can be set the start and the end which ANSS catalog started in 1960 to 2012.

5. Experiment with Indonesia and Japan Dataset

The proposed idea was examined with earthquake catalog from ANSS for Indonesia and Japan as the study case.

5.1. Indonesia

The first experiment used earthquake dataset in Indonesia. The dataset is starting date from January 01, 1960 to December 31, 2012, minimum latitude in Indonesia from -11 to 6, minimum longitude 95 to 14, and minimum magnitude 6.0 to 10 on the Richter.

The result obtained from centroid linkage is number 6 for the ideal cluster and different number of membership. This result is trustworthy because it has 2,43 of accuracy using centroid linkage method. Otherwise, this clustering has 0,000467 of variance.

Based on the result, earthquake points depicted on the Google maps are marked with difference colors from a cluster to others. Here are the visualization of earthquake cluster:

Figure 12. Earthquake visualization in Indonesia with 6 number of cluster

Table 3. describes the information of each cluster after we finished with automatic clustering. The number of point means that each cluster has different members because the grouping of dataset based on the closest distance of earthquake points. The variance shows the spread of points. The greater value of variance shows that the spread of point is greater too.

Table 3. Number of clusters and Variance in Indonesia

	Cluster 1	Cluster 2	Cluster 3
Number of Point	40 points	165 points	210 points
Variance	7,54751	13,1103	10,6519
	Cluster 4	**Cluster 5**	**Cluster 6**
Number of Point	140 points	111 points	47 points
Variance	17,1259	8,26487	7,42128

The measuring of earthquake density is on Table 4. The wide area of each cluster has by 1° of latitude and longitude or it has an approach to 12.248,7 km^2 on each grid cell.

Table 4. Earthquake density about Indonesia on each cluster

Cluster	Number of Points	Wide Area (km^2)	Density (n/ km^2)
1	40	293.976	1,36 x 10^{-4}
2	72	881.928	1.87 x 10^{-4}
3	66	808.434	**2.60 x 10^{-4}**
4	60	734.940	1.90 x 10^{-4}
5	45	551.205	2.01 x 10^{-4}
6	33	404.217	1.16 x 10^{-4}

The outcome from visualization earthquake points, here are Indonesia provinces included in each cluster, as shown at table 5.

Table 5. List of Indonesia Provinces on each cluster

Cluster 1 **Priority 5**	West Nusa Tenggara, South Sulawesi, Bali, East Java, East Nusa Tenggara
Cluster 2 **Priority 4**	East Nusa Tenggara, Maluku, South Sulawesi, Southeast Sulawesi, North Maluku, Special Region of West Papua
Cluster 3 **Priority 1**	North Maluku, North Sulawesi, Central Sulawesi, South Sulawesi, East Kalimantan, West Sulawesi, Gorontalo, Special Region of West Papua, Southeast Sulawesi
Cluster 4 **Priority 3**	Bengkulu, South Sumatra, Special Region of Aceh, Jambi, West Sumatra, North Sumatra, Lampung
Cluster 5 **Priority 2**	Special Region of Papua, Maluku, Special Region of West Papua
Cluster 6 **Priority 6**	Banten, Lampung, West Java, Central Java, East Java, South Sumatra, Special Region of Yogyakarta

The density measurement gives a result that cluster 3, it was on North Maluku, North Sulawesi, Central Sulawesi, South Sulawesi, East Kalimantan, West Sulawesi, Gorontalo, Special Region of West Papua, and Southeast Sulawesi province that have the highest density. This condition means that these areas have a huge potential of earthquake hazard.

As we know that Indonesia's government is prefer to give more attention on Special Region of Aceh and Yogyakarta because this area was occurred large earthquake before. But they have not given intention on Maluku and Sulawesi that have a huge potential of earthquake based on these research.

5.2. Japan

The second experiment used earthquake dataset in Japan. The difference from Indonesia is only the boundary of coordinate. It used minimum latitude from 30 to 46 and minimum longitude 128 to 149. It was different with Indonesia's experiment because we use complete linkage method for Japan.

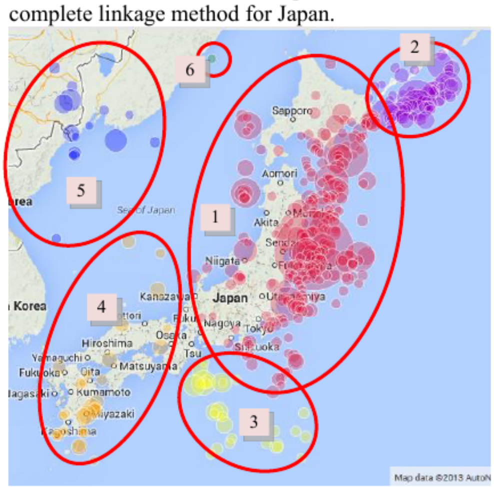


Figure 13. Earthquake visualization in Japan with 6 number of cluster

Complete linkage method obtained 6 of cluster number with 1,98 of accuracy. Otherwise, this clustering has 0,0019475 of variance.

Table 6. Number of clusters and variance in Japan

	Cluster 1	Cluster 2	Cluster 3
Number of Point	275 points	80 points	28 points
Variance	8,944	1,2125	4,4592
	Cluster 4	**Cluster 5**	**Cluster 6**
Number of Point	31 points	15 points	1 point
Variance	7,028	3,8643	1

Table 6. describes the information of each cluster after we finished with automatic clustering. The measuring of earthquake density is on Table 6. The wide

area of each cluster has by 1^o of latitude and longitude or it has an approach to 9.713,5 km^2 on each grid cell. The number of each 1o has is different with Indonesia, because the coordinate of Japan is in the north.

Table 7. Earthquake density about Japan on each cluster

Cluster	Number of Points	Wide Area (km^2)	Density (n/ km^2)
1	275	699.336	3.93 x 10^{-4}
2	12	116.556	6.86 x 10^{-4}
3	17	165.121	1.70 x 10^{-4}
4	27	262.251	1.18 x 10^{-4}
5	15	145.695	1.03 x 10^{-4}
6	1	9.713	1.03 x 10^{-4}

The outcome from visualization earthquake points, here are Japan prefectures included in each cluster:

Table 8 : List of Japan Prefectures on each cluster

Cluster 1 **Priority 2**	Hokkaido, Niigata, Gifu, Iwate, Chiba, Nagano, Akita, Aomori, Ibaraki, Miyagi, Yamagata, Ishikawa, Shizuoka, Yamanashi, Fukushima
Cluster 2 **Priority 1**	Hokkaido
Cluster 3	-
Cluster 4 **Priority 3**	Hyogo, Kyoto, Shiga, Ehime, Miyazaki, Oita, Kagoshima, Shimane, Fukuoka, Nagasaki, Kumamoto
Cluster 5	-
Cluster 6	-

Cluster 3, 5, and 6 do not have any list of prefecture because these areas are outside of Japan. The density measurement gives a result that cluster 2, it was on Hokkaido prefecture has the highest value. This condition means that these areas have a huge potential of earthquake hazard.

6. Conclusion

Indonesia and Japan are both located on tectonic plate boundaries. This condition makes both countries has a high priority of natural hazard and disaster prevention.

There are many methodologies to fine the earthquake-prone areas, such as intensity. This research proposes a new approach for measuring the density of earthquake. Earthquake density is different with earthquake intensity that only involves the number of earthquake even on an area. The density is measured after performing automatic clustering in areas affected by the earthquake.

This idea implements valley tracing method to detect the number of group spatial regions automatically from analysis of cluster moving variances. The Valley tracing method gives difference result for Indonesia and Japan. We choose centroid linkage for Indonesia and complete linkage for Japan as the highest value of accuracy.

Measuring the cluster density with a new approach gives the result as we measure the population density. First we have to count the member of each cluster as the population, and then we find the wide of cluster area. We have to convert the wide of cluster area by multiplying with 12.248,7 km^2 for Indonesia and 9.713,5 km^2 for Japan. The earthquake density was found out by dividing the population with the wide of cluster area.

The result of these research gives information that the areas which the highest potential of earthquake are in North Maluku, North Sulawesi, Central Sulawesi, South Sulawesi, East Kalimantan, West Sulawesi, Gorontalo, Special Region of West Papua, and Southeast Sulawesi province in Indonesia. Whereas in Japan, the highest potential is in Hokkaido prefecture. It means, the areas have the right to get concern from government about early warning of earthquake

References

[1] Sadeghian R., G. R. Jalali-Naini. 2008. A New Probability Density Function in Earthquake Occurences. Journal of Industrial Engineering International, Islamic Azad University, Vol. 4, No. 6, 1-6. South Tehran Branch

[2] Rusnardi Rahmat Putra, Junji Kiyono, and friends. 2012. Seismic Hazard Analysis for Indonesia. Journal of Natural Disaster Science, Vol. 33, No. 2, pp59-70. Japan

[3] Das S., C. Henry. 2003. Spatian Relation Between Main Earthquake Slip and Its Aftershock Distribution. Department of Earth Sciences, University of Oxford. Oxford, UK

[4] Masryur Irsyam, Donny T Dangkua, and friends. 2008. Proposed Seismic Hazard Maps of Sumatra and Java Islands and Microzonation Study for Jakarta City, Indonesia. U.S Geological Survey USGS-MS966. USA

[5] Fujiwara Hiroyuki, Sinichi Kawai and friends. 2006. National Seismic Hazard Maps of Japan. University of Tokyo. Japan

[6] ANNS Catalog search. Retrieved on June 1st, 2013 from ncedc.org. :http://www.ncedc.org/ anss/catalog-search.html

[7] Kenji Satake and Henry Harjono.Multi-Disciplinary Hazard Reduction from Earthquakes and Volcanoes in Indonesia. Journal of Disaster Reaserch Vol.7 No.1,2012.

[8] Ali Ridho Barakbah, Kohei Arai.2004. Identifying Moving Variance to Make Automatic Clustering for

ISBN : 978-602-9494-97-6

Normal Data Set. Proceedings of the IECI Japan Workshop. Musashi Institute of Technology, Japan.

[9] Ali Ridho Barakbah, Kohei Arai. 2007. Cluster construction method based on global optimum cluster determination with the newly defined moving variance. Reports of the Faculty of Science and Engineering, Saga University, Vol. 36, No.1. Japan

[10] MacQueen, J. B. 1967. Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297

[11] Mooi E, Sarstedit M. 2011. The Process, Data and Methods Using IBM SPSS Statistics. Springer-Verlag Berlin Heidelberg, ISBN: 978-3-642-12540-9