

## Cluster Oriented Spatio Temporal Multidimensional Data Visualization of Earthquakes in Indonesia

Mohammad Nur Shodiq<sup>1</sup>, Ali Ridho Barakbah<sup>2</sup>, Tri Harsono<sup>3</sup>

Postgraduate Program on Information Engineering and Computer,  
Electronic Engineering Polytechnic Institute of Surabaya

<sup>1</sup>mohnoershodiq@gmail.com, <sup>2</sup>ridho@pens.ac.id, <sup>3</sup>trison@pens.ac.id

### Abstract

Spatio temporal data clustering is challenge task. The result of clustering data are utilized to investigate the seismic parameters. Seismic parameters are used to describe the characteristics of earthquake behavior. One of the effective technique to study multidimensional spatio temporal data is visualization. But, visualization of multidimensional data is complicated problem. Because, this analysis consists of observed data cluster and seismic parameters. In this paper, we propose a visualization system, called as IES (Indonesia Earthquake System), for cluster analysis, spatio temporal analysis, and visualize the multidimensional data of seismic parameters. We analyze the cluster analysis by using automatic clustering, that consists of get optimal number of cluster and Hierarchical K-means clustering. We explore the visual cluster and multidimensional data in low dimensional space visualization. We made experiment with observed data, that consists of seismic data around Indonesian archipelago during 2004 to 2014.

**Keywords:** Clustering, visualization, multidimensional data, seismic parameters.

### 1. INTRODUCTION

Earthquake is the event of the earth due to release of energy in the earth suddenly. It was caused by the sudden breaking a layer of rock or plate fracture in the earth's crust [1]. Indonesian archipelago is located in the world's three mega plates; they are Australian plate, Eurasian plate, and Pacific plate. The interaction between these plates, placed Indonesia as the territory that has high activity of volcanoes and seismic [2].

The high seismic activity could be seen from the recording within the period 1897-2009. There are more than 14,000 earthquakes with magnitude  $M > 5.0$ . For example, within the period 2004 to 2010 there were a lot of seismic activity in Indonesia, i.e. Aceh earthquake with tsunami of 2004 ( $M_w = 9.2$ ), the Nias earthquake of 2005 ( $M_w = 8.7$ ), the 2006 Yogyakarta earthquake ( $M_w = 6.3$ ), Lake earthquake of 2009 ( $M_w = 7.4$ ) and the

earthquake in Padang in 2009 ( $M_w = 7.6$ ). That Earthquakes have caused thousands of casualties, the collapse and destruction of infrastructure and buildings, and trillions of funds for rehabilitation and reconstruction [3]. The various potential natural disasters we should have to know, so that the character of these natural hazard impacts can be minimized [2].

Seismic sequences are not formatted randomly, but they follow a spatial pattern with consequent triggering of events. In other words, This events produce non random grouping [8]. So, the scientists are developing a model to explain this grouping pattern. This grouping well-known by clustering. By using a clustering scheme, it is possible to retrieve spatio-temporal pattern that created by events [4]. Therefore, it is required modeling of earthquake clustering more accurate to develop a model that explains the pattern or grouping behavior [5].

To explore the relevance of the cluster obtained, the clustering must be empowered with visualization. Because, clustering never gives consistent. Visualization of cluster is considered as one of the most intuitive methods for cluster detection and validation, and especially is performing well for the representation of irregularly shaped clusters [6]. The result of clustering can be employed for predicting the nature of the events in the future [5][8]. Clustering as a tool for analyzing seismic parameters.

## 2. RELATED WORKS

There are many researchers who studied and proposed the clustering and visualization of earthquake. Yuen (2005) [7] present a web client-server service for analyzing and visualization of seismic data consisting of both small magnitude events and large earthquakes. This analysis consists of clustering, feature generation, feature extraction techniques and rendering algorithms form a computational framework of this environment. Yuen (2009) [8] proposed visualization the computed results from clustering the higher-dimensional data. this clustering analysis uses both synthetic and real data. Synthetic data consists of 7 parameters. six parameters are based on seismicity rates and one parameter combines a function of the event sizes with a function of the seismicity rate. Dzwinel [4] proposed a new approach employing clustering for multivariate analysis of seismic data. It showed using the Amira visualization package. Dzwinel [9] proposes a novel multi-resolutional approach, which combines local clustering techniques in the data space with a non-hierarchical clustering in the feature space. It uses multidimensional scaling procedures for visualization of multi-dimensional events in 3-D space.

## 3. ORIGINALITY

In this paper, we propose a new approach for visualization the multidimensional data of seismic based on cluster, called as IES (Indonesia Earthquake System). This system is used for cluster analysis, spatio temporal analysis, and visualize the multidimensional data of seismic parameters. We

cluster the earthquake data into small data groups based on the optimal number of cluster. Then, spatial data distribution based on clusters are utilized to investigate the other seismic parameters each clusters. These parameters generate the multidimensional data of earthquake parameters i.e. *a value*, *b value*, and *Tr* (recurrence time). Afterward, we use linked multiple views to visualize the clusters and seismic parameters. To realise a multidimensional visualization modelling, we use mathematical theories on multidimensional description.

#### 4. SYSTEM DESIGN

The procedural concept of our proposed system can be seen in Figure 1.

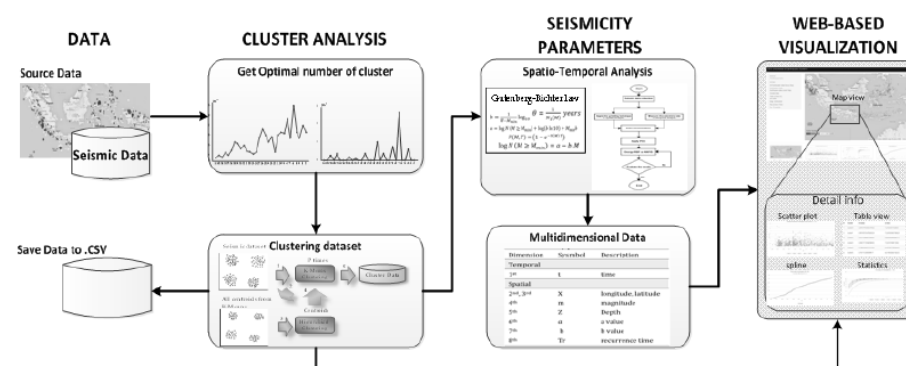


Figure 1. Block Diagram System

##### 4.1 Earthquake Data

The data representing seismic activities of the Indonesia. It was collected by the Meteorolgi, Climatology and Geophysics Agency (BMKG). In this research, the seismic dataset were collected during a 11 year time interval from 2004 to 2014. The earthquake data covers the entire territory of Indonesia, which is the boundary 6 north latitude - 11 south latitude and east longitude 95 - 141 east longitude, and depth of 0-650 km. The dataset consist of 2,737 seismic event with magnitude  $m \geq 3$ . The lowest magnitude were determined by using Magnitude of completeness ( $M_c$ ) of the Gutenberg-Richter frequency size distribution [6]. Magnitude of completeness ( $M_c$ ) also called as threshold or cut-off magnitude. So, seismic event with magnitude under the  $M_c$ , that will be eliminated [10] [11]. Incompleteness of seismic data will give the result in seismic risk parameters resulting into overestimated or underestimated. In this research, we use magnitude 5.1 for magnitude of completeness ( $M_c$ ) [10]. Table 1 and Tabel 2 show the Earthquake data specification and paramaters.

##### 4.2 Clustering Prosedure

Clustering is an effort to classify similar objects in the same groups [12]. Clustering analysis is a mathematical consept whose main usefull role is

to extract the most similar (or dissimilar) separated sets of objects according to a given similarity (dissimilarity) measure [6]. Cluster analysis constructs good cluster when the members of a cluster have a high degree of similarity to each other (internal homogeneity) and are not like members of other clusters (external homogeneity) [13].

Clustering procedure in this research consist of 2 phases : (1) to find global optimal number of cluster, (2) clustering the dataset. To find the global optimal number of clustering, we use Valley Tracing and Hill Climbing algorithm [13]. It analyzed the moving variance of clusters for each stage of cluster construction, then observed the pattern to find the global optimum as well as to avoid the local optima. The algorithm that to find optimal number of cluster is as follows:

1. Set as each data of  $A$ , where is attribute of  $n$ -dimensional vector.
2. Set  $K$  as the predefined number of clusters  $K_i \{k_i | i = 2, \dots, n-1\}$
3. Apply clustering algorithm with number of clusters  $K_i$
4. Record the calculation of variance as  $V_j \{v_j | j = 1, \dots, n-2\}$
5. Increment  $j=j+1$
6. Repeat from step 3 while  $j < n-2$ .
7. Calculate the moving variance with valley tracing -hill climbing  $\partial \{\partial_i | i = 1, \dots, n-1\}$
8. Put threshold  $\lambda$  on moving variance  $\partial$  for automatic clustering
9. Then, rank the moving variance  $\max(\partial) \geq \lambda$ , and  $i$  as optimal number of cluster

The second process is to cluster the dataset, in this research we use Single Linkage Hierarchical K-means clustering [14]. This clustering requires a number of cluster for real clustering seismic data. So that, a number of optimal cluster from first process becomes initial clusters for Hierarchical K-Means method [14]. Hierarchical K-means clustering is a combination of K-Means and hierarchical clustering [14]. The Hierarchical K-means clustering algorithm is as follows:

1. Set as each data of  $A$ , where is attribute of  $n$ -dimensional vector.
2. Set  $K$  as the predefined number of clusters.
3. Determine  $p$  as numbers of computation
4. Set  $i=1$  as initial counter
5. Apply K-means algorithm.
6. Record the centroids of clustering results as  $C_i \{c_{ij} | j = 1, \dots, K\}$
7. Increment  $i=i+1$
8. Repeat from step 5 while  $i < p$ .
9. Assume  $C_i \{c_i | i = 1, \dots, p\}$  as new data set, with  $K$  as predefined number of clusters
10. Apply hierarchical algorithm
11. Record the centroids of clustering result as  $D \{d_i | i = 1, \dots, K\}$
12. Then,  $D \{d_i | i = 1, \dots, K\}$  as initial cluster centers for K-means clustering.

**Table 1.** Total number of seismic events

Data source	$3 \leq m \leq 10$	$M_c \geq 5.1$
BMKG	2,737	1,334

**Table 2.** domain of physical paramaters

Dimension	Description	Interval and unit
X	-longitude -latitude	$6^0 \leq X_{(long)} \leq 11^0$ $95^0 \leq X_{(lat)} \leq -141^0$
Z	Depth	$0 < Z < 650$ km
t	Time	$2004 < t < 2014$
m	Earthquake magnitude	$5.1 \leq m \leq 10$

### 4.3 Seismicity Parameters

Seismicity parameters in this research is classified into two groups: spatial and temporal dimensions. Spatial and temporal variation in the seismicity parameters were reported and regarded as a precursory phenomenon for future major earthquake in different seismogenic areas [15]. Spatial dimensions are related to the size of earthquake magnitude, longitude, latitude, and depth [16]. Besides, in this paper, we cluster the earthquake data into small data groups based on the optimal number of cluster [5]. Afterward, spatial data distribution based on clusters are utilized to investigate the other parameters i.e *a value*, *b value*, and *Tr* (recurrence time). *a value* and *b value* are constants. The “a” describes the rate of seismic productivity for a given area, i.e., the higher the *a value*, the higher is the seismicity [15]. The *a value* can be defined as follows [5] [11].

$$a = \log N(M \geq M_{min}) + \log(b \ln 10) + M_{min} b \quad (1)$$

where,  $N(M \geq M_{min})$  is the cumulative number of earthquakes with magnitude equal or greater than  $M_{min}$  and “b” is parameter “b” or *b value*.

The parameter “b” characterized the relative size distribution of events and it widely varies from 0.5 to 1.8, depending on the stress regime and seismotectonic behavior of the region. The *b value* is calculated using the maximum likelihood technique [5][11][15][17][18]. The *b value* can be defined as follows:

$$b = \frac{1}{\bar{M} - M_{min}} \log_{10} e \quad (2)$$

where,  $\bar{M}$  denotes the mean magnitude and  $M_{min}$  is the minimum magnitude.

*Tr* is The probabilistic recurrence time. *Tr* for a shock with magnitude equal or greater than *M* [5] [15]. *Tr* can be defined as follows:

$$Tr = \Delta T / 10^{(a-bM)} \quad (3)$$

where, Where  $\Delta T$  is the length of observation period.

Temporal dimension is associated with natural frequencies of relevant behavior [19]. Whereas, temporal dimension in this paper is time. Table 3 show seismicity parameters and the dimension

**Table 3.** Seismicity parameters and the dimension

Dimension	Sysmbol	Description
Temporal		
1 <sup>st</sup>	t	time
Spatial		
2 <sup>nd</sup> , 3 <sup>rd</sup>	X	longitude, latitude
4 <sup>th</sup>	m	magnitude
5 <sup>th</sup>	Z	Depth
6 <sup>th</sup>	<b>a</b>	a value
7 <sup>th</sup>	<b>b</b>	b value
8 <sup>th</sup>	<b>Tr</b>	recurrence time

#### 4.3 Multidimensional Data Visualization

As shown in Fig. 1, the system consists of three phases. Phase 1: cluster analysis, this process is divided into 2 schemes. The first process is to find the global optimum of clustering using Valley Tracing [13]. It analyzed the moving variance of clusters for each stage of cluster construction, then observed the pattern to find the global optimum as well as to avoid the local optima. The second process is to cluster the dataset using Single Linkage Hierarchical K-means clustering [14]. This clustering requires a number of cluster for real clustering seismic catalog. So that, a number of optimal cluster from first process becomes initial clusters for Hierarchical K-Means method [14].

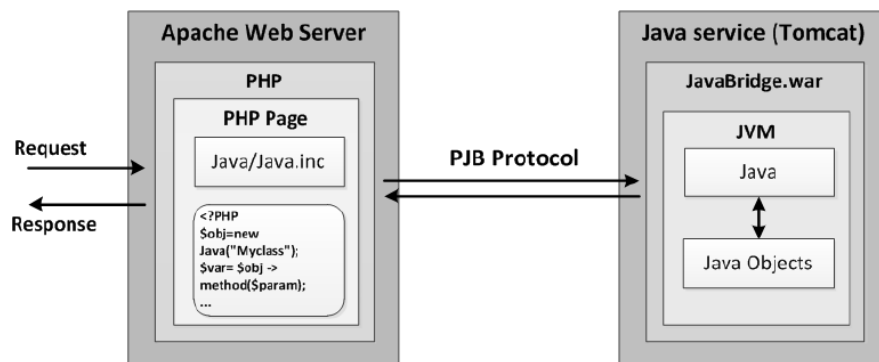
Phase 2: Seismicity parameter. Seismicity parameters could facilitate a better understanding of the earthquake mechanism and provide some prediction capability. In many cases, variation in the seismicity parameters might be correlated to earthquake occurrence and regarded as a precursory phenomenon for future strong main shocks [15]. In this phase, we use spatio-temporal analysis. This process is to analyze the spatial data distribution. In this case, we use the Gutenberg and Richter formula for get the parameters, that are: *a value* and *b value*, and *Tr*. Besides the others, we use dataset. i.e time, longitude, latitude, magnitude, and depth [5]. As shown in Table. 3, Seismicity parameter consists of 8 parameters, they are divided in to two groups, namely temporal dimension (1D) and spatial dimension (7D).

Phase 3: Visualization. The method of visualization in this phase, we use linking multiple views modelling. In this method, a user should be able to pose a query graphically, and a computer should be able to present the response graphically as well. Both query and response should occur in the same visual field [20]. To realise a multidimensional visualization modelling, we use mathematical theories on multidimensional description.

Multidimensional objects are built from their lower primitives, i.e. 3D volume object consists of 2D and 1D [21].

In Figure 2 shows the software infrastructure of the processing, namely PHP/JavaBridge. The PHP/Java Bridge is an implementation of a streaming, XML-based network protocol, which can be used to connect a native script engine, for example PHP, Scheme or Python, with a Java virtual machine. It is faster and more reliable than direct communication via the Java Native Interface, and it requires no additional components to invoke Java procedures from PHP or PHP procedures from Java.

We use PHP/JavaBridge package for analysing and visualization. This package consists of two programming scheme. Firstly, java program. This program including: to get the optimal number of cluster, earthquake clustering, spatial analysis (to get seismicity parameters), save data to CSV format, and save data to JSON format. Secondly, PHP program. This program focus on display the result. It provides the front end Graphical user interface under web page. Users can choose the navigation panel works like remote control for displaying the seismicity parameters. The visualization consists of maps, scatter plot, table views, and spline.



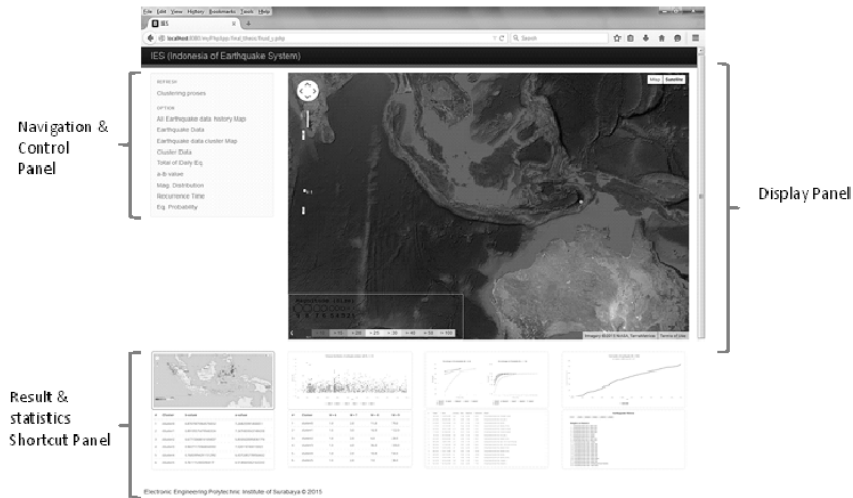
**Figure 2.** Desain System PHP/JavaBridge Software Scheme

## 5. EXPERIMENT AND ANALYSIS

In this paper, we use IES (Indonesian Earthquake System) toolkit. This toolkit provides cluster analysis and visualization system. IES can be seen in Figure 3.

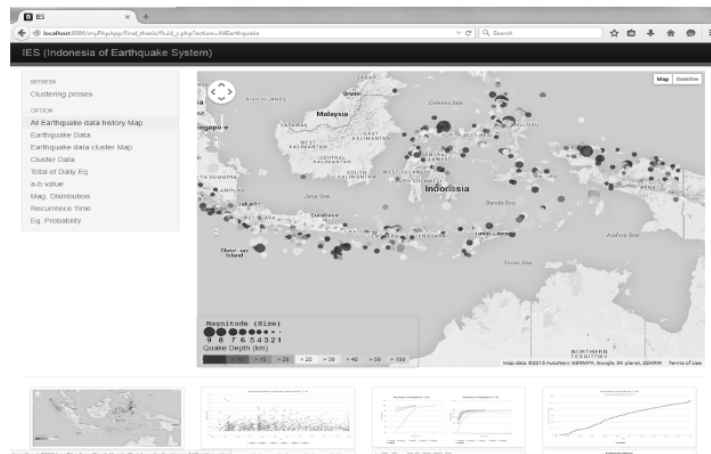
In IES toolkit there are 3 panels: 1) Navigation and control panel, this panel consists of clustering process and navigation. Clustering process allows user for cluster the dataset. Whereas, the navigation allows user for display the result of clustering, original data, seismic parameters (i.e. *a value*, *b value*, and *Tr*) and statistic prediction. This visualization displays as maps, *scatter plot*, *spline chart*, and *table views*. 2) Result and statistics shortcut panel, this panel allows user to observe in large view. 3) Display panel, This panel displays what the user choose in navigation and control panel (i.e. clustering process, All Earthquake data history Map, Earthquake Data,

Earthquake data cluster Map, Cluster Data, Total of Daily Eq, *a-b value*, Mag. Distribution, Recurrence Time, Eq. Probability).



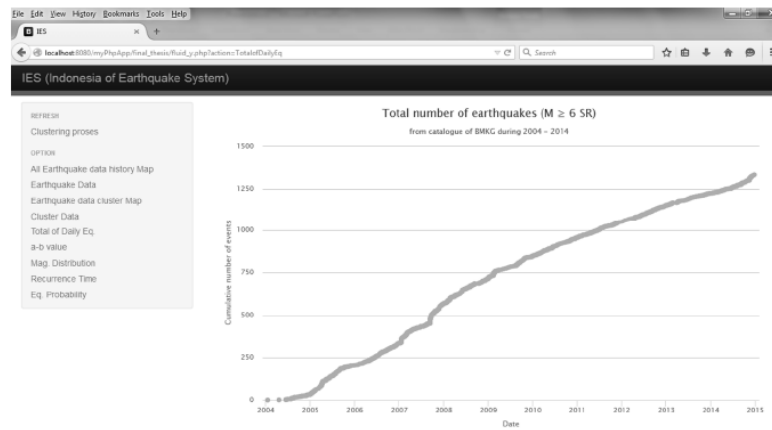
**Figure 3.** IES (Indonesia Earthquake System) toolkit

We analyze observed earthquake catalogs in 11 years time. The observed data represents seismic activities of Indonesia, that collected by the Meteorologi, Climatologi and Geophysics Agency (BMKG). The seismic events shown in Figure 4, were detected during time interval 2004 to 2014. The data set consists of 1,334 seismic events with magnitudes  $m$ , position in space (latitude  $X$ , longitude  $Y$ , depth  $z$ ), and occurrence time  $t$ . The magnitude of the earthquake and their depth are represented by differences of the radius of the circle and color. All seismic events information could be seen in Figure 6.

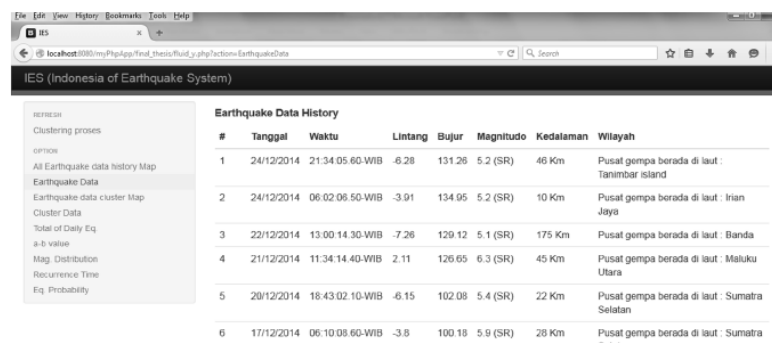


**Figure 4.** Seismic events around the Indonesian Archipelago from catalogue of BMKG during 2004 – 2014

The completeness of the catalog could be seen in Figure 5. We study the cumulative curve of the events with magnitud  $M \geq 5.1$  since 2004 until 2014 and the number of 1,334 seismic events. The result of curve show that the catalog is complete. Because, we assume that the cumulative curve is rectilinear. As support of that, we report in Figure 5.



**Figure 5.** Cumulative number of events versus time with magnitude  $\geq 5.1$  from catalogue of BMKG during 2004 - 2014



**Figure 6.** Total of earthquake data history with magnitude  $\geq 5.1$  from catalogue of BMKG during 2004 - 2014

Clustering results shown in Figure 7. In this visualization, colors represent clustering label and in space (X, Y coordinate) represent latitude and longitude. The seismic events are visualized with the PHP/JavaBridge and Google Map. In this paper, the seismic events are divided into 6 clusters. This clustering is based on the process of hierarchical kmeans clustering model with complete linkage scheme. The data members of each cluster can be seen in Figure 8. Naming the area of the seismic events is based on the distance offshore with the nearby area. Thus, the name of the region can be classified as more than one cluster.

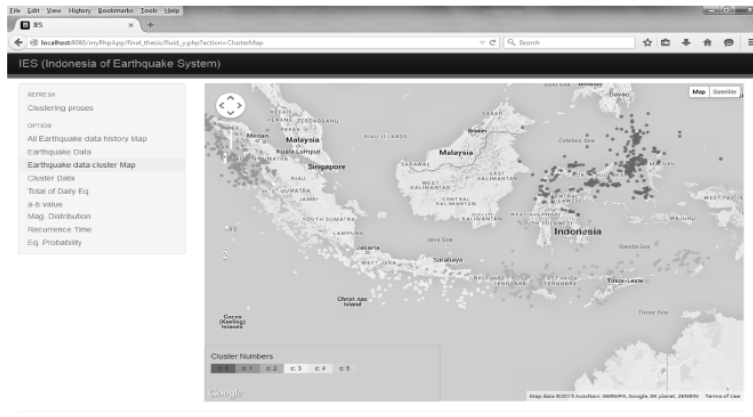


Figure 7. Seismic distribution based on cluster

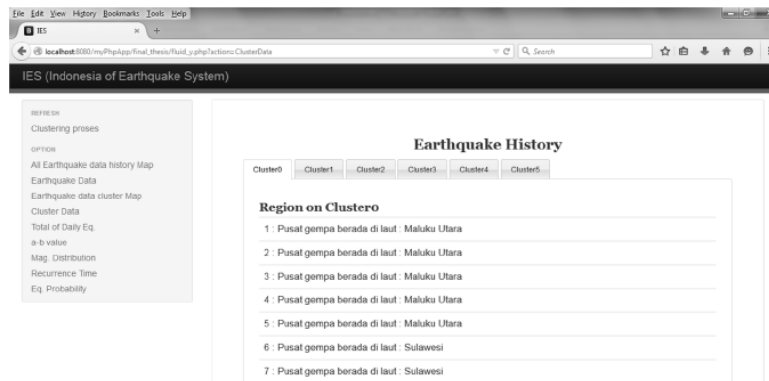


Figure 8. The data members each cluster

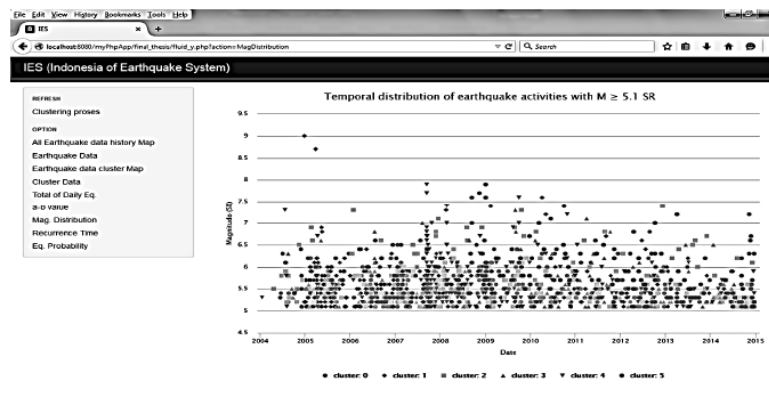
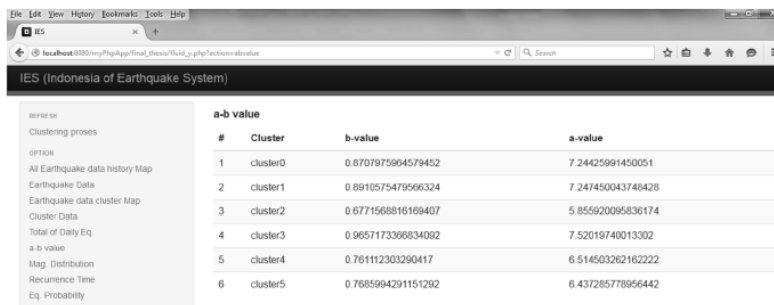


Figure 9. Temporal distribution of earthquake activities with magnitude  $\geq 5.1$  from 2004 to 2014 in Indonesia

The time sequence for earthquake in Indonesia with magnitud greather than 5.1 from 2004 to 2014 is shown in Figure 9. Seismic events that frequently occurs are on a scale of magnitude between 5.1 and 6.5, while on magnitude 6.5 to 7, the frequently will decreasing. Whereas, earthquakes are rarely occur at magnitud more than 7.

In Figure 10 show seismic parameters, *a value* and *b value* each cluster. The *a value* indicates seismicity in region and *b value* depends on the stress regime and tectonic character of the region. Variation of *b value* in a region is commonly correlated with characteristics of regional seismic activitis [22]. These parameters are importance for prediction of the hazard in the areas. In Figure 10 indicates "*b*" values view varying from 0.67 to 0.96. Whereas, "*a*" values view varying from 5.85 to 7.52. Cluster region which has lowest *b value* is cluster2, it indicates that, cluster2 region has a high levels of stress. So, cluster2 has higher potential that will be earthquake. Because, the area that has low *b value* means it has the accumulated stress that has not been released. And otherwise, if cluster has high *b value*. While, low *a value* indicates that the region has low seismic activities. Cluster region which has lowest *a value* is cluster2 that is 5.95, but the highest *a value* in cluster3 that is 7.52.

The result of return time period each cluster could be seen in Figure 11. Level of earthquake risk with a magnitude  $M \geq 6$  each cluster has return period of 1 year, it means that the occurrence of this earthquake will occur every year. In other word, every a year there will be an earthquake or more in that area. While, the earthquake with magnitude  $M \geq 7$  has return period of 2 years for cluster0, cluster2, cluster4, and cluster5. Then, cluster1 and cluster3 have return period of 3 years and 4 years. Level of earthquake risk with magnitude  $M \geq 8$  has different level of risk each cluster. on cluster0, cluster1, cluster2, cluster4, and cluster5 have a return period between 6 to 16 years, whereas in cluster3 has a return period of 36 years. While, level of earthquake risk with a magnitude  $M \geq 9$  on cluster2 has an earliest return period, that is 28 years, while the cluster that has the longest return period is cluster3 which has 330 years of return period.



a-b value			
#	Cluster	b-value	a-value
1	cluster0	0.8707975964579452	7.24425991450051
2	cluster1	0.8910575479566324	7.247450043748428
3	cluster2	0.6771568816169407	5.855920095836174
4	cluster3	0.9657173366834092	7.52019740013302
5	cluster4	0.761112303290417	6.514503262162222
6	cluster5	0.7685904291151292	6.437285778956442

Figure 10. *a value* and *b value* each cluster

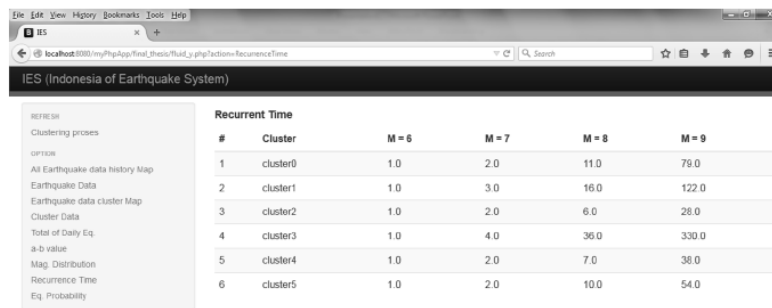


Figure 11. Tr (recurrent Time) each cluster

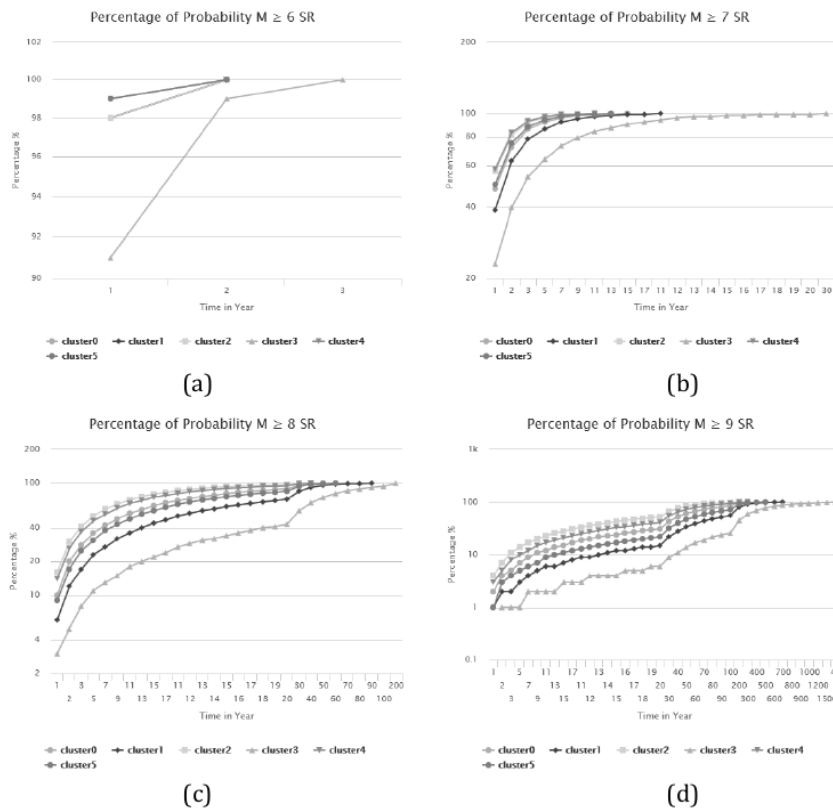


Figure 12. Percentage of probability earthquake occurrence: a) magnitude  $M \geq 6$ , b) magnitude  $M \geq 7$ , c) magnitude  $M \geq 8$ , d) magnitude  $M \geq 9$ .

Percentage of probability earthquake occurrence can be shown in Figure 12. In Figure 12a shows that, the percentage of probability with magnitude  $M \geq 6$  every clusters are different. In 1 year, the percentage of probability on cluster4 and cluster5 are 99%. Then, they will be 100% within 2 years. It means that, there will be one or more earthquake occurrences with magnitude greater than  $M \geq 6$  within 2 years. While, the

percentage of probability on cluster0, cluster1 and cluster2 in 1 year are 98%. Then, they will be 100% within 2 years. Whereas, percentage of probability on cluster3 is 91% in 1 year. Then, percentage of probability in 2 years is 99%. Afterword, it will be 100% in 3 years.

In Figure 12b. shows the percentage of probability with magnitude  $M \geq 7$ . Cluster0, cluster2, cluster4, and cluster 5 has percentage of probability in range time 8 years to 12 years to get the position on 100%. While, cluster3 has time 40 years for get 100%. In Figure 12c. Shows the percentage of probability with magnitude  $M \geq 8$ . The percentage of probability in 100%, it needs 50 years to 300 years. In Figure 12d. Shows the percentage of probability with magnitude  $M \geq 9$ . The percentage of probability in 100%, it needs 300 years to 3000 years.

## 6. CONCLUSION

We propose approach a visualization system for visualizing multidimensional data of seismic parameters and visual clustering. We have described visualization based on the concept of multiple linking views and data analysis. Data analysis means that data proccessing. It is consists of cluster analysis and spatio temporal analysis. The results of data proccessing will be visualized in multiple linking views. Multiple linking views mean that all of visualization are linked.

We demonstrate the visualization of cluster analysis and spatio temporal analysis. We show that user can operate clustering seismic data, then the result of cluster data could be visualized directly. Afterword, the result of clustering will be utilized to investigate the seismic parameters. We also show the probability of earthquake occurrences. This system uses 2-D space for visualize the multidimensional data, because human eye can decide whether the result is optimal or not.

Based on percentage of probability earthquake occurrence in Figure 12a through Figure 12d, there is no earthquake damage within a period of 8 to 12 years. In other words, there is no earthquake damage with magnitud more than 7 (SR) in 2022. While, earthquake occurrence with magnitude  $M \geq 8$  and ,  $M \geq 9$  will accur within the 50 and more 300 years.

cluster2 has a high level of earthquake risk , its mean that return period on this cluster is more short time than the others. In this cluster would accur earthquake with magnitud 6 (SR) every year. Whereas, magnitud 7 (SR), which can be categorized as highly damaging seismic event, would accur earthquake every 2 years. And also, it has a return time period 6 years on magnitude  $M \geq 8$  and 28 years on magnitude  $M \geq 9$ . Whereas, the cluster whose has lowest level of earthquake risk is cluster3. It has more long time of the return period than the others.

## REFERENCES

- [1] BMKG Sanglah Denpasar, **Geodinamika Informasi Meteorologi Klimatologi dan Geofisika**, *BMKG Sanglah Denpasar*, Vol.2, No.11, 2013.
- [2] \_\_\_\_\_, **Analisis Potensi Rawan Bencana Alam di Papua dan Maluku (Tanah Longsor - Banjir - Gempa Bumi - Tsunami)**, *Laporan Akhir*, Deputi Bidang Pembinaan Sarana Teknis dan Peningkatan Kapasitas, Kementerian Negara Lingkungan Hidup, Jakarta, 2007.
- [3] Irsyam, Masyhur. Dkk., **Ringkasan Hasil Studi Tim Revisi Peta Gempa Indonesia 2010**, *Tim Revisi Peta Gempa Indonesia*, Bandung, 2010.
- [4] Dzwinel, W. Yuen, D. Kaneko, Y. Boryczko, K. Ben-Zion, K., **Multi-resolution clustering analysis and 3-D visualization of multitudinous synthetic earthquakes**, *Vis Geosci*, Vol. 8, pp. 12-25, 2003.
- [5] Mohammad Nur Shodiq, Ali Ridho Barakbah, Tri Harsono, **Spatial Analysis of Earthquake Distribution with Automatic Clustering for Prediction of Earthquake Seismicity in Indonesia**, *The Fourth Indonesian-Japanese Conference on Knowledge Creation and Intelligent Computing (KCIC) 2015*, Surabaya/Bali, Indonesia, March 24-26, 2014.
- [6] Yuen, Dave A., Benamin J.K., Evan F.B., Dzwinel W., Zachary A.G., Cesar R.S., **Clustering and visualization of earthquake data in a grid environment**, *Visual GeoScience*, 2005.
- [7] Yuen, D.A. Benjamin, J.K. Bollig, E.F. Dzwinel, W. Garbow, Z.A. Silva, **Clustering and visualization of earthquake data in a grid environment**, *Vis Geosci*, Vol. 10, pp. 1-12, 2006.
- [8] Yuen, D.A., Dzwinel, W., Ben-Zion, Y., Kadlecd, B., **Visualization of Earthquake Clusters over Multidimensional Space**, *Encyclopedia of Complexity and Systems Science*, pp 2347-2371, 2009.
- [9] Dzwinel, Witold dkk., **Cluster Analysis, Data-Mining, Multi-dimensial Visualization of Earthquakes Over Spase, Time and Feature Spase**, *Earth and Planetary Sci. Letter*, 2003.
- [10] Rohadi, Supriyanto, **Studi Seismotektonik Sebagai Indikator Potensi Gempa bumi di Wilayah Indonesia**, *Jurnal Meteorologi dan Geofisika*, Vol. 10, No. 2, pp. 111-120, 2009.
- [11] Sunardi, Bambang, **Analisa Fraktal dan Rasio Slip Daerah Bali-NTB Berdasarkan Pemetaan Variasi Parameter Tektonik**, *Jurnal Meteorologi dan Geofisika*, Vol. 10, No.1, 2009.
- [12] Ali Ridho Barakbah, Kohei Arai, **Centronit: Initial Centroid Designation Algorithm for K-Means Clustering**, *EMITTER International Journal of Engineering Technology*, Vol. 02, No. 01, 2014.
- [13] Ali Ridho Barakbah, Kohei Arai, **Determining constraints of moving variance to find global optimum and make automatic clustering**,

- Proc. Industrial Electronics Seminar (IES) 2004*, Surabaya, Indonesia, pp. 409-413, October 12, 2004.
- [14] Kohei Arai, Ali Ridho Barakbah, **Hierarchical K-means: an algorithm for centroids initialization for K-means**, *Reports of the Faculty of Science and Engineering*, Saga University, Japan, Vol. 36, No. 1, 2007.
  - [15] Zamani, ahmad. Sorbi, Reza M., **Application of neurol network and ANFIS model for earthquake occurrence in Iran**, *Earth Sci Inform*, DOI:10.1007/s12145-013-0112-8, Springer-Verlag Berlin Heidelberg, 2003.
  - [16] N. Andrienko, G. Andrienko, P. Gatalsky, **Exploratory spatio-temporal visualization: an analytical review**, *Journal of Visual Languages and Computing*, Vol. 14, pp. 503–541, 2003.
  - [17] E. Florido, F. Martínez-Álvarez, A. Morales-Esteban, J. Reyes, J.L. Aznarte-Mellado, **Detecting precursory patterns to enhance earthquake prediction in Chile**, *Computers and Geosciences*, <http://dx.doi.org/10.1016/j.cageo.2014.12.002>, 2014.
  - [18] Rohadi, Supriyanto, dkk., **Studi Variasi Spatial Seismisitas Zona Subduksi Jawa**, *Jurnal Meteorologi Dan Geofisika*, Vol. 8, No.1, Juli, 2007
  - [19] Grijalva, S., **Multi-Dimensional, Multi-Scale Modeling and Algorithms for Integrating Variable Energy Resources in Power Networks: Challenges and Opportunities**, Published by Elsevier Inc., 2014.
  - [20] Buja, A. Cook, D. Swayne, D.F., **Interactive high dimensional data visualization**, *Journal of Computational and Graphical Statistics*, Vol. 5, No. 1, pp. 78-99, 1996.
  - [21] Oosterom, Peter van. Stoter, J., **5D Data Modelling: Full Integration of 2D/3D Space, Time, and Scale Dimension**, *GIScience*, LNCS 6292, pp. 310-324, 2010.
  - [22] Jamileh Vasheghani Farahani, J. V., **Monitoring the Variations of b-Value and Seismicity in the Makran Ranges, the Absence of a Notable Event in West of Makran Subduction Zone**, *Geodynamics Research International Bulletin (GRIB)*, Vol. 2, No. 02, SN:06, Summer, 2014.
  - [23] Rohadi, S., **Studi Seismotektonik Sebagai Indikator Potensi Gempa Bumi di Wilayah Indonesia**. Balai Besar Meteorologi dan Geofisika Wilayah II Jakarta, 2009.