Spatial and Temporal Clustering Analysis of Hotspot Pattern Distribution of Critical Land in Kalimantan, Indonesia

AP Kirana1*, IS Sitanggang2, L Syaufina3, Bhawiyuga A4
1 Department of Information Technology, State Polytechnic of Malang, Indonesia
2 Department of Computer Science, IPB University (Bogor Agricultural University) Bogor, Indonesia
3 Department of Silviculture, IPB University (Bogor Agricultural University) Bogor, Indonesia
4 Faculty of Computer Science, University of Brawijaya Malang, Indonesia

*Corresponding author: puspakirana@polinema.ac.id

Abstract. Geographic information systems (GIS) can play a vital function in fire prevention by predicting potential hotspot for fires. This paper presents the analysis of hotspot data in Kalimantan from 2008 to 2018 based on spatial and temporal aspects. Spatiotemporal cluster analysis with Kulldorff’s Scan Statistic (KSS) methods. KSS method is used to explore the spatial temporal patterns. This approach was originally designed to detect clusters and to determine their significance by Monte Carlo replication. the result shows that the provinces with the highest hotspot occurrence cluster are Central Kalimantan and West Kalimantan provinces. Based on critical land area, the cluster distributions of hotspot are dominated in ‘Rather Critical’ and ‘Potential to Critical’ area. Based on land used, the cluster distributions of hotspot are dominated ‘Swamp Shrub’ and ‘Mixed Dry Agriculture Land’. For temporal aspect, the result shows most hotspot occurred in August, September and October. In conclusion, the hotspot detected varied according to year, months and regions. Understanding hotspot patterns enables the allocation of resources for forest fire planning and management. Results showed that the approach is effective for detecting hotspot clusters and cluster accuracy is 91%.

1. Introduction

Forest and land fires are a serious problem in Indonesia, occurring almost every year and damaging to human life and the environment. Environment and human behaviour are the factors that cause fires in different studies [1]. Ceccato, et. al in [1] found that the anomaly of rainfall in Central Kalimantan is closely related to fire hotspot activity. According to Jaya et al. [2], human factor contributes 52% to fire risk and land cover 48%. Thoha et al. [3] also found that land cover had a high contribution to fire risk model in Kapuas District, namely 48%. Dry season in Kalimantan usually from June to October which is anomaly rainfall occurs. There are generally two kinds of factors triggering the occurrence of a fire that are both human and natural. In most cases, fires are caused by human factors, while a small fraction of the cases is caused by a natural factor [4]. The effects of the fires were not limited to areas of burning. Easter and south easterly winds spread haze from the fires and spread over a much larger area than where the fires occurred. The fires created visibility-limiting haze that caused transport slowdowns and casualties, school and business shutdowns, and health problems [5]. Fires are a major threat, in particular to peat swamp forests and lowland rain forests and, in particular, to the edges of oil palm plantation...
areas, where their 'accidental' nature must be questioned as they pave the way for further development of existing plantations [6].

Considering the bad impact and the multiple causes of forest fire and land fire, it is very important to develop an early warning system as a preventative action. We need to identify the pattern of distribution of hotspots as one of the forests and land fire indicators for developing an early warning system. By detecting the pattern of distribution of hot spots, we are able to detect areas with a high fire density so any early preventive action can be taken in that area [7]. The purpose of this study was to recognize the pattern of distribution of hot spot clusters in Kalimantan between 2008 and 2018 using the Kulldorff Scan Statistic (KSS) method especially in critical areas of land.

2. Material

2.1. Study area

This study was conducted in Kalimantan, Indonesia, with an area of approximately 743.330 km². We used hot spots dataset from January 2017 to December 2018. Kalimantan is the Indonesian part of Borneo Island. Kalimantan is divided into five provinces: Central Kalimantan, East Kalimantan, North Kalimantan, South Kalimantan, and West Kalimantan. The Indonesian territory covers 73% of the island by area and 69.5% (13,772,543 in the Indonesian Census of 2010) by population [8]. Approximately 37.7% of the 743,330 km² in Kalimantan are critical land and 55% of Kalimantan is covered by forest (intact and disturbed) and 31% of the 40 million ha is protected [8]. The map of study area shown in Figure 1. Green color in Figure 1 is the critical land area and red color is the distribution of districts in Kalimantan.

![Figure 1. Study area.](image.png)

2.2 Data

In this study, remote sensing satellite data from Fire Information for Resource Management System Moderate Resolution Imaging Spectroradiometer (FIRMS MODIS) Fire/Hotspot, NASA data were used. Datasets that we used in this study are the hotspot data we obtained from modis-catalog.lapan.go.id. Hot spots dataset we used especially in critical land area from 2017 to 2018 which consist of 267,825 hotspots with a confidence level ≥80%. Another dataset that we used are the administrative district boundary of Kalimantan, critical land and land use distribution in Kalimantan.

The administrative district boundary of Kalimantan was obtained from the Geospatial Information Agency of Indonesia (Badan Informasi Geospasial). Critical land area dataset of 2015 was obtained from the official site Ministry of Environment and Forestry, Indonesia (http://appgis.dephut.go.id/). In this study we want to analyze about the correlation between number of hotspot cluster distribution, land use coverage and type of critical land in Kalimantan. Degraded land in Indonesia may be defined as "critical land" on the basis of a standard as described in the Dirjen 4/V-SET/2013 Regulation on Technical Guidance for the Conduct of Critical Land Spatial Data by the Ministry of Forestry.
Critical land is a reduction in the physical, chemical or biological status of the land, which may also limit the productive capacity of the land's [9]. It encompasses not only soil degradation, but also vegetation degradation, which has been defined generically as “a temporary or permanent reduction in the density, structure, species composition or productivity of vegetation cover” [10]. There are four types of critical land which are not critical, potential for criticism, rather critical and critical. On the basis of the regulation critical land is the land that has been damaged thereby losing or reducing the function to the definition or intended level. The assessment of critical land in the area is often adjusted with the area function [11].

2.3 Tools
The tools used in this analysis were ArcGIS 10.6.1, PostGIS 2.1.8, PostgreSQL 9.2 (64-bit) and R Studio 1.2. ArcGIS were used for spatial data visualization, PostGIS for spatial analysis, PostgreSQL 9.2 for database management services, and R Studio for clustering processes. We used clustering library packages in R called 'spatialepi' for clustering and analysis. Spatialepi is a package of R developed by Kim et al. [12].

3. Methods
We process a hotspot dataset on critical area in Kalimantan by considering spatial and temporal factors. Spatial factor refers to the location of hot spots with longitude and latitude properties. Temporal refers to the occurrence time of fire to date attributes. Once the data are collected, we first perform a pre-processing hot spot data that divided into four steps [13, 14]. The first step is to select the most important attribute for the Kulldorff Scan Statistic (KSS) clustering method. The second step is to identify the hot spot in non-critical and critical area. The third step is to select the hotspot distribution in each district, and the last step is to load the hotspot data into database. In the implementation of this study, we used 'spatialepi' R package [12] for clustering hotspot process.

KSS formed a large collection of circular windows called a circular scanning window to detect clusters as shown in Figure 2. 'G' is the whole area of the study, 'Z' is a circular scanning window, 'μ(Z)' is the total population within the circular scanning window, 'μ(G)' is the total population within the study area 'n, 'NZ' is the number of cases within the circular scanning window 'Z', 'nG' is the number of cases in the study area of 'G', 'p' is the average occurrence rate in a circular window scan, and 'q' is the occurrence rate outside a circular scanning window 'Z'. The KSS method used to compare the number of cases inside and outside the scanning window. The comparison result is used to search for the distribution of hotspot clusters in the study area 'n'.

\[
\frac{N_z}{\mu(Z)} = \frac{ez}{nG} = \frac{NZ}{\mu(G)}
\]

In this research, we use the KSS method to determine the value of the possibility scanning window ratio (likelihood ratio) [16]. KSS used to compare the number of case occurrences in and out scanning window. Equation 1 is used to calculate the ratio of the possibility for each circular scanning window Z. ‘Nz’ is the number of cases in the circular scanning window ‘Z’, ‘ez’ is the estimated case in the circular scanning window ‘Z’, ‘nG’ is the number of cases in the study area ‘G’ [17].
The 'eZ' can be calculated using Equation 2 in which the total population of 'μ(Z)' is in the scanning window 'Z', 'nG' is the total number number of cases in study area 'G', 'μ(G)' is the total number of populations in study area 'G' [17].

\[
\lambda(Z) = \frac{n_Z}{e_Z} \cdot \left( \frac{n_G \cdot n_Z}{n_G \cdot e_Z} \right)^{n_G \cdot n_Z}, \text{ if } n_Z > e_Z \\
1, \text{ otherwise}
\] (1)

The 'eZ' can be calculated using Equation 2 in which the total population of 'μ(Z)' is in the scanning window 'Z', 'nG' is the total number number of cases in study area 'G', 'μ(G)' is the total number of populations in study area 'G' [17].

\[
e_Z = \mu(Z) \left( \frac{n_G}{\mu_G} \right)
\] (2)

The most likely cluster is estimated for each simulated dataset in exactly the same way as the actual data. Statistical significance is evaluated using the Monte Carlo Hypothesis test [15]. The Monte Carlo analysis was performed with at least 999 replications under the null hypothesis. If the significance cluster contains more than 95% of the replications, the cluster is said to be significant at the 95% level [16]. In the implementation of this research, we utilized clustering packages of data in R statistical tools, and PostgreSQL for database management. We use the ‘spatialepi’ R package developed by [12] for the cluster dataset.

4. Result and Discussion

The number of hotspots that occurred, particularly in the critical area of Kalimantan increased year by year from 2008 to 2018. The highest frequency held in 2015 amounted to 74,582 hotspots. If we compared with total hotspot in 2014, hotspot in 2015 reached almost twice from total entire number of hotspots in the previous year, which is 2014. The average number of hot spots for 11 years is 24,348 hotspots. The number of hotspot distributions in critical land from 2008 to 2018 in Kalimantan is shown in Figure 3.

![Figure 3. Hotspot in critical land area from 2008 until 2018 in Kalimantan.](image)

Monthly distribution of hotspot is shown in Figure 4. We divided our analysis into two periods. First period from January 2008 until December 2012 and second period from January 2013 until December 2013. In Figure 4, we can see that there is the same pattern of hotspot number of several months. Hotspot increased significantly in August to October and the highest density of hotspot occurred in September.
Hotspot distributed in 56 districts in Kalimantan with a total area of peatland distribution of 282,465 km\(^2\). The time period that we use in this analysis is from 2008 to 2018. Locations which have the most distribution of hotspot are Central Kalimantan province (Pulang Pisau, Kotawaringin Timur, Kapuas and Katingan district), West Kalimantan province (Ketapang, Sintang, Sanggau, Kubu Raya and Kapuas Hulu district), and East Kalimantan province (Kutai Kartanegara, East Kutai, and West Kutai) as shown in Table 1. Table 1 displays the distribution of hotspot clusters between 2008 and 2018 consisting of three clusters.

Table 1. Cluster distribution of hotspot 2008 to 2018.

| Cluster | District          | Province               |
|---------|-------------------|------------------------|
| P1      | Pulang Pisau      | Central Kalimantan     |
|         | Kotawaringin      |                        |
|         | Timur             |                        |
|         | Kapuas            |                        |
|         | Katingan          |                        |
| S1      | Ketapang          | West Kalimantan        |
|         | Sintang           |                        |
|         | Sanggau           |                        |
|         | Kubu Raya         |                        |
|         | Kapuas Hulu        |                        |
| S2      | Kutai Kartanegara | East Kalimantan        |
|         | Kutai Timur       |                        |
|         | Kutai Barat       |                        |

Figure 5 shown the cluster distribution of hotspot from 2008 to 2018. Clustering hotspots using the KSS method in the vital land area of Kalimantan can detect where cluster hot spots exist, where cluster hotspots are located, and the geographical size of cluster hotspots. KSS clusters and calculate their importance through Monte Carlo replication. Results have shown that the method is effective for detecting hotspot clusters with an accuracy of 91%. The highest density of hotspots, mostly found in the central Kalimantan province, particularly in the Pulang Pisau district. Percentage of distribution of hotspots from 2008 to 2018 as shown in Figure 6.
Figure 5. Cluster distribution of hotspot from 2008 to 2018.

Figure 6. Percentage of cluster hotspot distribution from 2008 to 2018.

Figure 7 shows the percentage of hotspot cluster distribution based on the type of critical land in Kalimantan. The distribution of hotspot clusters between 2008 and 2018 is dominated by ‘Rather Critical’ and ‘Potential to Critical’. 43% of hotspot from 2008 to 2018 were occurring in ‘Rather Critical’ land area and 33% of hotspot were occurred in ‘Potential to the critical’ area.
Figure 7. Cluster distribution of hotspot from 2008 to 2018 based on the type of critical land.

Figure 8 shows the percentage of hotspot cluster distribution based on the land coverage area in Kalimantan. The cluster distribution of hotspot between 2008 until 2018 are dominated in ‘Swamp Shrub’ and ‘Mixed Dry Agriculture Land’. 23.12% of hotspot from 2008 to 2018 were occurring in ‘Swamp Shrub’ land area and 17.9% of hotspot were occurred in ‘Mixed Dry Agriculture Land’ area.

Figure 8. Cluster distribution of hotspot from 2008 to 2018 based on land use coverage.

5. Conclusion
This research used the clustering approach to identify the distribution structure of the occurrence of a hotspot. In this research, a statistical approach is used to identify the pattern of distribution of hotspot in both spatial and temporal domains using the Kulldorff Scan Statistic (KSS) method. The dataset consists of the spatial aspect (longitude, latitude) of the hotspot, temporal aspect (date of occurrence) of hotspot, and peatland physical characteristics that may influence the distribution pattern of hotspot clusters in the study area Kalimantan, Indonesia between 2008 until 2018 Clustering hotspots in critical land areas in Kalimantan between 2008 and 2018, using the KSS method, discovers patterns of distribution of hotspot clusters. From a spatial aspect, the provinces with the occurrences of the highest hotspot cluster are located in Central Kalimantan province, especially in Pulang Pisau, Kotawaringin Timur, Kapuas
and Katingan district. Based on critical land area, the cluster distributions of hotspot dominate in ‘Rather Critical’ and ‘Potential to Critical’ area. Based on land used, the cluster distributions of hotspot are dominated ‘Swamp Shrub’ and ‘Mixed Dry Agriculture Land’. Based on temporal aspect, from 2008 to 2018 hotspot increased significantly in August to October, the highest cluster of hotspots occurred in 2015.

References

[1] Ceccato P, I N S Jaya, J H Qian, M K Tippet, A W Robertson and S Someshwar 2010 *Early Warning and Response to Fires in Kalimantan, Indonesia* (New York (US): IRI Technical Report 10-14, International Research Institute for Climate and Society) pp 23 2010

[2] Jaya I N S, R Boer, Samsuri, Fathurakhman 2008 *Development of wildfire vulnerability index in Central Kalimantan* (Bogor (ID): Research Report CARE)

[3] Thoha A S, Saharjo B H, Boer R, Ardiansyah M 2017 *Forest and land fires hazard level modeling: case study of Kapuas, Central Kalimantan*. In: Djalante R, Garschagen M, Thomalla F, Shaw R (eds) *Disaster Risk Reduction in Indonesia. Disaster Risk Reduction (Methods, Approaches, and Practices)* (Cham: Springer)

[4] Syaufina L 2008 *Kebakaran Hutan dan Lahan di Indonesia* (in Bahasa) (Malang: Bayumedia Publishing)

[5] E Frankenberg, D McKee, D Thomas 2004 *Health Consequences of Forest Fires in Indonesia* (Los Angeles: California Center for Population Research On-Line Working, University of California) Paper Series CCPR-030-04

[6] W. Stephan, G. Chris 2017 Executive Summary the Environmental Status of Borneo 2016 Report (Jakarta (ID): Heart of Borneo Program)

[7] Syaufina L, Sitanggang I S, Erman L M 2015 Challenges in Satellite-Based Research on Forest And Land Fires in Indonesia: Frequent Item Set Approach: The 2nd International Symposium on LAPAN-IPB Satellite for Food Security and Environmental Monitoring 2015 LISAT-FSEM 2015

[8] Badan Pusat Statistik 2010 *Hasil Sensus Penduduk 2010 dalam Agregat per Provinsi* (in Bahasa) (Jakarta (ID): Badan Pusat Statistik)

[9] Chartres C 1987 *Australia's land resources at risk. In: Chisholm, A., Dumsday, R. (Eds.), Land Degradation: Problems and Policies* (Cambridge: Cambridge University Press) pp 7e26

[10] Grainger A 1996 *The degradation of tropical rain forest in Southeast Asia: taxonomy and appraisal. In: Eden, M.E., Parry, J.T. (Eds.)* (London: Land Degradation in the Tropics) pp 61e75

[11] Information System of Critical Level of Land Degradation (Critical Land) Based on Agro-ecological Zone (AEZ) in Agricultural Areas with Recombination Method of Fuzzy Logic and Scoring Prasetyo S Y J, Hashilohan B, Hartomo K D, Paseleng M, Nuswantoro B 2013 *Geographic International Journal of Computer Science Issues* 10(1) 1694-0784

[12] Kim A Y, Wakefield J R 2010 *Data and Method for Spatial Epidemiology: The SpatialEpi Package* (Washington: Dept of Statistics, University of Washington)

[13] Kirana A P, Sitanggang I S, Syaufina L 2015 Hotspot Pattern Distribution in Peat Land Area in Sumatera based on Spatio Temporal Clustering *Procedia Environmental Sciences* 33 635-645

[14] Kirana A P, Sitanggang I S, Syaufina L 2015 Poisson Clustering Process on Hotspot in Peatland Area using Kulldorff’s Scan Statistics Method TELKOMNIKA (Telecommunication Computing Electronics and Control) 13(4) 1376-1383. https://doi.org/10.12928/telekommika.v13i4.2272

[15] Wen S and Kedem B 2009 A semiparametric Cluster Detection Method A Comprehensive Power Comparison with Kulldorff’s Method *International Journal of Health Geographics* (USA: BioMed Central)

[16] Kulldorff M 1997 A Spatial Scan Statistic. Communications in Statistics: Theory and Methods. 1997 26 1481-1496

[17] Kulldorff M 2014 SaTScan User Guide [downloaded at 2014 July 2] Available at http://www.satscan.org/.