https://doi.org/10.26525/jtfs2018.30.2.154162

PEATLAND FIRE DETECTION USING SPATIO-TEMPORAL DATA MINING ANALYSIS IN KALIMANTAN, INDONESIA

Syaufina L[1] & Sitanggang IS[2]

[1]*Department of Silviculture, Faculty of Forestry, Bogor Agricultural University (IPB), Jl Ulin Kampus IPB Darmaga, Bogor, Indonesia 16680*
[2]*Department of Computer Science, Faculty of Natural Science and Mathematics, Bogor Agricultural University (IPB), Jl. Meranti Wing 20 Level V, Kampus IPB Darmaga, Bogor, Indonesia 16680*

lailans@apps.ipb.ac.id

Submitted January 2017; accepted September 2017

Peatland fire has been an important environmental issue in Indonesia as well as in ASEAN region as it is strongly related with transboundary haze pollution. Hotspot has been used as an indicator of forest and land fires detection. Source of hotspot in Indonesia varies from a low to a relatively high degree of accuracy. Not all hotspots are strong indicators of forest and land fires. The incidence of hotspots in a sequence of at least three days in an adjacent location can be a strong indicator of land and forest fire occurrences. The use of spatio-temporal data mining approach to obtain sequence patterns of hotspots is still in question. This study aims to obtain sequence patterns of hotspot occurrence in Kalimantan peatland in 2015 by applying a spatio-temporal data mining approach. The algorithm used is Sequential Pattern Discovery Equivalent Classes (SPADE) that enables to discover all possible sequence patterns on hotspot datasets. Verification of hotspot sequences was conducted through ground truth and fire supression data. The study proved that real fire detection is indicated by three consecutive days of hotspot occurrences. The fire detection in peatland area is strongly affected by fire characteristics dominated by ground fire.

Keywords: Hotspot, sequential pattern mining, SPADE, fire characteristics, enviroment

INTRODUCTION

Peatland ecosystem play a very important role in maintaining the balance of the environment, especially in water regulation, biodiversity conservation, climate change mitigation and support of human welfare. Peatland area covers more than 400 million hectares in 180 countries, representing a third of global wetlands (Parish et al. 2008). Although tropical peatlands only cover 10–12% of total peatlands in the world, they have a significant role as a valuable and important natural resource towards global environment (Syaufina 2008). On the other hand, peatland fires, which are dominated by ground fire, is more difficult to control than that of non-peatland fires. It would be difficult to suppress the fire if it has reached the inner layer of the peat, especially if the water sources around the peatlands are not available and burned areas are difficult to reach. Therefore, the prevention of fires in peatlands should be prioritised in order to minimise the environmental impacts caused by peatland fires. Analysis and prediction of occurrence and spread of hotspots, as indicators of forest and land fires, is a requirement to build an early warning system of forest fires, either on dry land or in peatlands. Distribution pattern of hotspots can be analysed to classify the characteristics of the region where hotspots have potentially occurred.

Hotspot data are daily collected by some institutions such as the Ministry of Forestry and Environment Republic of Indonesia, National Institute of Aeronautics and Space (LAPAN) and the ASEAN Specialised Meteorological Centre (ASMC), that produce spatio-temporal datasets. However, the use of hotspot data is still limited for field fire detection. In addition to Geographic Information System (GIS) and remote sensing techniques, data mining is an approach that can be used to analyse hotspot data. Spatio-temporal data mining is used in analysing hotspots represented by spatio-temporal dataset and other data, regarding characteristics and location of the area where the hotspot occurred.

Hotspot distribution pattern can be used to detect the early appearance of hotspots in the study area, which is important for early fire suppression, hence minimising the impacts of forest and land fires. However, not all hotspots indicate real fires in the field. According to forest fire experts and fire fighter brigade under the Ministry of Forestry and Environment, Republic of Indonesia, hotspots that occur sequentially within 2 days or more in the same location has a high potential of becoming real fire in the area. Sequence pattern of hotspot occurrence can be obtained using sequential pattern mining, a technique in spatio-temporal data mining. This approach generates sequence patterns based on location and time of hotspot occurrence. Verification of sequence pattern is carried out to identify hotspots that become fires (fire spot). Verification is conducted in three ways, namely ground truth by field checking on burned area, utilising classified satellite imagery for burned area and comparing the sequence patterns to secondary data of fire suppression reports.

In data mining, sequential pattern is a sorted list of an item, data or event. Sequential pattern mining is a task of extracting certain sequential pattern where its support value exceeds minimum support. Minimum support is commonly defined by the user. Based on the value of minimum support, less interesting patterns can be ignored so that mining process becomes more efficient. Some sequential pattern mining algorithms include Generalized Sequential Pattern (GSP), Prefixspan and Clospan (Pei et al. 2004). Research related to application of sequential pattern mining on hotspot data have been performed by Agustina & Sitanggang (2014) and Nurulhaq & Sitanggang (2015). Clospan algorithm, which is available in the Sequential Pattern Mining Framework (SPMF) tool, was applied. Clospan works in two stages, namely generating a sequential pattern with Prefixspan algorithm and post-pruning of the sequences. The study applied Clospan on hotspots and precipitation data from 2001 to 2010 with minimum support of 1, 2, 3, 4, 5, 10, 15 and 20%. The study found that hotspot sequences commonly occurred in areas that had a temperature of 28.33, 28.89, 29.44 and 30 °C. (Han et al. 2003, Yan et al. 2003, Viger 2013). Another study used the Prefixspan algorithm. Prefixspan algorithm projects a data base by forming a prefix of the sequence. The study identified hotspots as a strong indicator of forest fires based on sequence patterns, where 16.95% hotspot occurrences became strong indicators of forest fires in 2014 (Han et al. 2005, Pei et al. 2004).

MATERIALS AND METHODS

Study area and data

The study area of the research was peatlands, Kalimantan. The peatland area in Kalimantan is 5,769,246 ha (Wahyunto-Ritung & Subagjo 2003) Figure 1 shows peatland distribution based on type and depth. The study used spatial and non-spatial data. Spatial data included (1) hotspot data in Kalimantan, 2015 that were collected from MODIS Fire FIRMS/Hotspot and NASA, University of Maryland, (2) a digital map of peatlands covering Kalimantan Island that was obtained from Wetlands International, Indonesia and (3) a boundary map of villages in Kalimantan Island in 2013 from Statistics Agency Indonesia. In addition to spatial data, this study utilised non-spatial data on location and size of land fires suppression in Central Kalimantan province in 2015. These secondary data were obtained from Environmental Management Agency and Sebangau National Park in Central Kalimantan province.

Methodology

There were three main tasks conducted to achieve the objectives of this study. The first task was pre-processing data in order to prepare task relevant data for mining process. The second task was generating sequence patterns on hotspot dataset by applying sequential pattern mining approach. The third task was sequence verification by visiting the 2015 peatland fire locations and comparing the patterns to the secondary data of peatland fires suppression reports.

Pre-processing data

Pre-processing data consists of several stages, namely data selection, data cleaning and data transformation. Selection data was carried out to select the data and its attributes that were used in the study and to select hotspots in the study area, namely peatlands of Kalimantan. In addition, hotspots data which had a confidence value of

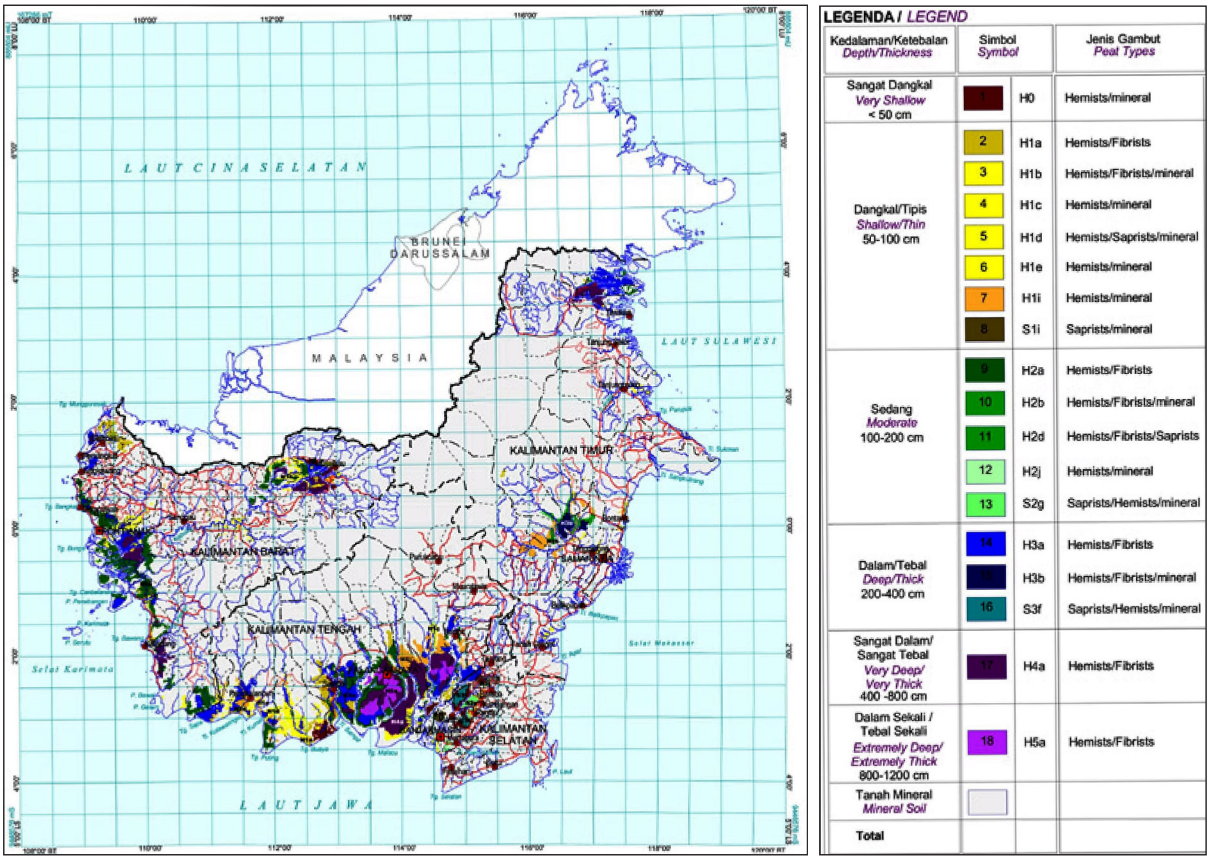


Figure 1 Peatland distribution in Kalimantan island (Wahyunto-Ritung & Subagjo 2003)

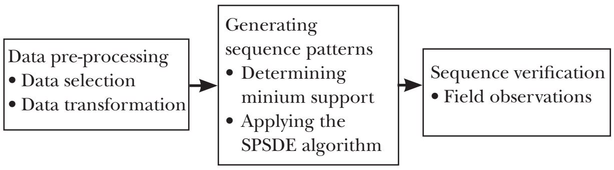
above or equal to 70% were also selected. Data transformation was conducted to prepare hotspot sequence datasets. The SPADE algorithm that is available in R software requires a certain format of input. The format of a sequence dataset needs four attributes namely sequence identifier (SID), transaction identifier (Tid), size and item(s). The SID represents longitude and latitude hotspot locations, Tid denotes date code of hotspot, size states the number of hotspot occurrences in the same location and date, and item(s) is a list of date codes. All steps in pre-processing data were performed using R software, Quantum GIS, Postgre Structured Query Language (SQL) and PostGIS.

Generating sequential patterns on hotspot dataset

Hotspot sequence patterns were generated using SPADE algorithm that was implemented in R software. In the first step of SPADE, frequent 1-sequences were determined. A frequent 1-sequence is a sequence consisting of an item or an event which has support value greater than or equal to minimum support. An event is a hotspot occurrence. Support of a sequence is the number of transactions within a dataset that contains the sequence. Support of a sequence is calculated based on the number of different SID included in the pair, sequence identifier-event identifier (SID-EID), which is also known as id-list of the item. In the next step of SPADE, all items in frequent 1-sequences were a parent class in the formation of frequent k-sequences, k = 2, 3, etc A k-frequent sequence is obtained by combining a k-1-frequent sequence to another k-1-frequent sequence. By applying the same procedure, the longest k-frequent sequence will be generated, until there is no possible sequence that can be discovered.

Sequence verification

Hotspot sequence patterns, as results of SPADE algorithm, were verified by two activities. The first was field observations of the burned peatland area in Pulang Pisau district and Palangkaraya city, Central Kalimantan province, which was conducted on 22–24 July 2016. The second activity was collecting secondary data on fire suppression in Pulang Pisau and Palangkaraya city in Central Kalimantan. These secondary data were then compared to the location of hotspot sequences, as the results of data mining, in order to identify real fire spots. Below is a flow chart of the study:



RESULTS AND DISCUSSION

Sequence of hotspot occurrence

Hotspot data analyses resulted in a number of hotspot sequences in Kalimantan in the year 2015 (Figure 2). These sequences showed that some hotspots occurred in a certain date and location, whereby these occurrences were followed by other hotspots at the same location. The duration between the first occurrence and the second occurrence was 2 to 13 days. Number of hotspot sequences in each province is as follows:

- West Kalimantan: 121
- South Kalimantan: 40
- Central Kalimantan: 801
- East Kalimantan: 4

Number of hotspot sequence in Central Kalimantan seems to be the highest among the provinces. Therefore, the study merely focused on Central Kalimantan. From peatland condition point of view, the largest degraded peatland area

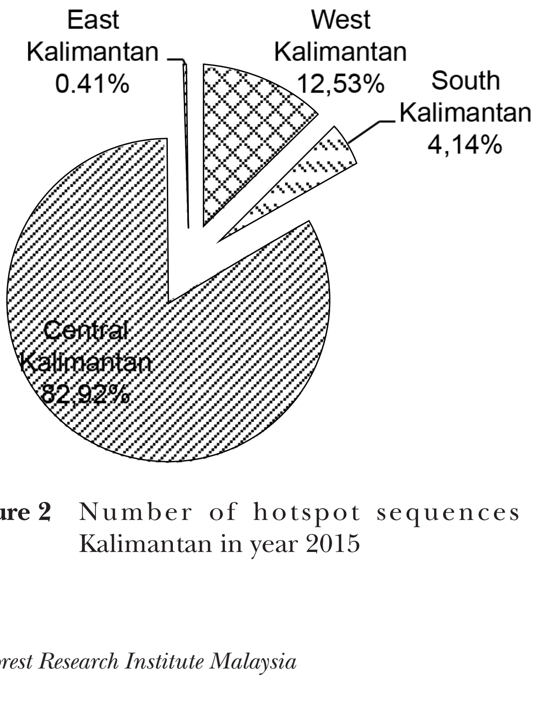


Figure 2 Number of hotspot sequences in Kalimantan in year 2015

is found in Central Kalimantan due to peatland mismanagement in the last few decades. Most of the primary peat swamp forest cover has dissapeared and taken over by secondary peat swamp forest vegetation, shrubs and ferns which are more vulnerable to fire, hence multiple fires occur for almost every year.

In vegetation, direct effects of fire may kill plants and cause injury, whereas indirect effects of fire includes open wound which attracts pest and disease attack. On the other hand, fire alters forest structure and composition. The magnitude of fire effects on tropical forest biodiversity is influenced by several factors, namely fire intensity, fire severity, soil types, post fire precipitation and burned area (Syaufina & Ainuddin 2011). A study conducted in Tumbang Nusa, Central Kalimantan in 1997 indicated that burned peat swamp forest area and multiple fire peat swamp forest area have different vegetation composition compared with secondary peat swamp forest (Tata & Pradjadinata 2013). *Calophyllum macrocarpum* was the commonest tree species in secondary peat swamp forest, whilst *Cratoxylum arborescens* was the commonest tree species in burned peat swamp and multiple fire peat swamp forest areas. Similarity Index of Sorensen (IS) between secondary peat swamp and burned peat swamp forest areas was relatively high (IS = 62.79%), meanwhile IS between secondary peat swamp and multiple fire peat swamp forest areas was considerably low (IS = 25.81%). Fire reduced tree diversity, which was indicated by low Shannon-Wiener diversity index (H'). The H' in secondary peat swamp forest was 3.30, whereas in burned peat swamp and multiple fire peat swamp forest areas were 2.61 and 1.75, respectively.

Fire incident, as indicated by hotspots found in Kalimantan in year 2015 from July to October, with the highest within 14–16 October and the lowest within 7–14 October (Figure 3), showed that most multiple fires occurred in Kalimantan, in October 2015. However, it was clear that number of hotspots in 2-sequences explained the real fire occurrence, but not the total number of fire.

Table 1 describes number of hotspot sequences in Kalimantan in terms of duration between the first and second date of occurrence, ranging from 2 to 13 days. Duration of hotspot sequences varied in each province: 2 to 8 days in West Kalimantan with highest of 2 days (20) followed by 3 days,

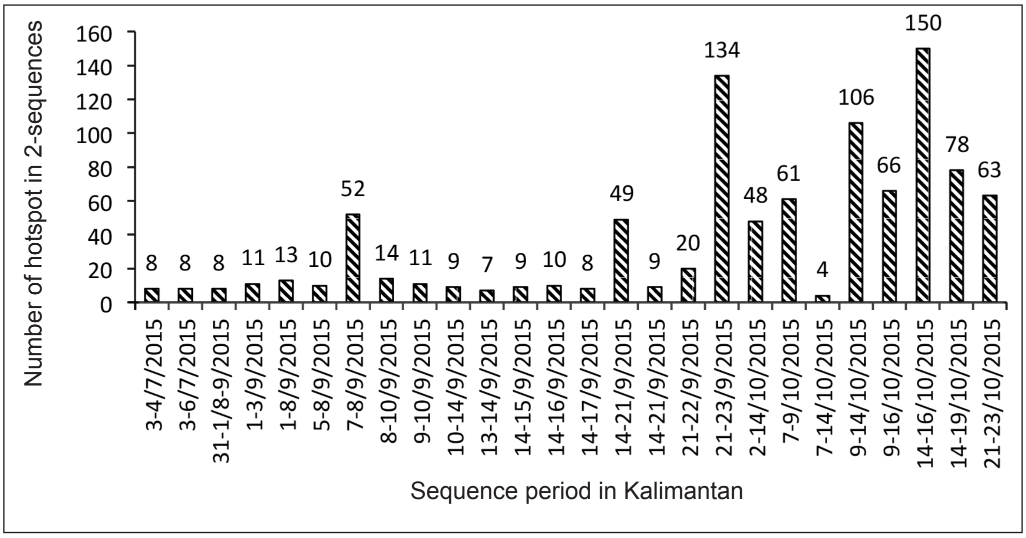


Figure 3 Number of 2-sequences in each period in Kalimantan, 2015

Table 1 Number of hotspot sequences in Kalimantan in year 2015

First date of occurrence	Second date of occurrence	Duration	West Kalimantan	South Kalimantan	Central Kalimantan	East Kalimantan	Total
7/3/2015	7/4/2015	2	8				8
7/3/2015	7/6/2015	4	8				8
8/31/2015	9/1/2015	2	8				8
9/1/2015	9/3/2015	3	11				11
9/1/2015	9/8/2015	8	13				13
9/5/2015	9/8/2015	4	10				10
9/7/2015	9/8/2015	2			52		52
9/8/2015	9/10/2015	3	14				14
9/9/2015	9/10/2015	2	11				11
9/10/2015	9/14/2015	5	9				9
9/13/2015	9/14/2015	2		7			7
9/14/2015	9/15/2015	2		9			9
9/14/2015	9/16/2015	3		10			10
9/14/2015	9/17/2015	4		8			8
9/14/2015	9/21/2015	8			49		49
9/14/2015	9/21/2015	8	9				9
9/21/2015	9/22/2015	2	20				20
9/21/2015	9/23/2015	3			134		134
10/2/2015	10/14/2015	13			48		48
10/7/2015	10/9/2015	3			61		61
10/7/2015	10/14/2015	8				4	4
10/9/2015	10/14/2015	6			106		106
10/9/2015	10/16/2015	8			66		66
10/14/2015	10/16/2015	3			150		150
10/14/2015	10/19/2015	6		6			6
10/14/2015	10/19/2015	6			72		72
10/21/2015	10/23/2015	3			63		63
		Total	**121**	**40**	**801**	**4**	**966**

2 to 6 days in South Kalimantan with highest of 3 days, 2 to 13 days in Central Kalimantan with highest of 3 days and 1 sequence of 8 days in East Kalimantan. Therefore, the highest number of hotspot sequences (150) for the whole of Kalimantan was found in the 3 days duration from 14–16 October 2015. The 3 days duration hotspot sequence corresponds to real fire occurence. However, discussion and verification need to be conducted for longer durations such as 8 to 13 days, to determine whether its the same fire occuring in an area or a new fire. Field verification is also strongly required.

Analysis of hotspot sequences in various districts of Central Kalimantan indicated the highest number in Pulang Pisau district, followed by Kapuas and Katingan districts (Figure 4). Peat swamp forest in Pulang Pisau district was included in the mega rice project 1995, when almost one million ha of pristine peat swamp forest was converted into agriculture land. The project led to extensive degraded peatland area which was susceptible to fire within the period of 1973 to 2005 (Hoscilo et al. 2011), and still remain until today. Meanwhile, land preparation for plantation development in Kapuas dan Katingan districts seems to be the cause of hotspot occurrences.

Table 2 shows the number of hotspot sequences in detail, found in districts of Barito Selatan, Katingan, Kota Palangka Raya, Kota Waringin Barat, Kota Waringin Timur, Pulang Pisau and Sukamara, mostly in 3 days duration, similar to previous temporal analyses. Similar

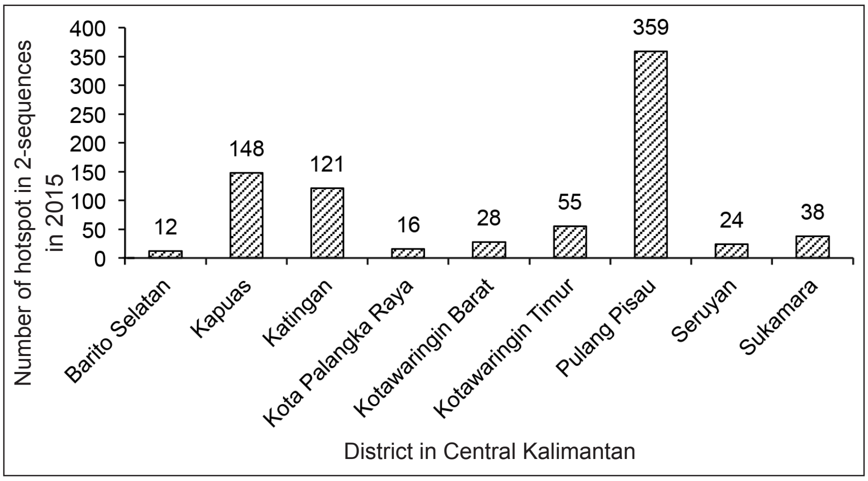


Figure 4 Number of hotspot sequences in Central Kalimantan

Table 2 Number of hotspot sequences in each district in Central Kalimantan

First date of occurrence	Second date of occurrence	Barito Selatan	Kapuas	Katingan	Kota Palangka Raya	Kota Waringin Barat	Kota Waringin Timur	Pulang Pisau	Seruyan	Sukamara	Total
9/7/2015	9/8/2015	1	6	3		5	6	24	6	1	52
9/14/2015	9/21/2015		8	5	1	1	9	23	1	1	49
9/21/2015	9/23/2015	5	22	8	1	3	13	79	2	1	134
10/2/2015	10/14/2015		10	4	2	1	3	26	2		48
10/7/2015	10/9/2015		17	14		1	3	24	2		61
10/9/2015	10/14/2015	1	28	19	3	3	3	37	4	8	106
10/9/2015	10/16/2015		9	25	1	2	3	18		8	66
10/14/2015	10/16/2015		14	40	8	4	5	68	3	8	150
10/14/2015	10/19/2015	5	30	3			10	17		7	72
10/21/2015	10/23/2015		4			8		43	4	4	63
	Total	**12**	**148**	**121**	**16**	**28**	**55**	**359**	**24**	**38**	**801**

trend was found in 2-sequences hotspot in the whole Kalimantan and Central Kalimantan (Figure 5). The highest number of hotspot sequences in Central Kalimantan was found between the period 14–16 October 2015 and 21–23 September 2015, indicating three days duration of hotspot occurences.

On the other hand, the study also found a number of 3-sequences within each period in Central Kalimantan, 2015 (Figure 6), showing hotspots that occurred sequentially on 9 October 2015, 14 October 2015 and 16 October 2015 within a radius of about 1 km. The sequences were found in districts of Kapuas, Katingan, Kota Palangkaraya, Kotawaringin Timur, Pulang Pisau and Seruyan, with the highest number found in Katingan followed by Pulang Pisau districts.

Real fire is possibly detected when peatland fire is dominated by ground fire and spread underground slowly for less then 3 cm hour^{-1} (De Bano et al. 1998). Smoldering process may remain for weeks or even months when peat materials are extremely dry. Irreversible drying is a unique peat characteristic that leads to increased fire susceptibility of peatland area.

Hotspot sequences were verified by visiting the 2015-burned area in Pulang Pisau and Palangkaraya city, in Central Kalimantan

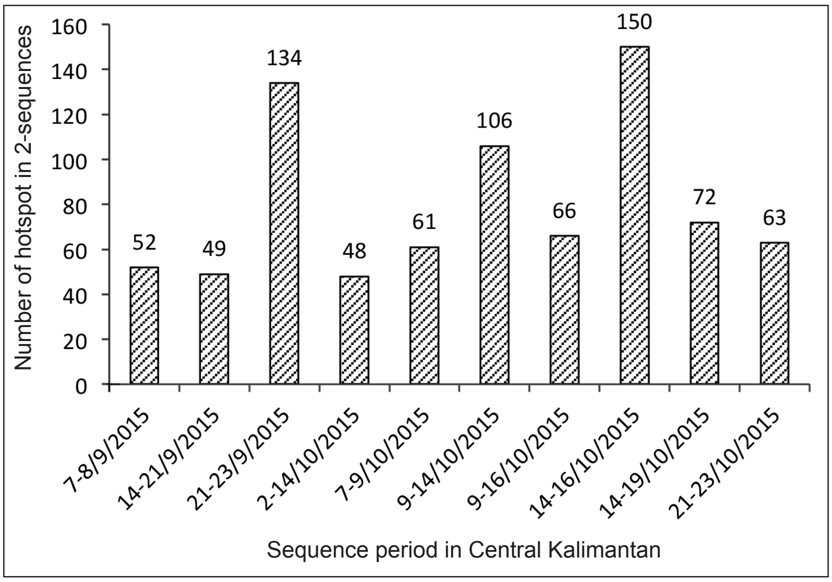


Figure 5 Number of 2-sequences in each period in Central Kalimantan, 2015

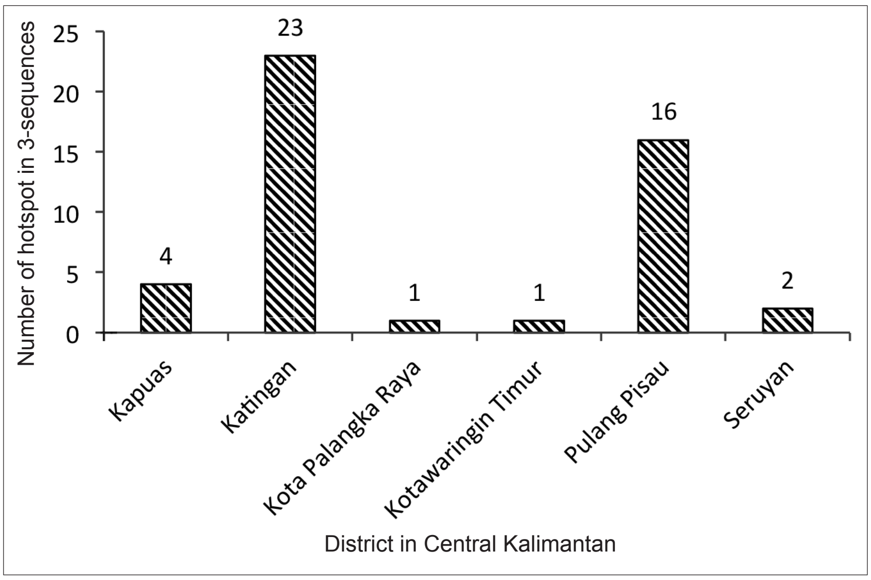


Figure 6 Number of 3-sequences in each period in Central Kalimantan, 2015

province. Verification in Central Kalimantan was performed between 22–24 July 2016. The hotspot sequences resulted from field verification spot and fire suppression data from local government in Palangkaraya city and Pulang Pisau district, and overlaid in an integrated map (Figure 7).

The study also identified hotspot sequence in Jabiren Raya, Pulang Pisau and Central Kalimantan in the following periods:

- 9/14/2015 followed by 9/21/2015
- 9/21/2015 followed by 9/23/2015
- 10/2/2015 followed by 10/14/2015
- 10/7/2015 followed by 10/9/2015
- 10/9/2015 followed by 10/14/2015
- 10/9/2015 followed by 10/16/2015
- 10/9/2015 followed by 10/14/2015 and 10/16/2015

Finally, the study found that hotspot sequences, as a result of sequential pattern mining, have been verified as real fire at Tumbang Nusa, Jabiren Raya, Pulang Pisau and Central Kalimantan province, in 2015 (Figure 8). The burned area was dominated by shrubs, ferns and small patches of secondary peat swamp forest, covering the degraded peatland area.

CONCLUSION

Among the provinces of Kalimantan, Central Kalimantan has the highest hotspot occurrences as a fire prone area, followed by Pulang Pisau district. Hotspot data can be used as an indicator of fire occurrence for peatland fire detection. However, only sequential hotspot seems to be real fire in the field. The study proved that real fire detection is indicated by three consecutive days of hotspot occurrences. The fire detection in peatland area is strongly affected by fire characteristics dominated by ground fire, where fire spread slowly underground. It is suggested that fire patrol need to be intensified when hotspots occurr within two consecutive days, at the same position.

ACKNOWLEDGEMENTS

The authors thank the Directorate General of Higher Education, Ministry of Research, Technology and Higher Education for funding the research through 'Penelitian Unggulan Perguruan Tinggi' (PUPT) 2016 scheme, and the Directorat Research and Innovation of Bogor Agricultural University (IPB) for publication grant.

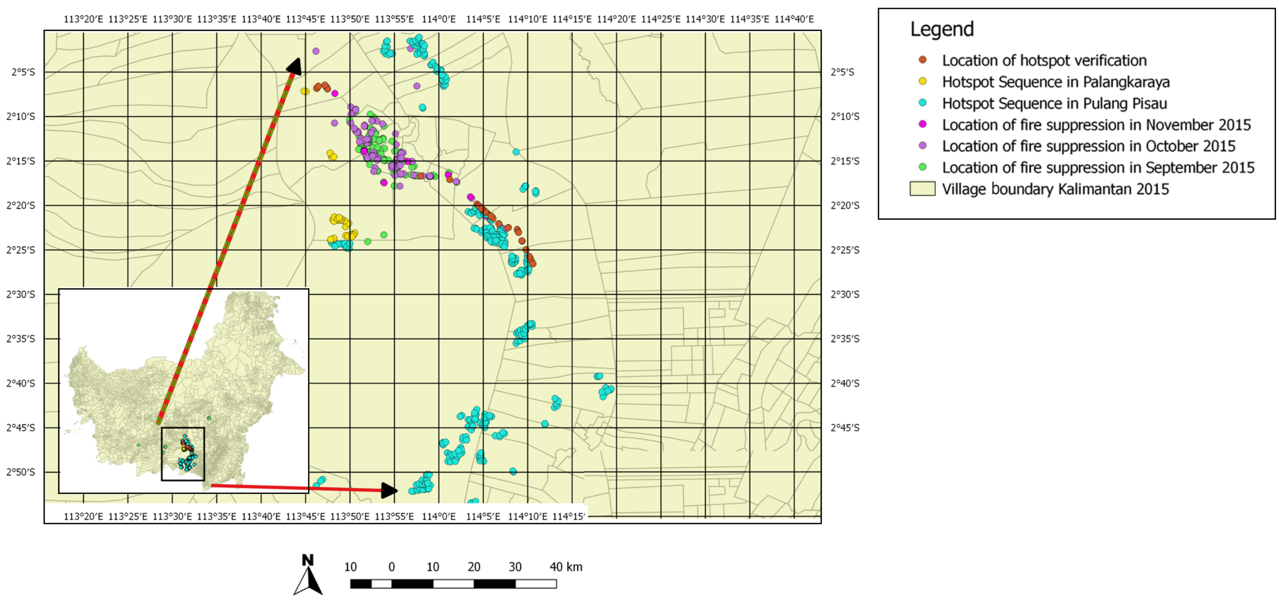


Figure 7 Hotspot sequence with field verification and fire suppression spots in Palangkaraya city and Pulang Pisau district

Figure 8 2015 burned area at Jabiren Raya, Pulang Pisau and Central Kalimantan

REFERENCES

Agustina T & Sitanggang IS. 2015. Sequential patterns for hotspot occurrences based weather data using clospan algorithm. Pp 301–305 in Romli M et al. (eds) *The 3rd International Conference on Adaptive and Intelligent Agroindustry (ICAIA)*, Bogor.

De Bano LF, Neary DG & Folliot PF. 1998. *Fire Effect on Ecosystems*. John Willey and Sons, New York.

Han J, Pei J & Yan X. 2005. Sequential pattern mining by pattern-growth: principles and extensions. *StudFuzz* 180: 83–220.

Han J, Yan X & Ashfar R. 2003. Clospan: mining closed sequential pattern in large dataset. *Data Mining* 16: 40–45.

Hoscilo A, Page SE & Tansey KJ & Rieley JO. 2011. Effect of repeated fires on land-cover change on peatland in Southern Central Kalimantan, Indonesia, form 1973 to 2005. International Journal of Wildland Fire 20: 578–588.

Nurulhaq NZ & Sitanggang IS. 2015. Sequential pattern mining on hotspot data in Riau Province using the PrefixSpan algorithm. Pp 313–316 Romli M et al. (eds) *The 3rd International Conference on Adaptive and Intelligent Agroindustry (ICAIA),* Bogor.

Parish F, Sirin A, Charman D, Joosten H, Minayeva T, Silvius M & Stringer L. 2008. *Assessment of Peatlands, Biodiversity and Climate Change: Main Report.* Global Enviroment Centre, Kuala Lumpur and Wetlands International, Wageningen.

Pei J, Han J, Mortazawi-Asl B et al. 2004. Mining sequential patterns by pattern-growth: the prefixspan approach. *IEEE Transactions on Knowledge and Data Engineering* 16: 1424–1440.

Syaufina L. 2008. *Forest and Land Fires in Indonesia: Fire Behaviour, Causes and Fire Impacts*. Bayumedia, Malang.

Syaufina L & Ainuddin AN. 2011. Impacts of fire on southeast Asia tropical forests biodiversity: a review. *Asian Journal of Plant Sciences* 10: 238–244.

Tata MHL & Pradjadinata S. 2013. Natural regeneration of burnt peat swamp forest and burnt peatland in Tumbang Nusa, Central Kalimantan and its implication on conservation. *Jurnal Penelitian Hutan dan Konservasi Alam* 10: 327–342.

Viger PF. 2013. Sequential pattern mining framework (SPMF) versi 0.94. http:// www.philippe-fournier-viger. com/spmf.

Wahyunto-Ritung S & Subagjo H. 2003. *Maps of Area of Peatland Distribution and Carbon Content in Sumatera, 1990–2002.* Wetlands International—Indonesia Programme & Wildlife Habitat Canada (WHC), Bogor.

Yan X, Han J & Ashfar R. 2003. Clospan: mining closed sequential pattern in large dataset. Pp 166–177 in Barbara D & Kamath C (eds) *Proceedings of the Third SIAM International Conference on Data Mining.* The Society for Industrial and Applied Mathematic, San Francisco.

Zaki MJ. 2001. *SPADE: An Efficient Algorithm for Mining Frequent Sequences. Machine Learning*. Kluwer Academic Publisher, Netherlands.