

Deteksi Pencilan Data Titik panas di Provinsi Riau Menggunakan Algoritme *Local Outlier Factor*

Shofyan(G64134009)*, Imas Sukaesih Sitanggang

Abstrak/Abstract

Indonesia mengalami kebakaran hutan yang signifikan. Pada tahun 2013 *World Resources Institute* (WRI) meneliti tren historis titik panas di Pulau Sumatera menggunakan data titik panas aktif *National Aeronautics and Space Administration* (NASA). Pada 13-30 Juni 2013 terjadi 2643 total jumlah peringatan titik panas. Tahun berikutnya Pada 20 Februari hingga 11 Maret tahun 2014 titik panas meningkat menjadi 3101 peringatan titik panas. Salah satu upaya untuk menangani kebakaran hutan ialah dengan menganalisis data titik panas yaitu dengan menganalisis pencilan titik panas sehingga dapat diidentifikasi wilayah yang beresiko terjadinya kebakaran hutan. Beberapa penelitian terkait deteksi pencilan yang sudah dilakukan diantaranya menggunakan algoritme *clustering* k-means dan juga menggunakan algoritme *clustering* berbasis medoids. Kedua penelitian tersebut mendeteksi pencilan berdasarkan frekuensi terjadinya titik panas dan belum mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas. Algoritme yang dapat mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas ialah algoritme *local outlier factor*. Dengan algoritme *local outlier factor* informasi mengenai wilayah yang berpotensi terjadi kebakaran hutan berdasarkan kepadatan penyebaran titik panas dapat dideteksi sehingga menjadi informasi tambahan untuk pengambilan keputusan oleh pihak terkait.

Kata Kunci

Kebakaran hutan; *local outlier factor*; pencilan; titik panas.

*Alamat Email: shofyanipb@gmail.com

PENDAHULUAN

Latar Belakang

Indonesia mengalami kebakaran hutan yang signifikan. Pada tahun 2013 *World Resources Institute* (WRI) meneliti tren historis titik panas di Pulau Sumatera menggunakan data titik panas aktif NASA pada 13-30 Juni 2013 terjadi 2643 total jumlah peringatan titik panas. Tahun berikutnya Pada 20 Februari hingga 11 Maret tahun 2014 titik panas meningkat menjadi 3101 peringatan titik panas (Sizer et al. 2014).

Kebakaran hutan dapat mengakibatkan pencemaran kabut asap, emisi karbon, degradasi dan deforesasi hutan yang mengakibatkan hilangnya hasil hutan dan berbagai jasa lingkungan yang diberikan hutan seperti kayu, hasil hutan non- kayu, dan keanekaragaman hayati, serta kerugian di sektor pedesaan contohnya dampak kabut asap pada hasil produksi pertanian (Tacconi 2003).

Data titik panas dapat dijadikan sebagai salah satu indikator tentang kemungkinan terjadinya kebakaran hutan (2005) sehingga dengan menganalisis data titik panas dapat diketahui langkah yang dapat diambil oleh pihak terkait. Diantara analisis yang dapat dilakukan ialah deteksi pencilan titik panas.

Beberapa penelitian terkait deteksi pencilan sudah dilakukan menggunakan algoritme *clustering* k-means (Baehaki 2014) dengan rata-rata pencilan yang terdeteksi adalah sebesar 481.22 titik panas. Frekuensi titik panas minimum yang terdeteksi sebagai pencilan sebesar 284 titik panas dan terbesar adalah 1118 titik panas. Penelitian kedua menggunakan *clustering* berbasis medoids yaitu PAM dan CLARA (Cahyadarena 2014). Hasil algoritme PAM pencilan titik panas terjadi pada nilai $k=17$ dengan cluster ke 13,14,15,16 dan 17. Algoritme CLARA pencilan titik panas terjadi pada nilai $k=19$ dengan cluster ke 14,15,17 dan 19.

Kedua penelitian tersebut meneliti pencilan titik panas berdasarkan frekuensi terjadinya titik panas, belum mendeteksi pencilan berdasarkan kepadatan penyebaran titik panas. Algoritme yang dapat mendeteksi pencilan dengan kriteria tersebut ialah algoritme *local outlier factor* (Beunig et al. 2000). *Local outlier factor* dapat mendeteksi pencilan lokal. Pencilan local ini tidak dapat dideteksi sebagai pencilan jika menggunakan pendekatan *clustering*.

Perumusan Masalah

Rumusan masalah dalam penelitian ini adalah bagaimana pencilan diidentifikasi dari data titik panas menggunakan metode *local outlier factor* dan informasi tentang karakteristik pencilan titik panas.

Tujuan

Tujuan penelitian ini adalah

1. Menentukan pencilan pada data titik panas di Provinsi Riau berdasarkan hasil algoritme *local outlier factor* data titik panas di Provinsi Riau, dan
2. Analisis pencilan data titik panas yang dihasilkan berdasarkan aspek lokasi dan waktu.

Ruang Lingkup

Ruang lingkup dari penelitian ini meliputi:

1. Pencilan yang dideteksi adalah pencilan lokal.
2. Implementasi menggunakan library DMwR package R.

Manfaat

Manfaat dari penelitian ini adalah mendapatkan informasi yang tersembunyi berupa pencilan data titik panas sebagai indikator kebakaran hutan. Penelitian ini juga bermanfaat untuk mengidentifikasi wilayah yang beresiko terjadi kebakaran hutan. Hasil yang diperoleh diharapkan dapat bermanfaat dalam pencegahan kebakaran hutan.

METODE PENELITIAN

Data Penelitian

Data yang digunakan dalam penelitian ini adalah data titik panas di Provinsi Riau dari tahun 2001 sampai 2015 yang diperoleh dari *Fire Information for Resource Management System (FIRMS) MODIS NASA*. Aspek yang diamati pada data titik panas adalah aspek spasial kemunculan titik panas bulanan. Field yang digunakan adalah field latitude, longitude, brightness, scan, track, acq-date, acqtime, satellite, confidence, version, brightt31, dan frp seperti yang ditunjukkan pada tabel 1.

Penjelasan *field* Tabel 1 yaitu: Latitude adalah Koordinat lintang Longitude adalah Koordinat bujur Brightness adalah Suhu kecerahan, diukur dalam Kelvin menggunakan saluran MODIS 21/22 dan saluran 31. Scan Track adalah Resolusi spasial yang sebenarnya dari pixel dipindai. Meskipun algoritma bekerja pada 1 km resolusi, piksel MODIS mendapatkan lebih besar ke tepi scan. AcqDate adalah Tanggal akuisisi pixel titik panas

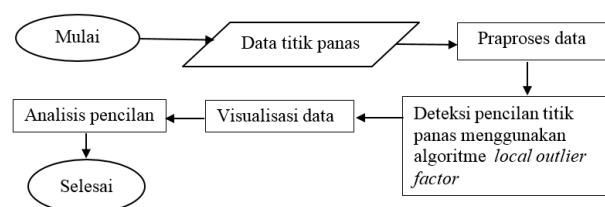
Tabel 1. *Field* data titik panas

Field	Value
LATITUDE	-0.747
LONGITUDE	100.915
BRIGHTNESS	312.2
SCAN	2.4
TRACK	1.5
ACQ_DATE	1/1/2012
ACQ_TIME	624
SATELLITE	A
CONFIDENCE	49
VERSION	5.1
BRIGHT_T31	295
FRP	29.6

aktif. AcqTime adalah Waktu layang satelit di UTC. Satellite adalah Jenis satelit yang mendeteksi. Satelit Terra atau Aqua. Confidence adalah Tingkat kepercayaan dari tiap pixel api yang aktif. Version adalah Versi yang mengidentifikasi koleksi data. BrightT31 adalah Saluran 31 suhu kecerahan (dalam Kelvin) dari pixel titik panas yang aktif. Frp adalah Fire Radiative Power, Menggambarkan kekuatan radiasi api pixel- terintegrasi dalam MW (megawat). FRP memberikan informasi mengenai output panas radiasi diukur dari kebakaran terdeteksi. Jumlah energi panas radiasi dibebaskan per satuan waktu (Fire Radiative Power) diduga terkait dengan tingkat dimana bahan bakar yang dikonsumsi

Tahapan Penelitian

Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan penelitian

a. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data titik panas di Provinsi Riau dari tahun 2001 sampai 2015 yang diperoleh dari FIRMS MODIS NASA. Data titik panas terdiri dari data titik panas tahun 2001 hingga tahun 2015 di wilayah Provinsi Riau. Data tersebut terdiri dari atribut latitude, longitude, brightness, scan,

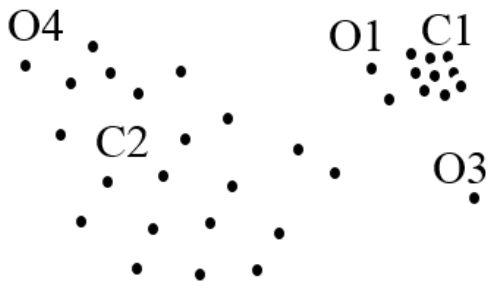
track, acqdate, acqtime, satellite, confidence, version, brightt31, frp. Setiap barisnya menjelaskan satu kemunculan titik panas yang diperoleh dari pengindraan jarak jauh menggunakan sensor MODIS.

b. Praproses Data

Menurut Han et al (2012) “dalam tahap praproses data, terdapat beberapa tahap utama, yaitu pembersihan data, pengintegrasian data, seleksi data, dan transformasi data”. Dalam penelitian ini dilakukan pembersihan dan transformasi data. Pembersihan data dilakukan untuk memilih data titik panas yang berada di Provinsi Riau juga memilih peta Provinsi Riau dari peta kabupaten dan kota se-Indonesia. Langkah ini dilakukan untuk menghilangkan data titik panas yang berada di luar Provinsi Riau. Tahap ini dilakukan menggunakan perangkat lunak PostgreSQL, PostGIS 2.0 *Shapefile* and *DBF Loader Eksporter*, dan Quantum GIS. Setelah data bersih, kemudian data titik panas pada tahun 2001 hingga 2015 dipilih menggunakan queri pada DBMS PostgreSQL dan dilakukan transformasi data yaitu agregasi data. Agregasi data adalah operasi penjumlahan jumlah kejadian titik panas menjadi data harian, bulanan ataupun tahunan.

c. Deteksi Pencilan Titik panas Menggunakan Algoritme *Local Outlier Factor*

Dalam tahapan ini diterapkan fungsi *local outlier factor* pada perangkat lunak R. Fungsi tersebut diberikan masukan atau argumen berupa data frekuensi titik panas harian dari tahun 2001 hingga 2015 juga nilai k sebesar 2 hingga 10. *Local outlier factor* adalah algoritme deteksi outlier berdasarkan jarak tetangga lokal (Beunig et al. 2000). *Local outlier factor* tidak menggunakan distribusi data *global*. Pada Gambar 2 dapat dilihat O1 dan O2 adalah *local outlier* C1, O3 adalah *global outlier*. Dengan pendekatan *clustering* O1 dan O2 tidak dapat terdeteksi sebagai *outlier*.



Gambar 2. Sebaran data

Ilustrasi *local outlier* dapat dilihat pada Gambar 2. *Local outlier factor* menghitung jarak maksimum dari

jarak tetangga dengan jumlah tetangga yang didefinisikan oleh pengguna. Notasinya sebagai berikut (Beunig et al. 2000):

$$reach-dist_k(p, o) = \max\{k - distance(o), d(p, o)\}. \quad (1)$$

dengan k adalah jumlah tetangga dan *distance* adalah fungsi jarak. Sebuah objek *local outlier factor* adalah rasio rata-rata dari o dengan jarak k -ketetanggaan lokal terdekat. Notasinya sebagai berikut (Beunig et al. 2000):

$$lrd_{MinPts}(p) = 1 / \left(\frac{\sum_{o \in N_{MinPts}(p)} reach-dist_{MinPts}(p, o)}{N_{MinPts}(p)} \right) \quad (2)$$

$$LOF_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts}(p)} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{N_{MinPts}(p)} \quad (3)$$

dengan lrd adalah *local reachability density* dan LOF adalah *local outlier factor*.

d. Visualisasi Data

Pada tahapan ini data yang diolah dengan algoritme *local outlier factor* divisualisasikan pada peta sehingga dapat terlihat dengan mudah data mana saja yang termasuk pencilan.

e. Analisis Pencilan

Pada tahap ini diperlihatkan objek-objek pencilan dari data penelitian. Data hasil deteksi pencilan dianalisis untuk mengetahui informasi yang terdapat pada data seperti ukuran pemusatan dan tanggal-tanggal yang terdeteksi pencilan.

Lingkungan Pengembangan

Spesifikasi perangkat keras dan perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Perangkat keras berupa komputer personal dengan spesifikasi
 - Prosesor Intel(R) Core(TM) i7-5500U 2.40GHz
 - RAM 12288 MB.
2. Perangkat lunak

- Komputasi statistika R versi 3.2.0.
- RStudio versi 0.98.1103.
- PostgreSQL dengan ekstensi PostGIS.
- Pengolah data spatial Quantum GIS 2.6.1.
- Microsoft Excel.
- PostGIS 2.0 *Shapefile* and *DBF Loader Ek-sporter*.
- Library DMwR pada perangkat lunak R.

Jadwal Kegiatan

Jadwal penelitian dimulai dari bulan April 2015 sampai dengan Desember 2015. Ilustrasi penjadwalan dapat dilihat pada Tabel 2.

DAFTAR PUSTAKA

- Adinugroho, WC, Suryadiputra INN, BH Saharjo, and L Siboro (2005). *Panduan Pengendalian Kebakaran Hutan dan Lahan Gambut. Proyek Climate Changes, Forests and Peatlands in Indonesia*. Bogor(ID): Wetlands International-Indonesia Programme and Wildlife Habitat Canada.
- Baehaki, D (2014). “Deteksi pencilan data titik panas di provinsi Riau menggunakan algoritme clustering K-Means [skripsi]”. Skripsi. Departemen Ilmu Komputer, Institut Pertanian Bogor. 20 pp.
- Beunig, MM, Kriegel HansPeter, Ng Raymond T, and Sander J, eds. (2000). *LOF: Identifying Density-Based Local Outliers*. ACM SIGMOD international conference on Management of data. Vol. ACM SIGMOD Volume 29 Issue Pages 93104. New York (USA).
- Cahyadarena, MB (2014). “Deteksi Pencilan Pada Data Titik Panas Menggunakan Clustering Berbasis Medoids [skripsi]”. Skripsi. Departemen Ilmu Komputer, Institut Pertanian Bogor. 31 pp.
- Guswanto and E Heriyanto (2009). “Operational Weather System for National Fire Danger Rating. Jurnal Meteorologi dan Geofisika” dalam: *Jurnal Meteorologi dan Geofisika* 10 (2), pp. 77–87.
- Han, J, M Kamber, and Pei J (2012). *Data mining: concepts and techniques*. Massachusetts (US) : Morgan Kaufmann.
- Hasan, MI (1999). *Pokok-Pokok Materi Statistik 1: Statistik Deskriptif*. Jakarta (ID): Bumi aksara.
- Sizer, N, J Anderson, F Stolle, S Minnemeyer, M Higgins, and A Utami A Leach A Alisjahbana (2014). *Kebakaran Hutan di Indonesia Mencapai Tingkat Tertinggi Sejak Kondisi Darurat Kabut Asap Juni 2013 [Internet]*. [diunduh 2015 17 Mei]. Tersedia pada <http://www.wri.org/blog/2014/03/kebakaran-hutan-di-indonesia-mencapai-tingkat-tertinggisejakkondisi-daruratkabut>.
- Suwarsono, Rokhmatuloh, and T Waryono (2013). “Pengembangan Model Identifikasi Daerah Bekas Kebakaran Hutan dan Lahan (Burned Area) Menggunakan Citra MODIS di Kalimantan [Model Development of Burned Area Identification Using MODIS Imagery in Kalimantan]” dalam: *Jurnal Penginderaan Jauh* 10 (2), pp. 99–112.
- Tacconi, L (2003). *Kebakaran Hutan di Indonesia: Penyebab, Biaya dan Implikasi Kebijakan [paper]*. Bogor(ID): Center For International Forestry Research.

Tabel 2. Rencana jadwal penelitian

Jadwal	Tahun 2015									
	April	Mei	Juni	Juli	Agus	Sept	Okt	Nov	Des	
Pengumpulan data										
Praproses data										
Kolokium										
Implementasi <i>Local outlier factor</i>										
Evaluasi hasil penelitian										
Penyusunan skripsi dan makalah seminar										
Seminar										
Sidang tugas akhir										
Revisi Skripsi dan penyelesaian										
Surat Keterangan Lulus (SKL)										