

Grouping Potential Forest and Land Fires Areas in Sumatera Island Based on Hotspot Using CLARA Method

Melda Safitri, Admi Salma*, Nonong Amalita, Fadhilah Fitri

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: admisalma1@fmipa.unp.ac.id

Submitted : 06 Juni 2024
Revised : 11 Agustus 2024
Accepted : 12 Agustus 2024

ABSTRACT

Sumatera Island is one of the areas with the potential for forest and land fires in Indonesia. Sumatera Island has the largest oil palm plantation in Indonesia. The vast land area of oil palm plantations in Indonesia can increase the risk of fires due to land expansion by burning. In addition, the burning of peatlands in Sumatera can exacerbate the impact of forest and land fires. Forest and land fires on the Sumatera Island that occur every year can cause various negative impacts, indicating the need for countermeasures and prevention efforts to minimize the impact of forest and land fires. Hotspots can be used to detect fires in a region and help with prevention and countermeasures to reduce the impact of land and forest fires. Clustering the hotspot data allows one to obtain information on the presence of a fire in a given area as well as its potential status high, medium, or low. The clustering method used is the CLARA method. The CLARA method is a clustering method that breaks the dataset into groups. The advantages of the CLARA method are robust to outliers and effective for large data sets. The results of this research show that the CLARA method can be used for hotspot clustering with a silhouette coefficient of 0.53 in the use of 2 clusters. The analysis of the clustering results shows that cluster 1 is a cluster with low fire potential while cluster 2 is a cluster with high fire potential.

Keywords: CLARA, Hotspot, Silhouette Coefficient, Sumatera Island



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Salah satu bencana alam yang sering terjadi di Indonesia adalah kebakaran hutan dan lahan. Kebakaran hutan dan lahan yang terjadi di Indonesia sudah dianggap sebagai bencana alam transnasional hal ini dikarenakan dampak yang ditimbulkan bukan hanya dirasakan oleh warga negara Indonesia saja tetapi juga dirasakan oleh warga negara tetangga seperti Malaysia dan Singapura (Muzaki dkk. 2021). Salah satu kebakaran hutan dan lahan terparah di Indonesia terjadi pada tahun 1997 hingga 1998 yang menghancurkan 11,7 juta hektar hutan dan lahan. Kebakaran tersebut menyebabkan 20 juta jiwa di Asia Tenggara terpapar asap berbahaya selama berbulan-bulan (Tacconi, 2003). Pada tahun 2015 kebakaran hutan dan lahan kembali terjadi, menyebabkan sekitar 600.000 jiwa menderita Infeksi Saluran Pernapasan Akut (ISPA) dan lebih dari 60 juta jiwa terpapar asap (Saharjo, 2018). Menurut World Bank (2019), Indonesia kembali mengalami kebakaran hutan dan lahan pada tahun 2019 yang menghancurkan sekitar 1,6 juta hektar hutan dan lahan dan menyebabkan kerugian ekonomi sebesar \$5,2 miliar.

Kebakaran hutan dan lahan diakibatkan oleh dua faktor yaitu faktor manusia dan faktor alam. Menurut Yusuf dkk. (2019), faktor manusia, seperti kebiasaan masyarakat untuk membuka lahan dengan dibakar merupakan penyebab utama kebakaran hutan dan lahan. Selain faktor manusia, faktor alam seperti *El-Nino* dapat menyebabkan kekeringan berkepanjangan sehingga meningkatkan risiko kebakaran hutan dan lahan.

Pulau Sumatera merupakan wilayah yang rentan terjadi kebakaran hutan dan lahan di Indonesia. Kebakaran hutan dan lahan yang terjadi di Pulau Sumatera sebagian besar diakibatkan oleh aktivitas manusia seperti pembukaan lahan dengan cara dibakar (Kumalawati, 2020). Berdasarkan data dari Direktorat Jenderal Perkebunan Kementerian Pertanian Republik Indonesia, pada tahun 2023 Sumatera memiliki lahan kelapa sawit terbesar di Indonesia, yaitu sebesar 10,2 juta hektar. Luasnya perkebunan kelapa sawit tersebut meningkatkan risiko terjadinya kebakaran hutan dan lahan akibat upaya perluasan lahan (Prabowo dkk. 2020). Indonesia memiliki 13,43 juta hektar lahan gambut dengan Sumatera sebagai wilayah terbesar yaitu sebesar 5,85 juta hektar (Anda, 2021). Lahan gambut ini sangat rentan terhadap kebakaran, terutama saat mengalami kekeringan, dan melepaskan asap serta gas beracun yang memperburuk dampak kebakaran (Agus & Subiksa, 2008). Hal tersebut dapat menimbulkan berbagai dampak negatif terhadap kesehatan,

transportasi, ekonomi, dan bidang lainnya. Oleh karena itu, untuk mengurangi kerusakan yang ditimbulkan oleh kebakaran hutan dan lahan, diperlukan upaya mitigasi dan pencegahan.

Kebakaran hutan dan lahan menimbulkan dampak yang luas terhadap lingkungan dan kehidupan manusia. Untuk menangani masalah ini, penggunaan teknologi canggih seperti penginderaan jauh dapat dimanfaatkan. Menurut Widodo dkk. (2017), kebakaran hutan dan lahan dapat diatasi dan dicegah dengan mendeteksi titik api menggunakan data penginderaan jauh dan algoritma tertentu. Data penginderaan jauh dapat digunakan untuk mendapatkan informasi suhu permukaan darat yang dapat menunjukkan sebaran titik panas yang mengindikasikan adanya kebakaran. Data titik panas membantu mengidentifikasi wilayah yang berpotensi terjadinya kebakaran hutan dan lahan. Titik panas atau *hotspot* adalah area dengan suhu permukaan yang lebih tinggi daripada area sekitarnya yang dapat diamati oleh satelit. Menurut Prasasti dkk. (2007), satelit Terra/Aqua MODIS memiliki kemampuan untuk mendeteksi titik panas. Satelit Terra/Aqua merupakan satelit observasi lingkungan yang dilengkapi sensor *Moderate Resolusi Imaging Spectroradiometer* (MODIS) yang mampu mengekstraksi data suhu permukaan regional maupun global.

Salah satu cara untuk mencegah kebakaran hutan dan lahan adalah dengan mengelompokkan area-area yang berpotensi terjadi kebakaran. Menurut Pramana dkk. (2023), *clustering* adalah metode pengelompokan data dengan karakteristik yang mirip, sehingga objek di dalam suatu kluster mempunyai karakteristik yang sama dengan objek lain di dalam kluster yang sama tetapi berbeda dari objek lain di dalam kluster yang berbeda. Menurut Simanjuntak & Khaira (2021), dengan melakukan pengelompokan pada data titik panas dapat memberikan informasi ada tidaknya kebakaran di suatu wilayah, ataupun informasi status probabilitas tinggi, sedang, atau rendahnya kebakaran di suatu wilayah.

Dalam mengelompokkan data titik panas akan digunakan metode CLARA. Menurut Kaufman & Rousseeuw (1990:126), metode CLARA adalah metode pengelompokan yang membagi data menjadi beberapa kelompok dengan menggunakan medoid atau objek yang representatif untuk setiap klasternya sebagai titik pusat kluster. Metode CLARA memiliki keunggulan yaitu *robust* terhadap pencilan dan efektif digunakan untuk himpunan data yang besar. Selain itu, CLARA memiliki efisiensi dalam waktu komputasi dan masalah penyimpanan untuk himpunan data yang besar.

Penelitian mengenai metode CLARA telah dilakukan oleh Rifa dkk. (2020), penelitian tersebut membandingkan metode K-Means dan CLARA untuk mengelompokkan data gempa bumi di Indonesia dari tahun 1973 hingga 2017. Hasil penelitian tersebut menyatakan bahwa metode CLARA adalah metode terbaik dengan menghasilkan *silhouette coefficient* sebesar 0,546, sedangkan metode K-Means memperoleh *silhouette coefficient* sebesar 0,516. Penelitian yang berkaitan dengan metode CLARA juga telah dilakukan oleh Jawale & Magar (2019), penelitian tersebut membandingkan antara metode K-Medoids dan CLARA untuk mengelompokkan dataset GeoJson yang terdiri dari 93.852 *records*. Hasil penelitian tersebut menyatakan bahwa metode CLARA adalah metode terbaik dengan menghasilkan *silhouette coefficient* sebesar 0,41 sedangkan metode K-Medoids menghasilkan *silhouette coefficient* sebesar 0,36. Selain itu, penelitian menggunakan metode CLARA juga dilakukan oleh Lindrawati dkk. (2022), untuk mengelompokkan titik panas di Australia dari tahun 2019 hingga 2020. Penelitian tersebut menghasilkan 2 kluster dengan nilai *silhouette coefficient* sebesar 0,89 yang menunjukkan bahwa struktur kluster kuat.

Pada penelitian ini menggunakan metode CLARA untuk mengelompokkan data titik panas di Pulau Sumatera sehingga dapat diketahui daerah di Pulau Sumatera yang berpotensi terjadi kebakaran hutan dan lahan. Data yang digunakan adalah data titik panas di Pulau Sumatera pada tahun 2023. Objek pada penelitian ini yaitu lokasi dari titik panas yang dapat diketahui melalui *latitude* dan *longitude* dari atribut titik panas. Pada penelitian ini digunakan 3 variabel yaitu *brightness*, *confidence*, dan *Fire Radiative Power* (FRP).

II. METODE PENELITIAN

A. Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini yaitu data sekunder berupa data titik panas di Pulau Sumatera tahun 2023 yang diperoleh dari situs NASA FIRMS, yang dikelola oleh *National Aeronautics and Space Administration* (NASA). Data tersebut diperoleh dari data satelit Terra/Aqua menggunakan sensor MODIS. Pada penelitian ini, objek penelitiannya adalah lokasi dari titik panas yang dapat diidentifikasi berdasarkan *latitude* dan *longitude* dari atribut titik panas. Adapun pada penelitian variabel yang digunakan ditunjukkan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Penjelasan	Skala	Satuan
<i>Confidence</i>	<i>Confidence</i> menunjukkan probabilitas terjadinya kebakaran di lapangan yang dipantau melalui citra satelit. Skala kualitas tingkat kepercayaan titik panas berkisar antara 0% hingga 100%. Semakin tinggi tingkat kepercayaannya, semakin besar kemungkinan titik panas tersebut benar-benar merupakan kebakaran.	Rasio	Persen

Variabel	Penjelasan	Skala	Satuan
Brightness	Suhu kecerahan piksel api yang diukur dalam satuan kelvin.	Rasio	Kelvin
Fire Radiative Power (FRP)	Kekuatan radiasi piksel api yang kemudian diintegrasikan ke dalam MW (MegaWatts).	Rasio	MegaWatts

B. Teknik Analisis Data

Dalam penelitian ini, analisis data menggunakan metode CLARA dengan memanfaatkan *software* R-Studio. Proses analisis pada penelitian ini yaitu sebagai berikut:

1. Mengumpulkan data titik panas di Pulau Sumatera pada tahun 2023 dari *website* NASA FIRMS.
2. Melakukan pemeriksaan mengenai keberadaan pencilan.

Pencilan (*outlier*) merupakan suatu pengamatan yang memiliki nilai yang cukup jauh berbeda dengan pengamatan lainnya (Hair dkk. 2019:85). Terdapat beberapa metode yang dapat digunakan untuk pendeteksian pencilan, diantaranya yaitu boxplot, grafis, maupun jarak kuadrat mahalanobis. Dalam kasus multivariat, jarak kuadrat mahalanobis dapat digunakan. Menurut Johnson & Wichern (2007:189), jarak kuadrat mahalanobis dapat dihitung menggunakan persamaan berikut.

$$d_{MD}^2(i) = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (1)$$

Keterangan:

$d_{MD}^2(i)$: Jarak kuadrat mahalanobis objek ke- i

\mathbf{x}_i : Vektor objek amatan ke- i

$\bar{\mathbf{x}}$: Vektor rata-rata

\mathbf{S} : Matriks kovarians

Objek ke- i teridentifikasi sebagai pencilan jika

$$d_{MD}^2(i) > \chi_{p,1-\alpha}^2$$

Dimana $\chi_{p,1-\alpha}^2$ merupakan *chi-square* dengan p adalah banyaknya variabel yang diteliti dengan probabilitas $1 - \alpha$.

3. Menggunakan metode *silhouette coefficient* untuk menentukan jumlah kluster optimal yang akan dibentuk.

Penentuan jumlah kluster optimal dapat ditentukan menggunakan metode *silhouette coefficient*. Menurut Larose & Larose (2015:583-584), metode *silhouette coefficient* menggabungkan dua metode perhitungan, yaitu metode *cohesion* untuk menilai kedekatan antar objek dalam kluster yang sama dan metode *separation* untuk menilai perbedaan sebuah kluster dengan kluster lainnya. *Silhouette coefficient* untuk setiap objek ke- i dihitung dengan menggunakan rumus berikut.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2)$$

Keterangan:

$s(i)$: Nilai *silhouette coefficient* pada objek ke- i

$a(i)$: Ukuran *cohesion* objek ke- i

$b(i)$: Ukuran *separation* objek ke- i

Untuk mendapatkan nilai *silhouette coefficient* dari seluruh hasil kluster dilakukan dengan menghitung rata-rata nilai *silhouette coefficient* dari setiap objeknya. Kaufman & Rousseeuw (1990:86) menyatakan bahwa nilai *silhouette coefficient* tertinggi di antara jumlah k lainnya adalah dasar untuk menentukan jumlah kluster yang optimal.

4. Melakukan pengelompokkan titik panas menggunakan metode CLARA.

Clustering Large Application (CLARA) merupakan pengembangan dari metode Partition Around Medoids (PAM) yang dikenalkan oleh Kaufman dan Rousseeuw pada tahun 1987. CLARA menggunakan pendekatan *sampling* dengan mengambil sampel acak dari kumpulan data, kemudian menerapkan algoritma PAM untuk menemukan medoid terbaik dari sampel tersebut. Dengan mengambil sampel acak dari dataset, diharapkan bahwa nilai medoid dari sampel akan mencerminkan nilai medoid dari seluruh dataset (Han, 2012:456). Menurut Kaufman & Rousseeuw (1990:144), CLARA memiliki sifat yang dapat mengatasi pencilan dan mampu menangani data dalam skala besar. Selain itu, metode CLARA lebih efisien daripada PAM dalam hal waktu komputasi dan masalah penyimpanan untuk data dalam jumlah besar. Berikut adalah algoritma dari metode CLARA:

- a. Menentukan banyak kluster yang akan dibentuk sebanyak k kluster.
- b. Menarik sampel sebanyak n' dimana $n' = 40 + 2k$, gugus sampel yang ditarik sebanyak lima set sampel.
- c. Menerapkan algoritma PAM pada setiap set sampel. Berikut langkah dalam algoritma PAM.
 - 1) Menentukan k , yaitu banyak kluster yang ingin dibentuk. Banyaknya kluster ditentukan oleh peneliti.
 - 2) Memilih secara acak k objek dari kumpulan objek yang akan dikelompokkan untuk dijadikan medoid awal.

- 3) Menggunakan jarak *Euclidean* untuk menghitung jarak antara objek non-medoid dengan medoid dalam setiap kluster.

$$d(i, j) = \sqrt{\sum_{f=1}^p (x_{if} - x_{jf})^2} \quad (3)$$

Keterangan:

$d(i, j)$: Jarak antara objek i dan objek j
 x_{ik} : Nilai data dari objek ke- i dalam variabel ke- f
 x_{jk} : Nilai data dari objek ke- j dalam variabel ke- f
 p : Jumlah variabel

- 4) Menempatkan objek berdasarkan jarak terdekat dengan medoid. Kemudian, menghitung total jarak yang diperoleh.
 5) Pada setiap kluster, pilih objek non-medoid secara acak sebagai kandidat medoid baru.
 6) Menghitung jarak setiap objek non-medoid terhadap kandidat medoid baru menggunakan jarak *Euclidean* dan tempatkan objek tersebut berdasarkan jarak terdekat dengan kandidat medoid baru tersebut. Selanjutnya, menghitung total jarak yang diperoleh dengan kandidat medoid baru.
 7) Setelah menghitung total jarak medoid baru dan total jarak medoid lama, selanjutnya yaitu menghitung selisih total jarak (S).

$$S = \text{total jarak medoid baru} - \text{total jarak medoid lama}$$

Jika didapatkan nilai $S < 0$, maka kandidat medoid baru tersebut menjadi medoid baru. Ulangi langkah 5 sampai 7 hingga tidak ada perubahan pada medoid. Medoid tidak berubah ketika diperoleh nilai $S > 0$.

- d. Menghitung jarak antara semua objek di dataset yang tidak menjadi medoid dengan objek yang menjadi medoid dengan menggunakan jarak *Euclidean*.
 e. Set sampel dengan jumlah jarak terkecil adalah yang dipilih.
 5. Memvalidasi hasil kluster dengan menggunakan metode *silhouette coefficient*
 Metode *silhouette coefficient* juga dapat digunakan untuk memvalidasi hasil kluster. Menurut Kaufman & Rousseeuw (1990:88), terdapat kategori nilai *silhouette coefficient* yang ditunjukkan pada Tabel 2 berikut.

Tabel 2. Kategori Nilai *Silhouette Coefficient*

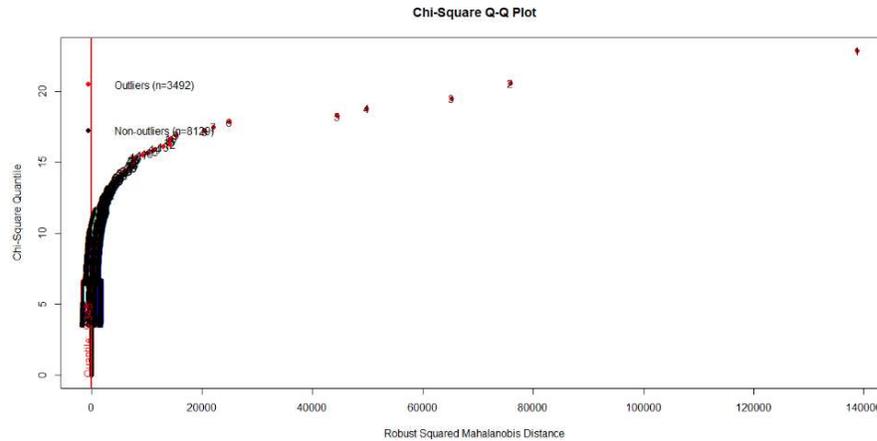
<i>Silhouette Coefficient</i>	Interpretasi
0,71 – 1,00	Struktur kluster kuat
0,51 – 0,70	Struktur kluster baik
0,26 – 0,50	Struktur kluster lemah
$\leq 0,25$	Tidak terstruktur

6. Menginterpretasikan hasil kluster.
 Interpretasi hasil kluster ditentukan berdasarkan nilai rata-rata (*mean*) variabel pada setiap objek yang terdapat di setiap kluster. Menurut Hair (2019:227), hasil kluster akan memberikan gambaran tentang karakteristik yang membedakan setiap kluster yang terbentuk melalui proses pengelompokan dengan metode CLARA.

III. HASIL DAN PEMBAHASAN

A. Mendeteksi Pencilan

Pendeteksian pencilan pada penelitian ini dilakukan dengan menggunakan jarak kuadrat mahalanobis. Pada Gambar 1 berikut menunjukkan sebaran pencilan pada data titik panas di Pulau Sumatera tahun 2023.

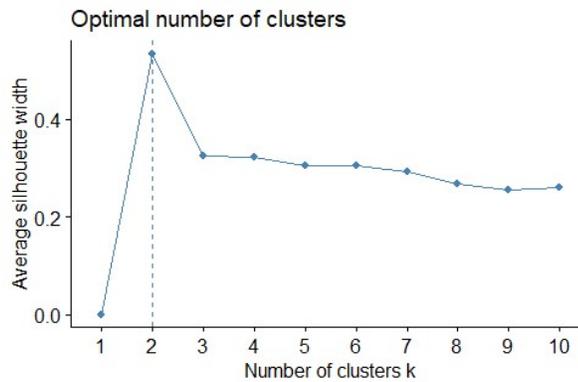


Gambar 1. Persebaran Penciran Titik Panas di Pulau Sumatera Tahun 2023

Sebaran penciran data titik panas di Pulau Sumatera tahun 2023 yang ditunjukkan pada Gambar 1 dapat diketahui bahwa terdapat data penciran yaitu sebanyak 3.492 titik panas. Oleh karena adanya penciran pada data titik panas di Pulau Sumatera maka analisis kluster yang dapat digunakan adalah metode CLARA karena memiliki sifat yang *robust* terhadap penciran dan efisien dalam mengelompokkan himpunan data yang besar.

B. Analisis Kluster dengan Metode CLARA

Pada metode CLARA, langkah awal yang perlu dilakukan yaitu menentukan jumlah kluster. Jumlah kluster optimal dapat ditentukan dengan menggunakan metode *silhouette coefficient*. Pada Gambar 2 ditampilkan jumlah kluster optimal pada metode CLARA.



Gambar 2. Plot *Silhouette Coefficient*

Hasil dari plot *silhouette coefficient* yang ditampilkan pada Gambar 2, menunjukkan bahwa jumlah kluster 2 memiliki nilai *silhouette coefficient* tertinggi di antara jumlah kluster lainnya. Oleh karena itu, titik panas di Pulau Sumatera akan dikelompokkan menjadi 2 kluster. Pengelompokkan titik panas di Pulau Sumatera menggunakan metode CLARA dengan 2 kluster diperoleh medoid terbaik yang ditunjukkan pada Tabel 3.

Tabel 3. Medoid Terbaik

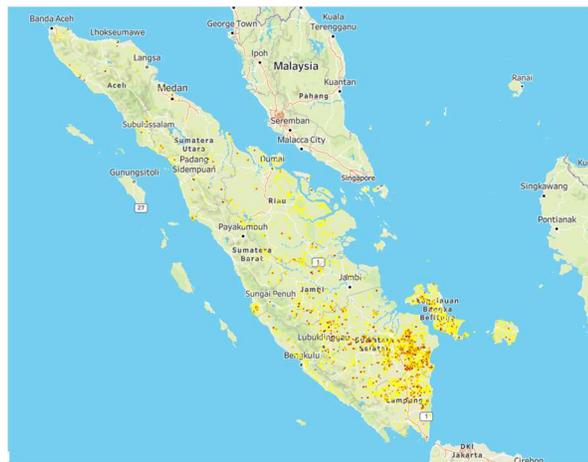
Kluster	Brightness	Confidence	FRP
1	311,92	59	12,54
2	351,2	95	56,51

Hasil dari pengelompokan menggunakan metode CLARA dengan menggunakan medoid terbaik yang ditunjukkan pada Tabel 3 diperoleh kluster 1 terdiri dari 9.432 titik panas dan kluster 2 terdiri dari 2.189 titik panas. Setelah diperoleh hasil pengelompokan menggunakan metode CLARA, pada Gambar 3 berikut akan ditampilkan validasi hasil kluster menggunakan *silhouette coefficient*.



Gambar 3. Validasi Hasil Kluster

Hasil validasi kluster menggunakan *silhouette coefficient* pada Gambar 3 menunjukkan nilai *silhouette coefficient* sebesar 0,53 yang mengindikasikan bahwa hasil pengelompokan baik. Selanjutnya, pada Gambar 4 berikut menunjukkan visualisasi pengelompokan dengan menggunakan metode CLARA pada data titik panas di Pulau Sumatera.



Gambar 4. Peta Sebaran Pengelompokan Menggunakan Metode CLARA (k=2)

Pada Gambar 4, titik kuning menunjukkan titik panas pada kluster 1 dan titik merah adalah titik panas pada kluster 2. Pada Tabel 4 disajikan karakteristik masing-masing kluster.

Tabel 4. Karakteristik Hasil Kluster

Kluster	Brightness	Confidence	FRP
1	316,5	58,54	15,282
2	340,3	89,54	75,39

Pada Tabel 4 ditunjukkan karakteristik masing-masing kluster yang diperoleh dari hasil rata-rata variabel untuk setiap kluster. Pada kluster 1 memiliki nilai rata-rata untuk setiap variabelnya lebih rendah dibandingkan dengan kluster 2. Sehingga dapat dikategorikan bahwa kluster 1 memiliki potensi yang lebih rendah untuk mengalami kebakaran hutan dan lahan. Sedangkan kluster 2 termasuk ke dalam kategori titik panas yang memiliki potensi yang tinggi untuk

mengalami kebakaran hutan dan lahan. Lokasi dari titik panas untuk masing-masing kluster ditunjukkan pada Gambar 5.



Gambar 5. (a) Peta Sebaran Pengelompokan kluster 1 (b) Peta Sebaran Pengelompokan kluster 2

Hasil visualisasi lokasi titik panas menunjukkan bahwa titik panas tersebar hampir di seluruh wilayah di Pulau Sumatera. Berdasarkan Gambar 5, kluster 2 menunjukkan titik panas yang memiliki potensi tinggi untuk kebakaran hutan dan lahan, sehingga informasi ini penting untuk ditindaklanjuti dengan pemantauan lebih lanjut terhadap lokasi titik panas. Dengan adanya titik panas yang menunjukkan potensi tinggi kebakaran hutan dan lahan, hasil pengelompokan ini dapat dipertimbangkan untuk pembentukan pos pemantauan kebakaran hutan dan lahan.

IV. Kesimpulan

Hasil analisis menunjukkan bahwa terbentuk 2 kluster menggunakan metode CLARA untuk mengelompokkan wilayah potensial kebakaran hutan dan lahan di Pulau Sumatera yang didasarkan pada titik panas pada tahun 2023. Berdasarkan hasil validasi *silhouette coefficient* diperoleh nilai *silhouette coefficient* sebesar 0,53 yang menandakan bahwa pembentukan kluster baik.

Metode CLARA menghasilkan 2 kluster dengan kategori kluster 1 merupakan kluster yang potensi terjadi kebakaran rendah, sedangkan kluster 2 merupakan kluster dengan potensi terjadi kebakaran tinggi. Daerah dengan potensi terjadi kebakaran yang tinggi dapat menjadi informasi penting untuk fokus pemantauan lebih lanjut terhadap lokasi titik panas.

Peneliti selanjutnya disarankan menggunakan metode *clustering* lainnya untuk mengelompokkan data titik panas di Pulau Sumatera sehingga dapat dibandingkan metode mana yang lebih efektif.

DAFTAR PUSTAKA

- Agus, F., & Subiksa, I. M. (2008). *Lahan Gambut: Potensi untuk Pertanian dan Aspek Lingkungan*. Bogor: Balai Penelitian Tanah dan World Agroforestry Centre (ICRAF).
- Anda, M., Ritung, S., Suryani, E., Hikmat, M., Yatno, E., Mulyani, A., & Subandiono, R. E. (2021). Revisiting tropical peatlands in Indonesia: Semi-detailed mapping, extent and depth distribution assessment. *Geoderma*, 402, 115235.
- Direktorat Jendral Perkebunan Kementerian Pertanian Republik Indonesia. (2023). *Statistik Perkebunan Unggulan Nasional 2021-2023*. Jakarta: Direktorat Jendral Perkebunan.
- Hair, J. F., Babin, B. J., Anderson, R. E., & Black, W. C. (2019). *Multivariate Data Analysis (8th Ed)*. England: Pearson Prentice.
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques (3rd ed.)*. Waltham: Morgan kaufman Publishers.

- Jawale, A., & Magar, G. (2019). Survey of Clustering Methods for Large Scale Dataset. *International Journal of Computer Sciences and Engineering (IJSE)*, 7(5), 1338-1344.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis (6th Ed)*. New Jersey: Prentice International Inc.
- Kaufman, L., & Rousseeuw, P.J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons Inc.
- Kumalawati, R., Nasruddin, N., Hendra Murliawan, K., Yuliarti, A., & Pangaribuan, A. N. (2020). Sebaran Hotspot Dan Komunikasi Masyarakat Dalam Menyikapi Bencana Kebakaran Lahan Gambut. *Prosiding Seminar Nasional Geomatika 2020: Informasi Geospasial untuk Inovasi Percepatan Pembangunan Berkelanjutan*. 223-243.
- Larose, D. T., & Larose, C. D. (2015). *Data Mining and Predictive Analytics (2nd ed.)*. New Jersey: John Wiley & Sons, Inc.
- Lindrawati, E., Bahri, S., Zubaeda, U. F., Carolina, V. P., Kusriani, & Maulina, D. (2022). Kebakaran Hutan Implementasi Metode CLARA Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot). *Journal of Computer System and Informatics (JoSYC)*, 3(4), 507-511.
- Muzaki, A., Pratiwi, R., & Zahro, S. R. A. (2021). Pengendalian Kebakaran Hutan Melalui Penguatan Peran Polisi Kehutanan Untuk Mewujudkan Sustainable Development Goals. *LITRA: Jurnal Hukum Lingkungan, Tata Ruang, dan Agraria*, 1(1), 22-44.
- Prabowo, M. R., Koesmaryono, Y., Faqih, A., & Sopaheluwakan, A. (2020). Karakteristik spasial dan temporal hotspot di Pulau Sumatera. *Jurnal Meteorologi dan Geofisika*, 21(1), 9-19.
- Pramana, S., Yuniarto, B., Santoso, I., Nooraeni, R., & Suadaa, L. H. (2023). *Data Mining dengan R Konsep dan Implementasi*. Bogor: IN MEDIA.
- Prasasti, I., Sambodo, K. A., Carolita, I., & Inderaja, P. P. (2007). Pengkajian Pemanfaatan Data Terra-Modis Untuk Ekstraksi Data Suhu Permukaan Lahan (SPL) Berdasarkan Beberapa Algoritma. *Lembaga Penerbangan dan Antariksa Nasional*, 4(1). 1-8.
- Rifa, H. H., Pratiwi, H., & Respatiwan. (2020). Clustering of Earthquake Risk in Indonesia Using K-Medoids and K-Means Algorithms. *Media Statistika*, 13(2), 194-205.
- Simanjuntak, K. P., & Khaira, U. (2021). Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 7-16.
- Tacconi, L., & Center for International Forestry Research (CIFOR). (2003). *Fires In Indonesia: Causes, Costs And Policy Implications*. (Vol.38). Bogor, Indonesia: Cifor.
- Widodo, A. M., Dulbahri, D., & Hartono, H. (2017). Penggunaan Data Penginderaan Jauh dan Sistem Informasi Geografis untuk Pembuatan Prototipe Perangkat Lunak Simulasi Penyebaran Kebakaran Hutan. *Majalah Geografi Indonesia*, 31(1), 12.
- World Bank. (2015). *Reforming Amid Uncertainty*. Washington, D.C: World Bank Group.
- World Bank. (2019). *Membangun Manusia*. Washington, D.C: World Bank Group.
- Yusuf, A., Hapsoh, H., Siregar, S. H., & Nurrochmat, D. R. (2019). Analisis Kebakaran Hutan Dan Lahan Di Provinsi Riau. *Dinamika Lingkungan Indonesia*, 6(2), 67-84.