

Pemetaan Zona Iklim di Sulawesi Utara Berdasarkan Kelembaban dan Presipitasi Menggunakan K-Means Clustering

Afrioni Roma Rio^{1*}

¹Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sam Ratulangi
Email: afrioni@unsrat.ac.id

ABSTRACT

This study aimed to map the rainfall patterns in Sulawesi Utara by clustering the relative humidity and precipitation data of 171 sub-districts in the province. Data from the period of 2020-2024 was used to perform k-means clustering to group the sub-districts based on their average relative humidity and precipitation. The analysis resulted in four clusters of sub-districts, where there was no significant difference in the average relative humidity among them. However, the distribution of precipitation data differed significantly among the clusters. Cluster 1 had the highest level of precipitation, followed by clusters 2 and 3, while cluster 4 had the lowest level of precipitation. This analysis can provide a basis for the government and related institutions to map the rainfall patterns in Sulawesi Utara and help in planning and managing water resources. The information on the distribution of precipitation data in each cluster can also guide the mapping of flood and landslide potential in the area. This study can serve as a reference for further research on rainfall patterns in Sulawesi Utara.

Keywords: Clusterization, Humidity, Precipitation

ABSTRAK

Klasterisasi kelembaban dan presipitasi di Sulawesi Utara dilakukan dalam rangka memetakan pola curah hujan di wilayah tersebut. Kelembaban relatif dan presipitasi adalah dua parameter penting dalam meteorologi yang sering digunakan untuk memprediksi pola curah hujan. Dalam penelitian ini, data yang digunakan adalah data dalam rentang lima tahun dari tahun 2020 hingga tahun 2024 yang diambil dari 171 kecamatan di Sulawesi Utara. Metode k-means clustering digunakan untuk mengelompokkan kecamatan-kecamatan di Sulawesi Utara berdasarkan rata-rata kelembaban relatif dan presipitasi. Dalam analisis ini, empat klaster kecamatan berhasil dihasilkan. Meskipun rata-rata kelembaban relatif tidak menunjukkan perbedaan yang signifikan antara keempat klaster, namun terdapat perbedaan signifikan dalam sebaran data presipitasi antara klaster-klaster tersebut. Klaster 1 merupakan klaster dengan tingkat presipitasi tertinggi, diikuti oleh klaster 2 dan klaster 3. Klaster 4 merupakan klaster dengan tingkat presipitasi paling rendah dibandingkan dengan klaster-klaster yang lain. Analisis ini dapat menjadi dasar bagi pemerintah dan lembaga terkait dalam memetakan pola curah hujan di Sulawesi Utara dan membantu dalam perencanaan dan pengelolaan sumber daya air. Selain itu, informasi tentang sebaran data presipitasi pada setiap klaster dapat menjadi pedoman dalam memetakan potensi bencana banjir dan tanah longsor di wilayah tersebut. Penelitian ini juga dapat menjadi acuan bagi penelitian lanjutan tentang pola curah hujan di Sulawesi Utara.

Kata Kunci: Klasterisasi, Kelembaban, Presipitasi

1. PENDAHULUAN

Sulawesi Utara, yang terletak di wilayah Indonesia bagian tengah, memiliki luas selebar 14.511 km² dan terdiri dari 11 kabupaten dan 4 kota madya (Kementerian Dalam Negeri-Dukcapil, 2024). Meskipun memiliki potensi sumber daya alam yang melimpah, provinsi ini tidak luput dari bencana alam seperti banjir, longsor, angin puting beliung, dan bencana alam lainnya. Menurut Data Informasi Bencana Indonesia (DIBI) (Badan Nasional Penanggulangan Bencana (BNPB), 2024), pada tahun 2022, di Sulawesi Utara tercatat sebanyak 74 bencana alam, sebagian besar di antaranya berasal dari banjir dan tanah longsor. Bencana ini menimbulkan kerugian yang besar bagi masyarakat setempat, termasuk kerugian material dan korban jiwa.

Salah satu faktor yang menyebabkan terjadinya banjir adalah curah hujan yang tinggi. Hal ini sering terjadi di Sulawesi Utara karena wilayah ini terletak di zona intertropis yang memiliki curah hujan tinggi sepanjang tahun. Pada saat musim hujan, intensitas hujan dapat meningkat secara drastis, yang dapat menyebabkan banjir dan tanah longsor. Dalam hal ini, keberadaan data presipitasi dan kelembaban sangat penting dalam memetakan potensi bencana di Sulawesi Utara.

Presipitasi, yang diukur dalam milimeter (mm), adalah hasil kondensasi uap air di atmosfer yang jatuh ke permukaan bumi (American Meteorological Society, 2024). Parameter ini sangat penting karena dampaknya yang langsung terhadap lingkungan dan kehidupan manusia. Selain presipitasi, parameter kunci lainnya adalah kelembaban udara, yaitu ukuran konsentrasi uap air di atmosfer yang dinyatakan dalam persen (%). Tingkat kelembaban berperan vital dalam siklus hidrologi karena memengaruhi proses pembentukan awan yang pada akhirnya menghasilkan presipitasi. Mengingat pentingnya kedua parameter tersebut, pengelompokan wilayah di Sulawesi Utara berdasarkan jumlah curah hujan dan tingkat kelembaban menjadi langkah penting untuk mitigasi bencana alam.

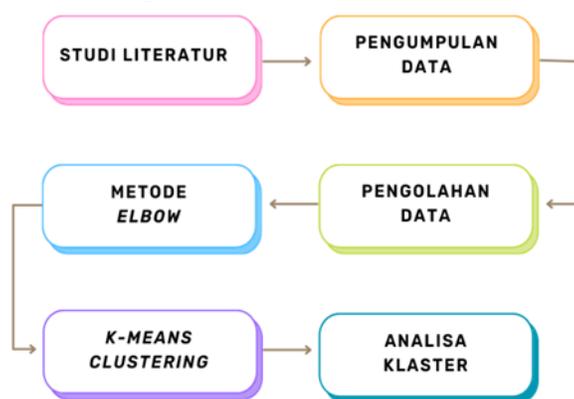
Pengelompokan 171 kecamatan di Sulawesi Utara dilakukan berdasarkan data curah hujan (presipitasi) dan kelembaban selama lima tahun (2020-2024) menggunakan metode K-Means. Dari analisis ini, terbentuk empat kluster atau kelompok wilayah dengan karakteristik yang berbeda. Kluster pertama memiliki tingkat presipitasi tertinggi, diikuti oleh kluster kedua dan ketiga, sementara kluster keempat memiliki tingkat presipitasi terendah dibandingkan dengan kluster lainnya. Sebaliknya, untuk kelembaban, tidak ada perbedaan berarti yang terdeteksi di antara keempat kluster tersebut.

Pengelompokan wilayah berdasarkan data presipitasi dan kelembaban di Sulawesi Utara memiliki dampak yang penting dalam pemetaan potensi bencana dan perencanaan pengelolaan sumber daya alam. Dalam konteks pengelolaan bencana, pemerintah dapat memetakan wilayah-wilayah yang berpotensi mengalami bencana dan meningkatkan kesiapsiagaan masyarakat dalam menghadapi bencana.

Sementara itu, pengelompokan wilayah berdasarkan data presipitasi dan kelembaban juga dapat membantu dalam perencanaan pengelolaan sumber daya alam. Dalam hal ini, pengelompokan wilayah dapat menjadi acuan dalam pengelolaan air dan pertanian di daerah yang memiliki karakteristik iklim dan tanah yang sama. Dengan mengelompokkan wilayah berdasarkan karakteristik iklim dan tanah, maka akan memudahkan dalam menentukan jenis tanaman yang cocok untuk ditanam di daerah tersebut. Selain itu, pengelompokan wilayah juga dapat menjadi dasar dalam penentuan kebijakan untuk mengoptimalkan penggunaan air dan menjaga ketersediaan air bagi masyarakat.

2. METODE PENELITIAN

Penelitian ini melalui beberapa tahapan, yaitu studi literatur, pengumpulan data, pengolahan data, metode *elbow* (siku), k-means clustering, dan analisa kluster (Gambar 1).



Gambar 1. Diagram alir penelitian.

2.1. Studi Literatur

Penelitian yang baik dan akurat selalu didasarkan pada kumpulan literatur dan referensi yang tepat. Oleh karena itu, tahap awal dalam sebuah penelitian adalah mengumpulkan literatur yang

berkaitan dengan topik yang akan diteliti. Dalam penelitian ini, literatur yang dikumpulkan meliputi beberapa aspek penting yang berhubungan dengan objek penelitian.

Pertama-tama, penulis mengumpulkan informasi geografis tentang Sulawesi Utara, yaitu provinsi di Indonesia yang menjadi objek penelitian. Hal ini dilakukan untuk memahami kondisi geografis daerah tersebut, seperti letak, topografi, dan karakteristik lainnya.

Selanjutnya, penulis mengumpulkan definisi-definisi meteorologi, yang meliputi berbagai istilah dan konsep penting dalam bidang meteorologi. Ini dilakukan untuk memperluas pemahaman penulis tentang topik yang akan diteliti dan untuk memastikan bahwa penulis memahami konsep-konsep dasar yang terkait dengan meteorologi.

Selain itu, penulis juga mempelajari tentang pembelajaran mesin atau *machine learning*, yaitu teknik atau algoritma yang digunakan untuk memproses data dalam penelitian ini. Hal ini penting karena *machine learning* menjadi dasar dalam klusterisasi data dalam penelitian ini.

Terakhir, penulis mempelajari teknik klusterisasi menggunakan metode k-means. Klusterisasi merupakan teknik untuk mengelompokkan data yang memiliki karakteristik yang sama ke dalam kelompok-kelompok tertentu. Dalam penelitian ini, klusterisasi dilakukan untuk memetakan pola kelembaban dan presipitasi di Sulawesi Utara.

2.2. Pengumpulan Data

Data yang digunakan dalam proses klusterisasi adalah data kelembaban dan presipitasi dari 171 kecamatan di provinsi Sulawesi Utara. Data diambil dari NASA Langley Research Center (LaRC) POWER (Center (LaRC), 2024). Untuk memperoleh data kelembaban dan presipitasi tersebut, diperlukan koordinat dari lokasi kecamatan, yaitu garis lintang (*latitude*) dan garis bujur (*longitude*). Koordinat diperoleh dengan *scrapping* data dari Google map untuk masing-masing kecamatan. Data kecamatan beserta dengan informasi koordinatnya kemudian disimpan dalam bentuk file csv.

Setelah data koordinat diperoleh, kemudian data kelembaban dan presipitasi dapat diambil menggunakan API (*application programming interface*) yang telah disediakan oleh LaRC. Pengumpulan data melalui API menggunakan program Python (Sanner & Michel, 1999), kemudian disimpan dalam file csv, untuk selanjutnya dapat diolah.

2.3. Pengolahan Data

Setelah data dari semua kecamatan diperoleh, langkah selanjutnya adalah mengolah dan eksplorasi data. Proses eksplorasi data menggunakan bahasa pemrograman Python (Sanner & Michel, 1999), dengan beberapa package dan library yang digunakan yaitu:

- Kalkulasi numerik dari suatu array atau matriks menggunakan Numpy (Harris dkk., 2020);
- Analisa dan pengolahan data dalam bentuk dataframe menggunakan Pandas (McKinney, 2011);
- Visualisasi grafik dan plot sebaran data pada peta menggunakan Matplotlib, Seaborn, dan Plotly (berserta API yang disediakan oleh Mapbox) (Barrett dkk., 2005; Mapbox, 2023; Plotly Technologies Inc., 2015; Waskom, 2021).

2.4. Metode Elbow

Metode *elbow* (siku) adalah sebuah teknik grafis yang digunakan untuk menentukan jumlah kluster yang optimal dalam proses klusterisasi data. Metode ini melibatkan pembuatan grafik yang menunjukkan seberapa baik setiap kluster menjelaskan variasi dalam data, berdasarkan jumlah kluster yang digunakan. Pada grafik tersebut, kita mencari titik "*elbow*" yang menandakan titik di mana penambahan kluster tidak lagi memberikan peningkatan yang signifikan dalam pemahaman kita terhadap variasi data (Bholowalia & Kumar, 2014).

2.5. K-means Clustering

Tahap selanjutnya adalah mengklusterisasi wilayah kecamatan yang ada di Sulawesi Utara berdasarkan data kelembaban dan presipitasi. Metode yang digunakan adalah metode k-means. Metode k-means adalah salah satu metode klusterisasi yang termasuk dalam kategori *unsupervised machine learning* (Mahesh, 2020). K-means adalah algoritma klusterisasi yang membagi data menjadi k kelompok (kluster) yang berbeda-beda. Proses klusterisasi dimulai dengan menentukan nilai k, yaitu jumlah kluster yang diinginkan.

Langkah selanjutnya adalah mengelompokkan data ke dalam kelompok-kelompok tersebut berdasarkan jarak antara titik data dan pusat kluster (*centroid*). Algoritma k-means melakukan iterasi untuk menentukan pusat kluster, yaitu titik yang paling dekat dengan titik data dalam setiap kelompok.

Proses iterasi ini dilakukan hingga tidak ada lagi perubahan dalam posisi titik pusat kluster atau telah mencapai batas iterasi tertentu (Madhulatha, 2012). *Library* Python yang digunakan untuk proses klusterisasi menggunakan k-means adalah Scikit-learn, yang merupakan *library* standar untuk *machine learning* (Pedregosa, 2011).

2.6. Analisa Kluster

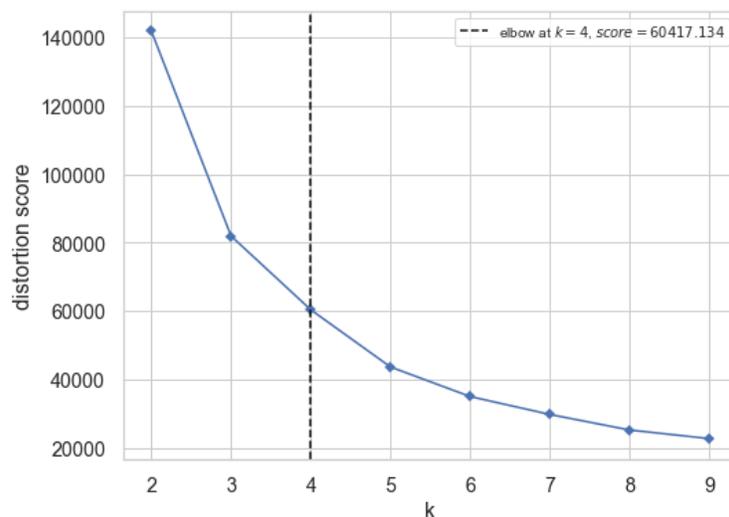
Selanjutnya, kluster-kluster yang dihasilkan akan dianalisis untuk mengidentifikasi pola atau kecenderungan yang terdapat dalam data. Analisis ini melibatkan pengecekan karakteristik dari setiap kluster, seperti nilai rata-rata variabel yang terkait, serta penggunaan grafik seperti *boxplot* untuk memvisualisasikan kluster tersebut.

3. HASIL DAN PEMBAHASAN

Kelembaban dan presipitasi adalah dua hal yang saling terkait dalam siklus hidrologi alam. Kelembaban adalah kandungan uap air dalam udara, sedangkan presipitasi adalah curahan air yang jatuh ke permukaan bumi. Kedua hal ini sangat penting dalam menentukan iklim dan cuaca di suatu wilayah.

Pola kelembaban dan presipitasi di suatu wilayah dipengaruhi oleh faktor-faktor seperti posisi geografis, topografi, dan sirkulasi udara. Wilayah yang berada di dekat garis khatulistiwa cenderung memiliki kelembaban dan presipitasi yang tinggi sepanjang tahun, sedangkan wilayah yang berada di belahan bumi utara dan selatan memiliki perbedaan yang jelas antara musim-musim yang lainnya.

Dalam artikel ini kelembaban dan presipitasi digunakan sebagai parameter untuk mengelompokkan (klusterisasi) wilayah kecamatan di Sulawesi Utara, kelembaban yang digunakan adalah kelembaban relatif, dimana kelembaban relatif (*relative humidity*) adalah rasio antara jumlah uap air yang terkandung dalam udara pada suhu tertentu dibandingkan dengan jumlah uap air maksimal yang dapat ditampung oleh udara pada suhu dan tekanan tersebut, dikalikan dengan 100%. Sedangkan pengukuran presipitasi menggunakan satuan milimeter (mm). Satuan ini digunakan untuk mengukur jumlah air hujan yang jatuh di suatu daerah dalam bentuk cairan. Dalam pengukuran presipitasi, satu milimeter presipitasi setara dengan satu liter air per meter persegi permukaan daerah yang terkena hujan.



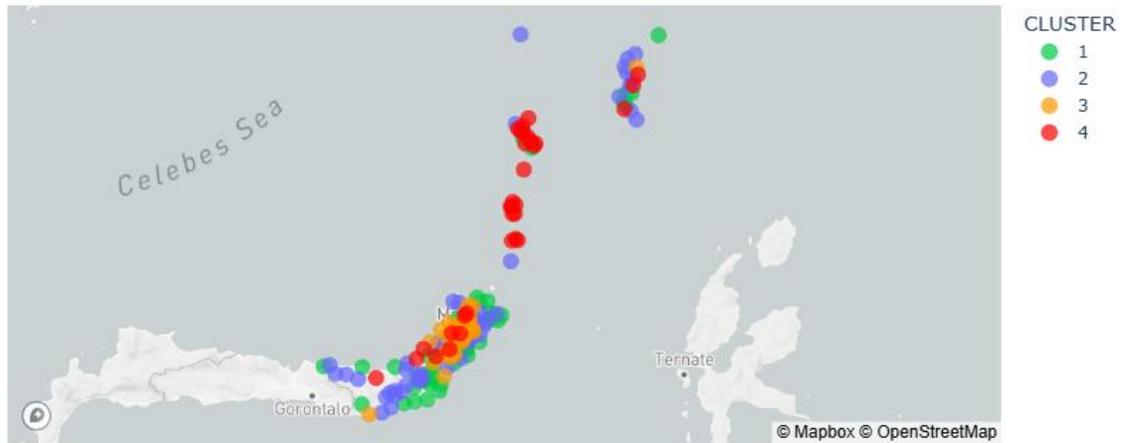
Gambar 2. Metode *elbow* untuk menentukan kluster optimum.

Untuk melakukan klusterisasi, langkah awal yang perlu dilakukan adalah menentukan jumlah kelompok atau kluster yang optimal. Dalam hal ini, metode *elbow* biasa digunakan sebagai alat bantu untuk menentukan jumlah kluster yang tepat. Metode *elbow* bekerja dengan mengukur variansi dari kluster terhadap jumlah kluster yang digunakan. Pada grafik *elbow*, akan terlihat kurva yang menunjukkan jumlah kluster di sumbu x dan variansi di sumbu y. Pada titik di mana kurva menunjukkan perubahan signifikan dan membentuk seperti siku (*elbow*), maka itulah jumlah kluster optimum yang sebaiknya digunakan dalam klusterisasi.

Dalam kasus klusterisasi wilayah kecamatan berdasarkan kelembaban relatif dan presipitasi menggunakan k-means clustering di Sulawesi Utara, metode *elbow* menghasilkan jumlah kluster optimum

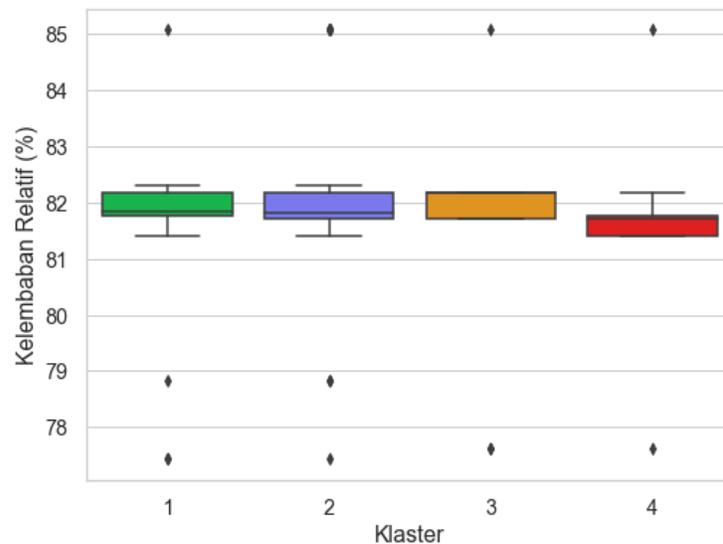
sebanyak 4 klaster (lihat Gambar 2). Hal ini menunjukkan bahwa wilayah kecamatan di Sulawesi Utara dapat dibagi menjadi 4 kelompok yang berbeda berdasarkan tingkat kelembaban relatif dan presipitasi.

Dengan adanya informasi ini, maka dapat dilakukan analisis lebih lanjut terhadap masing-masing kelompok, seperti menentukan karakteristik wilayah yang termasuk ke dalam kelompok tersebut, membandingkan nilai kelembaban relatif dan presipitasi antar kelompok, dan lain sebagainya. Hal ini dapat memberikan pemahaman yang lebih mendalam mengenai kondisi meteorologi di Sulawesi Utara, serta dapat digunakan sebagai bahan dasar dalam pengambilan keputusan terkait sektor-sektor yang terkait dengan kondisi lingkungan dan cuaca.



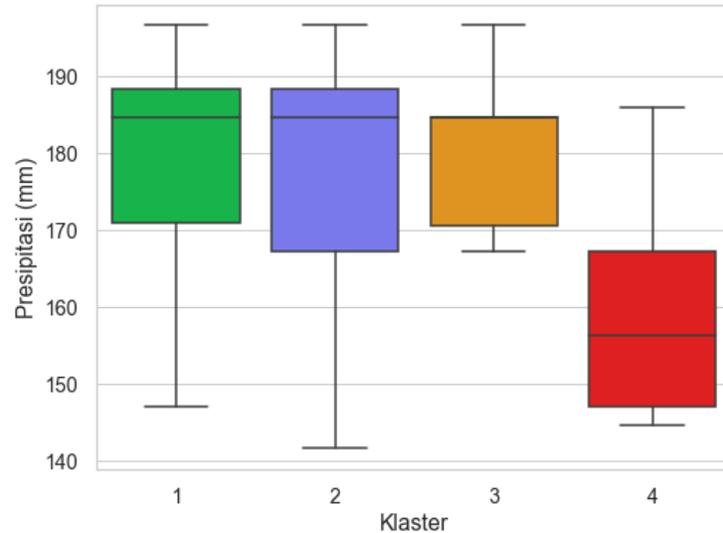
Gambar 3. Pemetaan klasterisasi kecamatan di Sulawesi Utara menggunakan *k-means clustering*.

Klasterisasi wilayah kecamatan berdasarkan kelembaban relatif dan presipitasi menggunakan metode *k-means clustering* di Sulawesi Utara dapat dilihat pada Gambar 3. Dari hasil klasterisasi diperoleh empat klaster kecamatan, yang dibedakan berdasarkan sebaran data presipitasi dan kelembaban relatif.



Gambar 4. Distribusi kelembaban relatif terhadap klaster.

Meskipun tidak ada perbedaan signifikan pada rata-rata kelembaban relatif dari keempat klaster (Gambar 4), namun sebaran data presipitasi memiliki perbedaan antar klaster (Gambar 5). Dimana klaster 1 memiliki sebaran data presipitasi yang paling tinggi, diikuti dengan klaster 2 dan 3, sementara klaster 4 memiliki tingkat presipitasi paling rendah dibandingkan dengan keempat klaster yang lain.



Gambar 5. Distribusi presipitasi terhadap kluster.

Terdapat perbedaan antara daerah daratan dan daerah kepulauan di Sulawesi Utara dalam hal sebaran kluster pada peta. Daerah daratan didominasi oleh kluster 1, 2, dan 3, sementara daerah kepulauan (daerah utara) cenderung berada pada kluster 4. Hal ini kemungkinan besar dipengaruhi oleh lokasi geografis masing-masing kecamatan di wilayah Sulawesi Utara.

Tabel 1. Sebaran kluster di Kabupaten/Kota Provinsi Sulawesi Utara.

Kabupaten/Kota	Jumlah Kluster 1	Jumlah Kluster 2	Jumlah Kluster 3	Jumlah Kluster 4
Kabupaten Kepulauan Sangihe	2	2	0	11
Kabupaten Siau Tagulandang Biaro	0	1	0	9
Kabupaten Kepulauan Talaud	6	9	1	3
Kabupaten Minahasa Selatan	5	4	5	3
Kabupaten Bolaang Mongondow	4	9	0	2
Kota Manado	3	2	4	2
Kabupaten Minahasa	8	7	9	1
Kabupaten Minahasa Tenggara	5	1	5	1
Kabupaten Bolaang Mongondow Selatan	4	2	1	0
Kabupaten Bolaang Mongondow Timur	5	1	1	0
Kabupaten Bolaang Mongondow Utara	2	4	0	0
Kabupaten Minahasa Utara	3	5	2	0
Kota Bitung	6	2	0	0
Kota Kotamobagu	0	4	0	0
Kota Tomohon	0	0	5	0

Tabel 1 menunjukkan pola sebaran kluster berdasarkan Kabupaten dan Kota di Sulawesi Utara. Pola yang menonjol ada pada kluster 4, yang merepresentasikan presipitasi (curah hujan) terendah, pada wilayah kepulauan. Hal ini terlihat jelas di Kabupaten Kepulauan Sangihe (11 kecamatan) dan Kabupaten Siau Tagulandang Biaro (9 kecamatan). Tingginya jumlah kecamatan pada kluster 4 di dua wilayah ini mengindikasikan adanya potensi kekeringan, serta kebutuhan akan manajemen sumber daya air dibandingkan daerah lainnya.

Sebaliknya, kluster 1 yang menandakan curah hujan tertinggi, terkonsentrasi di wilayah daratan. Kabupaten Minahasa tercatat sebagai wilayah dengan jumlah kecamatan terbanyak dalam kluster 1, yaitu sebanyak 8 kecamatan, yang menunjukkan bahwa sebagian besar wilayahnya memiliki karakteristik curah

hujan yang sangat tinggi. Hal yang juga patut menjadi perhatian khusus adalah Kota Manado. Sebagai ibu kota provinsi dengan tingkat kepadatan penduduk dan pembangunan yang tinggi, terdapat 3 kecamatan di Manado yang masuk dalam kluster 1. Kondisi ini menandakan adanya potensi bencana hidrometeorologi, seperti banjir perkotaan, yang dipicu oleh curah hujan ekstrem.

Hasil klusterisasi ini diharapkan dapat menjadi acuan bagi pemerintah daerah dalam merencanakan pengelolaan sumber daya air, seperti penyediaan infrastruktur dan distribusi air, serta dalam mengambil tindakan mitigasi bencana terkait cuaca ekstrem.

4. KESIMPULAN DAN SARAN

Dari hasil analisis klusterisasi kelembaban dan presipitasi di 171 kecamatan di Sulawesi Utara selama lima tahun, diperoleh empat kluster kecamatan dengan tingkat presipitasi yang berbeda-beda. Kecamatan pada kluster 1 memiliki tingkat presipitasi tertinggi, diikuti oleh kluster 2 dan 3, sedangkan kluster 4 memiliki tingkat presipitasi paling rendah.

Untuk analisis selanjutnya, terdapat beberapa hal yang dapat dipertimbangkan, seperti melakukan analisis berdasarkan faktor lain seperti suhu, kecepatan angin, dan sinar matahari, serta membandingkan dengan metode klusterisasi lain seperti *Hierarchical Clustering*, *Density-Based Clustering*, atau *Fuzzy C-Means Clustering*. Dengan mempertimbangkan hal-hal tersebut, analisis klusterisasi wilayah dapat memberikan hasil yang lebih akurat, lengkap, dan berguna bagi pihak-pihak terkait dalam pengambilan keputusan terkait dengan pengelolaan sumber daya alam, mitigasi bencana, dan pengembangan wilayah di Sulawesi Utara.

DAFTAR PUSTAKA

- American Meteorological Society. (2024). *Glossary of Meteorology*. <https://glossary.ametsoc.org/>
- Badan Nasional Penanggulangan Bencana (BNPB). (2024). *Pusat Data Informasi dan Komunikasi Kebencanaan (Pusdatinkom)*. <https://dibi.bnpb.go.id/xdibi2>
- Barrett, P., Hunter, J., Miller, J. T., Hsu, J.-C., & Greenfield, P. (2005). Matplotlib-A Portable Python Plotting Package. *Astronomical data analysis software and systems XIV*, 347, 91.
- Bholowalia, P., & Kumar, A. (2014). EBK-means: A clustering technique based on elbow method and k-means in WSN. *International Journal of Computer Applications*, 105(9).
- Center (LaRC), N. L. R. (2024, Mei). *Prediction of Worldwide Energy Resource (POWER)*. NASA Langley Research Center (LaRC).
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., & others. (2020). Array programming with NumPy. *Nature*, 585(7825), 357-362.
- Kementerian Dalam Negeri-Dukcapil. (2024). *Visualisasi Data Kependudukan*. <https://gis.dukcapil.kemendagri.go.id/peta/>
- Madhulatha, T. S. (2012). An overview on clustering methods. *arXiv preprint arXiv:1205.1117*.
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, 9, 381-386.
- Mapbox. (2023). *Map, Geocoding and Navigations APIs*. <https://docs.mapbox.com/api/maps/>
- McKinney, W. (2011). pandas: A foundational Python library for data analysis and statistics. *Python for high performance and scientific computing*, 14(9), 1-9.
- Pedregosa, F. (2011). Scikit-learn: Machine learning in python Fabian. *Journal of machine learning research*, 12, 2825.
- Plotly Technologies Inc. (2015). *Collaborative data science*. Plotly Technologies Inc. <https://plot.ly>
- Sanner, & Michel, F. (1999). Python: A programming language for software integration and development. *Journal of Molecular Graphics and Modelling*, 17(1), 57-61.
- Waskom, M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021.