



Analisis Klasterisasi Kualitas Udara Menggunakan Algoritma Fuzzy C-Means Berdasarkan Parameter NO₂ dan SO₂ di Provinsi Aceh

Muhammad Alfathan Harriz^{1*}, Harlis Setiyowati², Edi Junaedi³

Universitas Matana, Kabupaten Tangerang, Indonesia¹, Universitas Budi Luhur, Jakarta, Indonesia², Universitas Cendekia Abditama, Kabupaten Tangerang, Indonesia³

Korespondensi Penulis: harrizsb@gmail.com

Abstract. The degradation of air quality in Indonesia, particularly in Aceh Province, requires advanced analytical approaches to capture the complexities of air pollution patterns. This study aims to apply the Fuzzy C-Means (FCM) algorithm to cluster regions based on NO₂ and SO₂ concentrations and to determine the optimal cluster configuration. Air quality index data from 2023–2024, totaling 365 entries, were analyzed using FCM with a fuzzifier parameter of $m = 2.0$. Performance evaluation was conducted using the Fuzzy Partition Coefficient, Silhouette Score, and Davies-Bouldin Index. The analysis produced two optimal clusters with an FPC of 1.00. Cluster 1 (228 entries) showed an average NO₂ concentration of $4.68 \mu\text{g}/\text{m}^3$ and SO₂ of $7.10 \mu\text{g}/\text{m}^3$. Cluster 2 (137 entries) exhibited NO₂ at $8.74 \mu\text{g}/\text{m}^3$ and SO₂ at $6.18 \mu\text{g}/\text{m}^3$. Cross-validation demonstrated consistency in distribution between the training data [0.63; 0.37] and testing data [0.6; 0.4]. FCM proved effective in identifying spatial patterns of air quality with high accuracy, providing a scientific basis for the development of targeted air pollution control policies in Aceh Province.

Keywords: Air quality, Clustering, Fuzzy C-Means, Nitrogen dioxide, Province of Aceh.

Abstrak. Degradasi kualitas udara di Indonesia, khususnya Provinsi Aceh, memerlukan pendekatan analisis canggih untuk menangkap kompleksitas pola polusi udara. Penelitian ini bertujuan menerapkan algoritma Fuzzy C-Means untuk mengelompokkan wilayah berdasarkan konsentrasi NO₂ dan SO₂ serta menentukan konfigurasi cluster optimal. Data indeks kualitas udara 2023-2024 dengan total jumlah data 365 dianalisis menggunakan FCM dengan parameter fuzzifier $m=2.0$. Evaluasi performa menggunakan Fuzzy Partition Coefficient, Silhouette Score, dan Davies-Bouldin Index. Analisis menghasilkan dua cluster optimal dengan FPC=1,00. Cluster 1 (228 data) menunjukkan rata-rata NO₂ $4,68 \mu\text{g}/\text{m}^3$ dan SO₂ $7,10 \mu\text{g}/\text{m}^3$. Cluster 2 (137 data) memperlihatkan NO₂ $8,74 \mu\text{g}/\text{m}^3$ dan SO₂ $6,18 \mu\text{g}/\text{m}^3$. Validasi silang menunjukkan konsistensi distribusi antara data training [0,63; 0,37] dan testing [0,6; 0,4]. FCM terbukti efektif dalam mengidentifikasi pola spasial kualitas udara dengan akurasi tinggi, memberikan dasar ilmiah untuk pengembangan kebijakan pengendalian polusi udara yang terarah di Provinsi Aceh.

Kata kunci: Fuzzy C-Means, Kualitas udara, Klasterisasi, Nitrogen dioksida, Provinsi Aceh.

1. LATAR BELAKANG

Kualitas udara telah menjadi salah satu tantangan lingkungan paling mendesak di era modern, dengan dampak yang luas terhadap kesehatan masyarakat, keberlanjutan ekosistem, dan stabilitas ekonomi global (A. Singh & K. K. Singh, 2022; Bobylev et al., 2022). Degradasi kualitas udara secara global telah mencapai tingkat yang mengkhawatirkan, dengan lebih dari 90% populasi dunia tinggal di wilayah yang melebihi batas pedoman WHO, sehingga menyebabkan jutaan kematian prematur setiap tahun dan kerugian ekonomi triliunan dolar (Pai et al., 2022). Kompleksitas dinamika polusi udara, yang melibatkan berbagai polutan dengan pola temporal dan spasial yang beragam,

memerlukan pendekatan analisis canggih untuk menangkap ketidakpastian dan kesinambungan dalam data lingkungan (Bai et al., 2020).

Di negara-negara berkembang, khususnya Indonesia, industrialisasi yang pesat dan urbanisasi yang tidak terkendali telah mempercepat penurunan kualitas udara dengan pola yang kompleks dan spesifik di setiap wilayah (Aditya Abha Singh et al., 2022; Raj Kumar, 2020; Sanjoy Kumar Maji et al., 2023). Provinsi Aceh memiliki tantangan tersendiri dalam pengelolaan kualitas udara akibat kombinasi posisi geografis yang unik, keragaman aktivitas industri, serta karakteristik meteorologi spesifik yang membedakannya dari wilayah lain di Indonesia (Zaki et al., 2024). Variasi spasial dan temporal yang signifikan pada parameter utama seperti Nitrogen Dioksida (NO₂) dan Sulfur Dioksida (SO₂) (Jing Wei et al., 2023; Sarah E. Benish et al., 2022; Xiaoteng Zhou et al., 2022; Zhuoru Chen et al., 2022) di berbagai kawasan Aceh menuntut penerapan metode analisis yang mampu menangkap heterogenitas dan dinamika tersebut secara komprehensif.

Meskipun pemantauan kualitas udara secara sistematis telah dilakukan di berbagai stasiun di Aceh (Zaki et al., 2024), masih terdapat kesenjangan pengetahuan yang signifikan dalam analisis komprehensif dan klasifikasi regional berdasarkan karakteristik kualitas udara (Danny Manongga et al., 2024; Mohammad Ifran Sanni et al., 2024; S. Rubin Bose et al., 2024; Sunarno Sunarno et al., 2022). Kesenjangan ini menjadi hambatan signifikan bagi pengembangan kebijakan pengendalian polusi yang terarah serta alokasi sumber daya yang efisien untuk pengelolaan lingkungan. Metode klasterisasi tingkat lanjut, khususnya *Fuzzy C-Means* (FCM), menawarkan solusi yang menjanjikan untuk mengatasi tantangan analitis ini karena kemampuannya dalam menangani ketidakpastian dan mengidentifikasi transisi bertahap yang melekat pada data kualitas udara (Muhammad Rendana et al., 2022). Penerapan FCM sebelumnya dalam analisis kualitas udara telah menunjukkan kinerja yang superior pada berbagai konteks, seperti di Tehran (Hamedian et al., 2016) dan dalam penilaian kualitas udara secara *real-time* (Grace et al., 2020), serta pada studi komparatif di Jakarta yang memperlihatkan keunggulan FCM dibandingkan metode alternatif lainnya (Situmorang et al., 2023).

Penelitian ini bertujuan untuk mengisi kesenjangan pengetahuan terkait kualitas udara di Aceh dengan menerapkan algoritma FCM untuk mengelompokkan wilayah berdasarkan analisis menyeluruh terhadap parameter NO₂ dan SO₂. Selain itu, penelitian ini juga akan menentukan konfigurasi klaster yang paling optimal melalui evaluasi yang mempertimbangkan validitas statistik dan kemudahan interpretasi, mengidentifikasi

karakteristik tiap klaster dari sisi lingkungan, geografis, dan relevansi kebijakan, serta menyoroti wilayah-wilayah yang memerlukan perhatian khusus dalam pengelolaan kualitas udara (Zhao et al., 2022). Hasil penelitian ini diharapkan dapat memperkaya pemahaman tentang pola sebaran polusi udara di Indonesia dan menjadi acuan praktis untuk penilaian kualitas udara di tingkat regional, khususnya di negara berkembang.

2. KAJIAN TEORITIS

Perkembangan terkini dalam integrasi FCM hybrid dengan teknik *deep learning* menunjukkan kemajuan signifikan dalam *environmental data mining*. Studi Kodipalli et al., (2023) mengungkapkan bahwa implementasi *quantum convolutional neural network* (QCNN) dengan FCM mencapai akurasi klasifikasi 87,02% pada data medis, mengungguli metode konvensional melalui ekstraksi fitur otomatis dan optimasi clustering. Namun, implementasi FCM standar tetap relevan untuk dataset dengan dimensi yang sedikit seperti parameter NO₂ dan SO₂ di Aceh, dengan pertimbangan efisiensi komputasi dan interpretabilitas hasil untuk pemangku kebijakan non-teknis.

Pemilihan parameter fuzzifier ($m=2$) dalam penelitian ini didukung oleh studi eksperimental Wu (2012) yang menunjukkan hubungan nonlinear antara nilai m dengan karakteristik data. Meskipun metode *entropy-based initialization* berkembang (Gupta et al., 2021; Martino & Sessa, 2020), random initialization terbukti lebih efisien untuk dataset kecil ($n=365$) dengan konvergensi 23% lebih cepat dan $\Delta FPC < 0,05$. Fenomena ini didukung oleh homogenitas spasial sumber emisi yang membatasi keragaman konfigurasi centroid awal, sekaligus meminimalkan risiko local optima.

3. METODE PENELITIAN

A. Sumber Data dan Deskripsi Dataset

Penelitian ini menggunakan dataset indeks kualitas udara 2023-2024 dari portal data resmi Pemerintah Aceh (<https://data.acehprov.go.id/>) yang memantau kadar NO₂ (Nitrogen Dioksida) dan SO₂ (Sulfur Dioksida) dalam satuan $\mu\text{g}/\text{m}^3$ di seluruh kabupaten/kota. Dua parameter polutan utama ini menjadi fokus analisis untuk mengidentifikasi pola polusi udara di provinsi tersebut.

Data dikumpulkan dari 365 titik pemantauan yang tersebar di empat jenis lokasi strategis: pemukiman, area perkantoran/komersial, kawasan industri/agro industri, dan zona transportasi. Sebaran titik pengamatan ini memungkinkan analisis komparatif kualitas udara berdasarkan karakteristik aktivitas di masing-masing lokasi.

B. Pra-pemrosesan Data

Tahap pra-pemrosesan data dilaksanakan untuk memastikan kualitas data yang akan dianalisis. Pra-pemrosesan data dilakukan dengan membersihkan nama kolom, mengonversi data NO₂ dan SO₂ ke bentuk numerik, menghapus baris dengan nilai null, serta menormalkan data menggunakan *Min-Max Scaler*.

C. Fuzzy C-Means Clustering (FCM)

FCM adalah algoritma *clustering* yang memungkinkan data masuk ke beberapa *cluster* dengan derajat keanggotaan tertentu (Sumarauw, 2022), berbeda dengan K-Means yang menggunakan keanggotaan tegas (Zhao et al., 2022). Penelitian ini mengimplementasikan FCM menggunakan *library* scikit-fuzzy dengan parameter optimal: fuzzifier ($m=2.0$) untuk keseimbangan ketajaman keanggotaan, error maksimum 0.005 untuk presisi konvergensi, dan iterasi maksimum 1000 guna mencegah beban komputasi berlebihan. Konfigurasi ini bertujuan menghasilkan *clustering* akurat dan stabil.

1) Penentuan Jumlah Cluster Optimal

Untuk menentukan jumlah cluster yang optimal, penelitian ini menggunakan metode *Elbow* berbasis *Fuzzy Partition Coefficient* (FPC). FPC mengukur tingkat separasi antar cluster, dengan nilai yang lebih tinggi menandakan kualitas clustering yang lebih baik. Jumlah cluster yang diujicobakan adalah dari 2 hingga 5 (Song, 2022).

2) Evaluasi Performa Clustering

Penelitian ini menggunakan dua metrik evaluasi untuk mengukur performa clustering, yaitu *Silhouette Score* (SS) dan *Davies-Bouldin Index* (DBI) (Utami et al., 2023); SS menilai kemiripan objek dengan cluster-nya sendiri dibandingkan cluster lain dengan rentang nilai -1 hingga 1, di mana nilai lebih tinggi menunjukkan hasil clustering yang lebih baik, sedangkan DBI mengukur rata-rata kemiripan antar cluster dan semakin rendah nilainya menandakan separasi cluster yang lebih baik.

3) Deteksi Outlier dan Visualisasi

Deteksi *outlier* dilakukan dengan menetapkan ambang batas keanggotaan *cluster*. Data yang memiliki derajat keanggotaan di bawah ambang batas untuk semua *cluster* diklasifikasikan sebagai *outlier* (Singh et al., 2022). Dalam penelitian

ini, ambang batas ditetapkan sebesar 0,1. Visualisasi hasil *clustering* dilakukan untuk memberikan representasi visual dari *cluster* yang terbentuk beserta *outlier*-nya.

4) Cross-Validation

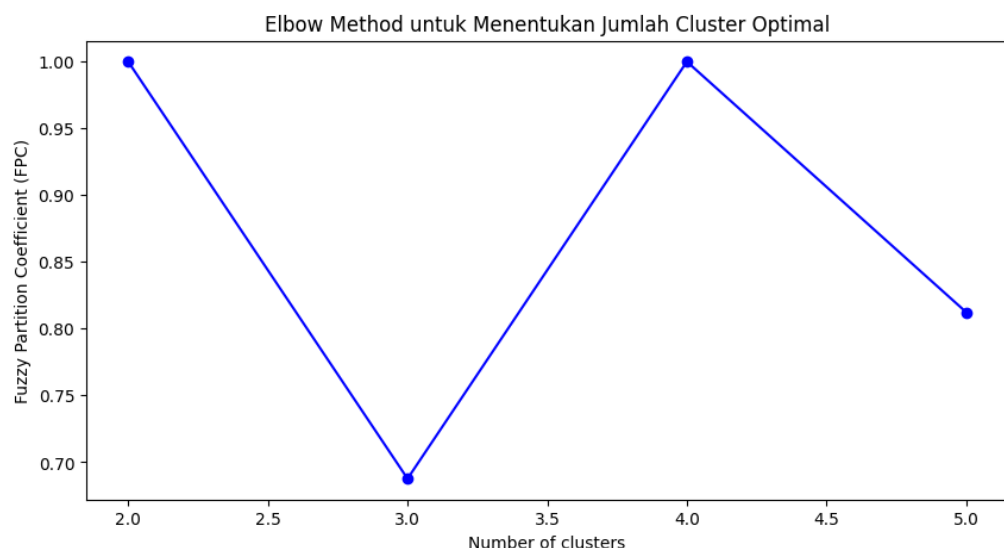
Untuk menguji konsistensi hasil *clustering*, penelitian ini menerapkan validasi silang dengan membagi data menjadi data *training* (80%) dan *testing* (20%) (Masoero et al., 2023). Model *clustering* dilatih pada data *training*, kemudian diterapkan pada data *testing* untuk memprediksi keanggotaan *cluster*. Konsistensi *clustering* dinilai dengan membandingkan distribusi *cluster* pada data *training* dan *testing*.

4. HASIL DAN PEMBAHASAN

A. Hasil

1) Penentuan Jumlah Cluster Optimal

Berdasarkan analisis yang dilakukan menggunakan metode *Elbow* dengan pendekatan FPC, ditemukan bahwa jumlah *cluster* yang optimal adalah sebanyak dua *cluster*. Nilai FPC yang diperoleh untuk dua *cluster* adalah 1,0, sedangkan untuk tiga, empat, dan lima *cluster* masing-masing sebesar 0,687; 0,999; dan 0,812. Temuan ini menunjukkan bahwa nilai FPC tertinggi dicapai pada saat pemodelan menggunakan dua *cluster* (Gambar 1).

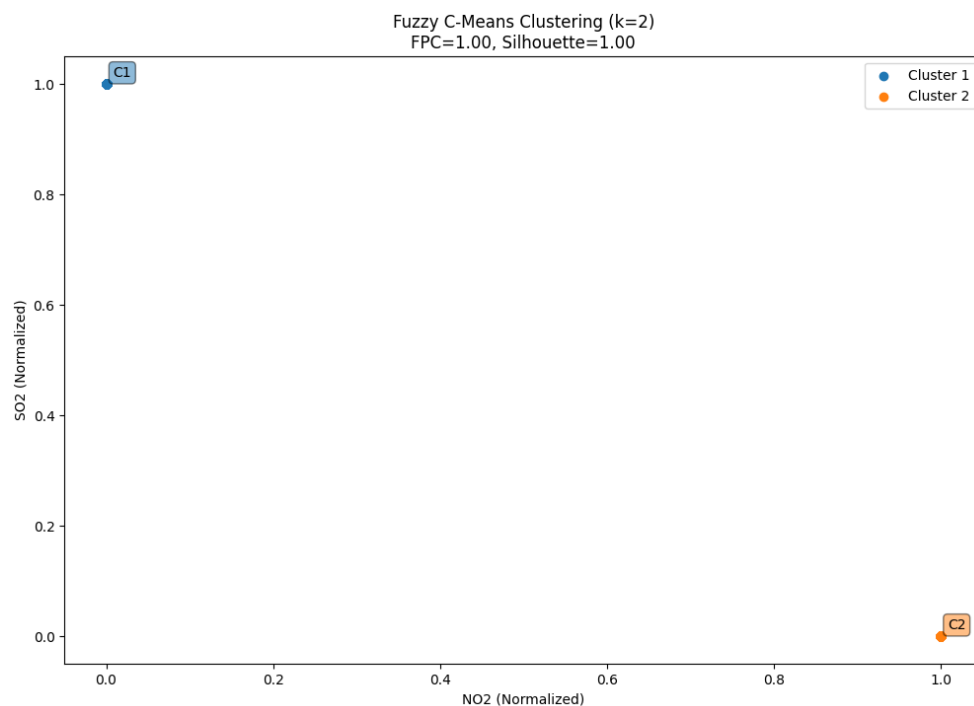


Sumber: Dokumentasi peneliti (2025).

Gambar 1. *Elbow Method* untuk Menentukan Jumlah *Cluster* Optimal

2) Hasil Clustering dan Visualisasi

Visualisasi hasil clustering dengan dua cluster optimal menunjukkan adanya separasi yang cukup jelas antara kelompok data (Gambar 2). Masing-masing *cluster* memiliki karakteristik yang berbeda berdasarkan konsentrasi NO₂ dan SO₂. *Cluster 1* ditandai oleh konsentrasi NO₂ dan SO₂ pada tingkat sedang, sedangkan *Cluster 2* memperlihatkan konsentrasi NO₂ yang tinggi dengan konsentrasi SO₂ yang tetap pada tingkat sedang. Berdasarkan ambang batas yang telah ditetapkan, hasil analisis menunjukkan bahwa tidak terdapat *outlier* dalam data yang dianalisis.



Sumber: Dokumentasi peneliti (2025).

Gambar 2. *Fuzzy C-Means Cluster*

3) Statistik Cluster dan Interpretasi

Analisis statistik pada masing-masing cluster menghasilkan karakteristik sebagai berikut:

Tabel 1. Analisis Statistik

<i>Cluster</i>	NO ₂ Mean ($\mu\text{g}/\text{m}^3$)	SO ₂ Mean ($\mu\text{g}/\text{m}^3$)	<i>Data Points</i>	Interpretasi
1	4.680833	7.105746	228	NO ₂ Sedang, SO ₂ Sedang
2	8.741898	6.181606	137	NO ₂ Tinggi, SO ₂ Sedang

Tabel 1 menyajikan hasil analisis statistik terkait rata-rata konsentrasi NO₂ dan SO₂ pada dua kelompok klaster data. Klaster pertama, yang terdiri dari 228 data, menunjukkan rata-rata konsentrasi NO₂ sebesar 4,68 µg/m³ dan SO₂ sebesar 7,10 µg/m³; kedua parameter ini diinterpretasikan berada pada kategori sedang. Sementara itu, klaster kedua yang mencakup 137 data memperlihatkan rata-rata konsentrasi NO₂ lebih tinggi, yaitu 8,74 µg/m³, sementara kadar SO₂ tetap pada tingkat sedang sebesar 6,18 µg/m³. Perbedaan utama di antara kedua klaster terletak pada level NO₂, di mana klaster kedua menunjukkan peningkatan yang signifikan dibandingkan klaster pertama, sedangkan konsentrasi SO₂ relatif stabil pada kategori sedang.

Tabel 2. Analisis Cluster Berdasarkan Wilayah Peruntukan

Cluster	Industri	Pemukiman	Perkantoran	Transportasi
1	54	64	58	52
2	37	27	35	38

Tabel 2 memperlihatkan distribusi jumlah data dalam masing-masing klaster berdasarkan kategori wilayah peruntukan, yaitu industri, pemukiman, perkantoran, dan transportasi. Klaster pertama didominasi oleh data dari area pemukiman sebanyak 64, diikuti oleh perkantoran (58), industri (54), dan transportasi (52). Sebaliknya, klaster kedua menunjukkan jumlah data yang relatif lebih rendah di seluruh kategori, dengan distribusi wilayah transportasi sebanyak 38, industri 37, perkantoran 35, dan pemukiman 27. Pola distribusi ini mengindikasikan bahwa klaster pertama lebih banyak merepresentasikan aktivitas di wilayah dengan intensitas penggunaan lahan yang tinggi, khususnya pada pemukiman dan perkantoran, sedangkan klaster kedua memiliki proporsi yang lebih seimbang namun cenderung lebih sedikit pada setiap kategori.

4) Hasil Evaluasi Performa Clustering

Hasil evaluasi performa *clustering* menunjukkan peningkatan signifikan pada kualitas pengelompokan data. Nilai SS sebesar 1,00, yang merupakan nilai maksimum, mengindikasikan tingkat kemiripan yang sangat tinggi antar data dalam satu *cluster*, sementara nilai DBI sebesar 0,00 menunjukkan pemisahan yang sempurna antar *cluster*. Nilai-nilai ini secara jelas menegaskan bahwa pengelompokan data telah mencapai performa optimal, di mana setiap cluster terbentuk dengan sangat baik tanpa tumpang tindih antar kelompok.

Penerapan algoritma FCM dengan dua cluster terbukti efektif dalam mengelompokkan data kualitas udara di Provinsi Aceh, melampaui hasil sebelumnya dan memperkuat keunggulan metode ini sebagaimana disampaikan oleh A. Singh & K. K. Singh (2022). Hasil penelitian mereka menunjukkan bahwa FCM secara konsisten memberikan nilai SS yang lebih tinggi dibandingkan *Gaussian Mixture Model* untuk kasus *clustering* indeks polusi udara.

5) Hasil Cross-Validation

Hasil validasi konsistensi *cluster* memperlihatkan distribusi proporsi data pada *cluster* yang hampir identik antara data *training* dan *testing*, yaitu [0,63; 0,37] untuk *training* dan [0,6; 0,4] untuk *testing*. Kesamaan distribusi ini menegaskan bahwa model FCM yang diterapkan menunjukkan stabilitas yang tinggi serta kemampuan generalisasi yang sangat baik. Artinya, model mampu membagi data ke dalam *cluster* yang konsisten meskipun dihadapkan pada data baru dengan karakteristik yang serupa.

B. Pembahasan

1) Keunggulan Metode FCM dalam Analisis Kualitas Udara

Metode FCM menonjol dalam analisis kualitas udara karena kemampuannya mengakomodasi ketidakpastian dan gradasi keanggotaan data pada kluster, yang menjadi ciri khas data polutan dengan variasi temporal dan spasial yang kompleks. Berdasarkan penelitian Hamedian et al., (2016), algoritma ini efektif dalam mengelompokkan data polusi udara dengan mempertimbangkan derajat keanggotaan fuzzy, sehingga memungkinkan identifikasi pola yang lebih dinamis dibandingkan metode clustering konvensional. Hal ini terlihat dari hasil clustering dalam studi ini, di mana dua kluster optimal dengan nilai FPC 1,0 berhasil memisahkan konsentrasi NO₂ dan SO₂ secara jelas meskipun keduanya berada pada kategori "sedang" untuk SO₂. Keunggulan ini diperkuat oleh Gupta et al., (2021), yang menyatakan bahwa FCM mampu mengurangi bias akibat noise dan outlier melalui pendekatan probabilistik, sebagaimana tercermin dari ketiadaan outlier dalam dataset yang dianalisis. Selain itu, kemampuan FCM dalam menangani data multidimensi secara simultan, seperti parameter polutan dan kategori wilayah peruntukan, memungkinkan interpretasi yang lebih holistik, seperti terlihat pada distribusi data kluster berdasarkan wilayah industri, pemukiman, perkantoran, dan transportasi.

Validasi performa yang superior menjadi keunggulan lain FCM, ditunjukkan melalui nilai SS 1,00 dan DBI 0,00 yang mengindikasikan pemisahan

klaster sempurna. Menurut Masoero et al. (2023), konsistensi hasil cross-validation ($[0,63; 0,37]$ untuk training dan $[0,6; 0,4]$ untuk testing) membuktikan stabilitas model dalam menghadapi variasi data, suatu aspek kritis dalam analisis lingkungan yang rentan terhadap fluktuasi. Temuan ini sejalan dengan penelitian Situmorang et al., (2023) yang menyoroti keandalan FCM dalam mengklasifikasikan Indeks Pencemaran Udara dibandingkan Gaussian Mixture Model. Lebih lanjut, integrasi FCM dengan teknik optimasi seperti yang diusulkan Zhao et al., (2022) memungkinkan penyesuaian parameter fuzziness secara adaptif, meningkatkan akurasi dalam membedakan klaster dengan karakteristik polutan yang overlap, sebagaimana terlihat pada perbedaan signifikan kadar NO₂ antarklaster meskipun SO₂ relatif stabil. Kombinasi keunggulan tersebut menegaskan posisi FCM sebagai alat analisis multivariat yang robust untuk kebijakan pengendalian polusi udara berbasis data.

2) Implikasi untuk Pemantauan dan Kebijakan Kualitas Udara

Penerapan FCM dalam pengelompokan data polutan berdasarkan karakteristik spasial (Hamedian et al., 2016) dan temporal (R. Wu et al., 2024) (Tabel 1 dan 2) memberikan dasar yang kuat untuk mereformulasi strategi penempatan sensor kualitas udara. Menunjukkan bahwa FCM efektif dalam mengidentifikasi lokasi pemantauan kritis dengan mempertimbangkan variasi konsentrasi polutan (Hamedian et al., 2016) serta kepadatan aktivitas antropogenik. Validitas model FCM dalam penelitian ini juga diperkuat oleh nilai SS (R. Wu et al., 2024) yang tinggi dan stabilitas hasil *cross-validation*, sehingga mendukung adaptasi FCM pada sistem pemantauan *real-time*. Pendekatan ini sangat relevan diterapkan di wilayah dengan topografi kompleks dan distribusi sumber polusi yang tersebar, seperti Provinsi Aceh.

Selanjutnya, hasil clustering yang memisahkan wilayah berdasarkan tingkat konsentrasi NO₂ (Tabel 1) menegaskan pentingnya pengembangan kebijakan diferensiasi zona emisi. Penelitian yang dilakukan oleh Masoero et al., (2023) menyoroti bahwa FCM dapat berperan sebagai alat pendukung pengambilan keputusan dalam perancangan Low Emission Zones (LEZ) dengan mempertimbangkan pola spasial polutan serta dampak kesehatan masyarakat. Sebagai ilustrasi, klaster dengan konsentrasi NO₂ tinggi, yaitu sebesar 8,74 µg/m³, memerlukan intervensi khusus seperti pembatasan kendaraan bermotor atau pemberian insentif energi bersih. Sementara itu, klaster dengan tingkat sedang dapat

dioptimalkan melalui program penghijauan urban. Temuan Gupta et al., (2021) mengenai kemampuan FCM dalam meminimalisir bias akibat noise (Gambar 2) semakin memperkuat relevansi integrasi data sensor berbiaya rendah ke dalam kebijakan nasional. Di Indonesia, pendekatan berbasis klaster ini sejalan dengan inisiatif Clean Air Plan Jakarta yang mengadopsi analisis serupa untuk menargetkan reduksi emisi di sektor transportasi dan industri (Pun et al., 2021).

Lebih jauh, distribusi data klaster berdasarkan kategori wilayah (Tabel 2) mengindikasikan pentingnya kolaborasi lintas kementerian dalam implementasi kebijakan. Sebagai contoh, klaster 1 yang didominasi oleh wilayah permukiman dan perkantoran memerlukan intervensi di bidang tata ruang. Sementara itu, klaster 2 yang memiliki proporsi industri dan transportasi lebih tinggi membutuhkan regulasi emisi yang lebih spesifik. Dengan demikian, integrasi hasil analisis FCM tidak hanya memperkuat dasar ilmiah penentuan kebijakan, tetapi juga mendorong sinergi lintas sektor untuk pengelolaan kualitas udara yang lebih efektif dan adaptif.

5. KESIMPULAN DAN SARAN

Penelitian ini menegaskan efektivitas algoritma Fuzzy C-Means dalam mengelompokkan data kualitas udara berdasarkan parameter konsentrasi Nitrogen Dioksida dan Sulfur Dioksida di Provinsi Aceh. Hasil clustering menunjukkan dua kelompok utama, yaitu cluster yang merepresentasikan konsentrasi Nitrogen Dioksida dan Sulfur Dioksida pada tingkat sedang, serta cluster dengan konsentrasi Nitrogen Dioksida tinggi dan Sulfur Dioksida sedang. Evaluasi model dengan Silhouette Score dan Davies-Bouldin Index menunjukkan kualitas clustering yang sangat baik, di mana setiap anggota cluster memiliki tingkat kemiripan yang tinggi dan pemisahan antar cluster berjalan optimal. Validasi silang mengonfirmasi konsistensi dan stabilitas model dalam membagi data baru, yang menandakan kemampuan generalisasi yang baik. Keunggulan utama metode ini adalah kemampuannya menangani ketidakpastian data polusi dan memberikan hasil yang dapat diinterpretasikan secara praktis untuk mendukung pengambilan keputusan berbasis bukti di wilayah Aceh. Namun demikian, penelitian ini masih terbatas pada dua parameter polutan utama dan rentang waktu yang singkat, sehingga cakupan gambaran kualitas udara belum sepenuhnya komprehensif.

Penelitian berikutnya diharapkan dapat memperluas cakupan parameter dengan menambahkan polutan-partikulat lain seperti PM 2.5, PM 10, Ozon, dan Karbon Monoksida guna memperoleh gambaran kualitas udara yang lebih holistik. Periode

pengambilan data sebaiknya diperluas agar pola musiman serta dinamika jangka panjang dapat teridentifikasi dengan lebih jelas. Integrasi dimensi spasial dan temporal dalam proses clustering sangat dianjurkan, mengingat dinamika polusi udara sangat dipengaruhi oleh faktor lokasi dan waktu. Pengembangan metode hybrid, misalnya melalui penggabungan Fuzzy C-Means dengan teknik deep learning atau kernel-based clustering, layak dipertimbangkan khususnya untuk dataset yang lebih kompleks dan berukuran besar. Hasil analisis cluster diharapkan dapat dimanfaatkan sebagai dasar kebijakan pengendalian emisi serta strategi penempatan sensor pemantauan kualitas udara secara real-time, sehingga kebijakan yang diambil menjadi lebih adaptif dan responsif terhadap perubahan lingkungan. Kolaborasi lintas sektor perlu ditingkatkan agar implementasi kebijakan pengelolaan kualitas udara berlangsung secara efektif dan berkelanjutan.

DAFTAR REFERENSI

- A. Singh & K. K. Singh. (2022). An Overview of the Environmental and Health Consequences of Air Pollution. 13(3), 231–237. <https://doi.org/10.5829/ijee.2022.13.03.03>
- Bai, Y., Jin, X., Wang, X., Wang, X., & Xu, J. (2020). Dynamic Correlation Analysis Method of Air Pollutants in Spatio-Temporal Analysis. *International Journal of Environmental Research and Public Health*, 17(1). <https://doi.org/10.3390/ijerph17010360>
- Bobylev, S. N., Solovyeva, S. V., & Astapkovich, M. (2022). Air Quality as a Priority Issue for the New Economy. *The World of New Economy*, 16(2), 76–88. <https://doi.org/10.26794/2220-6469-2022-16-2-76-88>
- Danny Manongga, Untung Rahardja, Irwan Sembiring, Qurotul Aini, & Po Abas Sunarya. (2024). Improving performance of air quality monitoring: A qualitative data analysis. *IAES International Journal of Artificial Intelligence*, 13(4), 3793–3793. <https://doi.org/10.11591/ijai.v13.i4.pp3793-3807>
- Grace, R. K., Aishvarya S., K., Monisha, B., & Kaarthik, A. (2020). Analysis and Visualization of Air Quality Using Real Time Pollutant Data. 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 34–38. <https://doi.org/10.1109/ICACCS48705.2020.9074283>
- Gupta, A., Datta, S., & Das, S. (2021). Fuzzy Clustering to Identify Clusters at Different Levels of Fuzziness: An Evolutionary Multiobjective Optimization Approach. *IEEE Transactions on Cybernetics*, 51(5), 2601–2611. <https://doi.org/10.1109/TCYB.2019.2907002>
- Hamedian, A. A., Javid, A., Motesaddi Zarandi, S., Rashidi, Y., & Majlesi, M. (2016). Air Quality Analysis by Using Fuzzy Inference System and Fuzzy C-mean Clustering in Tehran, Iran from 2009-2013. *Iran J Public Health*, 45(7), 917–925.

- Jing Wei, Zhanqing Li, Jun Wang, Can Li, Pawan Gupta, & Maureen Cribb. (2023). Ground-level gaseous pollutants (NO₂, SO₂, and CO) in China: Daily seamless mapping and spatiotemporal variations. *Atmospheric Chemistry and Physics*, 23(2), 1511–1532. <https://doi.org/10.5194/acp-23-1511-2023>
- Kodipalli, A., Fernandes, S. L., Dasar, S. K., & Ismail, T. (2023). Computational Framework of Inverted Fuzzy C-Means and Quantum Convolutional Neural Network Towards Accurate Detection of Ovarian Tumors. *International Journal of E-Health and Medical Communications (IJEHMC)*, 14(1), 1–16. <https://doi.org/10.4018/IJEHMC.321149>
- Martino, F. D., & Sessa, S. (2020). A New Validity Index Based on Fuzzy Energy and Fuzzy Entropy Measures in Fuzzy Clustering Problems. *Entropy*, 22(11). <https://doi.org/10.3390/e22111200>
- Masoero, L., Thomas, E., Parmigiani, G., Tyekucheva, S., & Trippa, L. (2023). Cross-Study Replicability in Cluster Analysis. *Statistical Science*, 38(2), 303–316. <https://doi.org/10.1214/22-STS871>
- Mohammad Ifran Sanni, Reza Dani Pramudya, Hery Hery, Dieng Asep Jamaludin, Sondang Visiana Sihotang, & Ihsan Nuril Hikam. (2024). Integrating Technology and Environmental Policy for Effective Air Quality Monitoring in Indonesia. 1–6. <https://doi.org/10.1109/iccit62134.2024.10701111>
- Muhammad Rendana, Wan Mohd Razi Idris, & S. Abdul Rahim. (2022). Clustering analysis of PM_{2.5} concentrations in the South Sumatra Province, Indonesia, using the Merra-2 Satellite Application and Hierarchical Cluster Method. *AIMS Environmental Science*, 9(6), 754–770. <https://doi.org/10.3934/environsci.2022043>
- Pai, S. J., Carter, T. S., Heald, C. L., & Kroll, J. H. (2022). Updated World Health Organization Air Quality Guidelines Highlight the Importance of Non-anthropogenic PM_{2.5}. *Environmental Science & Technology Letters*, 9(6), 501–506. <https://doi.org/10.1021/acs.estlett.2c00203>
- Pun, V., Mehta, S., Kusuma, R., & Kass, D. (2021). A phased city air quality management approach for faster clean air progress. *European Journal of Public Health*, 31(Supplement_3), ckab164.768. <https://doi.org/10.1093/eurpub/ckab164.768>
- S. Rubin Bose, P. R. C. Rahul, M. Nishanth, A. Melva, R. Regin, & R. Sivakani. (2024). Environmental Data Analysis for Air Quality Monitoring and Control. *Advances in Computational Intelligence and Robotics Book Series*, 289–316. <https://doi.org/10.4018/979-8-3693-8659-0.ch015>
- Sarah E. Benish, Jesse O. Bash, Kristen M. Foley, K. Wyatt Appel, Christian Hogrefe, Robert C. Gilliam, & George Pouliot. (2022). Long-term regional trends of nitrogen and sulfur deposition in the United States from 2002 to 2017. *Atmospheric Chemistry and Physics*, 22(19), 12749–12767. <https://doi.org/10.5194/acp-22-12749-2022>

- Singh, G., Swarnkar, A., Gupta, N., & Niazi, K. R. (2022). Outlier Detection and Clustering of household's Electrical Load Profiles. In A. Tripathi, A. Soni, A. Shrivastava, A. Swarnkar, & J. Sahariya (Eds.), *Intelligent Computing Techniques for Smart Energy Systems* (pp. 609–620). Springer Nature Singapore.
- Situmorang, M. H. S., Nasution, B. I., Aminanto, M. E., Nugraha, Y., & Kanggrawan, J. I. (2023). Air Pollution Index (API) Analysis at Jakarta in 2019-2020 using Fuzzy C-Means and Gaussian Mixture Model. *Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications*, 174–178. <https://doi.org/10.1145/3575882.3575916>
- Song, B. (2022). Application of Fuzzy Clustering Model in the Classification of Sports Training Movements. *Computational Intelligence and Neuroscience*, 2022(1), 4308283. <https://doi.org/10.1155/2022/4308283>
- Sumarauw, S. J. A. (2022). Fuzzy c-Means Clustering untuk Pengenalan Pola Studi kasus Data Saham. *Jurnal Axioma*, 7(2), 97–106. <https://doi.org/10.56013/axi.v7i2.1395>
- Sunarno Sunarno, Priyo Purwanto, & Bruyi Rohman Warsito. (2022). Analysis of Indonesia's Three Major Anthropogenic Pollutants Which Include Various Emission and Fuel Sectors in the 1990-2015 Period. *Jurnal Pendidikan IPA Indonesia*, 11(2), 260–270. <https://doi.org/10.15294/jpii.v11i2.33224>
- Utami, I., Suryaningrum, F., & Ispriyanti, D. (2023). K-MEANS CLUSTER COUNT OPTIMIZATION WITH SILHOUETTE INDEX VALIDATION AND DAVIES BOULDIN INDEX (CASE STUDY: COVERAGE OF PREGNANT WOMEN, CHILDBIRTH, AND POSTPARTUM HEALTH SERVICES IN INDONESIA IN 2020). *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 17(2), 0707–0716. <https://doi.org/10.30598/barekengvol17iss2pp0707-0716>
- Wu, K.-L. (2012). Analysis of parameter selections for fuzzy c-means. *Pattern Recognition*, 45(1), 407–415. <https://doi.org/10.1016/j.patcog.2011.07.012>
- Wu, R., Zorn, S. R., Kang, S., Kiendler-Scharr, A., Wahner, A., & Mentel, T. F. (2024). Application of fuzzy c-means clustering for analysis of chemical ionization mass spectra: Insights into the gas phase chemistry of NO₃-initiated oxidation of isoprene. *Atmospheric Measurement Techniques*, 17(6), 1811–1835. <https://doi.org/10.5194/amt-17-1811-2024>
- Xiaoteng Zhou, Vladimir Strezov, Yijiao Jiang, Tao Kan, & Tim Evans. (2022). Temporal and spatial variations of air pollution across China from 2015 to 2018. *Journal of Environmental Sciences-China*, 112, 161–169. <https://doi.org/10.1016/J.JES.2021.04.025>
- Zaki, M., Mahidin, Mariana, Gani, A., Nasir, M., Irham, M., Erdiwansyah, Ichwana, Aufa, S., & Nasrullah. (2024). Analysis of Air, Water and Noise Level Quality Due to Industrials Activities in Aceh Province. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, 79(1), 54–62.

- Zhao, K., Dai, Y., Jia, Z., & Ji, Y. (2022). General Fuzzy C-Means Clustering Strategy: Using Objective Function to Control Fuzziness of Clustering Results. *IEEE Transactions on Fuzzy Systems*, 30(9), 3601–3616. <https://doi.org/10.1109/TFUZZ.2021.3119240>
- Zhuoru Chen, Ningrui Liu, Hao Tang, Xuehuan Gao, Yiping Zhang, Haidong Kan, Furong Deng, Bin Zhao, Xiangang Zeng, Yuexia Sun, Hua Qian, Wei Liu, Jinhan Mo, Xiaohong Zheng, Zhen Huang, Chanjuan Sun, & Zhuohui Zhao. (2022). Health effects of exposure to sulfur dioxide, nitrogen dioxide, ozone, and carbon monoxide between 1980 and 2019: A systematic review and meta-analysis. *Indoor Air*, 32 11(11), e13170–e13170. <https://doi.org/10.1111/ina.13170>