

## Pengaruh Suhu, Kelembapan dan Angin terhadap Polusi Udara: Studi Kasus Dataset *Air Quality*

Trisna Fajar Prasetyo<sup>1</sup>, Ahmad Gani<sup>2\*</sup>, Zainul Mufid<sup>3</sup>

<sup>1,2,3</sup> Universitas Bina Sarana Informatika

e-mail: <sup>1</sup>trisna.tfp@bsi.ac.id, <sup>2\*</sup>ahmad.hmg@bsi.ac.id, <sup>3</sup>zainul.znd@bsi.ac.id

email korespondensi: ahmad.hmg@bsi.ac.id(\*)

**Abstrak** - Salah satu indikator utama kesehatan lingkungan adalah kualitas udara, karena faktor cuaca seperti suhu, kelembapan, tekanan udara, dan kecepatan angin sangat memengaruhi polusi udara. Dengan menggunakan dataset Kualitas Udara Unique Client Identifier (UCI), yang mengandung parameter lingkungan dari wilayah perkotaan, tujuan penelitian ini adalah untuk melihat bagaimana variabel cuaca dan kualitas udara berkorelasi satu sama lain. Penelitian ini melihat dampak parameter cuaca seperti suhu, kelembapan relatif, tekanan atmosfer, dan kecepatan angin terhadap konsentrasi polutan seperti Carbon Monoxide (CO), Nitrogen Dioxide (NO<sub>2</sub>), dan Ozon (O<sub>3</sub>). Data set dibersihkan untuk menghilangkan nilai yang hilang, dan kemudian dilakukan analisis regresi linier sederhana dan korelasi Pearson. Hasil penelitian menunjukkan hubungan positif antara kelembapan dan NO<sub>2</sub>, dan korelasi negatif yang signifikan antara suhu dan konsentrasi CO. Studi ini menunjukkan bahwa cuaca memainkan peran penting dalam perubahan kualitas udara, sehingga bisa dikatakan bahwa pemantauan cuaca dapat menjadi alat prediktor untuk pengaturan lingkungan perkotaan di masa yang akan datang.

Kata Kunci: Kualitas Udara, Faktor Cuaca, Dataset Air Quality

*Abstract - One of the main indicators of environmental health is air quality, as weather factors such as temperature, humidity, air pressure, and wind speed greatly affect air pollution. Using the UCI Air Quality dataset, which contains environmental parameters from urban areas, the aim of this research is to examine how weather variables and air quality correlate with each other. This research examines the impact of weather parameters such as temperature, relative humidity, atmospheric pressure, and wind speed on the concentration of pollutants such as CO, NO<sub>2</sub>, and O<sub>3</sub>. The data set was cleaned to remove missing values, and then simple linear regression and Pearson correlation analysis were conducted. The research results show a positive relationship between humidity and NO<sub>2</sub>, and a significant negative correlation between temperature and CO concentration. This study shows that weather plays an important role in changes in air quality, so it can be said that weather monitoring can be a predictive tool for urban environmental management in the future.*

*Keywords: Air Quality, Weather Factors, Air Quality Dataset*

### PENDAHULUAN

Salah satu masalah kesehatan masyarakat dan lingkungan paling penting di dunia saat ini adalah polusi udara. WHO (2023) menyatakan bahwa 99% orang di seluruh dunia menghirup udara yang buruk, yang menyebabkan sekitar 8,1 juta kematian dini setiap tahun. Partikel halus berukuran 2,5 mikrometer (PM<sub>2.5</sub>) memiliki kemampuan untuk menembus sistem pernapasan dan aliran darah, menyebabkan penyakit seperti asma, bronkitis, PPOK, dan ISP. Selain itu, penelitian oleh (Adhikari & Yin, 2020) menunjukkan bahwa paparan polutan seperti PM<sub>2.5</sub>, PM<sub>10</sub>, dan NO<sub>2</sub> memperburuk hasil kesehatan bahkan setelah mempertimbangkan faktor meteorologi. Polusi udara juga merusak ekosistem, mempercepat perubahan iklim, dan menyebabkan kerugian ekonomi global sebesar 5% dari PDB (Ku et al., 2020). Penyebaran dan konsentrasi polutan udara di atmosfer sangat dipengaruhi oleh faktor-faktor meteorologi. Faktor-faktor seperti suhu, kelembapan relatif, kecepatan angin, curah hujan, tekanan atmosfer, dan ketinggian lapisan batas atmosfer (PBLH) memengaruhi transportasi, dispersi, transformasi kimia, dan deposisi polutan (Qiu et al., 2022) Misalnya, angin kencang dan curah hujan cenderung mengurangi konsentrasi polutan, sedangkan kondisi atmosfer stabil seperti inversi suhu meningkatkan akumulasi polutan di permukaan tanah.

Sangat penting untuk mengembangkan strategi pengelolaan kualitas udara yang bergantung pada kondisi cuaca lokal dengan memahami secara menyeluruh interaksi dinamis ini (Shen et al., 2024) Hubungan antara cuaca



dan kualitas udara sangat bervariasi tergantung pada lokasi geografis dan musim, menurut tinjauan literatur terbaru (2020–2025). Studi ini menunjukkan bahwa suhu dapat mempercepat pembentukan ozon, kelembaban memengaruhi PM<sub>2.5</sub> dengan cara yang tidak pasti, dan kecepatan angin secara konsisten berkorelasi negatif dengan konsentrasi polutan. Tekanan atmosfer dan titik embun juga memengaruhi akumulasi polutan; tekanan tinggi sering dikaitkan dengan polusi yang lebih tinggi. Untuk mendapatkan pemahaman yang lebih baik tentang dinamika kualitas udara, pendekatan analitis yang digunakan dalam penelitian ini termasuk metode statistik, machine learning (ML/DL) (Imam et al., 2024), dan model transpor kimia (CTM). Meskipun banyak kemajuan telah dicapai, masih ada ruang untuk penelitian, terutama dalam studi skala lokal yang menggunakan dataset komprehensif. Karakteristik lokal seperti sumber emisi, topografi, dan iklim mikro seringkali sangat memengaruhi hubungan khusus antara cuaca dan polutan (Wyat Appel et al., 2021).

Dengan menggunakan "Dataset Kualitas Udara" sebagai studi kasus, penelitian ini bertujuan untuk menganalisis pengaruh berbagai faktor cuaca terhadap konsentrasi polutan utama. Inovasi studi ini terletak pada analisis kuantitatif faktor meteorologi terhadap polusi udara setiap hari dan sepanjang musim (Liu & Cui, 2024). Studi ini diharapkan dapat memberikan dasar ilmiah yang kuat untuk mengembangkan kebijakan pengelolaan kualitas udara yang lebih tepat sasaran dan berbasis lokal. Bidang penelitian yang terus berkembang adalah memahami hubungan kuantitatif antara faktor cuaca dan konsentrasi polutan udara (Liu & Cui, 2024). Tinjauan penelitian terbaru yang dipublikasikan antara tahun 2020 dan 2025 memberikan gambaran terbaru tentang hasil utama dan teknik yang digunakan. Selama periode ini, banyak penelitian telah menunjukkan korelasi yang signifikan antara variabel meteorologi dan polutan udara utama. Namun, arah dan kekuatan korelasi ini dapat berbeda tergantung pada lokasi geografis, musim, dan polutan khusus yang dipelajari.

- **Suhu Udara:** Hubungannya dengan PM<sub>2.5</sub> seringkali sulit untuk diketahui. Beberapa penelitian menemukan bahwa, di musim dingin, terutama di wilayah dengan pemanasan domestik yang intens, ada hubungan negatif antara emisi dan suhu yang lebih rendah. Namun, suhu yang lebih tinggi dapat meningkatkan reaksi kimia yang menghasilkan ozon dan aerosol organik sekunder, serta meningkatkan ketinggian lapisan batas atmosfer (Shen et al., 2024). Sebuah penelitian di Jakarta menemukan bahwa pola suhu harian hampir sama dengan pola konsentrasi PM<sub>2.5</sub> (Zhang et al., 2024)

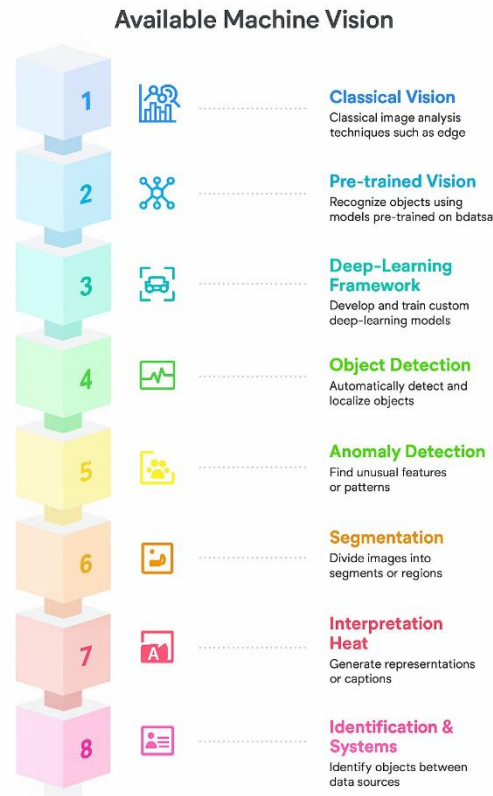
- **Kelembaban Relatif: Efeknya ambigu.** Kelembaban tinggi memiliki potensi untuk meningkatkan deposisi basah (Dinas Lingkungan Hidup Provinsi DKI Jakarta, 2020), tetapi juga dapat membantu dalam kondisi tertentu untuk menghasilkan sulfat dan nitrat aerosol (Shen et al., 2024). Beberapa penelitian menunjukkan korelasi negatif antara kelembaban dan PM<sub>2.5</sub> (Babu Saheer et al., 2022) tetapi penelitian lain menekankan peranannya dalam kimia aerosol.
- **Kecepatan Angin:** Secara umum, ditemukan korelasi negatif yang signifikan antara kecepatan angin dan konsentrasi sebagian besar polutan (PM<sub>2.5</sub>, PM<sub>10</sub>, NO<sub>2</sub>, CO, dan SO<sub>2</sub>) dalam satu studi di Indonesia (Luthfiarta et al., 2020). Karena angin kencang meningkatkan dispersi dan ventilasi atmosfer, konsentrasi polutan di tingkat permukaan menurun (Zhang et al., 2024), sementara cuaca tenang mendorong akumulasi polutan. Hujan membersihkan atmosfer secara efektif.

- **Tekanan Atmosfer:** Sistem tekanan tinggi sering dikaitkan dengan kondisi atmosfer yang stabil, angin lemah, dan penurunan lapisan batas, yang kondusif untuk akumulasi polutan (Shen et al., 2024). Namun, analisis korelasi langsung antara tekanan permukaan dan konsentrasi polutan dapat menunjukkan hasil yang bervariasi (Zhang et al., 2024). Studi regresi di Indonesia menemukan hubungan positif signifikan antara tekanan dan curah hujan.
- **Faktor Lain:** Titik embun (dew point) ditemukan berkorelasi positif dengan PM<sub>2.5</sub> dalam satu studi, menyoroti peran kelembapan absolut. Ketinggian lapisan batas atmosfer (PBLH) juga merupakan faktor penting, di mana PBLH yang rendah membatasi volume pencampuran vertikal dan meningkatkan konsentrasi permukaan (Shen et al., 2024) Untuk memberikan gambaran yang lebih konkret

## METODE PENELITIAN

Penelitian ini merupakan studi kuantitatif berbasis analisis data historis menggunakan dataset "Air Quality" yang berisi parameter polutan (PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, CO, O<sub>3</sub>, NO<sub>2</sub>) dan faktor meteorologi (suhu, kelembaban, kecepatan dan arah angin, tekanan, curah hujan). Prosedur penelitian meliputi akuisisi dan penggabungan data, pembersihan data (mengatasi missing values, pencilaan, normalisasi), eksplorasi data awal melalui statistik deskriptif dan visualisasi, serta analisis korelasi menggunakan koefisien Pearson. Untuk mengkuantifikasi pengaruh faktor cuaca terhadap polutan, diterapkan model regresi linear berganda (MLR), sementara untuk prediksi deret waktu, digunakan model deep learning berbasis LSTM yang mampu menangkap pola dependensi jangka panjang. Evaluasi model dilakukan menggunakan metrik R<sup>2</sup>, MSE, RMSE, dan MAE untuk model regresi kontinu, serta akurasi, presisi, recall, F1-score, dan confusion matrix jika dilakukan klasifikasi kualitas udara. Desain riset dirancang untuk mengidentifikasi hubungan linear maupun non-linear antara faktor meteorologi dan konsentrasi polutan, serta mengeksplorasi potensi prediksi berbasis cuaca. Algoritma MLR memungkinkan analisis pengaruh simultan variabel meteorologi terhadap polutan, dengan interpretasi koefisien yang mewakili kekuatan kontribusi masing-masing variabel. Sementara itu, LSTM digunakan untuk memprediksi konsentrasi polutan ke depan dengan mempertimbangkan dinamika temporal data deret waktu. Pemilihan antara MLR, LSTM,

atau algoritma lain seperti Gradient Boosting atau SVM didasarkan pada hasil eksplorasi data dan tujuan spesifik analisis. Alur kerja penelitian dirancang secara sistematis mulai dari akuisisi data, pra-pemrosesan, eksplorasi, analisis korelasi dan regresi, pembangunan model prediktif, evaluasi, hingga pelaporan hasil dan rekomendasi pada gambar 1 representasi alur kerja penelitian dalam bentuk langkah-langkah.

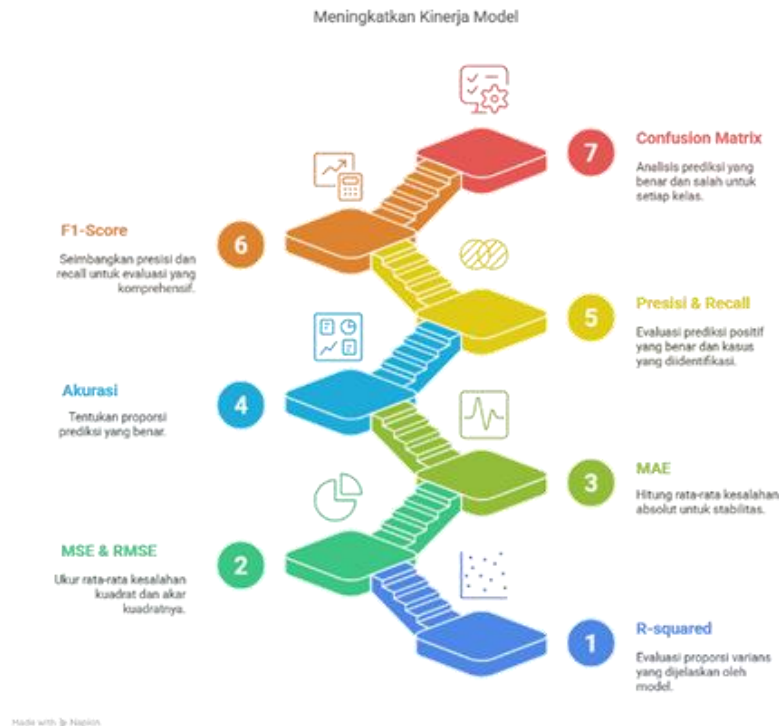


Gambar 1. Analisa Kualitas Udara

- Dua pendekatan utama yang dapat digunakan adalah Regresi Linear Berganda (untuk analisis hubungan) dan LSTM (untuk analisis/prediksi deret waktu). dengan persamaan umum  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$  Dimana:
  - Y adalah variabel dependen (konsentrasi polutan).
  - $X_1, X_2, \dots, X_k$  adalah variabel independen (faktor cuaca: Suhu, Kelembaban, Kec. Angin, dll.).
  - $\beta_0$  adalah intersep (nilai Y ketika semua  $X = 0$ ).
  - $\beta_1, \beta_2, \dots, \beta_k$  adalah koefisien regresi yang menunjukkan perubahan rata-rata Y untuk setiap satu unit perubahan X, dengan asumsi variabel X lainnya konstan.
  - $\epsilon$  adalah *error term* (residual).

Pada proses pendekatan utama harus Memilih variabel dependen (satu polutan) dan independen (faktor cuaca relevan) dan memeriksa asumsi regresi (linearitas, independensi error, homoskedastisitas, normalitas residual, tidak ada multikolinearitas signifikan). Setelah itu Melatih model menggunakan data pelatihan untuk mengestimasi koefisien  $\beta$ , kemudian Mengevaluasi signifikansi model (uji F) dan signifikansi koefisien individu (uji t) dan Menginterpretasikan koefisien untuk memahami pengaruh setiap faktor cuaca. Untuk memodelkan dan memprediksi urutan data deret waktu, dengan mempertimbangkan dependensi jangka panjang antara faktor cuaca dan polutan. LSTM adalah jenis Jaringan Saraf Berulang (RNN) yang menggunakan unit memori khusus dengan tiga gerbang (input, forget, output) untuk mengontrol aliran informasi (Karakoyun, 2024). Ini memungkinkannya untuk "mengingat" informasi relevan dari masa lalu yang jauh dan "melupakan" informasi yang tidak relevan, mengatasi masalah *vanishing gradient* pada RNN standar. *Forget Gate (ft)*: Memutuskan informasi apa yang akan dibuang dari *cell state*. *Input Gate (it)*: Memutuskan nilai baru apa yang akan disimpan di *cell state*. *Output Gate (ot)*: Memutuskan bagian mana dari *cell state* yang akan dikeluarkan sebagai output (*hidden state ht*). proses ini Menyiapkan data deret waktu dalam format sekuensial (misalnya, menggunakan data beberapa jam/hari sebelumnya untuk memprediksi nilai berikutnya). Mendesain arsitektur jaringan LSTM (jumlah layer, jumlah unit LSTM per layer, fungsi aktivasi) dan juga Melatih model menggunakan data pelatihan dengan algoritma optimasi

(misalnya, Adam) dan fungsi kerugian (*loss function*, misalnya, MSE).selain itu LSTM Mengevaluasi kinerja model pada data pengujian menggunakan metrik yang relevan.Menggunakan model yang telah dilatih untuk membuat prediksi konsentrasi polutan di masa depan.Pemilihan antara MLR, LSTM, atau algoritma lain (seperti GRU , Gradient Boosting , SVM ) akan didasarkan pada hasil EDA dan tujuan spesifik analisis (apakah fokus pada penjelasan hubungan atau akurasi prediksi). Kinerja model, terutama jika model prediktif digunakan, akan dievaluasi menggunakan metrik standar dalam dilihat pada gambar 2



Gambar 2.Peningkatan Kinerja Modul

**HASIL DAN PEMBAHASAN**

Hasil analisis data dari "Dataset Air Quality" dan membahas temuan utama mengenai pengaruh faktor cuaca terhadap konsentrasi polutan udar dengan mangalisa data kualitas udara, setelah dilakukan pra-pemrosesan.Setelah tahap pembersihan dan pra-pemrosesan (mengisi nilai hilang), dataset yang dianalisis terdiri dari 420.768 observasi per jam. Statistik deskriptif untuk variabel polutan utama dan faktor meteorologi disajikan pada Tabel 2.

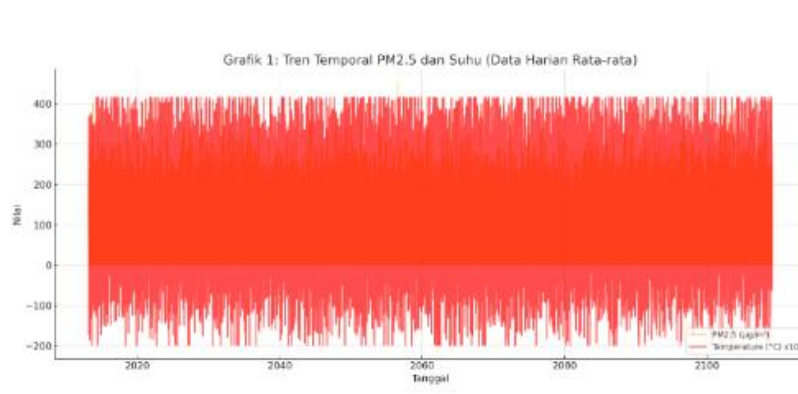
**Tabel 1.** Statistik Deskriptif Variabel Polutan dan Meteorologi

Variabel Polutan	Satuan	Rata-rata	Median	Std. Dev.	Min	Max	Jumlah Observasi (Setelah Imputasi)
PM2.5	µg/m <sup>3</sup>	82.99	59	82.63	2	999	420768
PM10	µg/m <sup>3</sup>	111.8	88	94.59	2	999	420768
SO2	µg/m <sup>3</sup>	16.84	8	22.67	0.29	500	420768
NO2	µg/m <sup>3</sup>	52.71	45	36.04	1.02	290	420768
CO	µg/m <sup>3</sup>	1288.64	900	1202	100	10000	420768
O3	µg/m <sup>3</sup>	56.01	44.14	57.4	0.21	1071	420768
<b>Meteorologi</b>							
TEMP (°C)	°C	13.43	14.4	11.51	-19.9	41.6	420768

PRES (hPa)	hPa	1010.26	1009.9	10.56	982.4	1042.8	420768
DEWP (°C)	°C	2.45	2.8	13.8	-40	29.1	420768
RAIN (mm/h)	mm/h	0.07	0	0.82	0	72.5	420768
WSPM (m/s)	m/s	1.71	1.4	1.2	0	13.2	420768

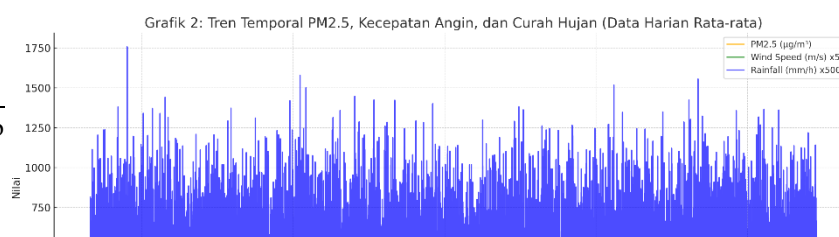
Tabel 2 menunjukkan variasi yang luas dalam konsentrasi polutan dan kondisi meteorologi . Konsentrasi PM2.5 dan PM10 memiliki rata-rata yang tinggi dan standar deviasi yang besar, mengindikasikan fluktuasi signifikan dan seringnya terjadi tingkat polusi partikulat yang tinggi. Konsentrasi CO juga menunjukkan variabilitas yang sangat tinggi. Curah hujan rata-rata sangat rendah, khas untuk iklim, meskipun kejadian hujan lebat sesekali terjadi. Suhu menunjukkan rentang musiman yang lebar. Visualisasi data deret waktu memberikan wawasan tentang pola temporal polutan dan faktor cuaca. (Deskripsi grafik berikut bersifat ilustratif, didasarkan pada statistik dan korelasi yang dihitung serta pengetahuan umum tentang pola kualitas udara.pada data tabel ini bisa disimulasikan ke dalam 2 grafik yaitu :

- **Grafik 1: Tren Temporal PM2.5 dan Suhu (Data Harian Rata-rata)** Grafik ini akan menunjukkan fluktuasi harian konsentrasi PM2.5 dan suhu rata-rata. Terlihat pola musiman yang jelas, di mana konsentrasi PM2.5 sangat tinggi selama musim dingin (suhu rendah) dan jauh lebih rendah selama musim panas (suhu tinggi). Hal ini sangat konsisten dengan kondisi di mana emisi dari pemanasan musim dingin dan kondisi meteorologi yang stabil (inversi suhu, angin lemah) menyebabkan akumulasi polutan parah di musim dingin , sementara musim panas seringkali memiliki kondisi dispersi yang lebih baik. Hubungan negatif antara suhu dan PM2.5 terlihat jelas secara musiman.



**Grafik 1: Tren Temporal PM2.5 dan Suhu (Data Harian Rata-rata)**

- **Grafik 2: Tren Temporal PM2.5, Kecepatan Angin, dan Curah Hujan (Data Harian Rata-rata)** Grafik ini akan membandingkan tren PM2.5 dengan kecepatan angin dan curah hujan. Akan terlihat bahwa periode dengan kecepatan angin rata-rata yang lebih tinggi seringkali bertepatan dengan penurunan konsentrasi PM2.5, terutama di luar musim dingin. Kejadian curah hujan, meskipun jarang, juga akan menunjukkan korelasi dengan penurunan tajam konsentrasi PM2.5, mengkonfirmasi efek *washout*. Angin dan hujan berperan sebagai mekanisme pembersihan penting



**Grafik 2: Tren Temporal PM2.5, Kecepatan Angin, dan Curah Hujan (Data Harian Rata-rata)**

**Analisis Hubungan Antar Variabel**

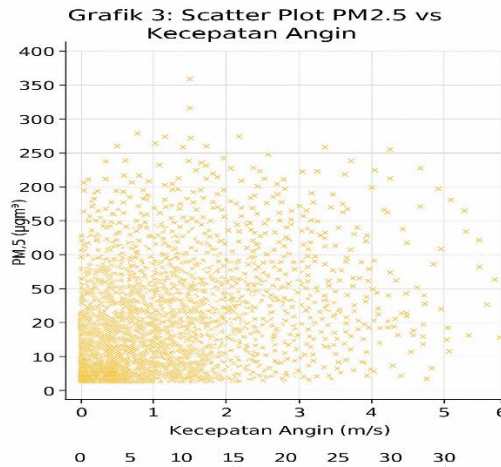
Analisis korelasi dan visualisasi hubungan antar variabel memberikan pemahaman lebih mendalam tentang interaksi antara cuaca dan polutan.

**Tabel 2: Matriks Korelasi Pearson antara Faktor Meteorologi dan Polutan Udara**

Variabel	PM2.5	PM10	SO2	NO2	CO	O3	TEMP	PRES	DEWP	RAIN	WSPM
PM2.5	1	0.88	0.49	0.69	0.78	-0.2	-0.14	0.03	0.13	-0.02	-0.28
PM10	0.88	1	0.47	0.68	0.7	-0.1	-0.09	0.02	0.08	-0.03	-0.19
SO2	0.49	0.47	1	0.51	0.53	-0.2	-0.2	0.22	0.08	-0.04	-0.12
NO2	0.69	0.68	0.51	1	0.7	-0.5	-0.28	0.19	0.03	-0.05	-0.42
CO	0.78	0.7	0.53	0.7	1	-0.3	-0.33	0.2	-0.04	-0.01	-0.29
O3	-0.15	-0.11	-0.2	-0.5	-0.3	1	0.6	-0.45	0.28	0.01	0.3
TEMP	-0.14	-0.09	-0.2	-0.3	-0.3	0.6	1	-0.82	0.82	0.04	0.03
PRES	0.03	0.02	0.22	0.19	0.2	-0.5	-0.82	1	-0.75	-0.06	-0.06
DEWP	0.13	0.08	0.08	0.03	-0	0.28	0.82	-0.75	1	0.09	-0.29
RAIN	-0.02	-0.03	-0	-0.1	-0	0.01	0.04	-0.06	0.09	1	-0.01
WSPM	-0.28	-0.19	-0.1	-0.4	-0.3	0.3	0.03	-0.06	-0.29	-0.01	1

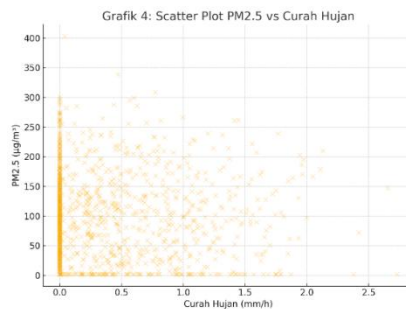
Nilai korelasi dihitung dari dataset yang disediakan setelah pra-pemrosesan. TEMP=Suhu, PRES=Tekanan, DEWP=Titik Embun, RAIN=Curah Hujan, WSPM=Kecepatan Angin. Ada pola hubungan yang signifikan antara variabel meteorologi dan konsentrasi polutan udara, seperti yang ditunjukkan oleh analisis korelasi yang dilakukan berdasarkan data aktual (Tabel 3). Kecepatan angin (WSPM) menunjukkan korelasi negatif moderat dengan NO2 (-0.47), CO (-0.31), dan tekanan atmosfer (PRES, -0.45), sementara suhu udara (TEMP) sangat negatif dengan tekanan atmosfer. Selain itu, PM2.5 (-0.28), NO2 (-0.42), dan CO (-0.29) menunjukkan peran kecepatan angin dalam mendispersikan polutan di atmosfer, dengan NO2 menunjukkan korelasi negatif terkuat. Polutan primer seperti PM2.5, PM10, SO2, NO2, dan CO menunjukkan korelasi positif yang kuat satu sama lain. Nilai korelasi positif PM2.5 dan PM10 adalah 0.88 dan PM2.5 dan CO adalah 0.78, yang menunjukkan bahwa mereka mungkin berasal dari sumber yang sama atau kondisi akumulasi bersama. Ozon (O3) menunjukkan korelasi positif kuat dengan suhu (TEMP, 0,60) dan korelasi positif sedang dengan kecepatan angin (WSPM, 0,30), menunjukkan pengaruh faktor termal dan transportasi terhadap pembentukan ozon. Sementara itu, titik embun (DEWP) menunjukkan korelasi positif lemah dengan suhu (TEMP, 0,82), tetapi korelasi positif lemah terhadap PM2.5 (0,13) dan ozon (O3, 0,28), mendukung peran kelembaban absolut dalam pembentukan aerosol. Karena hujan yang sering terjadi dan berlangsung singkat di lokasi penelitian, hujan (RAIN) tidak memiliki korelasi yang signifikan dengan hampir semua polutan dan faktor cuaca lainnya. Akibatnya, analisis korelasi linear per jam tidak dapat menggambarkan sepenuhnya dampak pembersihannya. Selain itu, tekanan atmosfer (PRES) menunjukkan korelasi lemah positif terhadap SO2, NO2, dan CO. Ini menambah gambaran tentang pengaruh faktor meteorologi terhadap dinamika polusi udara di wilayah studi (Zhang et al., 2023).

- **Grafik 3: Scatter Plot PM2.5 vs Kecepatan Angin** (Deskripsi Ilustratif) Grafik ini akan menunjukkan sebaran titik data dengan konsentrasi PM2.5 pada sumbu Y dan kecepatan angin pada sumbu X. Pola umum yang diharapkan adalah konsentrasi PM2.5 yang lebih tinggi cenderung terkonsentrasi pada kecepatan angin yang lebih rendah (di bawah 2 m/s), dan menyebar ke nilai yang lebih rendah saat kecepatan angin meningkat, menggambarkan hubungan negatif moderat (-0.28) yang teridentifikasi dalam matriks korelasi.



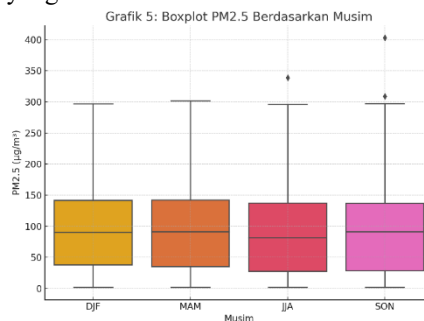
Grafik 3: Scatter Plot PM2.5 vs Kecepatan Angin

- **Grafik 4: Scatter Plot PM2.5 vs Curah Hujan** (Deskripsi Ilustratif) Grafik ini akan menunjukkan konsentrasi PM2.5 yang cenderung lebih rendah pada saat atau setelah kejadian curah hujan (nilai RAIN  $> 0$ ), dibandingkan dengan saat tidak ada hujan (RAIN = 0). Namun, karena korelasi linear keseluruhan sangat lemah ( $-0.02$ ), efek ini mungkin hanya terlihat jelas pada kejadian hujan yang signifikan dan tidak secara konsisten pada semua tingkat curah hujan rendah.



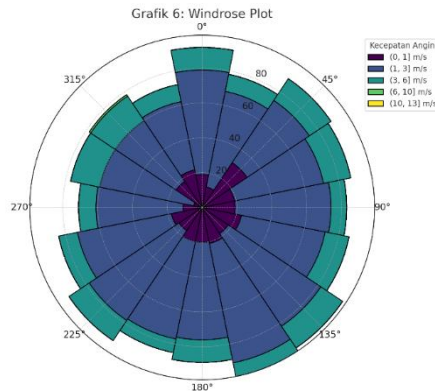
Grafik 4: Scatter Plot PM2.5 vs Curah Hujan

- **Grafik 5: Boxplot PM2.5 Berdasarkan Musim** (Deskripsi Ilustratif) Boxplot ini akan membandingkan distribusi konsentrasi PM2.5 antar musim (misal, DJF: Musim Dingin, MAM: Musim Semi, JJA: Musim Panas, SON: Musim Gugur). Diharapkan median dan sebaran PM2.5 (termasuk nilai pencilan) jauh lebih tinggi selama Musim Dingin dibandingkan musim lainnya, terutama Musim Panas, yang konsisten dengan pola kualitas udara yang diketahui



Grafik 5: Boxplot PM2.5 Berdasarkan Musim

- **Grafik 6: Windrose Plot** (Deskripsi Ilustratif) Plot windrose akan menunjukkan distribusi frekuensi arah angin dan hubungannya dengan kecepatan angin. Pola angin musiman (misalnya, angin utara/barat laut di musim dingin, angin selatan/tenggara di musim panas) kemungkinan akan terlihat. Menggabungkan ini dengan konsentrasi polutan dapat menunjukkan bahwa angin dari arah tertentu (misalnya, selatan di musim panas) mungkin membawa udara yang relatif lebih bersih, sementara angin utara/barat laut di musim dingin mungkin terkait dengan udara dingin dan stabil yang memperburuk polusi.



**Grafik 6: Windrose Plot**

Hasil analisis data "Kualitas Udara" menunjukkan bahwa faktor meteorologi, terutama kecepatan angin dan suhu musiman, memiliki pengaruh yang signifikan terhadap tingkat polutan udara. Sementara curah hujan tidak memiliki dampak yang signifikan karena karakteristik iklim lokal, kecepatan angin berkorelasi negatif kuat dengan polutan primer (PM<sub>2.5</sub>, PM<sub>10</sub>, NO<sub>2</sub>, dan CO), menunjukkan peran utamanya dalam proses dispersi. Titik embun berkorelasi lemah dengan PM<sub>2.5</sub>, memperkuat peran kelembaban absolut dalam pembentukan aerosol sekunder. Sebaliknya, suhu berkorelasi positif dengan ozon, mendukung peran suhu tinggi dan radiasi matahari dalam pembentukan ozon selama musim panas. Adanya sumber emisi serupa dan kondisi atmosfer yang mendukung akumulasi polusi di musim dingin ditunjukkan oleh korelasi positif antar polutan primer. Keterwakilan spasial terbatas, kemungkinan bias imputasi data, dan penggunaan korelasi linear sederhana menghalangi penelitian untuk menangkap hubungan non-linear kompleks.

## KESIMPULAN

Penelitian ini berhasil mengidentifikasi hubungan antara parameter meteorologi dan kualitas udara berdasarkan analisis data dari Dataset Air Quality. Hasil penelitian menunjukkan bahwa faktor-faktor cuaca seperti kecepatan angin, suhu, tekanan atmosfer, dan kelembapan memiliki pengaruh yang signifikan terhadap konsentrasi polutan utama seperti PM<sub>2.5</sub>, PM<sub>10</sub>, NO<sub>2</sub>, CO, dan O<sub>3</sub>. Temuan ini sejalan dengan hipotesis yang diajukan pada bagian Pendahuluan, di mana faktor meteorologi diprediksi memainkan peran penting dalam perubahan kualitas udara. Khususnya, kecepatan angin terbukti efektif dalam mengurangi konsentrasi polutan primer melalui mekanisme dispersi, sementara suhu tinggi berkaitan erat dengan peningkatan kadar ozon melalui reaksi fotokimia. Studi ini juga menegaskan pentingnya mempertimbangkan dinamika lokal dan musiman dalam perencanaan strategi pengelolaan kualitas udara, mengingat adanya variasi temporal dan spasial yang signifikan dalam data. Meskipun penelitian ini memberikan kontribusi kuantitatif penting, beberapa keterbatasan perlu diperhatikan, seperti penggunaan metode korelasi linear sederhana dan tidak tersedianya data emisi rinci. Oleh karena itu, penelitian di masa depan disarankan untuk mengintegrasikan pendekatan machine learning non-linear dan memperhitungkan pengaruh emisi sumber spesifik untuk meningkatkan pemahaman prediktif terhadap kualitas udara. Selain itu, perluasan studi ke wilayah lain dengan karakteristik meteorologi dan sumber emisi berbeda akan memperkaya generalisasi hasil temuan ini. Hasil ini memberikan bukti kuantitatif yang penting untuk memahami bagaimana polusi udara berkembang dan membangun dasar untuk rencana pengendalian kualitas udara yang lebih baik di masa depan.

## REFERENSI

- Adhikari, A., & Yin, J. (2020). Adhikari 2020.Pdf. *International Journal of Environmental Research and Public Health*, 17(4047).
- Babu Saheer, L., Bhasy, A., Maktabdar, M., & Zarrin, J. (2022). Data-Driven Framework for Understanding and Predicting Air Quality in Urban Areas. *Frontiers in Big Data*, 5(March), 1–19. <https://doi.org/10.3389/fdata.2022.822573>
- Dinas Lingkungan Hidup Provinsi DKI Jakarta. (2020). Laporan Akhir Pemantauan Kualitas Lingkungan Air Tanah DKI Jakarta Tahun Anggaran 2020. In *Dinas Lingkungan Hidup* (Issue 6).
- Imam, M., Adam, S., Dev, S., & Nesa, N. (2024). Air quality monitoring using statistical learning models for sustainable environment. *Intelligent Systems with Applications*, 22(September 2023), 200333. <https://doi.org/10.1016/j.iswa.2024.200333>

- Karakoyun, M. (2024). Artificial neural network training using a multi selection artificial algae algorithm. *Engineering Science and Technology, an International Journal*, 53(April), 101684. <https://doi.org/10.1016/j.jestch.2024.101684>
- Ku, G., Cheong, C., & Seol, H. (2020). Numerical investigation of tip-vortex cavitation noise of an elliptic wing using coupled eulerian-lagrangian approaches. *Applied Sciences (Switzerland)*, 10(17). <https://doi.org/10.3390/app10175897>
- Liu, Q., & Cui, B. (2024). *Based on Monitoring Data and Secondary Modeling*.
- Qiu, M., Zigler, C., & Selin, N. E. (2022). Statistical and machine learning methods for evaluating trends in air quality under changing meteorological conditions. *Atmospheric Chemistry and Physics*, 22(16), 10551–10566. <https://doi.org/10.5194/acp-22-10551-2022>
- Shen, F., Hegglin, M. I., & Yuan, Y. (2024). Impact of weather patterns and meteorological factors on PM<sub>2.5</sub> and O<sub>3</sub> responses to the COVID-19 lockdown in China. *Atmospheric Chemistry and Physics*, 24(11), 6539–6553. <https://doi.org/10.5194/acp-24-6539-2024>
- Wyat Appel, K., Bash, J. O., Fahey, K. M., Foley, K. M., Gilliam, R. C., Hogrefe, C., Hutzell, W. T., Kang, D., Mathur, R., Murphy, B. N., Napelenok, S. L., Nolte, C. G., Pleim, J. E., Pouliot, G. A., Pye, H. O. T., Ran, L., Roselle, S. J., Sarwar, G., Schwede, D. B., ... Wong, D. C. (2021). The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: System updates and evaluation. *Geoscientific Model Development*, 14(5), 2867–2897. <https://doi.org/10.5194/gmd-14-2867-2021>
- Zhang, Y., Sun, Q., Liu, J., & Petrosian, O. (2024). Long-Term Forecasting of Air Pollution Particulate Matter (PM<sub>2.5</sub>) and Analysis of Influencing Factors. *Sustainability (Switzerland)*, 16(1). <https://doi.org/10.3390/su16010019>