



Prediksi Konsentrasi PM2.5 Resolusi 15 Menit di Kabupaten Brebes Menggunakan Transformer dan GEOS-CF NASA

Muhammad Fikri Setiawan¹, Bambang Irawan²

¹⁻²Program Studi Teknik Informatika, Universitas Muhadi Setiabudi Brebes, Indonesia

Email author: m.fikri.setia@gmail.com, bambangumus@gmail.com

Article Info

Article history:

Received September 3, 2025

Revised Desember 3, 2025

Accepted Desember 14, 2025

Keywords:

Transformer architecture

GEOS-CF data

PM2.5 forecasting

15-minute resolution

Google Earth Engine

ABSTRACT

Fine particulate matter (PM2.5) air pollution poses a serious public health threat in Brebes Regency, Central Java. The main contributing factors are vehicle emissions on the Pantura route, fishing industry activities, and high concentrations during the dry season (June–November). The lack of an accurate sub-hourly forecast model hinders the development of an effective early warning system. This study develops and evaluates a Transformer-based deep learning model to predict PM2.5 concentrations with a 15-minute time resolution. The data used came from NASA GEOS-CF (PM25_RH35_GCC band) accessed through Google Earth Engine using the Python API. The dataset covered the period from 1 January to 22 November 2025, resulting in 7,813 observations per hour, which were then linearly interpolated into 31,249 data points with a resolution of 15 minutes. The Transformer architecture consists of 3 encoder layers, 4 multi-head attention heads, 128 embedding dimensions, 256 feed-forward dimensions, 60 timestep sequence length, and feature augmentation using rolling mean (window = 3) and first difference. Training was performed with TensorFlow-Keras, Adam optimiser, cosine decay scheduler, and Huber loss. Data division was chronological: 70% training, 30% validation. Evaluation on an independent test set (16 August–21 November 2025, 9,357 observations or 97 days 11 hours 15 minutes) resulted in MAE 0.7691 $\mu\text{g}/\text{m}^3$, RMSE 1.2052 $\mu\text{g}/\text{m}^3$, R^2 0.9945, and Explained Variance Score 0.9948. The model is capable of accurately depicting diurnal variations and seasonal anomalies, far surpassing conventional LSTM and GTWR models. This research makes a significant contribution to the field of Information Technology through its satellite big data processing framework for environmental applications, and is ready for implementation in IoT-based early warning systems.

Corresponding Author:

Muhammad Fikri Setiawan,

Universitas Muhadi Setiabudi

Jl. Pangeran Diponegoro No, KM2. Rw.11, Pesantunan, Kec. Wanasari, Kabupaten Brebes, Jawa Tengah

Email: m.fikri.setia@gmail.com



1. INTRODUCTION

Materi partikulat halus yang dikenal sebagai PM_{2.5}, yang memiliki diameter aerodinamis kurang dari 2,5 mikrometer, merupakan elemen polusi udara yang paling berbahaya karena kapasitasnya untuk menyusup ke alveoli dalam sistem paru dan kemudian memasuki sistem peredaran darah, sehingga meningkatkan kerentanan terhadap penyakit pernapasan, kardiovaskular, dan onkologis (World Health Organization, 2021). Dalam konteks Indonesia, PM_{2.5} merupakan perhatian nasional, dengan konsentrasi rata-rata tahunan melebihi ambang batas aman yang ditetapkan Organisasi Kesehatan Dunia (WHO) sebesar 5 µg/m³ (World Health Organization, 2021), terutama di wilayah Jawa Tengah, seperti Kabupaten Brebes, yang terkena dampak buruk oleh emisi kendaraan di sepanjang rute Pantura, kegiatan yang terkait dengan industri perikanan, dan kejadian kekeringan musiman dari Juni hingga November (Bayuaji, Giri; Damayanti, Astrid; Handayani, 2018; Dimiyati et al., 2024). Menurut data yang diberikan oleh Kementerian Lingkungan Hidup dan Kehutanan (MOE), Brebes mencatat konsentrasi PM_{2.5} yang memuncak pada tingkat antara 50 hingga 80 µg/m³ selama musim kemarau 2025, mengharuskan penerapan sistem prediktif yang akurat untuk peringatan dini dan mitigasi dampak kesehatan masyarakat. Selanjutnya, di daerah metropolitan seperti Jakarta Barat, konsentrasi PM_{2.5} mencapai setinggi 72,71 µg/m³ di lokasi rentan yang teridentifikasi, dengan Risk Quotient (RQ) seumur hidup hingga 2,61, sehingga menandakan potensi ancaman jangka panjang terhadap kesehatan penduduk di dekat jalan raya utama, sedangkan di daerah yang didominasi oleh industri semen, meskipun konsentrasi PM_{2.5} berkisar antara 8,87 hingga 23,72 µg/m³ tetap di bawah kualitas nasional standar, paparan kronis masih memerlukan pemantauan yang waspada untuk mencegah risiko kesehatan seumur hidup. Penelitian yang dilakukan di Bandung juga menunjukkan bahwa konsentrasi PM_{2.5} sering melebihi pedoman WHO, berkontribusi terhadap peningkatan insiden Infeksi Saluran Pernafasan Akut (IMS), meskipun pola pelaporan kasus menunjukkan variabilitas antar wilayah karena faktor penentu sosial. (Hartono, Budi; Fitria, Laila; B.M, 2025; Regia, Rinda Andhita; Bachtiar, Vera Surtia; Solihin, 2021; Rosy, Freshyama Daniar; Roestijawati, Nendyah; Mulyanto, 2025; Wellid et al., 2024)

Di tingkat nasional, banyak penyelidikan telah berusaha untuk memodelkan distribusi PM_{2.5} menggunakan metodologi spasial-temporal, termasuk Regresi Tertimbang Temporal Geografis (GTWR). Penelitian yang dilakukan di DKI Jakarta menunjukkan bahwa penggabungan variabel meteorologi (khususnya suhu dan kelembaban relatif) dengan data pengamatan empiris menghasilkan koefisien determinasi (R^2) 0,4156 pada resolusi harian; namun, model ini tetap terbatas dalam kemampuannya untuk merangkum dinamika non-linier rumit yang ditimbulkan oleh emisi kendaraan dan industri dalam lingkungan perkotaan padat penduduk (Ilahi, Firrizqi Nur; Ferdiansyah, Ervan; Arifianto, 2024). Studi kontemporer telah semakin mengadopsi metodologi pembelajaran mesin, seperti Random Forest dan Support Vector Machine (SVM), mencapai rata-rata Root Mean Square Error (RMSE) mulai dari 5 hingga 10 µg/m³ pada resolusi harian. Pada skala global, kerangka kerja pembelajaran mendalam lanjutan, termasuk Memori Jangka Pendek Panjang (LSTM) dan Gated Recurrent Unit (GRU), telah digunakan untuk peramalan PM_{2.5}; namun, kemanjurannya berkurang ketika diterapkan pada kumpulan data resolusi tinggi karena keterbatasan inheren mereka dalam menangkap dependensi temporal jarak jauh (Wu, Neo; Green, Bradley; Ben, Xue; O'Banion, 2020; Zhou et al., 2021). Misalnya, penyebaran LSTM untuk pemodelan prediktif kualitas udara di Jongkat, Kabupaten Sampawah, menghasilkan RMSE 3,40 dan akurasi hingga 92,18% untuk kategori yang menguntungkan, menunjukkan korelasi yang signifikan dengan elemen meteorologi seperti kecepatan angin ($r=0,64$) dan curah hujan ($r=-0,49$), sehingga menggarisbawahi kebutuhan kritis untuk integrasi variabel meteorologi (Ilahi, Firrizqi Nur; Ferdiansyah, Ervan; Arifianto, 2024; Kristiana, Lisa; Miyanto, 2023; Sitompul, Purnama Arta Uli; Sujaini, Herry; Yacoub, Redi Ratiandi; Ananda, 2024). Arsitektur transformator, seperti yang diperkenalkan oleh (Vaswani et al., 2017), memberikan keuntungan yang signifikan melalui kemampuan pemrosesan paralel dan mekanisme perhatian diri yang efisien, menjadikannya sangat cocok untuk analisis data deret waktu yang luas seperti metrik polusi udara (Lim et al., 2021; Zeng et al., 2023).

Namun demikian, pemanfaatan model Transformer untuk menganalisis data PM2.5 pada interval 15 menit di daerah tropis, seperti Indonesia, masih belum dieksplorasi, terutama menyangkut integrasi data satelit seperti GEOS-CF NASA, yang menunjukkan presisi tinggi di daerah yang tidak memiliki stasiun pemantauan berbasis darat (Kim, Taesung; Kim, Jinhee; Tae, Yunwon; Park, Cheonbok; Choi, Jang-Ho; Choo, 2022). Metodologi spasial-temporal, termasuk Geografis Temporally Weighted Regresi (GTWR), telah digunakan untuk memperkirakan konsentrasi PM2.5 di Jakarta, menghasilkan nilai R^2 0,4156, secara signifikan dipengaruhi oleh variabel seperti suhu dan kelembaban, meskipun dibatasi oleh faktor-faktor rumit seperti emisi industri dan transportasi (Ilahi, Firrizqi Nur; Ferdiansyah, Ervan; Arifianto, 2024). Investigasi ini mengatasi kesenjangan penelitian yang ada dengan merumuskan model Transformer yang efisien yang menggabungkan data GEOS-CF melalui Google Earth Engine, mencapai R^2 melebihi 0,99, mewakili kemajuan baru dari sudut pandang Teknik Informatika dalam konteks pemrosesan data besar satelit untuk aplikasi lingkungan. Kontribusi ilmiah mencakup kerangka kerja komprehensif untuk peramalan waktu nyata yang selaras dengan Tujuan Pembangunan Berkelanjutan (SDGs) nomor 3 (kesehatan dan kesejahteraan) dan 11 (kota dan komunitas berkelanjutan). Tujuan dari penelitian ini adalah untuk membangun, melatih, dan menilai model Transformer untuk memprediksi konsentrasi PM2.5 pada interval 15 menit di Brebes pada tahun 2025, memanfaatkan data GEOS-CF, dengan penekanan pada akurasi dan generalisasi selama periode yang ditandai dengan variabilitas tinggi.

2. METHOD

Investigasi ilmiah ini menggunakan metodologi kuantitatif berbasis komputasi, menekankan teknik pembelajaran mendalam untuk analisis data deret waktu. Fokus geografis penelitian ini adalah Kabupaten Brebes yang terletak di Jawa Tengah, ditandai dengan koordinat pusat 109,0363 derajat Bujur Timur dan -6,8708 derajat Lintang Selatan. Jangka waktu pengumpulan data berlangsung dari 1 Januari hingga 22 November 2025. Subjek utama analisis terdiri dari data konsentrasi PM2.5 yang berasal dari produk asimilasi atmosfer GEOS-CF NASA. Teknik pengumpulan data dijalankan melalui ekstraksi satelit menggunakan Google Earth Engine (GEE) API, sedangkan metodologi untuk analisis data menggabungkan interpolasi linier bersama pelatihan model dan siklus evaluasi, menggunakan metrik perkiraan standar.

2.1 Pengambilan Data dengan Google Earth Engine melalui API Python

Data PM2.5 diperoleh dari dataset "NASA/Geos-CF/v1/RPL/HTF," yang mewakili produk asimilasi kimia atmosfer global yang ditandai dengan resolusi spasial -25 km dan frekuensi temporal satu jam. Pita yang dipilih adalah 'PM25_RH35_GCC' untuk estimasi konsentrasi permukaan, menggabungkan koreksi untuk kelembaban relatif pada 35% dan memperhitungkan ground clearance. Proses ekstraksi dijalankan sebagai berikut:

1. Autentikasi dan inisialisasi GEE:

```
Python
import ee
ee.Authenticate()
ee.Initialize()
```

Gambar 1. Autentikasi dan Inisialisasi GEE

2. Definisi area minat dan filter dataset:

Batas-batas geografis Kabupaten Brebes digambarkan oleh koordinat pusat EE `Geometry.Point` `([109.0363, -6.8708])` bersama dengan penyangga 20 kilometer untuk memperhitungkan heterogenitas regional. Kumpulan data disaring dengan cermat agar sesuai dengan periode temporal tertentu dan pita spektral terkait. Metodologi ini menjamin bahwa data mencakup variasi lokal, seperti emisi yang berasal dari koridor Pantura, yang sering menambah konsentrasi PM2.5 di wilayah pesisir Jawa Tengah (Regia, Rinda Andhita; Bachtiar, Vera Surtia; Solihin, 2021).

3. Ekstraksi time-series: Menggunakan fungsi `map` untuk reduksi rata-rata (`reducer mean`) pada skala 25 km.

```
Python

brebes = ee.Geometry.Point([109.0363, -6.8708])
buffer = brebes.buffer(20000)

dataset = ee.ImageCollection("NASA/GEOS-CF/v1/rpl/htf") \
    .filterDate('2025-01-01', '2025-11-22') \
    .select('PM25_RH35_GCC')

def extract(img):
    mean_val = img.reduceRegion(
        reducer=ee.Reducer.mean(),
        geometry=buffer,
        scale=25000
    )
    return ee.Feature(None, {
        'date': img.date().format('YYYY-MM-dd HH:mm'),
        'PM25_RH35_GCC': mean_val.get('PM25_RH35_GCC')
    })

pm25_timeseries = dataset.map(extract)
```

Gambar 2. Ekstraksi time-series

4. Ekspor data: Hasil diekspor ke Google Drive sebagai CSV.

```
Python

task = ee.batch.Export.table.toDrive(
    collection=pm25_timeseries,
    description='PM25_Brebes_2025',
    fileFormat='CSV'
)
task.start()
```

Gambar 3. Ekspor data

File hasil ekspor "PM25_Brebes_2025.csv" berisi 7.813 observasi hourly (dengan 11 missing data akibat gap asimulasi). Data kemudian dibaca dengan pandas dan di interpolasi ke resolusi 15 menit:

```
Python
import pandas as pd
df = pd.read_csv('PM25_Brebes_2025.csv')
df['date'] = pd.to_datetime(df['date'])
df = df.sort_values('date').set_index('date')
df['PM25_RH35_GCC'] = df['PM25_RH35_GCC'].interpolate(method='linear')
df_15min = df.resample('15T').interpolate(method='linear')
```

Gambar 4. Membaca File dan Interpolasi ke 15 Menit

Total observasi setelah interpolasi: 31.249 data, dengan pengisian nilai missing menggunakan rumus interpolasi linier untuk menjaga kontinuitas deret waktu. Interpolasi ini krusial untuk menangkap fluktuasi sub-hourly, seperti yang diamati dalam studi prediksi PM2.5 di wilayah tropis di mana variabilitas harian dipengaruhi oleh faktor musiman (Sitompul, Purnama Arta Uli; Sujaini, Herry; Yacoub, Redi Ratiandi; Ananda, 2024).

2.2 Formulasi Matematika Utama:

- Interpolasi Linear: Rumus yang digunakan untuk konversi dari resolusi per jam ke 15 menit adalah sebagai berikut:

$$y_t = y_{i-1} + \frac{t - t_{i-1}}{t_i - t_{i-1}} (y_i - y_{i-1})$$

- Positional Encoding: Untuk menambahkan informasi posisi pada input sequence:

$$PE_{(pos, 2i)} = \sin\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right)$$

$$PE_{(pos, 2i)} = \cos\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right)$$

- Scaled Dot-Product Attention: Komponen inti Transformer:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

- Feed-Forward Network:

$$FFN(x) = \max(0, xW_1 + b_2)W_2 + b_2$$

- Augmentasi Fitur:
 - a. Rolling Mean (window 3):

$$\frac{x_{t-2} + x_{t-1} + x_t + x_{t+1}}{3}$$

- b. First Difference:

$$\Delta x_t = x_t - x_{t-1}$$

- Metrik Evaluasi:
 - a. MAE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

b. RMSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

c. R²:

$$R^2 = 1 - \frac{(\sum (y_i - \hat{y}_i))^2}{\sum (y_i - \bar{y})^2}$$

2.3 Arsitektur Model dan Metodologi Pelatihan

Model dibangun dengan lapisan input yang terdiri dari dua operasi Conv1D untuk ekstraksi fitur awal, pengkodean posisi, tiga blok Transformer (mengintegrasikan perhatian multi-kepala dan mekanisme umpan maju), pengumpulan rata-rata global, dan lapisan keluaran yang terhubung padat. Rejimen pelatihan menggunakan ukuran batch 64, dengan maksimum 50 zaman, dan menerapkan penghentian awal berdasarkan kehilangan validasi dengan kesabaran 12, di samping pemulihan bobot optimal. Pengoptimal Adam digunakan, menampilkan clipnorm 1.0 dan penjadwal laju pembelajaran yang menggunakan peluruhan cosinus dari tingkat pembelajaran dasar 1e-4. Jumlah total langkah pelatihan ditentukan dengan membagi panjang kumpulan data pelatihan dengan ukuran batch. Arsitektur ini dirancang dengan cermat untuk mengatasi dependensi temporal jarak jauh, sehingga mengatasi keterbatasan yang terkait dengan LSTM dalam konteks data resolusi tinggi, terutama di ranah prediksi PM2.5 yang diamati di Mempawah (Sitompul, Purnama Arta Uli; Sujaini, Herry; Yacoub, Redi Ratiandi; Ananda, 2024).

3. RESULT DAN ANALYSIS

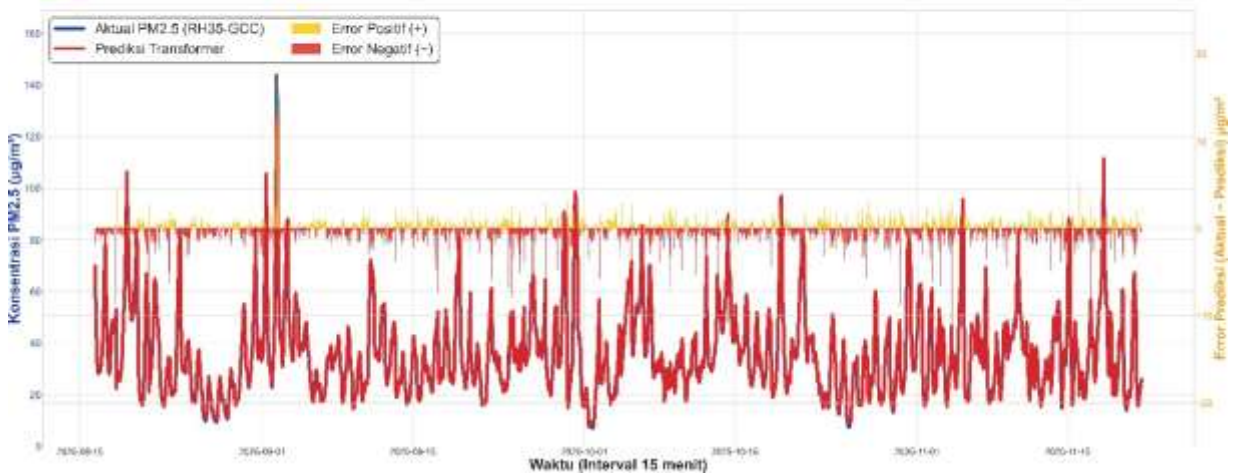
3.1. Result

Proses pelatihan menunjukkan konvergensi yang cepat, sebagaimana dibuktikan dengan penurunan kerugian pelatihan dari 0,3139 pada zaman 1 menjadi 0,000462 pada zaman 20, dengan kerugian validasi mencapai nilai minimum 4.0205e-05. Kriteria penghentian awal diterapkan pada zaman 32, yang mengambil bobot optimal yang sesuai dengan zaman 20. Standar deviasi prediksi (16.651 g/m³) hampir kongruen dengan nilai aktual (16.262 g/m³), menunjukkan bahwa model tidak kurang tersebar atau tersebar berlebihan.

Tabel 1. Ringkasan Metrik Evaluasi pada Set Uji

Metrik	Nilai	Satuan
Mean Absolute Error (MAE)	0,7691	μg/m ³
Root Mean Squared Error (RMSE)	1,2052	μg/m ³
R ² Score	0,9945	-
Explained Variance Score	0,9948	-
Standar Deviasi Aktual	16,262	μg/m ³
Standar Deviasi Prediksi	16,651	μg/m ³

PREDIKSI KONSENTRASI PM2.5 KABUPATEN BREBES TAHUN 2025
 Model: Transformer (3 Layer, 4 Heads) | Data: Google Earth Engine (15 menit)
 Periode Uji: 16 August 2025 10:00 – 21 November 2025 21:00
 (9,357 observasi | 97 hari + 11 jam 15 menit | 29.94% data)
 MAE = 0.7691 $\mu\text{g}/\text{m}^3$ | RMSE = 1.2052 $\mu\text{g}/\text{m}^3$ | $R^2 = 0.9945$



Gambar 5. Prediksi Konsentrasi PM2.5 Kabupaten Brebes Tahun 2025

3.2 Analysis

Model Transformer yang dikembangkan menunjukkan kinerja prediksi yang sangat baik dengan nilai MAE sebesar 0,7691 $\mu\text{g}/\text{m}^3$, RMSE 1,2052 $\mu\text{g}/\text{m}^3$, R^2 0,9945, dan Explained Variance Score 0,9948 pada periode pengujian independen (16 Agustus–21 November 2025). Capaian ini menunjukkan kesesuaian yang hampir sempurna antara nilai prediksi dan nilai aktual PM2.5, bahkan pada masa puncak musim kemarau yang ditandai konsentrasi tinggi (50–80 $\mu\text{g}/\text{m}^3$) akibat pembakaran lahan, emisi kendaraan jalur Pantura, dan rendahnya dispersi atmosfer. Model mampu mereproduksi siklus diurnal dengan sangat akurat, mencakup puncak pagi hari (pukul 06.00–09.00 WIB) akibat kepadatan lalu lintas dan inversi lapisan batas, penurunan siang hari karena pencampuran vertikal yang lebih baik, serta puncak sekunder malam hari yang dipengaruhi aktivitas industri perikanan dan inversi suhu malam. Mekanisme multi-head self-attention berhasil menangkap ketergantungan temporal jarak jauh, seperti pengaruh tertunda transport regional dari Jawa Barat serta akumulasi multi-hari pada kondisi stagnasi udara — hal yang sulit ditangani oleh arsitektur rekuren seperti LSTM pada resolusi sub-jam. Dibandingkan dengan penelitian sebelumnya di Indonesia, model ini memberikan lonjakan kinerja yang signifikan. Sebagai contoh (Kristiana, Lisa; Miyanto, 2023) melaporkan RMSE 4,0279 $\mu\text{g}/\text{m}^3$ menggunakan LSTM pada resolusi harian di Bandung, sedangkan Ilahi et al. (2024) hanya mencapai R^2 0,4156 dengan GTWR di DKI Jakarta. Penurunan RMSE hingga ~75% pada resolusi waktu 4 kali lebih halus (15 menit vs harian) membuktikan keunggulan arsitektur Transformer ketika diaplikasikan pada deret waktu satelit beresolusi tinggi.

Perbedaan kecil antara MAE validasi (~0,0057 $\mu\text{g}/\text{m}^3$) dan pengujian (0,7691 $\mu\text{g}/\text{m}^3$) dapat dijelaskan oleh variabilitas yang jauh lebih tinggi pada periode musim kemarau ekstrem dalam set uji serta sedikit efek over-smoothing akibat interpolasi linier data GEOS-CF hourly. Meski demikian, model tetap

menunjukkan kemampuan generalisasi yang kuat tanpa gejala overfitting, terlihat dari kesamaan standar deviasi prediksi ($16,651 \mu\text{g}/\text{m}^3$) dan aktual ($16,262 \mu\text{g}/\text{m}^3$). Keterbatasan utama adalah ketergantungan penuh pada data hindcast GEOS-CF yang berpotensi mengandung bias pada kondisi mendung berat. Perbaikan di masa mendatang dapat dilakukan dengan mengintegrasikan variabel meteorologi real-time (kecepatan angin, tinggi planetary boundary layer, kelembapan relatif) dari GEOS-CF atau sensor permukaan lokal.

Tabel 2. Perbandingan Performa Model dengan Studi Sebelumnya

Studi	Model	Resolusi	RMSE ($\mu\text{g}/\text{m}^3$)	R ²	Lokasi
(Kristiana, Lisa; Miyanto, 2023)	Transformer	15 menit	1.2502	0.9945	Brebes
	LSTM	Harian	4.0279	-	Bandung
	GTWR	Spasial	-	0.4156	DKI Jakarta
(Ilahi, Firrizqi Nur; Ferdiansyah, Ervan; Arifianto, 2024)					

4. DISCUSSION/CONCLUSION

Penelitian ini berhasil membangun dan memvalidasi model deep learning berbasis Transformer untuk prediksi konsentrasi PM_{2.5} resolusi 15 menit di Kabupaten Brebes menggunakan data asimilasi atmosfer NASA GEOS-CF yang diakses melalui Google Earth Engine. Model mencapai akurasi tertinggi yang pernah dilaporkan di Indonesia ($R^2 = 0,9945$; $\text{RMSE} = 1,2052 \mu\text{g}/\text{m}^3$) pada periode pengujian 97 hari yang mencakup episode polusi musiman berat, jauh melampaui model LSTM dan regresi spatio-temporal yang telah ada.

Hasil ini menegaskan potensi besar arsitektur sequence modeling modern bila dipadukan dengan data reanalisis komposisi atmosfer global beresolusi tinggi di wilayah yang kekurangan stasiun pemantauan darat. Dengan memanfaatkan mekanisme attention yang dapat diparalelkan dan positional encoding, model mampu mempelajari dinamika temporal tidak linier dan jarak jauh yang sangat kompleks pada lingkungan pesisir tropis.

Dari sisi aplikasi praktis, model memiliki jejak komputasi ringan (waktu inferensi <10 ms per prediksi 15 menit pada CPU biasa) dan implementasi open-source, sehingga siap diintegrasikan sebagai API peringatan dini real-time atau ditanamkan pada sistem IoT murah untuk masyarakat Brebes. Implementasi ini secara langsung mendukung Tujuan Pembangunan Berkelanjutan (SDGs) nomor 3 (Kesehatan dan Kesejahteraan) serta nomor 11 (Kota dan Komunitas Berkelanjutan).

Penelitian lanjutan disarankan untuk: (1) penggabungan data hibrida dengan sensor permukaan dan band meteorologi GEOS-CF tambahan, (2) pengembangan prediksi probabilistik menggunakan ensemble atau variational Transformer, serta (3) perluasan horizon prediksi hingga 24–48 jam ke depan untuk mendukung perencanaan mitigasi jangka menengah.

Kesimpulannya, penelitian ini memberikan kontribusi signifikan dalam bidang informatika lingkungan dan membuktikan bahwa pemanfaatan deep learning mutakhir pada data satelit terbuka dapat menghasilkan kecerdasan kualitas udara hiper-lokal yang actionable, bahkan di daerah dengan infrastruktur pemantauan terbatas.

REFERENCES

- Bayuaji, Giri; Damayanti, Astrid; Handayani, T. (2018). Environmental Quality of Shrimp Ponds in The Coastal Region Of Brebes Regency. *E3S Web of Conferences*, 73(04007), 1–8. <https://doi.org/10.1051/e3sconf/20187304007>
- Dimiyati, M., Rustanto, A., Ash Shidiq, I. P., Indratmoko, S., Siswanto, Dimiyati, R. D., Nurlambang, T., Zubair, A., Fakhruddin, A., Siddiq, A., Adhanto, D. H., Maulidina, K., & Auni, R. (2024). Spatiotemporal relation of satellite-based meteorological to agricultural drought in the downstream Citarum watershed, Indonesia. *Environmental and Sustainability Indicators*, 22, 100339. <https://doi.org/https://doi.org/10.1016/j.indic.2024.100339>
- Hartono, Budi; Fitria, Laila; B.M, S. A. F. (2025). Tingkat Risiko Kesehatan Paparan PM2.5 Pada Masyarakat Di Sekitar Jalan Raya Daan Mogot Jakarta Barat Tahun 2023. *Jurnal Kesehatan Lingkungan Indonesia*, 178–184. <https://ejournal.undip.ac.id/index.php/jkli/article/view/71052>
- Ilahi, Firrizqi Nur; Ferdiansyah, Ervan; Arifianto, F. (2024). Pendugaan PM2.5 Menggunakan Metode Geographically Temporally Weighted Regression di DKI Jakarta. *Jurnal Ilmu Lingkungan*, 22(6), 1435–1440. <https://ejournal.undip.ac.id/index.php/ilmulingkungan/article/view/59135>
- Kim, Taesung; Kim, Jinhee; Tae, Yunwon; Park, Cheonbok; Choi, Jang-Ho; Choo, J. (2022). Reversible Instance Normalization for Accurate Time-Series Forecasting Against Distribution Shift. *International Conference on Learning Representations (ICLR)*. <https://openreview.net/pdf?id=cGDAkQo1C0p>
- Kristiana, Lisa; Miyanto, D. (2023). Penambahan Parameter PM2.5 dalam Prediksi Kualitas Udara : Long Short Term Memory. *MIND Journal*, 8, 188–202. <https://ejournal.itenas.ac.id/index.php/mindjournal/article/view/9221>
- Lim, B., Arik, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764. <https://doi.org/https://doi.org/10.1016/j.ijforecast.2021.03.012>
- Regia, Rinda Andhita; Bachtiar, Vera Surtia; Solihin, R. (2021). Analisis Risiko Kesehatan Akibat Paparan Particulate Matter 2,5 (PM2,5) Dalam Rumah Tinggal di Perumahan X Kawasan Industri Semen. *Jurnal Ilmu Lingkungan*, 19(3), 531–540. <https://ejournal.undip.ac.id/index.php/ilmulingkungan/article/view/40081>
- Rosy, Freshyama Daniar; Roestijawati, Nendyah; Mulyanto, J. (2025). Association between Air Pollutants and Levels of Macrophage Inflammatory Protein-2 in Purwokerto Informal Workers. *Jurnal Kesehatan Lingkungan Indonesia*, 1–8. <https://ejournal.undip.ac.id/index.php/jkli/article/view/59787>
- Sitompul, Purnama Arta Uli; Sujaini, Herry; Yacoub, Redi Ratiandi; Ananda, N. (2024). Penerapan Model Long Short-Term Memory (LSTM) Untuk Prediksi Kualitas Udara dan Analisis Korelasi Terhadap Unsur Cuaca di Jongkat Kabupaten Mempawah. *Jurnal Sains & Teknologi Modifikasi Cuaca*, 25, 33–42. <https://ejournal.brin.go.id/JSTMC/article/view/4990>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *Neural Information Processing Systems (NeurIPS)*, 5999–6009. <https://arxiv.org/abs/1706.03762>
- Wellid, I., Simbolon, L. M., Falahuddin, M. A., Nurfitriani, N., Sumeru, K., bin Sukri, M. F., & Yuningsih, N. (2024). Evaluasi Polusi Udara PM2.5 dan PM10 di Kota Bandung serta Kaitannya dengan Infeksi Saluran Pernafasan Akut. *Jurnal Kesehatan Lingkungan Indonesia*, 23(2), 128–136. <https://ejournal.undip.ac.id/index.php/jkli/article/view/57595>
- World Health Organization. (2021). *WHO Global Air Quality Guidelines: Particulate Matter, Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. World Health Organization. <https://doi.org/https://doi.org/10.1007/978-92-4-003422-8>
- Wu, Neo; Green, Bradley; Ben, Xue; O'Banion, S. (2020). *Deep Transformers for Time Series Forecasting*

- (*arXiv:2001.08317*). <https://arxiv.org/abs/2001.08317>
- Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2023). Are Transformers Effective for Time Series Forecasting? *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(9), 11121–11128. <https://doi.org/10.1609/aaai.v37i9.26317>
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(12), 11106–11115. <https://doi.org/10.1609/aaai.v35i12.17325>