

Variational Mode Decomposition and Deep Learning for Geomagnetic K-Index Prediction

Asmadi Djasman^{1*}, Ahmad Musyafa^{2**}, Kahfi Heryandi Suradiradja^{3*}

* Program Studi Teknik Informatika, Pascasarjana, Universitas Pamulang | Badan Meteorologi, Klimatologi dan Geofisika

** Program Studi Teknik Informatika, Pascasarjana, Universitas Pamulang

*** Program Studi Teknik Informatika, Pascasarjana, Universitas Pamulang

asmadi.djasman@bmkg.go.id¹, ahmad.musyafa@unpam.ac.id², dosen01514@unpam.ac.id³

Article Info

Article history:

Received 2026-07-08

Revised 2026-07-30

Accepted 2026-08-12

Keyword:

Convolutional Neural Network,
Geomagnetic Storm,
K-Index Prediction,
Long Short-Term Memory,
Variational Mode
Decomposition.

ABSTRACT

Geomagnetic K-index characterizes local disturbance intensity on a quasi-logarithmic scale and serves as a critical input for space weather early warning systems. Predicting this index is challenging due to its non-stationary, chaotic nature and severely imbalanced class distribution. This study develops and compares two univariate hybrid deep learning models, VMD-CNN 1D and VMD-LSTM, for K-index prediction using data from BMKG Tuntungan Observatory from January 2020 to June 2025 (16,064 samples, 3-hour resolution). The preprocessing pipeline applies Variational Mode Decomposition (VMD) with six intrinsic mode functions, followed by an eight-timestep sliding window. An ablation study confirms that VMD substantially contributes to predictive performance, reducing RMSE by 55.4% and 52.4% for CNN 1D and LSTM respectively compared to their non-VMD counterparts with comparable parameter capacity. Both models were evaluated using RMSE, MAE, and R^2 , with statistical significance confirmed via Wilcoxon signed-rank tests. VMD-CNN 1D achieved superior overall performance (RMSE 0.3608, R^2 0.8846) compared to VMD-LSTM (RMSE 0.3889, R^2 0.8660), and both substantially outperformed a Persistence baseline (RMSE 0.9643). However, under storm conditions ($K \geq 5$, $p = 0.024$), VMD-LSTM outperformed VMD-CNN 1D (R^2 0.6632 versus 0.5430) and achieved higher storm-detection recall (0.891 versus 0.812). These findings indicate that architecture choice should reflect the target operational context, with VMD-CNN 1D suited for routine monitoring and VMD-LSTM for storm-period forecasting.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Cuaca antariksa mengacu pada kondisi dinamis ruang antariksa yang didorong oleh aktivitas Matahari dan interaksinya dengan magnetosfer Bumi. Gangguan geomagnetik yang ditimbulkan dapat mengganggu berbagai sistem teknologi kritis, termasuk operasi satelit komunikasi, sistem navigasi berbasis *Global Navigation Satellite System* (GNSS), dan jaringan transmisi tenaga listrik [1]. Untuk mengukur intensitas gangguan lokal tersebut, para ilmuwan geofisika menggunakan Indeks K, yaitu indikator kuasi-logaritmik dengan skala 0 hingga 9 yang merepresentasikan tingkat fluktuasi komponen horizontal medan magnet bumi pada suatu observatorium dalam interval pengukuran tiga jam

[2]. Nilai Indeks $K \geq 5$ menandakan kondisi badai geomagnetik yang berpotensi mengganggu sistem operasional, sehingga kemampuan memprediksi indeks ini secara akurat menjadi komponen penting dalam sistem peringatan dini cuaca antariksa [3].

Meskipun prediksi indeks Kp geomagnetik global telah banyak diteliti dengan berbagai pendekatan *machine learning* dan *deep learning* [3], penelitian yang berfokus pada prediksi Indeks K lokal di wilayah Indonesia masih sangat terbatas [4]. Salah satu studi terbaru yang secara langsung relevan memprediksi Indeks K stasiun BMKG Tuntungan menggunakan model LSTM *multivariat* dengan mekanisme *multi-head attention*, menggabungkan data geomagnetik dari

BMKG dengan parameter angin surya dari NASA [4]. Meskipun pendekatan *multivariat* tersebut berpotensi menangkap interaksi antara kondisi lokal dan pendorong eksternal aktivitas geomagnetik, ketergantungan gannya pada data angin surya real-time menghadirkan batasan operasional bagi observatorium lokal yang tidak selalu memiliki akses ke sumber data tersebut. Penelitian ini mengambil posisi metodologis yang berbeda dengan mengeksplorasi pendekatan univariat berbasis dekomposisi sinyal untuk memprediksi Indeks K tanpa memerlukan input data eksternal, sekaligus menganalisis karakteristik kinerja kedua arsitektur pada beragam tingkat aktivitas geomagnetik. Pendekatan univariat pada penelitian ini dipilih sebagai bagian dari desain metodologi, bukan karena keterbatasan data. Meskipun berbagai parameter *space weather*, seperti kecepatan angin surya, komponen medan magnet antarplanet (IMF Bz), kerapatan proton, dan indeks F10.7, telah banyak dimanfaatkan untuk meningkatkan akurasi prediksi aktivitas geomagnetik, penggunaan variabel tersebut meningkatkan ketergantungan terhadap sumber data satelit eksternal yang ketersediaannya secara *real-time* tidak selalu terjamin. Atas dasar itu, penelitian ini mengevaluasi kemampuan dekomposisi VMD dalam mengekstraksi informasi temporal dari deret waktu historis Indeks K sehingga model tetap dapat menghasilkan prediksi yang kompetitif hanya dengan memanfaatkan data lokal.

Data Indeks K geomagnetik memiliki karakteristik yang bersifat non-linear, *non-stationary*, dan kacau (*chaotic*), sehingga model *deep learning* tunggal seringkali kesulitan menangkap struktur frekuensi yang tersembunyi di balik fluktuasi sinyal tersebut [5]. Untuk mengatasi masalah ini, *Variational Mode Decomposition* (VMD) yang diperkenalkan oleh Dragomiretskiy dan Zosso [6] digunakan sebagai tahap pra-pemrosesan. Berbeda dengan *Empirical Mode Decomposition* (EMD) yang bersifat rekursif dan rentan terhadap *mode mixing* serta sensitivitas *noise*, VMD mentransformasikan masalah dekomposisi menjadi sebuah optimasi variasional yang secara serentak menguraikan sinyal menjadi sejumlah komponen *Intrinsic Mode Functions* (IMF) dengan frekuensi pusat masing-masing. Pendekatan ini telah terbukti efektif meningkatkan akurasi prediksi pada deret waktu geofisika yang kompleks dan *non-stationary* [7].

Convolutional Neural Network satu dimensi (CNN 1D) dan *Long Short-Term Memory* (LSTM) merupakan dua arsitektur *deep learning* yang memiliki keunggulan komplementer dalam pemodelan deret waktu. CNN 1D unggul dalam mengekstraksi pola lokal dan fitur temporal jangka pendek dari data sekuensial melalui operasi konvolusi yang bergerak melintasi dimensi waktu [8], sementara LSTM dirancang khusus untuk menangkap dependensi temporal jangka panjang melalui mekanisme gerbang yang secara selektif menyimpan dan melupakan informasi historis [5]. Kombinasi keduanya dalam kerangka hibrida berbasis VMD telah menunjukkan peningkatan akurasi yang signifikan pada berbagai domain prediksi deret waktu geofisika yang kompleks [7],[9].

Sejumlah penelitian terdahulu telah mengeksplorasi berbagai arsitektur *deep learning* untuk prediksi indeks geomagnetik. Prediksi pada kondisi geomagnetik aktif terbukti secara inheren lebih sulit dibandingkan kondisi tenang [3]. Pendekatan berbasis citra matahari EUV juga telah digunakan untuk memprediksi Kp beberapa hari ke depan, menunjukkan potensi sumber data non-angin surya dalam memperpanjang cakrawala prediksi [10]. Lebih lanjut, variasi lokal dalam aktivitas geomagnetik yang tidak tertangkap oleh model global dapat diprediksi lebih baik melalui pendekatan berbasis observatorium individual [11]. Pada domain dekomposisi sinyal hibrida, kombinasi VMD dengan arsitektur *deep learning* telah menunjukkan efektivitasnya pada prediksi deret waktu keuangan dan beban daya, memperkuat generalitas pendekatan hibrida berbasis VMD lintas domain [12], [13]. Meskipun pendekatan hibrida berbasis VMD telah banyak diterapkan pada berbagai permasalahan prediksi deret waktu, penerapannya pada prediksi Indeks K geomagnetik lokal, khususnya di wilayah ekuatorial, masih belum banyak dilaporkan. Sebagian besar penelitian sebelumnya berfokus pada domain non-geomagnetik, sementara studi prediksi Indeks K lokal [4] belum mengintegrasikan dekomposisi sinyal sebagai tahap ekstraksi fitur. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut melalui penerapan VMD pada pemodelan prediksi Indeks K geomagnetik.

Penelitian ini memberikan dua kontribusi utama. Pertama, dikembangkan dua arsitektur hibrida untuk prediksi Indeks K secara univariat: VMD-CNN 1D *multi-channel* yang memproses seluruh IMF secara serentak sebagai satu input tergabung, dan VMD-LSTM yang memproses setiap IMF secara independen melalui cabang paralel. Kedua, penelitian ini mengungkap *trade-off* performa yang relevan secara operasional antara kedua arsitektur, di mana keunggulan masing-masing model bergantung pada kondisi aktivitas geomagnetik yang dihadapi. Temuan ini memberikan panduan praktis bagi pemilihan arsitektur yang sesuai dengan konteks operasional sistem peringatan dini cuaca antariksa lokal.

Meskipun kombinasi VMD dengan CNN maupun LSTM telah dilaporkan pada berbagai aplikasi prediksi deret waktu di domain lain [9], [13], penelitian ini memberikan validasi empiris terhadap penerapannya pada prediksi Indeks K geomagnetik lokal. Melalui *ablation study*, penelitian ini memperlihatkan bahwa penambahan tahap dekomposisi VMD menghasilkan peningkatan performa yang konsisten dibandingkan arsitektur yang identik tanpa VMD. Dengan demikian, kontribusi utama penelitian ini terletak pada pembuktian bahwa representasi frekuensi yang dihasilkan VMD memberikan nilai tambah yang nyata pada pemodelan Indeks K, sekaligus memungkinkan prediksi yang kompetitif tanpa memerlukan data *space weather* eksternal.

II. METODE

Rancangan eksperimen dalam penelitian ini bersifat komparatif, membandingkan dua arsitektur hibrida untuk

prediksi Indeks K, yaitu VMD-CNN 1D Multi-channel dan VMD-LSTM. Kedua model dirancang dengan kondisi eksperimen identik, mencakup input yang sama, *split* data kronologis yang sama, dan prosedur evaluasi yang sama, sehingga perbedaan performa yang teramati dapat dikaitkan secara langsung dengan perbedaan strategi ekstraksi fitur antara kedua arsitektur.

A. Data

Penelitian ini menggunakan data historis Indeks K geomagnetik yang bersumber dari Observatorium BMKG Tuntungan, Sumatera Utara, Indonesia, yang secara khusus memantau variasi geomagnetik di wilayah ekuatorial. Dataset mencakup periode 1 Januari 2020 hingga 30 Juni 2025 dengan resolusi temporal 3 jam, menghasilkan 8 titik pengamatan per hari dan total 16.064 sampel univariat. Hasil verifikasi integritas menunjukkan dataset bersifat 100% lengkap tanpa adanya nilai yang hilang serta konsisten pada interval 3 jam di sepanjang 16.063 selisih waktu antar pengamatan.

Distribusi nilai Indeks K bersifat sangat tidak seimbang (*highly imbalanced*), dengan 92,16% sampel berada pada kondisi geomagnetik tenang hingga sedang ($K = 0-3$), sementara hanya 1,81% sampel merepresentasikan kondisi badai ($K \geq 5$). Ketidakseimbangan distribusi kelas menjadi salah satu tantangan utama dalam penelitian ini, khususnya pada evaluasi performa model untuk kejadian badai geomagnetik yang jumlahnya relatif sedikit. Teknik penyeimbangan data, seperti oversampling, undersampling, dan weighted loss function, tidak diterapkan karena berpotensi mengubah karakteristik temporal deret waktu Indeks K. Sebagai gantinya, performa model dianalisis berdasarkan setiap kategori kondisi geomagnetik serta menggunakan metrik deteksi khusus untuk kejadian badai pada Bagian Hasil dan Pembahasan, sehingga kemampuan model pada kelas minoritas dapat dievaluasi secara lebih representatif.

B. Pra Pemrosesan Data

Dataset dibagi secara kronologis tanpa proses pengacakan dengan rasio 80:10:10, sehingga diperoleh *training set* sebanyak 12.851 sampel (1 Januari 2020–25 Mei 2024), *validation set* sebanyak 1.606 sampel (25 Mei 2024–12 Desember 2024), dan *testing set* sebanyak 1.607 sampel (12 Desember 2024–30 Juni 2025). Pembagian secara berurutan ini dipilih untuk menjaga urutan temporal data serta mencegah kebocoran informasi temporal (*lookahead bias*) antar subset.

Setelah proses pembagian data, setiap subset dinormalisasi menggunakan metode *Min-Max scaling* untuk memetakan nilai Indeks K ke dalam rentang $[0,1]$. Parameter normalisasi dihitung hanya dari training set ($x_{\min} = 0$ dan $x_{\max} = 9$), kemudian diterapkan tanpa perubahan pada *validation set* dan *testing set*. Karena nilai Indeks K secara teoretis berada pada rentang 0–9 dan seluruh rentang tersebut telah tercakup dalam *training set*, proses normalisasi dapat diterapkan secara

konsisten pada seluruh subset tanpa menghasilkan nilai di luar interval $[0,1]$.

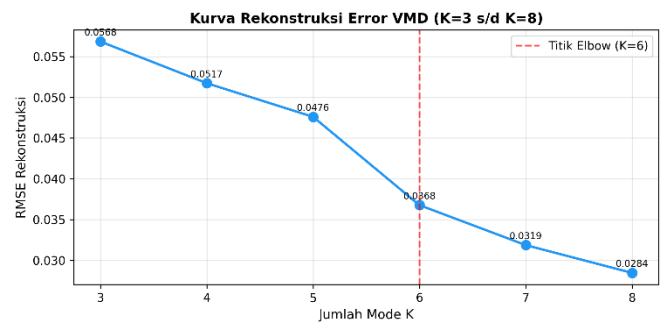
C. Variational Mode Decomposition

Variational Mode Decomposition (VMD) menguraikan sinyal deret waktu menjadi K komponen *Intrinsic Mode Functions* (IMF) dengan menyelesaikan masalah optimasi variasional terbatas yang dirumuskan sebagai berikut [6].

$$\min_{\{u_k\}, \{\omega_k\}} \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \quad \text{s.t.} \quad \sum_{k=1}^K u_k(t) = f(t)$$

Di mana $u_k(t)$ menyatakan komponen IMF ke- k , ω_k merupakan frekuensi pusat dari komponen tersebut, $f(t)$ adalah sinyal masukan, $\delta(t)$ menyatakan fungsi *impuls Dirac*, j merupakan unit imajiner, sedangkan operator $*$ menunjukkan operasi konvolusi. Permasalahan optimasi ini kemudian diselesaikan menggunakan metode *Alternating Direction Method of Multipliers* (ADMM) dengan memperkenalkan pengali Lagrange λ dan parameter penalti kuadratik α , sehingga formulasi optimasi terbatas dapat ditransformasikan menjadi permasalahan optimasi tak terbatas [9].

Dalam penelitian ini, VMD dikonfigurasi menggunakan jumlah mode $K = 6$, faktor penalti $\alpha = 2.000$, dan toleransi konvergensi $\tau = 1 \times 10^{-7}$. Nilai $K = 6$ dipilih secara empiris berdasarkan kriteria *elbow* pada kurva RMSE rekonstruksi yang dievaluasi pada rentang $K = 3$ hingga $K = 8$. Kurva RMSE rekonstruksi terhadap variasi jumlah mode ditunjukkan pada Gambar 1.



Gambar 1. Kurva pemilihan jumlah mode (K) optimal berdasarkan RMSE rekonstruksi

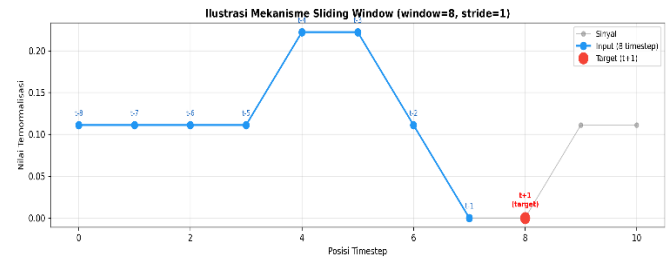
Berdasarkan Gambar 1, nilai RMSE rekonstruksi menurun secara monoton seiring bertambahnya jumlah mode K . Namun, penurunan paling signifikan terjadi pada transisi dari $K = 5$ ke $K = 6$, yaitu sebesar 0,01083, sedangkan pada transisi berikutnya dari $K = 6$ ke $K = 7$ laju penurunan mengecil menjadi 0,00491 atau sekitar 54,7% lebih kecil. Titik siku (*elbow*) pada $K = 6$ tersebut menandakan keseimbangan optimal antara akurasi rekonstruksi dan kualitas dekomposisi, sehingga $K = 6$ dipilih sebagai jumlah mode pada penelitian ini.

Selain jumlah mode K , pengaruh parameter penalti bandwidth α dan toleransi konvergensi τ terhadap hasil dekomposisi VMD juga dievaluasi menggunakan *training set*. Variasi nilai α pada rentang 500 hingga 4.000 menunjukkan kenaikan RMSE rekonstruksi secara bertahap dari 0,0212 menjadi 0,0477, yang mencerminkan adanya *trade-off* antara ketepatan rekonstruksi dan kualitas pemisahan frekuensi antar mode. Nilai α yang lebih rendah menghasilkan rekonstruksi yang lebih akurat, tetapi meningkatkan risiko *mode mixing* antar komponen IMF. Oleh karena itu, $\alpha = 2.000$ dipilih sebagai kompromi yang seimbang antara akurasi rekonstruksi dan kualitas dekomposisi, serta sejalan dengan rentang nilai yang umum digunakan pada penelitian VMD sebelumnya [6]. Sebaliknya, perubahan nilai τ pada rentang 0 hingga 1×10^{-4} hanya menghasilkan variasi RMSE rekonstruksi yang sangat kecil (kurang dari 0,001), sehingga mengindikasikan bahwa parameter ini terutama memengaruhi presisi numerik proses optimasi tanpa memberikan pengaruh yang berarti terhadap kualitas dekomposisi.

Untuk menghindari *lookahead bias*, proses dekomposisi VMD dilakukan secara terpisah pada setiap subset (*subset-wise decomposition*) setelah pembagian data secara kronologis dan normalisasi *Min-Max*. Dengan pendekatan ini, proses dekomposisi pada *training set* tidak memanfaatkan informasi statistik dari *validation set* maupun *testing set*, sehingga model dilatih menggunakan komponen IMF yang sepenuhnya bersifat kausal terhadap horizon prediksi [14].

D. Pembentukan Sampel Sliding Window

Setelah proses dekomposisi VMD selesai, setiap komponen IMF pada masing-masing subset diubah menjadi pasangan data masukan dan target menggunakan teknik *sliding window* dengan ukuran jendela 8 *timestep* dan langkah geser 1 *timestep*. Dengan resolusi temporal data sebesar 3 jam per *timestep*, setiap jendela masukan merepresentasikan riwayat data selama 24 jam. Pemilihan jendela 24 jam ini sejalan dengan praktik dalam pemodelan prediksi geomagnetik, karena konteks historis selama satu hari dinilai memadai untuk menangkap dinamika temporal jangka pendek aktivitas geomagnetik [10]. Penerapan *sliding window* dengan ukuran jendela 8 *timestep* menyebabkan berkurangnya jumlah sampel yang dapat dibentuk pada setiap subset, karena delapan pengamatan pertama hanya berfungsi sebagai riwayat masukan tanpa memiliki target prediksi. Selain itu, panjang deret pada setiap subset disesuaikan menjadi bilangan genap sebagai prasyarat dekomposisi VMD. Setelah kedua proses tersebut, *testing set* yang semula berisi 1.607 pengamatan menghasilkan 1.598 sampel siap model. Ilustrasi mekanisme pembentukan sampel *sliding window* dengan ukuran jendela 8 *timestep* dan langkah geser 1 *timestep* ditunjukkan pada Gambar 2.

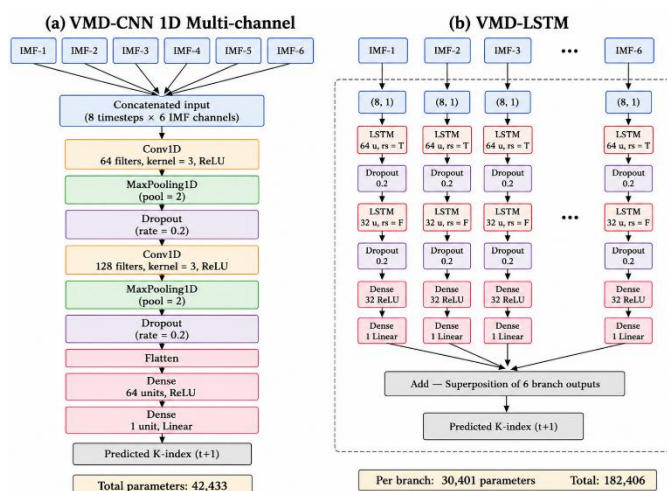


Gambar 2. Ilustrasi mekanisme *sliding window*

E. Arsitektur Model

Kedua arsitektur hibrida dirancang dengan strategi ekstraksi fitur yang berbeda. VMD-CNN 1D Multi-channel memproses keenam komponen IMF secara bersamaan dalam satu jaringan melalui input *multi-channel* berdimensi (8, 6). Arsitektur ini terdiri atas dua blok konvolusi bertingkat dengan jumlah filter yang meningkat (64 dan 128 filter, kernel berukuran 3), yang masing-masing diikuti oleh lapisan *MaxPooling1D* (*pool size* 2) untuk mengurangi dimensi temporal serta lapisan *Dropout* (*rate* 0,2) sebagai regularisasi. Selanjutnya, peta fitur diratakan menggunakan lapisan *Flatten*, kemudian diteruskan ke lapisan *Dense* dengan 64 unit beraktivasi ReLU sebelum lapisan keluaran linear menghasilkan prediksi Indeks K pada timestep berikutnya. Pendekatan *multi-channel* ini memungkinkan kernel konvolusi mempelajari korelasi antar komponen frekuensi secara bersamaan dalam satu proses, sebagaimana telah ditunjukkan efektif pada penelitian sebelumnya yang menerapkan arsitektur hibrida VMD-CNN-LSTM [15].

Berbeda dengan VMD-CNN 1D *Multi-channel*, VMD-LSTM mendistribusikan setiap komponen IMF ke dalam cabang LSTM yang terpisah dan diproses secara paralel. Setiap cabang menerima input berdimensi (8, 1) dan terdiri atas dua lapisan LSTM bertingkat, yaitu lapisan pertama dengan 64 unit memori dan parameter *return sequences* bernilai *True* untuk mempertahankan seluruh urutan *hidden state*, diikuti lapisan kedua dengan 32 unit dan *return sequences* bernilai *False* untuk mengompresi representasi temporal. Selanjutnya, setiap cabang dilengkapi dengan lapisan *Dense* berukuran 32 unit beraktivasi ReLU dan lapisan keluaran linear. Hasil prediksi dari keenam cabang LSTM kemudian digabungkan melalui operasi superposisi (penjumlahan) untuk menghasilkan prediksi akhir. Pendekatan cabang independen ini memungkinkan setiap unit LSTM mempelajari dinamika temporal dari masing-masing komponen frekuensi tanpa dipengaruhi oleh komponen lainnya, sebagaimana diterapkan pada arsitektur hibrida berbasis VMD untuk prediksi deret waktu non-stasioner [16].



Gambar 3. Arsitektur *dual-panel* VMD-CNN 1D *Multi-channel* dan VMD-LSTM

Gambar 3 memperlihatkan perbedaan topologi antara kedua arsitektur. Pada panel (a), VMD-CNN 1D *Multi-channel* menggunakan satu jalur pemrosesan, di mana keenam komponen IMF disusun sebagai *channel* paralel dalam satu *tensor* masukan dan diproses secara bersamaan oleh kernel konvolusi yang sama. Sebaliknya, panel (b) menunjukkan arsitektur multi-cabang VMD-LSTM, di mana setiap komponen IMF diproses melalui jalur LSTM yang terpisah dengan bobot pelatihan yang independen. Selanjutnya, hasil prediksi dari setiap cabang digabungkan melalui operasi penjumlahan untuk menghasilkan estimasi akhir Indeks K. Perbedaan topologi tersebut menjadi dasar bagi karakteristik komputasi yang berbeda pada kedua arsitektur, sebagaimana dirangkum pada Tabel I.

TABEL I
KOMPARASI ARSITEKTUR KEDUA MODEL

Aspek	VMD-CNN 1D	VMD-LSTM
Bentuk input	(8, 6) — satu <i>tensor</i> tergabung	(8, 1) × 6 — enam <i>tensor</i> paralel
Strategi pemrosesan IMF	Serentak (<i>multi-channel</i>)	Independen per cabang
Layer ekstraksi fitur	2× Conv1D (64, 128 filter, kernel 3)	2× LSTM per cabang (64, 32 unit)
Fungsi aktivasi	ReLU + Linear (<i>output</i>)	<i>tanh</i> internal + ReLU + Linear (<i>output</i>)
Reduksi dimensi	MaxPooling1D (pool size 2)	LSTM dengan <i>return_sequences = False</i>
Regularisasi	Dropout 0,2 (2 lapisan)	Dropout 0,2 (2 lapisan per cabang)
Mekanisme agregasi	Flatten + Dense	Superposisi 6 <i>output</i> cabang
Parameter per cabang	-	30.401
Total parameter	42.433	182.406

Berdasarkan Tabel I, VMD-CNN 1D *Multi-channel* memiliki total 42.433 parameter, sedangkan VMD-LSTM memiliki 182.406 parameter (30.401 parameter per cabang × 6 cabang). Meskipun jumlah parameter VMD-LSTM sekitar empat kali lebih besar daripada VMD-CNN 1D, hal tersebut tidak secara otomatis menunjukkan performa yang lebih unggul, karena kompleksitas model tidak selalu berbanding lurus dengan kemampuan generalisasi pada deret waktu non-stasioner.

F. Pelatihan dan Evaluasi Model

Kedua model dilatih menggunakan konfigurasi *hyperparameter* yang sama untuk memastikan bahwa perbedaan performa yang diperoleh berasal dari perbedaan arsitektur, bukan dari perbedaan konfigurasi pelatihan. Rincian konfigurasi tersebut disajikan pada Tabel II.

TABEL II
KONFIGURASI *HYPERPARAMETER* PELATIHAN

<i>Hyperparameter</i>	Nilai
<i>Optimizer</i>	Adam
<i>Learning rate</i>	0,001
<i>Batch size</i>	64
<i>Epoch maksimum</i>	100
<i>Early Stopping patience</i>	15
<i>Loss function</i>	MSE
<i>Dropout rate</i>	0,2

Berdasarkan Tabel II, *optimizer* Adam dipilih karena kemampuannya menyesuaikan *learning rate* secara adaptif pada setiap parameter bobot, sehingga cocok diterapkan pada arsitektur *multi-channel* (VMD-CNN 1D) maupun *multi-branch* (VMD-LSTM) [4]. Sementara itu, mekanisme *Early Stopping* dengan *patience* 15 *epoch* yang memantau *validation loss* digunakan untuk menghentikan pelatihan pada titik generalisasi terbaik serta mencegah terjadinya *overfitting* [15].

Kinerja kedua model dievaluasi secara eksklusif pada *testing set* menggunakan tiga metrik statistik, yaitu *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan Koefisien Determinasi (R^2). RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi ekstrem, MAE mengukur rata-rata penyimpangan secara linier dan lebih *robust* terhadap distribusi data yang tidak seimbang, sedangkan R^2 mengukur proporsi variansi data aktual yang dapat dijelaskan oleh model. Ketiga metrik tersebut dirumuskan sebagai berikut:

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

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

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

Di mana y_i menyatakan nilai aktual, $\hat{y}_i$ merupakan nilai prediksi yang dihasilkan model, $\bar{y}$ adalah rata-rata nilai aktual, dan n menyatakan jumlah sampel pada *testing set*.

Sebagai pembandingan tambahan, kedua model dievaluasi terhadap model *Persistence*, yaitu *baseline* naif yang mengasumsikan nilai Indeks K pada *timestep* berikutnya sama dengan nilai pada *timestep* saat ini tanpa melalui proses pembelajaran [10]. Model ini banyak digunakan sebagai acuan dalam penelitian prediksi indeks geomagnetik untuk merepresentasikan tingkat kesulitan dasar permasalahan prediksi secara independen dari kompleksitas model yang digunakan. Signifikansi statistik dari perbedaan performa antar model dianalisis menggunakan uji *Wilcoxon signed-rank* terhadap *absolute error* pada taraf signifikansi $\alpha = 0,05$. Uji non-parametrik ini dipilih karena tidak mengharuskan data mengikuti distribusi normal serta sesuai untuk membandingkan dua model pada sampel pengujian yang sama. Selain itu, kemampuan kedua model dalam mendeteksi kejadian badai geomagnetik ($K \geq 5$) dievaluasi menggunakan empat metrik klasifikasi, yaitu *Precision*, *Recall* (*Hit Rate*), *F1-Score*, dan *Critical Success Index* (CSI). *Precision* mengukur proporsi prediksi badai yang benar, *Recall* menunjukkan proporsi kejadian badai aktual yang berhasil dideteksi, *F1-Score* merepresentasikan keseimbangan antara *Precision* dan *Recall*, sedangkan CSI mengukur tingkat keberhasilan deteksi dengan mempertimbangkan *false alarm* dan *missed detection*.

Untuk mengkaji kontribusi dekomposisi VMD terhadap peningkatan performa model, dilakukan *ablation study* dengan melatih dua model pembandingan yang mengadopsi arsitektur VMD-CNN 1D dan VMD-LSTM, tetapi menggunakan sinyal Indeks K mentah sebagai masukan tanpa melalui proses dekomposisi VMD. Pada konfigurasi ini, kedua model menerima masukan *single-channel* berdimensi (8, 1), berbeda dengan VMD-CNN 1D yang menggunakan masukan *multi-channel* berdimensi (8, 6) serta VMD-LSTM yang memproses enam cabang paralel. Seluruh konfigurasi pelatihan, termasuk *hyperparameter* pada Tabel II, dipertahankan sama sehingga setiap perbedaan performa yang diamati dapat diatribusikan pada keberadaan proses dekomposisi VMD.

III. HASIL DAN PEMBAHASAN

Seluruh evaluasi kinerja model dilakukan secara eksklusif pada *testing set* yang terdiri atas 1.598 sampel pengamatan periode 12 Desember 2024 hingga 30 Juni 2025. Dataset ini telah dipisahkan sejak awal dan tidak pernah digunakan

dalam proses pelatihan maupun penyesuaian parameter. Pembahasan diawali dengan hasil dekomposisi VMD sebagai tahap pra-pemrosesan, dilanjutkan dengan analisis kinerja kedua model berdasarkan metrik agregat dan kondisi geomagnetik, serta ditutup dengan perbandingan terhadap penelitian terdahulu yang relevan.

A. Hasil Dekomposisi VMD

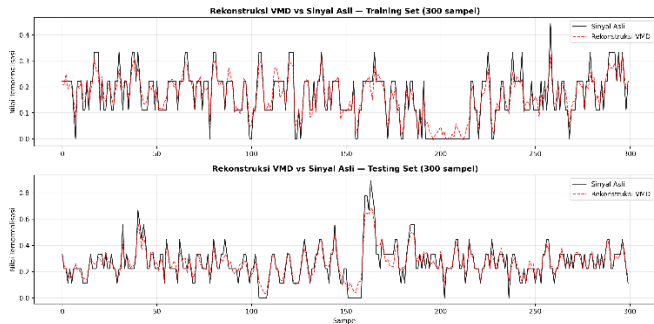
Proses dekomposisi VMD dengan parameter $K = 6$, $\alpha = 2.000$, dan $\tau = 1 \times 10^{-7}$ diterapkan secara terpisah pada masing-masing subset (*subset-wise decomposition*) setelah pembagian data secara kronologis. Pendekatan ini bertujuan mencegah kebocoran informasi temporal dari periode pengujian ke periode pelatihan yang dapat menghasilkan estimasi performa model yang terlalu optimistis [14]. Konsistensi frekuensi pusat pada setiap subset disajikan pada Tabel III.

TABEL III
FREKUENSI PUSAT TIAP IMF PER SUBSET

IMF	Training	Validation	Testing
IMF-1	0,000133	0,000051	0,000073
IMF-2	0,037761	0,036824	0,042854
IMF-3	0,124526	0,124915	0,124159
IMF-4	0,247804	0,251714	0,247232
IMF-5	0,371811	0,374089	0,327278
IMF-6	0,465140	0,478402	0,409396

Berdasarkan Tabel III, frekuensi pusat setiap komponen IMF menunjukkan pola hierarki yang konsisten pada ketiga subset, dimulai dari IMF-1 yang merepresentasikan komponen berfrekuensi paling rendah hingga IMF-6 yang merepresentasikan komponen berfrekuensi paling tinggi. Kemiripan nilai frekuensi pusat antar subset, terutama pada IMF-3 dan IMF-4 yang memiliki selisih kurang dari 0,005, menunjukkan bahwa strategi *subset-wise decomposition* mampu menghasilkan pemisahan frekuensi yang stabil tanpa memanfaatkan informasi dari subset lain. Sementara itu, perbedaan yang lebih besar pada IMF-5 dan IMF-6 antara *training set* dan *testing set* mencerminkan adanya variasi karakteristik aktivitas geomagnetik pada masing-masing periode pengamatan, meskipun pola hierarki frekuensi tetap konsisten di seluruh subset.

Untuk mengevaluasi kualitas dekomposisi secara keseluruhan, sinyal hasil rekonstruksi yang diperoleh dari superposisi enam komponen IMF dibandingkan dengan sinyal asli, sebagaimana ditunjukkan pada Gambar 4.

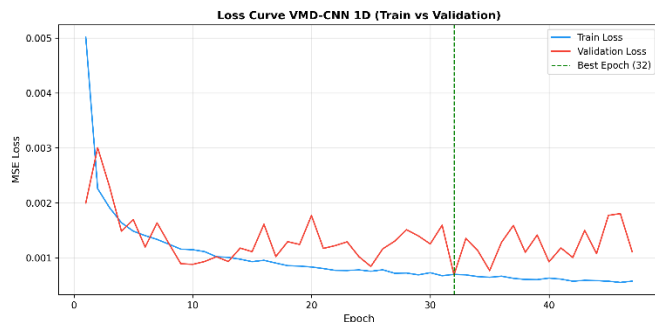


Gambar 4. Rekonstruksi VMD vs sinyal asli pada *training set* dan *testing set*

Gambar 4 menunjukkan tingkat kesesuaian yang tinggi antara sinyal asli (garis hitam) dan sinyal hasil rekonstruksi (garis merah putus-putus) pada 300 sampel pertama *training set* dan *testing set*, dengan nilai RMSE rekonstruksi sebesar 0,03676 pada *training set* untuk konfigurasi $K = 6$. Perbedaan yang terlihat terutama terjadi pada lonjakan nilai K yang bersifat diskontinu, sebagai konsekuensi dari karakteristik algoritma VMD yang lebih menitikberatkan pada pemisahan komponen frekuensi daripada ketepatan rekonstruksi pada titik-titik transisi yang tajam. Hasil rekonstruksi tersebut menunjukkan bahwa enam komponen IMF yang diperoleh telah mampu merepresentasikan sinyal dengan baik sehingga layak digunakan sebagai data masukan bagi kedua arsitektur *deep learning* pada tahap pemodelan selanjutnya.

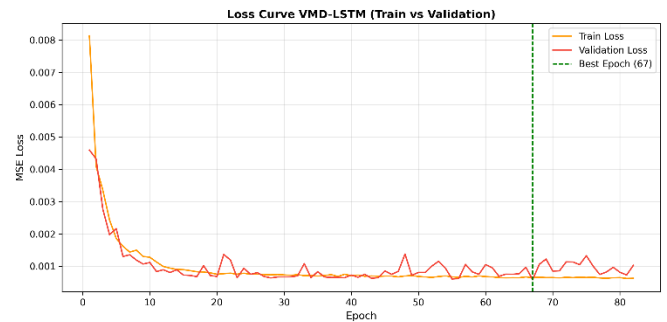
B. Kurva Pelatihan Model

Dinamika konvergensi selama proses pelatihan kedua model divisualisasikan melalui kurva *train loss* dan *validation loss*. Kurva pelatihan VMD-CNN 1D ditunjukkan pada Gambar 5.



Gambar 4. Kurva pelatihan VMD-CNN 1D

Berdasarkan Gambar 5, model VMD-CNN 1D yang memiliki 42.433 parameter menjalani pelatihan selama 47 *epoch* sebelum dihentikan oleh mekanisme *Early Stopping*. Nilai *validation loss* terendah sebesar 0,000692 dicapai pada *epoch* ke-32, dengan total waktu pelatihan 0,24 menit. Kurva *train loss* dan *validation loss* menunjukkan pola penurunan yang berdekatan dan hampir berimpit sepanjang proses pelatihan, sehingga mengindikasikan bahwa model tidak mengalami *overfitting* yang signifikan pada saat mencapai konvergensi optimal.



Gambar 5. Kurva pelatihan VMD-LSTM

Adapun Gambar 6 menunjukkan model VMD-LSTM yang memiliki 182.406 parameter menjalani pelatihan selama 82 *epoch* hingga mencapai konvergensi, dengan nilai *validation loss* terendah sebesar 0,000583 pada *epoch* ke-67 dan waktu pelatihan 1,68 menit. Kurva pelatihan menunjukkan penurunan yang tajam pada 15 *epoch* pertama, kemudian diikuti fase penurunan yang lebih bertahap hingga mencapai titik konvergensi optimal. Jumlah *epoch* yang lebih banyak serta waktu pelatihan yang sekitar tujuh kali lebih lama dibandingkan VMD-CNN 1D mencerminkan kompleksitas komputasi yang lebih tinggi pada arsitektur *multi-branch*, di mana enam cabang LSTM dioptimasi secara sekuensial menggunakan mekanisme *backpropagation through time*.

Perbedaan karakteristik konvergensi tersebut memberikan implikasi dalam pemilihan arsitektur. VMD-CNN 1D menawarkan efisiensi komputasi yang lebih baik untuk skenario yang memerlukan pelatihan ulang secara berkala, sedangkan VMD-LSTM membutuhkan sumber daya komputasi yang lebih besar, meskipun durasi pelatihannya masih tergolong efisien untuk implementasi operasional.

C. Evaluasi Agregat

Perlu diperhatikan bahwa *testing set* pada penelitian ini memiliki tingkat kesulitan yang lebih tinggi dibandingkan *training set*, ditunjukkan oleh proporsi kejadian badai geomagnetik sebesar 3,98% pada *testing set* dibandingkan 1,31% pada *training set*. Dengan kondisi tersebut, hasil evaluasi yang diperoleh mencerminkan kemampuan generalisasi model dalam menghadapi data dengan karakteristik yang lebih menantang. Rekapitulasi metrik evaluasi agregat kedua model pada 1.598 sampel *testing set* disajikan pada Tabel IV.

TABEL IV
KOMPARASI METRIK EVALUASI AGREGAT

Metrik	VMD-CNN 1D	VMD-LSTM	Persistence
RMSE	0,3608	0,3889	0,9643
MAE	0,2652	0,3068	0,6671
R ²	0,8846	0,8660	0,1757
Mean Error	0,0005	-0,0271	-
Min Error	-1,2280	-1,4596	-

<i>Max Error</i>	1,5207	1,4948	-
<i>p-value</i> (<i>Wilcoxon</i> , <i>CNN 1D</i> <i>vs LSTM</i>)	< 0,001	< 0,001	-

Berdasarkan Tabel IV, VMD-CNN 1D memberikan performa yang lebih baik daripada VMD-LSTM pada seluruh metrik utama, dengan RMSE lebih rendah sebesar 0,0281, MAE lebih rendah sebesar 0,0416, dan nilai R^2 lebih tinggi sebesar 0,0186. Nilai R^2 kedua model sama-sama berada pada kategori baik karena melebihi 0,85, yang menunjukkan bahwa keduanya mampu menjelaskan lebih dari 86% variasi nilai Indeks K selama periode pengujian. Selisih MAE yang lebih besar dibandingkan selisih RMSE mengindikasikan bahwa keunggulan VMD-CNN 1D tidak hanya berasal dari berkurangnya beberapa kesalahan prediksi ekstrem, tetapi juga terjadi secara konsisten pada sebagian besar sampel pengujian. Uji Wilcoxon *signed-rank* terhadap selisih *error* absolut antara kedua model menghasilkan nilai $p < 0,001$, yang mengonfirmasi bahwa keunggulan VMD-CNN 1D secara agregat bersifat signifikan secara statistik.

Ditinjau dari distribusi *error*, kedua model menunjukkan karakteristik yang berbeda. VMD-CNN 1D memiliki *mean error* yang sangat mendekati nol (0,0005), sehingga menunjukkan keseimbangan antara *overestimation* dan *underestimation*. Sebaliknya, VMD-LSTM menghasilkan *mean error* negatif (-0,0271), yang menunjukkan kecenderungan melakukan *underestimation*. Selain itu, rentang *error* VMD-LSTM pada sisi negatif lebih besar dengan *Min Error* sebesar -1,4596, sedangkan VMD-CNN 1D memiliki *Min Error* sebesar -1,2280. Pada sisi positif, kedua model menunjukkan *Max Error* yang relatif serupa, yaitu 1,5207 pada VMD-CNN 1D dan 1,4948 pada VMD-LSTM.

Sebagai evaluasi tambahan, kedua model hibrida dibandingkan dengan model *Persistence* sebagai *baseline* sederhana yang mengasumsikan nilai Indeks K pada *timestep* berikutnya sama dengan nilai pada *timestep* saat ini tanpa melalui proses pembelajaran. Berdasarkan Tabel IV, model *Persistence* menghasilkan RMSE sebesar 0,9643 dan R^2 sebesar 0,1757, yang menunjukkan performa jauh di bawah kedua model hibrida. Dibandingkan dengan *Persistence*, VMD-CNN 1D dan VMD-LSTM masing-masing menurunkan RMSE sebesar 62,6% dan 59,7%. Selain itu, uji Wilcoxon *signed-rank* terhadap selisih *absolute error* antara VMD-CNN 1D dan *Persistence* menghasilkan nilai $p < 0,001$, yang menunjukkan bahwa peningkatan performa tersebut signifikan secara statistik dan tidak terjadi akibat variasi acak. Hasil ini mengindikasikan bahwa integrasi dekomposisi VMD dengan arsitektur *deep learning* memberikan peningkatan kemampuan prediksi yang nyata dibandingkan pendekatan *naive* yang hanya mengandalkan nilai observasi sebelumnya.

Meskipun hasil evaluasi secara agregat menunjukkan keunggulan VMD-CNN 1D, metrik tersebut belum sepenuhnya menggambarkan perilaku kedua model pada setiap tingkat aktivitas geomagnetik yang memiliki distribusi kelas sangat tidak seimbang [3]. Oleh karena itu, diperlukan analisis lebih lanjut terhadap performa kedua model pada setiap kelas kondisi geomagnetik untuk mengungkap karakteristik yang tidak tercermin pada metrik agregat.

Untuk menguji secara spesifik apakah peningkatan performa berasal dari tahap dekomposisi VMD dan bukan semata dari kapasitas arsitektur *deep learning*, dilakukan studi ablasi dengan melatih ulang CNN 1D dan LSTM menggunakan sinyal Indeks K mentah tanpa dekomposisi VMD, dengan konfigurasi pelatihan yang identik. Hasil studi ablasi disajikan pada Tabel V.

TABEL V
HASIL ABLATION STUDY: KONTRIBUSI VMD TERHADAP PERFORMA MODEL

Model	RMSE	MAE	R^2	Parameter
CNN 1D tanpa VMD	0,8092	0,6197	0,4193	41.473
LSTM tanpa VMD	0,8175	0,6265	0,4074	30.401

Berdasarkan Tabel V, CNN 1D tanpa dekomposisi VMD memperoleh RMSE sebesar 0,8092 dan R^2 sebesar 0,4193, jauh lebih rendah dibandingkan VMD-CNN 1D pada Tabel IV yang menghasilkan RMSE 0,3608 dan R^2 0,8846. Penghapusan tahap dekomposisi VMD menyebabkan RMSE meningkat sebesar 124,3%. Pola yang sama juga terlihat pada arsitektur LSTM, di mana RMSE meningkat dari 0,3889 pada VMD-LSTM (Tabel IV) menjadi 0,8175 tanpa VMD, sementara nilai R^2 menurun dari 0,8660 menjadi 0,4074. Pada arsitektur CNN, jumlah parameter model tanpa VMD (41.473) relatif sebanding dengan VMD-CNN 1D (42.433; Tabel I), sehingga penurunan performa yang signifikan tersebut tidak dapat dijelaskan hanya oleh perbedaan kapasitas model. Temuan ini menunjukkan bahwa dekomposisi VMD memberikan kontribusi yang penting terhadap peningkatan kemampuan prediksi. Sementara itu, pada arsitektur LSTM, meskipun VMD-LSTM memiliki jumlah parameter yang jauh lebih besar (182.406 dibandingkan 30.401) sebagai konsekuensi penggunaan enam cabang paralel, peningkatan performa yang tetap besar mengindikasikan bahwa keunggulan model tidak semata-mata berasal dari bertambahnya jumlah parameter, tetapi juga dari representasi frekuensi yang dihasilkan melalui proses dekomposisi VMD.

D. Analisis Per-Kelas Kondisi Geomagnetik

Untuk mengidentifikasi karakteristik performa yang belum tergambarkan oleh metrik agregat, dilakukan evaluasi lanjutan dengan mengelompokkan *testing set* ke dalam tiga kategori kondisi geomagnetik berdasarkan nilai Indeks K, yaitu tenang ($K = 0-2$), sedang ($K = 3-4$), dan badai ($K \geq 5$). Selisih performa antar model pada setiap kategori turut diuji

signifikansinya menggunakan uji *Wilcoxon signed-rank*. Hasil evaluasi per kelas disajikan pada Tabel VI.

TABEL VI
KOMPARASI PERFORMA PER-KELAS KONDISI GEOMAGNETIK

Metrik	Tenang	Sedang	Badai
RMSE - CNN	0,3655	0,3480	0,4530
RMSE - LSTM	0,3814	0,3955	0,3889
MAE - CNN	0,2575	0,2651	0,3548
MAE - LSTM	0,3006	0,3132	0,2967
R ² - CNN	0,5107	0,3412	0,5430
R ² - LSTM	0,4673	0,1494	0,6632
<i>p</i> -value (Wilcoxon)	< 0,001	< 0,001	0,024

Catatan: Tenang ($K=0-2$, $N=727$), Sedang ($K=3-4$, $N=807$), Badai ($K \geq 5$, $N=64$)

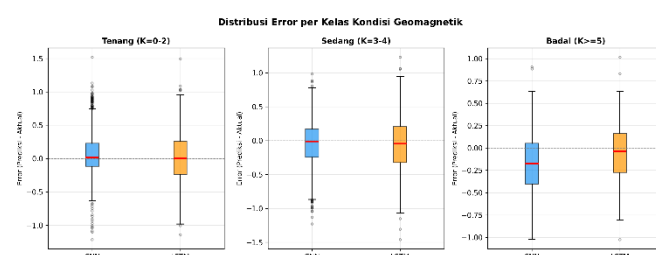
Berdasarkan Tabel VI, terlihat adanya pembalikan pola keunggulan yang tidak tampak pada evaluasi agregat. VMD-CNN 1D menunjukkan performa yang lebih baik pada kelas tenang dan kelas sedang, yang masing-masing mencakup 45,2% dan 50,2% dari *testing set*. Sebaliknya, VMD-LSTM memberikan hasil yang lebih baik pada kelas badai, meskipun kategori ini hanya mencakup sekitar 4,0% dari total sampel pengujian.

Pada kelas tenang, VMD-CNN 1D menghasilkan RMSE sebesar 0,3655 dan MAE sebesar 0,2575, lebih rendah dibandingkan VMD-LSTM yang memperoleh RMSE 0,3814 dan MAE 0,3006. Nilai R² kedua model pada kategori ini relatif rendah (VMD-CNN 1D sebesar 0,5107 dan VMD-LSTM sebesar 0,4673), yang dapat dipahami karena rentang variasi nilai K pada kondisi tenang sangat sempit sehingga R² menjadi kurang sensitif sebagai indikator. Pada kelas sedang, keunggulan VMD-CNN 1D semakin terlihat dengan RMSE sebesar 0,3480 dan R² sebesar 0,3412, sedangkan VMD-LSTM hanya mencapai R² sebesar 0,1494, yang menunjukkan kemampuan yang lebih terbatas dalam menjelaskan variasi data pada kategori tersebut.

Keunggulan VMD-CNN 1D pada kedua kategori dominan tersebut dapat dipahami dari kesesuaian antara karakteristik operasi konvolusi dan pola data geomagnetik pada kondisi normal. Kondisi tenang dan sedang, yang mencakup sekitar 96% *testing set*, memiliki pola temporal yang cenderung teratur dan berulang setiap 24 jam mengikuti rotasi harian Bumi terhadap Matahari. Karakteristik tersebut lebih mudah dipelajari oleh operasi konvolusi yang bekerja secara lokal pada jendela *timestep* terbatas, terutama melalui pemrosesan *multi-channel* yang mampu menangkap korelasi antar komponen frekuensi secara simultan.

Sebaliknya, pada kelas badai, VMD-LSTM memberikan performa yang lebih baik dengan RMSE sebesar 0,3889, MAE sebesar 0,2967, dan R² sebesar 0,6632, sedangkan VMD-CNN 1D memperoleh RMSE sebesar 0,4530, MAE sebesar 0,3548, dan R² sebesar 0,5430. Perbedaan RMSE sebesar 0,0641 dan R² sebesar 0,1202 menunjukkan

keunggulan VMD-LSTM yang cukup jelas pada kondisi gangguan geomagnetik. Meskipun kelas badai hanya memiliki 64 sampel, uji *Wilcoxon signed-rank* terhadap selisih *error* absolut kedua model pada kategori ini menghasilkan nilai $p = 0,024$ (Tabel VI), yang mengonfirmasi bahwa selisih performa tersebut signifikan secara statistik dan bukan merupakan artefak dari ukuran sampel yang kecil. Temuan ini sejalan dengan karakteristik LSTM yang dirancang untuk mempelajari dependensi temporal jangka panjang [5], sehingga pola perubahan yang mendahului terjadinya badai dapat dimodelkan lebih baik dibandingkan pendekatan konvolusi lokal pada VMD-CNN 1D. Untuk memberikan gambaran visual mengenai distribusi *error* kedua model pada ketiga kategori tersebut, *boxplot error* per kelas disajikan pada Gambar 7.



Gambar 6. Box plot distribusi *error* per-kelas kondisi geomagnetik

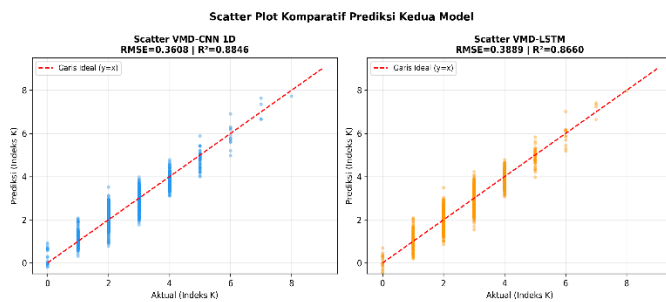
Berdasarkan Gambar 7, VMD-CNN 1D (kotak biru) memiliki rentang interkuartil yang lebih sempit dibandingkan VMD-LSTM (kotak oranye) pada kelas tenang dan sedang, konsisten dengan keunggulan VMD-CNN 1D pada kedua kategori tersebut. Namun, pada kelas badai, meskipun rentang interkuartil kedua model relatif sama, sebaran *error* VMD-CNN 1D menunjukkan *whisker* yang lebih panjang serta beberapa *outlier* ekstrem yang lebih jauh dari *median*, sedangkan VMD-LSTM mempertahankan sebaran *error* yang lebih terkendali. Pola ini mendukung hasil pada Tabel VI bahwa VMD-LSTM memberikan performa yang lebih stabil pada kondisi gangguan geomagnetik, meskipun secara agregat masih berada di bawah VMD-CNN 1D.

Pembalikan keunggulan pada kelas badai memberikan implikasi yang penting bagi penerapan operasional. Dalam sistem peringatan dini badai geomagnetik, akurasi prediksi pada kondisi $K \geq 5$ memiliki nilai praktis yang lebih tinggi dibandingkan pada kondisi tenang, karena pada kondisi tersebut potensi gangguan terhadap infrastruktur satelit, GNSS, dan jaringan listrik menjadi lebih besar [1]. Meskipun demikian, hasil pada kelas badai perlu diinterpretasikan secara hati-hati mengingat jumlah sampelnya hanya 64, sehingga estimasi metrik pada kategori ini berpotensi memiliki variasi yang lebih tinggi dibandingkan kategori dengan jumlah sampel yang lebih besar.

E. Analisis Visual Prediksi

Untuk memberikan gambaran menyeluruh mengenai tingkat akurasi prediksi kedua model terhadap nilai observasi

aktual pada setiap titik data, hasil prediksi VMD-CNN 1D dan VMD-LSTM divisualisasikan dalam bentuk *scatter plot* komparatif pada Gambar 8.



Gambar 7. *Scatter plot* komparatif prediksi kedua model terhadap nilai aktual Indeks K

Berdasarkan Gambar 8, sumbu horizontal menunjukkan nilai Indeks K aktual, sedangkan sumbu vertikal menunjukkan nilai prediksi model. Garis ideal (garis merah putus-putus) merepresentasikan kondisi ketika nilai prediksi sama dengan nilai aktual. Oleh karena itu, semakin dekat sebaran titik terhadap garis tersebut, semakin tinggi tingkat akurasi prediksi model. *Scatter plot* disajikan dalam dua panel terpisah agar sebaran masing-masing model dapat dievaluasi tanpa saling menutupi, dengan VMD-CNN 1D pada panel kiri dan VMD-LSTM pada panel kanan.

Pada panel kiri, VMD-CNN 1D menampilkan sebaran titik prediksi (biru) yang terkonsentrasi di sekitar garis ideal pada rentang $K = 0$ hingga $K = 4$, yang menunjukkan kemampuan prediksi yang baik pada kondisi geomagnetik tenang dan sedang. Namun, pada $K \geq 5$ sebaran titik mulai melebar, dengan beberapa titik berada cukup jauh di bawah garis ideal yang menunjukkan kecenderungan *underestimation* pada nilai puncak. Pola ini terlihat, misalnya, pada beberapa sampel dengan nilai aktual $K = 6$ yang diprediksi mendekati $K = 5$, serta pada satu sampel dengan nilai aktual $K = 8$ yang diprediksi sekitar $K = 7$.

Pada panel kanan, VMD-LSTM memperlihatkan pola sebaran titik (oranye) yang secara umum serupa pada rentang $K = 0$ hingga $K = 4$, meskipun dengan variasi yang sedikit lebih besar dibandingkan VMD-CNN 1D pada kelas sedang. Sebaliknya, pada $K \geq 5$ sebaran titik VMD-LSTM tampak lebih rapat dan lebih dekat dengan garis ideal, terutama pada rentang $K = 5$ hingga $K = 7$. Pola ini sejalan dengan hasil pada Tabel VI yang menunjukkan keunggulan VMD-LSTM pada kelas badai ($K \geq 5$), sekaligus memperkuat bahwa keunggulan tersebut tidak hanya tercermin pada metrik evaluasi, tetapi juga pada sebaran prediksi masing-masing sampel pada kondisi gangguan geomagnetik.

F. Kemampuan Deteksi Kondisi Badai Geomagnetik

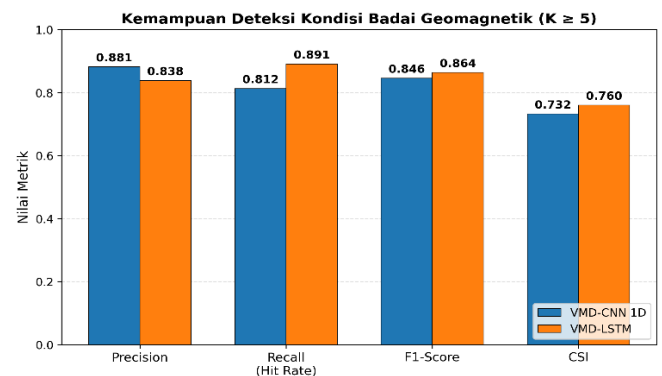
Selain evaluasi berbasis RMSE, MAE, dan R^2 , kemampuan kedua model dalam mendeteksi kejadian badai geomagnetik ($K \geq 5$) turut dievaluasi menggunakan metrik yang relevan secara operasional untuk sistem peringatan dini, yaitu

Precision, *Recall*, *F1-Score*, dan *Critical Success Index* (CSI). Hasil evaluasi disajikan pada Tabel VII.

TABEL VII
METRIK DETEKSI BADAI GEOMAGNETIK ($K \geq 5$)

Metrik	VMD-CNN 1D	VMD-LSTM
<i>Precision</i>	0,881	0,838
<i>Recall (Hit Rate)</i>	0,812	0,891
<i>F1-Score</i>	0,846	0,864
CSI	0,732	0,760

Berdasarkan Tabel VII, VMD-CNN 1D memperoleh nilai *Precision* yang lebih tinggi, yaitu 0,881 dibandingkan 0,838 pada VMD-LSTM, yang menunjukkan bahwa proporsi prediksi badai yang benar lebih besar pada model tersebut. Sebaliknya, VMD-LSTM menunjukkan kinerja yang lebih baik pada metrik *Recall*, *F1-Score*, dan *Critical Success Index* (CSI), masing-masing sebesar 0,891, 0,864, dan 0,760, dibandingkan VMD-CNN 1D yang memperoleh nilai 0,812, 0,846, dan 0,732. Nilai *Recall* yang lebih tinggi mengindikasikan bahwa VMD-LSTM mampu mengidentifikasi kejadian badai geomagnetik secara lebih menyeluruh, sehingga jumlah kejadian badai yang tidak terdeteksi (*false negative*) lebih sedikit dibandingkan VMD-CNN 1D. Untuk memberikan gambaran yang lebih jelas mengenai perbandingan keempat metrik tersebut, hasil evaluasi divisualisasikan pada Gambar 9.



Gambar 8. Kemampuan deteksi kondisi badai geomagnetik ($K \geq 5$)

Berdasarkan Gambar 9, terlihat adanya *trade-off* yang jelas antara kedua model. VMD-CNN 1D menunjukkan nilai *Precision* yang lebih tinggi, sedangkan VMD-LSTM memperoleh nilai yang lebih baik pada *Recall*, *F1-Score*, dan *Critical Success Index* (CSI). Dalam konteks sistem peringatan dini geomagnetik, kemampuan mendeteksi sebanyak mungkin kejadian badai aktual umumnya lebih penting secara operasional dibandingkan sekadar mengurangi jumlah peringatan palsu, karena dampak dari kejadian badai yang tidak terdeteksi (*false negative*) berpotensi lebih besar daripada konsekuensi yang ditimbulkan oleh peringatan yang keliru (*false positive*). Temuan ini semakin menguatkan peran VMD-LSTM sebagai model pelengkap yang dapat

dimanfaatkan ketika teridentifikasi indikasi peningkatan aktivitas geomagnetik.

G. Kontekstualisasi Terhadap Penelitian Terdahulu

Sebagaimana telah dikemukakan pada bagian pendahuluan, penelitian yang paling relevan dengan studi ini adalah Saputra et al. [4], yang mengembangkan model prediksi Indeks K geomagnetik lokal di Observatorium BMKG Tuntungan menggunakan pendekatan *multivariat* dengan memanfaatkan data historis Indeks K, parameter angin surya dari satelit eksternal, serta mekanisme *multi-head attention*. Berbeda dengan penelitian tersebut, studi ini mengadopsi paradigma *univariat* yang hanya memanfaatkan data historis Indeks K lokal melalui dekomposisi Variational Mode Decomposition (VMD), tanpa melibatkan variabel eksternal sebagai masukan model. Perbedaan karakteristik kedua pendekatan tersebut dirangkum pada Tabel VIII.

TABEL VIII
KONTEKSTUALISASI PENDEKATAN TERHADAP PENELITIAN TERDAHULU

Aspek	Saputra et al. [4]	Penelitian ini
Pendekatan	Multivariat	Univariat
Data masukan	Indeks K + angin surya satelit	Indeks K lokal
Metode	LSTM + <i>multi-head attention</i>	VMD + CNN 1D / VMD + LSTM
Kebutuhan data eksternal	Ya	Tidak
Orientasi tujuan	Akurasi maksimal dengan informasi eksternal	Kemandirian operasional berbasis data lokal

Perbandingan pada Tabel VIII tidak dimaksudkan sebagai evaluasi langsung terhadap performa kedua penelitian, karena konfigurasi data, variabel masukan, dan paradigma pemodelan yang digunakan tidak sepenuhnya identik. Sebaliknya, tabel tersebut digunakan untuk menunjukkan posisi metodologis penelitian ini terhadap penelitian terdahulu yang paling relevan. Pendekatan *multivariat* pada Saputra et al. memanfaatkan informasi tambahan dari data satelit untuk memperkaya representasi masukan model, sedangkan penelitian ini mengeksplorasi sejauh mana informasi temporal yang terkandung pada deret waktu Indeks K lokal mampu dimanfaatkan secara optimal tanpa bantuan variabel eksternal. Dengan demikian, kedua penelitian menjawab pertanyaan penelitian yang berbeda dan sebaiknya dipandang sebagai dua pendekatan yang saling melengkapi, bukan sebagai metode yang saling menggantikan.

Hasil penelitian ini menunjukkan bahwa pendekatan univariat berbasis VMD mampu menghasilkan performa yang baik meskipun hanya menggunakan data historis Indeks K sebagai masukan. Temuan tersebut mengindikasikan bahwa

proses dekomposisi sinyal adaptif melalui VMD dapat mengekstraksi informasi temporal yang cukup representatif dari sinyal Indeks K lokal. Dari sisi operasional, pendekatan ini berpotensi menjadi alternatif yang lebih mandiri pada lingkungan yang ketersediaan data angin surya satelit secara real-time tidak selalu terjamin, sehingga tetap mendukung pengembangan sistem prediksi dan peringatan dini geomagnetik berbasis data lokal.

IV. KESIMPULAN

Penelitian ini mengembangkan dan membandingkan dua arsitektur hibrida univariat, yaitu VMD-CNN 1D Multi-channel dan VMD-LSTM, untuk prediksi Indeks K geomagnetik lokal menggunakan data Observatorium BMKG Tuntungan periode Januari 2020 hingga Juni 2025. Penerapan *subset-wise decomposition* melalui *Variational Mode Decomposition* (VMD) dengan enam komponen *Intrinsic Mode Function* (IMF) menghasilkan representasi frekuensi yang stabil tanpa menimbulkan *lookahead bias* pada data pengujian. Hasil *ablation study* menunjukkan bahwa dekomposisi VMD berkontribusi signifikan terhadap peningkatan performa, ditunjukkan oleh penurunan RMSE sebesar 55,4% pada CNN 1D dan 52,4% pada LSTM dibandingkan model tanpa VMD dengan kapasitas parameter yang sebanding. Evaluasi menunjukkan bahwa VMD-CNN 1D memperoleh kinerja terbaik pada metrik agregat dengan RMSE 0,3608, MAE 0,2652, dan R^2 0,8846, serta secara signifikan mengungguli model baseline *Persistence* (uji *Wilcoxon signed-rank*, $p < 0,001$). Sebaliknya, VMD-LSTM menunjukkan performa yang lebih baik pada kondisi badai geomagnetik dan berbeda secara signifikan ($p = 0,024$), dengan R^2 sebesar 0,6632 dan *storm-detection recall* sebesar 0,891, dibandingkan VMD-CNN 1D yang memperoleh R^2 sebesar 0,5430 dan recall sebesar 0,812. Dari sepuluh aspek komparasi yang dievaluasi, VMD-CNN 1D unggul pada tujuh aspek, sedangkan VMD-LSTM unggul pada tiga aspek yang berkaitan dengan kondisi badai geomagnetik. Secara keseluruhan, hasil penelitian menunjukkan bahwa pendekatan *univariat* berbasis VMD mampu menghasilkan prediksi Indeks K yang akurat, andal secara statistik, dan tidak bergantung pada data angin surya dari satelit eksternal.

Dari perspektif operasional, karakteristik kedua model membuka peluang penerapan skema komplementer untuk mendukung sistem pemantauan geomagnetik BMKG. VMD-CNN 1D, yang memiliki efisiensi komputasi lebih tinggi, berpotensi digunakan sebagai model utama untuk menghasilkan prediksi rutin pada kondisi operasional normal. Sementara itu, ketika nilai Indeks K yang teramati mendekati atau melampaui ambang $K=3$ sebagai indikasi awal peningkatan aktivitas geomagnetik, VMD-LSTM dapat dipertimbangkan sebagai model pendukung karena menunjukkan kemampuan yang lebih baik dalam mendeteksi kejadian badai (Tabel VII). Pendekatan ini berpotensi diwujudkan melalui mekanisme peralihan berbasis ambang (*threshold-based switching*) yang bekerja sebagai lapisan

logika di atas kedua model tanpa memerlukan pelatihan ulang secara *real-time*. Meskipun demikian, implementasi dalam lingkungan operasional BMKG masih memerlukan validasi lebih lanjut, termasuk pengujian menggunakan data *real-time* serta integrasi dengan mekanisme diseminasi peringatan yang berlaku, sehingga aspek tersebut berada di luar cakupan penelitian ini.

Penelitian ini masih memiliki beberapa keterbatasan. Distribusi kelas yang tidak seimbang, dengan hanya 64 sampel badai pada *testing set*, menyebabkan estimasi metrik pada kelas tersebut berpotensi memiliki variasi yang lebih tinggi. Selain itu, penelitian ini belum menerapkan teknik penanganan ketidakseimbangan data, seperti *oversampling* maupun *weighted loss function*. Penggunaan data yang hanya berasal dari satu observatorium di wilayah ekuator juga membatasi validasi kemampuan generalisasi model pada wilayah dengan karakteristik geomagnetik yang berbeda, seperti lintang menengah dan tinggi. Di samping itu, penelitian ini hanya mengevaluasi prediksi satu langkah ke depan dengan interval waktu 3 jam, sehingga kinerja model pada skenario prediksi multi-langkah belum dikaji.

Penelitian lanjutan direkomendasikan mengarah pada tiga aspek pengembangan yang saling terkait. Pertama, menerapkan teknik penanganan *class imbalance* untuk meningkatkan kemampuan model dalam mendeteksi kejadian badai geomagnetik. Kedua, memanfaatkan data dari beberapa observatorium dengan variasi lintang geografis untuk menguji sekaligus meningkatkan kemampuan generalisasi model. Ketiga, mengembangkan skema prediksi multi-langkah dengan horizon 6 hingga 24 jam ke depan agar lebih sesuai dengan kebutuhan sistem peringatan dini geomagnetik operasional di Indonesia.

Selain ketiga arah pengembangan tersebut, temuan mengenai keunggulan yang saling melengkapi dari kedua arsitektur membuka peluang untuk pengembangan sistem *ensemble*. Dalam skema ini, VMD-CNN 1D dapat digunakan sebagai model utama untuk pemantauan geomagnetik rutin, sedangkan VMD-LSTM berperan sebagai model pendukung yang diaktifkan ketika terdeteksi indikasi awal peningkatan aktivitas menuju kondisi badai.

DAFTAR PUSTAKA

- [1] P. Thaker, A. Desai, P. Panchal, D. Sutaria, M. Shah, and M. Prajapati, "Evaluating AI approaches for space weather prediction: Strengthening satellite resilience and technology systems," *Sp. Habitat.*, vol. 1, no. 3, p. 100032, 2025, doi: 10.1016/j.spaceh.2025.100032.
- [2] J. Matzka, C. Stolle, Y. Yamazaki, O. Bronkalla, and A. Morschhauser, "The Geomagnetic Kp Index and Derived Indices of Geomagnetic Activity," *Sp. Weather*, vol. 19, no. 5, pp. 1–21, 2021, doi: 10.1029/2020SW002641.
- [3] Y. Y. Shprits, R. Vasile, and I. S. Zhelavskaya, "Nowcasting and Predicting the Kp Index Using Historical Values and Real-Time Observations," *Sp. Weather*, vol. 17, no. 8, pp. 1219–1229, 2019, doi: 10.1029/2018SW002141.
- [4] A. R. Saputra, A. Saikhu, and B. Amaliah, "Forecasting the Geomagnetic K Index with LSTM Networks: Insights from Indonesian Data," in *2025 15th International Conference on Information & Communication Technology and System (ICTS)*, Denpasar, Indonesia: IEEE, Nov. 2025, pp. 1–6, doi: 10.1109/ICTS67612.2025.11369590.
- [5] W. Yu, H. Li, M. Liu, and L. Song, "A geomagnetic variation prediction method based on chaotic VMD-LSTM neural network," *Acta Seismol. Sin.*, vol. 46, no. 1, pp. 92–105, 2024, doi: 10.11939/jass.20220132.
- [6] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.
- [7] Y. Xie *et al.*, "Enhancing GNSS Deformation Monitoring Forecasting with a Combined VMD-CNN-LSTM Deep Learning Model," *Remote Sens.*, vol. 16, no. 10, pp. 1–21, 2024, doi: 10.3390/rs16101767.
- [8] S. Guessoum *et al.*, "The Short-Term Prediction of Length of Day Using 1D Convolutional Neural Networks (1D CNN)," *Sensors*, vol. 22, no. 23, pp. 1–17, 2022, doi: 10.3390/s22239517.
- [9] H. Han *et al.*, "Hybrid Climate Forecasting: Variational Mode Decomposition and Convolutional Neural Network with Long-Term Short Memory," *Polish J. Environ. Stud.*, vol. 33, no. 2, pp. 1121–1134, 2024, doi: 10.15244/pjoes/172756.
- [10] G. Bernoux, A. Brunet, É. Buchlin, M. Janvier, and A. Sicard, "Forecasting the Geomagnetic Activity Several Days in Advance Using Neural Networks Driven by Solar EUV Imaging," *J. Geophys. Res. Sp. Phys.*, vol. 127, no. 10, pp. 1–22, 2022, doi: 10.1029/2022JA030868.
- [11] G. Kervalishvili, I. Michaelis, M. Korte, J. Rauberg, and J. Matzka, "A Novel Model for Forecasting Geomagnetic Indices Using Machine Learning," *Geophys. Res. Lett.*, vol. 52, no. 8, 2025, doi: 10.1029/2025GL114848.
- [12] D. H. Kim, D. J. Kim, and S. Y. Choi, "A Variational-Mode-Decomposition-Cascaded Long Short-Term Memory with Attention Model for VIX Prediction," *Appl. Sci.*, vol. 15, no. 10, pp. 1–28, 2025, doi: 10.3390/app15105630.
- [13] H. Sun, Z. Yu, and B. Zhang, "Research on short-term power load forecasting based on VMD and GRU," *PLoS One*, vol. 19, no. 7 JULY, pp. 1–21, 2024, doi: 10.1371/journal.pone.0306566.
- [14] X. Yang, J. Li, and X. Jiang, "Research on information leakage in time series prediction based on empirical mode decomposition," *Sci. Rep.*, vol. 14, no. 1, pp. 1–18, 2024, doi: 10.1038/s41598-024-80018-9.
- [15] L. N. Boucetta, Y. Amrane, A. Chouder, S. Arezki, and S. Kichou, "Enhanced Forecasting Accuracy of a Grid-Connected Photovoltaic Power Plant: A Novel Approach Using Hybrid Variational Mode Decomposition and a CNN-LSTM Model," *Energies*, vol. 17, no. 7, 2024, doi: 10.3390/en17071781.
- [16] R. Feng *et al.*, "A Hybrid VMD-SSA-LSTM Framework for Short-Term Wind Speed Prediction Based on Wind Farm Measurement Data," *Energies*, vol. 19, no. 3, pp. 1–20, 2026, doi: 10.3390/en19030520.