

Prediksi Jarak Pandang Menggunakan *Long Short-Term Memory* pada Data *Spatio-Temporal Multivariat*

<http://dx.doi.org/10.28932/jutisi.v12i2.13421>

Riwayat Artikel

Received: 29 September 2025 | Final Revision: 15 April 2026 | Accepted: 18 April 2026

Creative Commons License 4.0 (CC BY – NC)



Anjang Qismatul Islam^{✉*1}, Alfani Alfarisy^{#2}, Shanas Septy Prayuda^{#3}, Khalid^{*4}, Dwi Rolliawati^{*5}

^{*}Sistem Informasi, UIN Sunan Ampel Surabaya, Jl. Dr. Ir. H. Soekarno No.682, Gn. Anyar, Gn. Anyar, Surabaya, 60294, Indonesia

¹anjangqismatulislam@gmail.com

⁴khalid@uinsa.ac.id

⁵dwi_roll@uinsby.ac.id

[#]Stasiun Meteorologi Kelas I Juanda, Jl. Raya Bandara Juanda, Pranti, Sedati, Sidoarjo, 61253, Indonesia

²alfanalfarisy@gmail.com

³shanas.prayuda@bmk.go.id

✉Corresponding author: anjangqismatulislam@gmail.com

Abstrak — Prediksi jarak pandang di bandara merupakan faktor krusial dalam menjaga keselamatan dan efisiensi operasional penerbangan. Penelitian ini bertujuan mengembangkan model prediksi jarak pandang di Bandara Internasional Juanda menggunakan data *Automatic Weather Observation System* (AWOS) multivariat berbasis *Long Short-Term Memory* (LSTM) dengan pendekatan spatio-temporal. Data berupa observasi cuaca per menit sepanjang tahun 2022 dari tiga titik pengamatan, yaitu titik 10, titik 10–28, dan titik 28. Proses penelitian meliputi pra-pemrosesan data, imputasi nilai hilang dengan interpolasi kontinu dan vektor, reduksi dimensi menggunakan *Principal Component Analysis* (PCA), pembentukan *sequence* sepanjang 168 langkah waktu, serta pelatihan model LSTM dua lapisan dengan regularisasi dropout. Evaluasi model dilakukan menggunakan metrik *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Hasil menunjukkan bahwa model menghasilkan prediksi dengan tingkat kesalahan yang relatif rendah, dengan nilai MAE sebesar 0.0617 (validasi) dan 0.0618 (uji), MSE sebesar 0.0170 (validasi) dan 0.0150 (uji), serta RMSE sebesar 0.1305 (validasi) dan 0.1225 (uji). Nilai MAPE sebesar 27,58% (validasi) dan 14,24% (uji) menunjukkan variasi kesalahan relatif terhadap nilai aktual. Perbedaan kecil antara data validasi dan uji mengindikasikan kemampuan generalisasi model yang baik tanpa indikasi *overfitting* yang signifikan, sementara penggunaan PCA terbukti meningkatkan kinerja dengan menyoroti fitur utama yang paling relevan. Kinerja ini didukung oleh penggunaan data resolusi tinggi, pendekatan multivariat, serta penerapan PCA yang meningkatkan efisiensi pembelajaran. Secara keseluruhan, pendekatan LSTM multivariat berbasis AWOS mampu memodelkan pola spatio-temporal pada data jarak pandang secara efektif dan berpotensi untuk dikembangkan lebih lanjut dalam studi prediksi meteorologi.

Kata kunci— AWOS; *Deep learning*; Jarak Pandang; LSTM; PCA.

Visibility Prediction Using Long Short-Term Memory on Multivariate Spatio-Temporal Data

Abstract — *Visibility prediction at airports is a crucial factor in ensuring flight safety and operational efficiency. This study aims to develop a visibility prediction model at Juanda International Airport using multivariate Automatic Weather Observation System (AWOS) data based on Long Short-Term Memory (LSTM) with a spatio-temporal approach. The dataset consists of minute-level weather observations throughout 2022 from three observation points, namely point 10, point 10–28, and point 28. The research process includes data preprocessing, missing value imputation using continuous and vector interpolation methods, dimensionality reduction using Principal Component Analysis (PCA), sequence construction with 168 time steps, and training a two-layer LSTM model with dropout regularization. Model performance is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). The results indicate that the model achieves relatively low prediction errors, with MAE values of 0.0617 (validation) and 0.0618 (test), MSE values of 0.0170 (validation) and 0.0150 (test), and RMSE values of 0.1305 (validation) and 0.1225 (test). The MAPE values of 27.58% (validation) and 14.24% (test) reflect the relative error variability compared to actual values. The small difference between validation and test results indicates good generalization capability without significant overfitting. Additionally, the use of PCA improves model performance by highlighting the most relevant features. Overall, the multivariate LSTM approach based on AWOS data is effective in modeling spatio-temporal patterns of visibility data and shows strong potential for further development in meteorological prediction studies.*

Keywords— *AWOS; Deep learning; LSTM; PCA; Visibility.*

I. PENDAHULUAN

Perkembangan sistem informasi cuaca berbasis data telah menjadi bagian penting dalam mendukung keselamatan dan efisiensi transportasi udara. Dalam operasional penerbangan, cuaca merupakan salah satu faktor paling kritis yang mempengaruhi keputusan teknis dan strategis, baik dalam perencanaan rute, manuver pesawat, hingga proses lepas landas dan pendaratan [1]. Oleh karena itu, kemampuan untuk memantau dan memprediksi kondisi atmosfer secara *real-time* menjadi hal yang sangat dibutuhkan, terutama di area strategis seperti bandar udara internasional yang memiliki intensitas lalu lintas penerbangan tinggi.

Di Indonesia, Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) telah mengoperasikan sistem monitoring cuaca *real-time* di bandara-bandara utama, termasuk Bandara Internasional Juanda [2]. BMKG Stasiun Meteorologi Juanda, memiliki peran penting dalam memantau dan menyediakan layanan informasi meteorologi penerbangan yang akurat dan tepat waktu guna mendukung keselamatan, efisiensi, dan keteraturan penerbangan di Bandara Internasional Juanda [3]. Dalam menjalankan tugas tersebut, berbagai instrumen dan sensor otomatis canggih digunakan untuk merekam kondisi atmosfer secara *real-time*. Salah satu sistem yang digunakan adalah *Automatic Weather Observation System* (AWOS) yang mampu mencatat berbagai parameter meteorologi secara kontinu dengan resolusi temporal tinggi. Meskipun sistem sensor canggih telah diimplementasikan, pengembangan model prediktif berbasis pembelajaran mesin tetap relevan untuk meningkatkan akurasi prakiraan dan mendukung sistem peringatan dini.

Dalam bidang meteorologi dan penerbangan, jarak pandang merupakan salah satu parameter cuaca yang penting karena berkaitan dengan keselamatan dan efisiensi operasional [4]. Jarak pandang atau *visibility* didefinisikan sebagai jarak maksimum di mana objek dapat terlihat dan diidentifikasi dengan mata telanjang [5]. Penurunan jarak pandang secara tiba-tiba, seperti akibat kabut atau hujan lebat, dapat menyebabkan keterlambatan, pembatalan, bahkan kecelakaan jika tidak diantisipasi secara akurat [6]. Dalam konteks tersebut, kebutuhan terhadap sistem prediksi jarak pandang yang handal menjadi semakin penting, terlebih dengan ketersediaan data meteorologi yang semakin melimpah dan beresolusi tinggi dari berbagai sensor otomatis. Prediksi jarak pandang sendiri merupakan tantangan kompleks karena dipengaruhi oleh berbagai faktor, tidak hanya *temporal* (waktu) tetapi juga *spasial* (kondisi di lokasi-lokasi terdekat). Oleh karena itu, pendekatan *spatio-temporal multivariate* yang mempertimbangkan data dari berbagai titik pengamatan dan berbagai variabel meteorologi dapat memberikan prediksi yang lebih komprehensif dan akurat [7].

Untuk menjawab tantangan tersebut, dibutuhkan metode analisis yang mampu menangkap hubungan *non-linear* dan pola jangka Panjang dalam data deret waktu. Disinilah pendekatan *deep learning* menunjukkan keunggulannya, karena memiliki kemampuan dalam mengenali pola kompleks dan hubungan tersembunyi pada data berskala besar [8]. Data AWOS dari ketiga titik pengamatan merupakan data *time series* multivariat yang memiliki karakteristik spasial dan temporal yang saling terkait [9]. Untuk menangani kompleksitas tersebut, penelitian ini menggunakan arsitektur *deep learning* yang dirancang khusus untuk data *time series*, yaitu *Long Short-Term Memory* (LSTM) pertama kali diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997 [10].

Penelitian tentang penggunaan model LSTM untuk memprediksi parameter cuaca telah banyak dilakukan. Simanjuntak et al [11] menunjukkan bahwa model LSTM mampu memberikan hasil prediksi cuaca harian lebih baik dibandingkan SARIMA, dengan nilai RMSE prediksi suhu sebesar 0,78 dan tekanan udara 0,65, lebih rendah dibandingkan SARIMA yang masing-masing sebesar 1,07 dan 0,91, sehingga LSTM lebih unggul dalam menangkap pola nonlinier pada data *time series*. Penelitian Khumaidi et al [12] menunjukkan bahwa pemodelan LSTM dengan empat hidden layer, batch size 32, dan epoch 1000 mampu menghasilkan prediksi kualitas udara dan cuaca di Bandung dengan performa cukup baik pada tiga parameter utama,

yaitu suhu (RMSE 3,15), kelembaban (RMSE 6,97), dan ISPU (RMSE 1,85), sedangkan prediksi PM10 masih menunjukkan hasil kurang optimal (RMSE 1,76 lebih besar dari standar deviasi 1,75). Dalam penelitian Sitompul et al [13] diterapkan model LSTM dengan preprocessing normalisasi *Min-Max* serta evaluasi menggunakan RMSE. Model diuji pada data PM2,5 dengan pembagian data training dan testing 90%:10% dan menghasilkan performa terbaik dengan nilai RMSE sebesar 3,40. Akurasi prediksi mencapai 92,18% pada kategori baik dan 87,75% pada kategori sedang hingga tidak sehat.

Dalam penelitiannya, Supriadi [14] melaporkan bahwa validasi prediksi parameter cuaca menggunakan LSTM menghasilkan RMSE yang relatif kecil, yakni 0,576 untuk suhu dan 1,0647 untuk tekanan udara, sehingga metode ini dinilai efektif dalam menangani data time series meteorologi. Akbar et al [15] menunjukkan bahwa model LSTM dengan konfigurasi terbaik (70% data latih, 128 unit LSTM, learning rate 0,009, batch size 16, dan 800 epochs) mampu memprediksi temperatur Kota Semarang dengan tingkat akurasi tinggi, ditunjukkan oleh nilai MAPE sebesar 1,896% dan RMSE sebesar 0,725, yang relatif kecil dibandingkan interval data uji sebesar 5,8.

Meskipun berbagai penelitian telah menunjukkan keberhasilan model LSTM dalam memprediksi parameter meteorologi, sebagian besar penelitian tersebut masih menggunakan pendekatan univariate time series atau hanya memanfaatkan data dari satu titik pengamatan. Pendekatan tersebut belum sepenuhnya merepresentasikan karakteristik sistem atmosfer yang bersifat multivariat dan memiliki keterkaitan spasial antar lokasi pengamatan. Selain itu, sebagian besar penelitian menggunakan data dengan resolusi temporal yang relatif rendah, seperti data harian atau per jam.

Dalam konteks operasional penerbangan, khususnya pada lingkungan bandar udara, kondisi atmosfer dapat berubah secara cepat dalam skala menit. Oleh karena itu, pemanfaatan data meteorologi resolusi tinggi dari AWOS yang merekam parameter cuaca setiap menit memiliki potensi besar untuk meningkatkan akurasi model prediksi. Namun, penelitian yang mengintegrasikan data AWOS multivariat dari beberapa titik pengamatan dengan pendekatan spatio-temporal berbasis deep learning masih sangat terbatas, khususnya pada studi kasus bandara di Indonesia.

Penelitian ini memanfaatkan data AWOS yang merekam parameter cuaca setiap satu menit selama satu tahun di tiga titik lokasi strategis di Bandara Juanda, yaitu titik 10, titik 28, dan titik tengah 10-28. Data AWOS ini menyediakan dataset komprehensif yang mencakup berbagai parameter cuaca seperti suhu, kelembaban, kecepatan dan arah angin, dan tekanan udara, curah hujan dan jarak pandang yang sangat berharga untuk eksplorasi metode pembelajaran mendalam [2].

Kontribusi utama penelitian ini terletak pada pemanfaatan data AWOS multivariat dengan resolusi temporal tinggi dalam kerangka pemodelan *spatio-temporal* menggunakan pendekatan deep learning. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan data dari satu lokasi atau parameter terbatas, penelitian ini mengintegrasikan data meteorologi dari tiga titik pengamatan AWOS untuk membangun model prediksi jarak pandang berbasis LSTM. Selain itu, penelitian ini juga menerapkan PCA untuk mereduksi dimensi data serta mengidentifikasi fitur meteorologi yang paling berpengaruh terhadap prediksi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun model prediksi jarak pandang berbasis data AWOS menggunakan pendekatan *spatio-temporal multivariate* LSTM, serta mengevaluasi performa model dalam memprediksi jarak pandang di Bandara Internasional Juanda berdasarkan data observasi dari tiga titik pengamatan. Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam eksplorasi pemanfaatan pendekatan deep learning untuk pemodelan data meteorologi berbasis spatio-temporal, khususnya pada studi prediksi jarak pandang menggunakan data observasi AWOS.

II. METODE PENELITIAN

Metodologi penelitian mencakup serangkaian langkah yang meliputi rancangan penelitian, deskripsi data, serta prosedur analisis yang digunakan dalam membangun dan mengevaluasi model prediksi jarak pandang.

A. Rancangan Penelitian

Penelitian ini menggunakan pendekatan supervised learning berbasis *deep learning* dengan jenis tugas berupa regresi *spatio-temporal multivariate*. Istilah ini merujuk pada data atau pemodelan yang melibatkan lebih dari satu variabel (*multivariate*) yang saling bergantung, diukur pada berbagai lokasi pengamatan (*spatial*) dan periode waktu (*temporal*), sehingga membentuk struktur ketergantungan kompleks lintas variabel, ruang, dan waktu [7]. Pendekatan ini dipilih karena variabel target yang diprediksi, yaitu *visibility* pada titik 28, merupakan data kontinu yang tersedia dalam bentuk historis berlabel, sehingga sesuai untuk diproses dalam kerangka supervised learning.

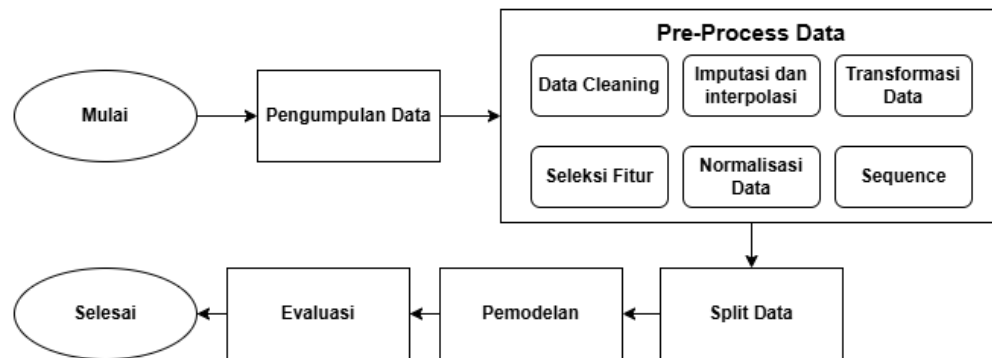
Model prediksi dikembangkan menggunakan arsitektur LSTM, yang termasuk ke dalam keluarga *Recurrent Neural Networks* (RNN). Arsitektur ini dipilih karena memiliki kemampuan dalam mempelajari ketergantungan jangka panjang pada data sekuensial, sehingga sesuai untuk memodelkan data meteorologis berbasis deret waktu. Untuk mengakomodasi dimensi spasial, data yang berasal dari tiga titik pengamatan AWOS titik 10, 10-28, dan 28 digabungkan dalam bentuk input multivariat yang selanjutnya diproses sebagai deret waktu supervised. Model yang dibangun menggunakan arsitektur LSTM dengan dukungan seleksi fitur melalui PCA untuk meningkatkan kinerja prediksi.

B. Data Penelitian

Dataset yang digunakan dalam penelitian ini berasal dari data AWOS tahun 2022 yang beroperasi di Bandara Internasional Juanda, Sidoarjo. AWOS merupakan sistem otomatis yang merekam kondisi cuaca secara *real-time* dari berbagai sensor meteorologi, dengan frekuensi pencatatan setiap satu menit sekali. AWOS tersebut dipasang pada tiga titik lokasi di sepanjang landas pacu bandara, sehingga pengamatan cuaca dilakukan secara spasial di titik 10, titik 10–28, dan titik 28. Dataset yang digunakan mencakup periode 1 Januari 2022 hingga 31 Desember 2022, dengan frekuensi pencatatan setiap satu menit. Total jumlah baris data mencapai 525.497 entri. Struktur penyimpanan data diklasifikasikan berdasarkan tiga titik lokasi pengamatan serta format waktu pengamatan dalam bentuk yyyy-mm-dd, hh:mm:ss.

C. Prosedur Penelitian

Prosedur penelitian dilakukan melalui tahapan pada Gambar 1.



Gambar 1. Diagram Alur Penelitian

1. Pengumpulan Data

Data historis AWOS 2022 diperoleh dari tiga titik pengamatan. Ketiga titik tersebut dipasang di sepanjang landas pacu Bandara Internasional Juanda (titik 10, titik 10–28, dan titik 28) untuk merekam parameter cuaca secara *real-time* dengan interval pencatatan setiap satu menit.

2. Preproses Data

Tahap pre-processing dilakukan untuk memastikan kualitas data observasi yang digunakan dalam penelitian berada dalam kondisi baik serta siap diolah pada tahap analisis dan pemodelan. Proses ini melibatkan beberapa langkah sistematis sebagai berikut.

Pertama, tahap *data cleaning* dilakukan penyeragaman format waktu, penghapusan atribut yang tidak relevan, deteksi duplikasi data, serta pemeriksaan *missing values* untuk mengetahui sebaran dan tingkat kelengkapan data sebelum masuk ke tahap imputasi. Kedua, tahap *imputasi dan interpolasi*. Nilai hilang pada variabel kontinu, seperti kelembaban, suhu udara, jarak pandang, intensitas cahaya, tekanan udara, dan ketinggian dasar awan, ditangani dengan interpolasi linear dua arah. Metode interpolasi linear dua arah digunakan untuk variabel kontinu, sesuai dengan pendekatan interpolasi deret waktu seperti dalam penelitian Lepo et al [16].

Pada variabel arah dan kecepatan angin digunakan pendekatan interpolasi berbasis vektor. Metode ini dilakukan dengan mengubah terlebih dahulu parameter angin ke dalam bentuk komponen kartesian x dan y , sehingga proses interpolasi dapat dilakukan secara linear tanpa terpengaruh oleh sifat siklikal sudut arah angin. Dalam penelitian ini, transformasi komponen angin ke dalam koordinat kartesian mengikuti konvensi meteorologi. Berbeda dengan sistem koordinat matematika umum, sudut arah angin tidak diukur dari sumbu- x positif, melainkan dari arah utara sumbu- y positif dan meningkat searah jarum jam [17]. Oleh karena itu, komponen kartesian didefinisikan sebagai persamaan (1).

$$x_i = u_i \cdot \sin(\theta_i), \quad y_i = u_i \cdot \cos(\theta_i) \quad (1)$$

dengan u_i menyatakan kecepatan angin pada titik ke- i dan θ_i menyatakan arah angin dalam derajat. Hasil interpolasi berupa (x_p, y_p) selanjutnya dikonversi kembali ke bentuk polar untuk memperoleh kecepatan dan arah angin pada titik target p dalam persamaan (2)

$$u_p = \sqrt{x_p^2 + y_p^2}, \quad \theta_p = \tan^{-1} \left(\frac{y_p}{x_p} \right) \quad (2)$$

Penentuan sudut arah angin θ_p mengikuti konvensi meteorologi, di mana sudut diukur terhadap sumbu-y positif (arah utara) dan meningkat searah jarum jam. Nilai sudut bergantung pada tanda (*sign*) dari komponen x_p dan y_p , sehingga representasi arah mempertimbangkan kuadran dalam koordinat kartesian.

Selain itu, pada titik pengamatan 10–28 dilakukan imputasi tambahan dengan mengisi nilai jarak pandang dan cahaya menggunakan rata-rata dari titik 10 dan titik 28, sedangkan lapisan awan ditentukan berdasarkan nilai maksimum (*dominasi*). Untuk interpolasi dominasi, metode ini dipandang sebagai strategi imputasi berbasis aturan atau *rule-based imputation* dengan memilih nilai maksimum dari dua titik pengamatan.

Ketiga, tahap transformasi data. Setelah proses imputasi selesai, data dari ketiga titik pengamatan digabungkan ke dalam satu dataset terintegrasi. Selanjutnya, data di-*resample* menjadi rata-rata setiap 30 menit untuk memperoleh representasi temporal yang lebih stabil dan sesuai dengan kebutuhan analisis.

Tahap berikutnya seleksi fitur menggunakan metode *Principal Component Analysis* (PCA) untuk mereduksi dimensi data. PCA merupakan teknik statistik multivariat yang mentransformasi sekumpulan variabel saling berkorelasi menjadi variabel baru yang tidak berkorelasi, yang disebut komponen utama (*principal components*). Transformasi ini dilakukan dengan tujuan mempertahankan sebagian besar informasi dari data asli namun dalam dimensi yang rendah [18]. Secara matematis, proses transformasi PCA dapat dinyatakan sebagai (3):

$$Z = TU = T(V_1, V_2, \dots, V_I) \quad (3)$$

di mana T merupakan matriks data asli dan U merupakan matriks eigenvector hasil dekomposisi matriks kovarians Transformasi ini menghasilkan matriks baru Z yang merepresentasikan data dalam ruang komponen utama dengan dimensi yang lebih rendah [19].

Dengan cara ini, PCA mampu menyaring informasi utama dari data berdimensi tinggi serta menghilangkan redundansi antar variabel sehingga hanya komponen yang paling berkontribusi terhadap variabilitas data yang digunakan dalam pemodelan [19], [20]. Pemilihan jumlah komponen utama dilakukan berdasarkan nilai *cumulative explained variance* sehingga sebagian besar variabilitas data tetap dipertahankan dalam dimensi yang lebih rendah [19]. Dari hasil PCA, dipilih sembilan komponen utama PC1–PC7 yang secara kumulatif menjelaskan sekitar 95% total variansi, yang disajikan pada Tabel 1.

TABEL 1
FITUR-FITUR UTAMA

Komponen Utama	Varian yang Dijelaskan	Kontribusi Tertinggi	Nilai Kontribusi
PC1	45.08%	Suhu Udara titik 10–28	0.2904
PC2	14.42%	Tekanan Udara titik 28	0.4440
PC3	10.60%	Lapisan Awan titik 10–28	0.4141
PC4	8.83%	Arah Angin titik 10–28	0.3715
PC5	8.21%	Kecepatan Angin titik 28	0.4055
PC6	3.75%	Kelembaban titik 10	0.4231
PC7	3.31%	Jarak Pandang titik 10	0.6118

Ketujuh fitur dominan dari masing-masing komponen utama tersebut dipilih sebagai representasi input akhir untuk tahap pemodelan. Pemilihan fitur ini tidak hanya membantu mengurangi kompleksitas model, tetapi juga mempertahankan informasi penting yang merepresentasikan kondisi meteorologis di seluruh titik pengamatan.

Setelah seleksi fitur, data kemudian dinormalisasi menggunakan metode *Min-Max Scaling* ke dalam rentang [0, 1] untuk menyetarakan skala antar fitur dan mencegah dominasi variabel dengan rentang nilai yang lebih besar dengan persamaan (4) [21].

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4)$$

Pada persamaan tersebut, x menunjukkan nilai asli dari data sebelum dinormalisasi, x_{min} adalah nilai minimum dalam data, sedangkan x_{max} adalah nilai maksimum dalam data. Hasil dari perhitungan ini disebut $x_{normalized}$, yaitu nilai yang sudah dinormalisasi.

Tahap terakhir dalam *preprocessing* adalah pembentukan *sequence* data. data *time series* diubah menjadi bentuk *sequence* dengan menggunakan *Time Series Generator* dari pustaka Keras. Tujuannya adalah agar model pembelajaran mesin dapat menangkap pola historis dalam rentang waktu tertentu untuk memprediksi nilai di masa depan. Proses pembentukan *sequence* dilakukan dengan menentukan panjang jendela waktu (*sequence length*) sebesar 168 langkah, yang merepresentasikan 7 hari data dengan interval 30 menit. Artinya, setiap *sequence* terdiri dari 168 data historis yang digunakan sebagai masukan (*input*), sedangkan target (*output*) adalah nilai variabel yang ingin diprediksi pada langkah waktu berikutnya, yaitu “Jarak_Pandang_titik28”

Secara matematis, jika deret waktu dinyatakan pada persamaan (5)

$$D = \{x_1, x_2, x_3, \dots, x_t\} \quad (5)$$

maka pembentukan *sequence* dengan panjang $L = 168$ akan menghasilkan pasangan data (6)

$$X_i = \{x_i, x_{i+1}, \dots, x_{i+167}\}, \quad y_i = x_{i+168} \quad (6)$$

dengan X_i sebagai urutan data input dan y_i sebagai target prediksi. Hasil dari proses ini adalah tensor tiga dimensi untuk input (X) dengan bentuk:

$$X \in R^{N \times 168 \times F} \quad (7)$$

dan vektor target y dengan bentuk:

$$y \in R^N \quad (8)$$

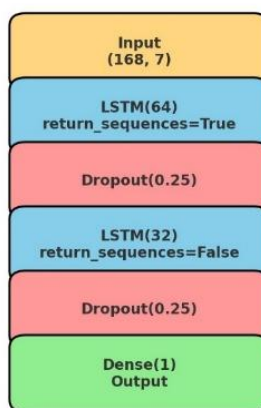
3. Split Data

Dataset hasil transformasi dibagi menjadi tiga subset, yakni data latih (80%), data validasi (10%), dan data uji (10%). Pembagian ini dilakukan setelah pembentukan *sequence* agar kontinuitas temporal tetap terjaga. Semua subset kemudian disimpan dalam format .npy untuk efisiensi pemuatan data pada tahap pelatihan model.

4. Pemodelan

Arsitektur model yang digunakan dalam penelitian ini berbasis *Long Short-Term Memory* (LSTM), yaitu salah satu jenis *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi permasalahan *vanishing gradient* pada data sekuensial [10]. LSTM memiliki kemampuan untuk menyimpan, memperbarui, dan menghapus informasi pada jangka waktu panjang sehingga sesuai untuk pemodelan data *time series* seperti data cuaca.

Pada penelitian ini, model LSTM disusun dalam bentuk arsitektur bertingkat (*stacked LSTM*), yang terdiri dari dua lapisan LSTM yang diikuti oleh lapisan *dropout* dan lapisan *dense* sebagai keluaran model. Pendekatan *stacked LSTM* banyak digunakan dalam pemodelan deret waktu karena mampu menangkap pola temporal yang lebih kompleks dibandingkan dengan satu lapisan LSTM saja. Arsitektur serupa juga digunakan dalam penelitian analisis iklim berbasis pembelajaran mendalam yang memanfaatkan dua lapisan LSTM untuk memodelkan dependensi temporal pada data meteorologi dengan panjang sekuens 168 langkah waktu [22]. Panjang *sequence* sebesar 168 langkah waktu dipilih karena merepresentasikan pola temporal selama 7 hari dengan interval 30 menit. Pemilihan ini didasarkan pada asumsi bahwa fenomena meteorologi, seperti jarak pandang, memiliki pola periodik mingguan yang dipengaruhi oleh siklus atmosfer. Dengan demikian, model dapat menangkap dependensi temporal jangka pendek hingga menengah secara lebih optimal [22]. Arsitektur model terdiri atas beberapa lapisan utama sebagaimana ditunjukkan pada Gambar 2, dengan rincian sebagai berikut:



Gambar 2. Arsitektur LSTM

- 1) Lapisan Input (LSTM Pertama)
Lapisan awal berupa LSTM dengan 64 unit neuron dan fungsi aktivasi tanh. Parameter `return_sequences=True` digunakan agar lapisan ini mengeluarkan keluaran pada setiap time step dari sekuens, bukan hanya keluaran akhir. Hal ini memungkinkan hasil dari setiap time step diteruskan ke lapisan LSTM berikutnya. Bentuk input yang diberikan adalah (*timesteps, features*), yaitu 168 *timesteps* setara 84 jam dengan interval 30 menit dan 7 fitur meteorologi.
- 2) Lapisan Dropout Pertama
Setelah LSTM pertama, diterapkan lapisan dropout dengan tingkat 0,25. Tujuan dari lapisan ini adalah sebagai teknik regularisasi, yakni menonaktifkan secara acak 25% neuron pada saat pelatihan untuk mencegah *overfitting* dan meningkatkan kemampuan generalisasi model.
- 3) Lapisan LSTM Kedua
Lapisan berikutnya berupa LSTM dengan 32 unit neuron dan fungsi aktivasi tanh. Pada lapisan ini digunakan pengaturan `return_sequences=False`, sehingga hanya keluaran dari time step terakhir yang digunakan. Keluaran ini berupa representasi vektor tunggal yang merangkum informasi dari seluruh sekuens masukan.
- 4) Lapisan Dropout Kedua
Sama seperti sebelumnya, lapisan dropout dengan tingkat 0,25 ditempatkan setelah LSTM kedua. Hal ini dimaksudkan untuk memberikan regularisasi tambahan guna mengurangi risiko *overfitting*.
- 5) Lapisan Output (Dense Layer)
Bagian akhir dari arsitektur adalah lapisan dense dengan 1 unit neuron tanpa fungsi aktivasi (linear). Lapisan ini menghasilkan satu nilai kontinu yang merepresentasikan prediksi target, yaitu jarak pandang (*visibility*).

Setiap sel LSTM memiliki struktur khusus yang dirancang untuk mengatur aliran informasi melalui tiga gerbang utama, yaitu *forget gate*, *input gate*, dan *output gate*, yang berfungsi untuk mengontrol informasi yang dipertahankan, diperbarui, atau dikeluarkan pada setiap langkah waktu. Mekanisme ini memungkinkan LSTM mempelajari ketergantungan temporal jangka panjang pada data sekuensial [23].

5. Training Data

Proses pelatihan model dilakukan dengan maksimum 100 epoch menggunakan batch size sebesar 32. Validasi dilakukan terhadap data `X_val` dan `y_val` untuk memantau performa model serta mengevaluasi kemampuan generalisasi model. Selama proses pelatihan, metrik evaluasi berupa loss dan *Mean Absolute Error* (MAE) dipantau secara berkala baik pada data pelatihan maupun validasi. Pemantauan ini bertujuan untuk memastikan konvergensi model sekaligus mencegah terjadinya *overfitting*. Model dikompilasi menggunakan konfigurasi sebagai berikut:

1. Loss function: *Mean Absolute Error* (MAE), sesuai dengan tujuan regresi.
2. Metric: *Mean Squared Error* (MSE), digunakan sebagai metric evaluasi tambahan.
3. Optimizer: Adam dengan `learning_rate=0.001` dan `clipnorm=1.0` untuk stabilisasi gradien.

Selama proses pelatihan, digunakan dua callback utama: *EarlyStopping*, Untuk menghentikan pelatihan lebih awal jika validasi loss tidak membaik selama 10 epoch berturut-turut, serta mengembalikan bobot terbaik. *ModelCheckpoint*, Menyimpan model terbaik berdasarkan metrik validasi loss (`val_loss`).

6. Evaluasi Model

Performa model dievaluasi menggunakan metrik MAE, MSE, RMSE dan MAPE.

Mean Absolute Error (MAE) mengukur rata-rata nilai absolut selisih antara observasi dan prediksi dengan rumus pada persamaan (9) [24].

$$MAE = \frac{1}{s} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (9)$$

MAE menimbang semua kesalahan secara linear sehingga lebih robust terhadap outlier dan lebih mudah diinterpretasikan (satuan sama dengan variabel). MAE cocok untuk ukuran kesalahan yang mudah dibaca dan tidak ingin menghukum kesalahan besar secara berlebihan.

Mean Squared Error (MSE) mengukur rata-rata kuadrat selisih antara nilai observasi y_i dan prediksi $\hat{y}_i$. Rumus untuk sampel berukuran n dengan persamaan (10) [25].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Root Mean Square Error (RMSE) digunakan untuk mengukur seberapa jauh rata-rata hasil prediksi menyimpang dari nilai sebenarnya. Semakin kecil nilai RMSE, semakin baik kualitas model karena prediksi semakin mendekati data aktual. Namun, karena kesalahan (*error*) dikuadratkan dalam perhitungannya, RMSE menjadi lebih sensitif terhadap outlier atau kesalahan yang berukuran besar [26]. Secara sistematis dituliskan dalam persamaan (11).

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_i - y_i)^2} \quad (11)$$

Mean Absolute Percentage Error (MAPE) digunakan untuk mengukur rata-rata persentase kesalahan absolut antara nilai prediksi dan nilai aktual [19]. MAPE menyajikan kesalahan dalam bentuk persentase sehingga memudahkan interpretasi tingkat kesalahan model secara relatif terhadap nilai aktual. Secara sistematis dituliskan dalam persamaan (12).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (12)$$

III. HASIL DAN PEMBAHASAN

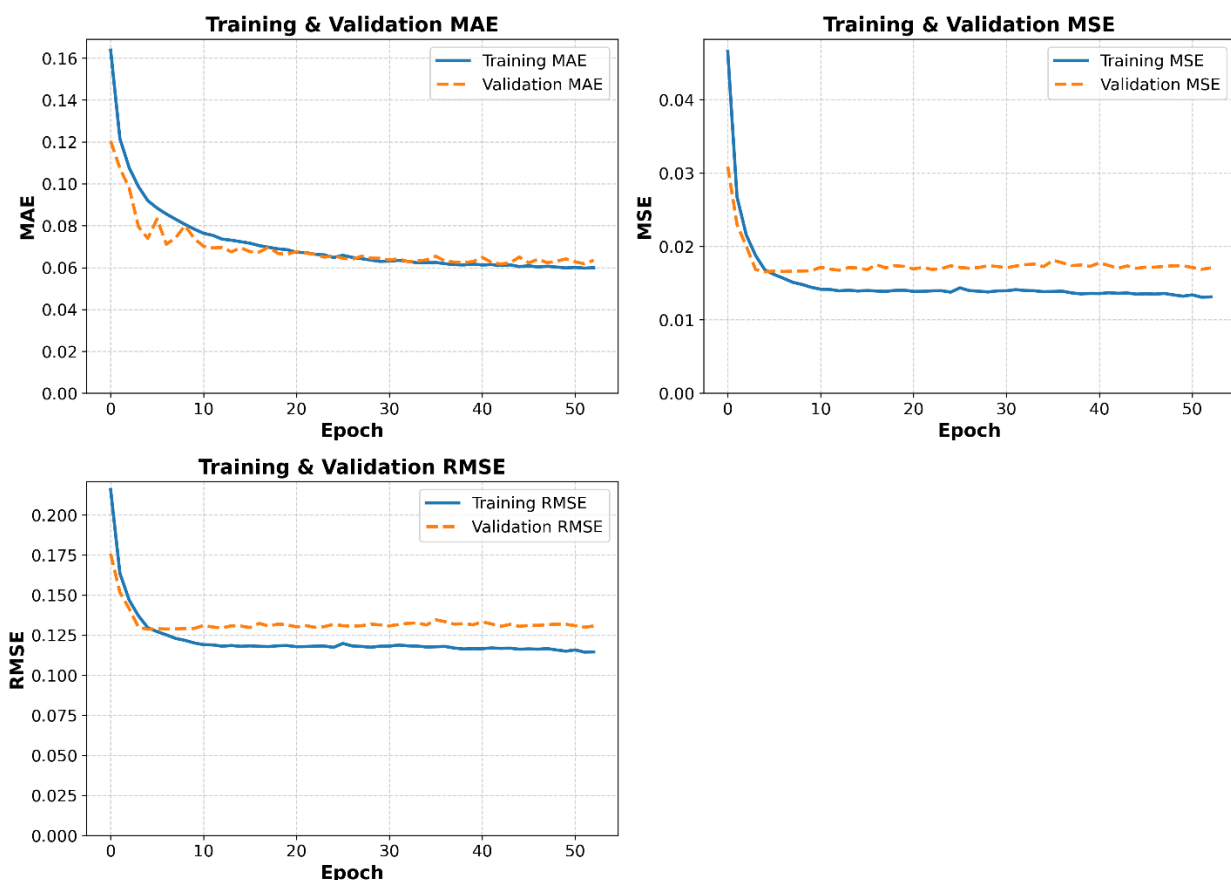
A. Evaluasi Pelatihan Model LSTM

Untuk mengevaluasi performa model selama proses pelatihan, dilakukan pemantauan terhadap tiga metrik utama, yaitu *Mean Squared Error* (MSE), *Mean Absolute Error* (MAE), dan *Root Mean Square Error* (RMSE), baik pada data pelatihan (*train*) maupun data validasi. Gambar 3 memperlihatkan kurva loss train terhadap jumlah epoch yang dilalui selama proses pelatihan.

Pada Gambar 3 sebelah kiri, ditampilkan analisis kurva MAE. Grafik menunjukkan pola konvergensi yang cukup stabil antara data pelatihan dan validasi. Pada epoch awal sekitar epoch 0–4, nilai MAE mengalami penurunan yang cukup tajam dari sekitar 0.16 pada data pelatihan dan sekitar 0.12 pada data validasi hingga mendekati 0.08. Penurunan yang signifikan pada fase awal ini mengindikasikan bahwa model dengan cepat mampu menangkap pola dasar hubungan antara variabel meteorologi multivariat dan nilai jarak pandang. Setelah fase tersebut, kurva MAE pelatihan terus menurun secara gradual hingga mencapai kisaran sekitar 0.06 pada epoch akhir. Sementara itu, kurva MAE validasi cenderung stabil di kisaran 0.062–0.065 dengan fluktuasi kecil. Kedekatan nilai antara kurva pelatihan dan validasi menunjukkan bahwa model memiliki kemampuan generalisasi yang baik serta tidak menunjukkan indikasi *overfitting* yang signifikan. Fluktuasi kecil pada kurva validasi kemungkinan disebabkan oleh variasi alami pada data meteorologi yang bersifat dinamis dan non-linear. Ripple yang terlihat pada kurva validasi sekitar epoch ke-4 hingga ke-9 kemungkinan terjadi karena model masih berada pada fase awal penyesuaian parameter, di mana proses optimasi dengan mini-batch gradient serta penggunaan dropout menyebabkan variasi kecil pada bobot model di setiap epoch sehingga nilai error validasi tidak menurun secara monoton.

Gambar 3 sebelah kanan diperlihatkan analisis kurva MSE. Pola yang tampak relatif sejalan, di mana *training* MSE menurun cukup signifikan hingga sekitar epoch ke-20 sebelum akhirnya stabil pada nilai rendah mendekati 0.013. Di sisi lain, *validation* MSE memang menurun pada fase awal, tetapi setelah epoch ke-10 mulai menunjukkan fluktuasi ringan dengan kisaran nilai antara 0.015 hingga 0.017 yang cenderung datar meski sedikit bergelombang. Pola ini menunjukkan bahwa sebagian besar penurunan kesalahan telah terjadi pada fase awal pelatihan ketika model mulai menangkap hubungan utama antara variabel meteorologi dan jarak pandang. Setelah pola dasar tersebut dipelajari, peningkatan performa menjadi semakin kecil sehingga kurva validasi cenderung stabil. Bentuk kurva yang tetap sedikit melengkung (*convex*) mencerminkan adanya variasi alami pada data validasi yang bersifat dinamis dan non-linear, sehingga nilai kesalahan kuadrat tidak menurun secara monoton pada setiap epoch. Namun demikian, perbedaan antara training dan validation MSE tetap relatif kecil, sehingga belum memberikan indikasi adanya *overfitting* yang signifikan.

Adapun pada Gambar 3 di bawah kiri, kurva RMSE menunjukkan pola yang selaras dengan MSE. Nilai training RMSE menurun tajam, awalnya berada di atas 0.20 hingga stabil pada kisaran 0.11 setelah sekitar epoch ke-15. Kurva validation RMSE juga turun pada fase awal kemudian cenderung datar di kisaran 0.125 – 0.13 dengan sedikit fluktuasi. Yang menunjukkan bahwa sebagian besar proses pembelajaran model terjadi pada fase awal pelatihan ketika model mulai menangkap hubungan utama antara variabel meteorologi multivariat dan nilai jarak pandang. Hal ini sejalan dengan karakteristik arsitektur LSTM yang dirancang untuk mempelajari dependensi temporal pada data deret waktu, sehingga informasi historis dalam jendela sequence mampu membantu model mengidentifikasi pola perubahan atmosfer secara lebih efektif [27]. Setelah pola dasar tersebut dipelajari, penurunan error pada data validasi menjadi sangat terbatas sehingga kurva cenderung stabil. Bentuk kurva yang sedikit melengkung tersebut mencerminkan variasi alami pada data meteorologi yang bersifat dinamis dan non-linear, sehingga nilai kesalahan tidak menurun secara monoton pada setiap epoch. Kondisi ini menguatkan hasil analisis MSE bahwa meskipun model semakin optimal pada data latih, peningkatan performa pada data validasi cenderung stagnan setelah mencapai titik tertentu. Meskipun demikian, selisih antara training dan validation RMSE tetap relatif kecil, yang menunjukkan bahwa model memiliki kemampuan generalisasi yang baik tanpa indikasi overfitting yang signifikan.



Gambar 3. Grafik Loss Train

Secara keseluruhan, hasil dari ketiga metrik MAE, MSE, dan RMSE menunjukkan bahwa model berhasil menurunkan error secara signifikan sejak awal pelatihan dan mempertahankan nilai *error* yang relatif rendah. Tidak terdapat indikasi *overfitting* yang parah, meskipun fluktuasi ringan pada *validation error* menandakan adanya gejala awal *underfitting* atau keterbatasan model dalam menyerap keragaman data validasi. Dengan demikian, model dapat dikatakan stabil dan cukup andal, namun tetap memiliki ruang perbaikan melalui optimasi arsitektur maupun penambahan fitur yang lebih *representative*.

B. Evaluasi Performa Model

Evaluasi performa model dilakukan dengan menggunakan 4 metrik, yaitu MSE, MAE, RMSE, dan MAPE baik pada data validasi maupun data uji. Metrik MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual, Metrik MSE dipakai untuk mengukur seberapa besar rata-rata kesalahan kuadrat antara nilai prediksi model dengan nilai actual. sedangkan RMSE digunakan untuk mengukur akar dari rata-rata kuadrat kesalahan prediksi, yang lebih sensitif terhadap kesalahan besar. Adapun metrik MAPE digunakan untuk mengukur rata-rata persentase kesalahan absolut antara nilai prediksi dan nilai aktual, sehingga memberikan gambaran tingkat kesalahan model dalam bentuk persentase yang lebih mudah diinterpretasikan. Prediksi dilakukan terhadap data validasi dan data uji menggunakan model yang telah dilatih. Hasil evaluasi metrik evaluasi model dapat dilihat pada Tabel 2.

TABEL 2
HASIL EVALUASI METRIK

Metrik Evaluasi	Nilai Validasi	Nilai Tes
MAE	0.0617	0.0618
MSE	0.0170	0.0150
RMSE	0.1305	0.1225
MAPE	27.58%	14.24%

Model yang telah dilatih kemudian diuji pada data validasi dan data uji untuk melihat kemampuan generalisasi terhadap data yang belum pernah dilihat sebelumnya. Berdasarkan hasil evaluasi, diperoleh nilai MAE sebesar 0.0617 pada data validasi dan 0.0618 pada data uji. Sementara itu, nilai MSE yang diperoleh adalah 0.0170 pada data validasi dan 0.0150 pada data uji, sedangkan nilai RMSE sebesar 0.1305 pada data validasi dan 0.1225 pada data uji. Perbedaan yang sangat kecil antara data validasi dan data uji menunjukkan bahwa model memiliki performa yang stabil dan tidak mengalami *overfitting*. Nilai RMSE yang hanya sedikit lebih tinggi dari MAE juga mengindikasikan bahwa model tidak menghasilkan kesalahan besar secara signifikan, sehingga dapat disimpulkan bahwa model memiliki akurasi dan konsistensi yang baik dalam melakukan prediksi. Selain itu, nilai MAPE yang diperoleh sebesar 27.58% pada data validasi dan 14.24% pada data uji menunjukkan tingkat kesalahan prediksi relatif model dalam bentuk persentase. Nilai tersebut mengindikasikan bahwa secara rata-rata prediksi model memiliki deviasi sekitar 27% pada data validasi dan sekitar 14% pada data uji terhadap nilai aktual. Perbedaan nilai MAPE antara kedua dataset tersebut kemungkinan dipengaruhi oleh perbedaan distribusi data serta keberadaan beberapa nilai aktual yang relatif kecil pada data validasi, mengingat metrik MAPE sangat sensitif terhadap nilai aktual yang mendekati nol. Meskipun demikian, hasil ini menunjukkan bahwa model masih mampu memberikan estimasi yang cukup representatif terhadap variasi nilai jarak pandang.

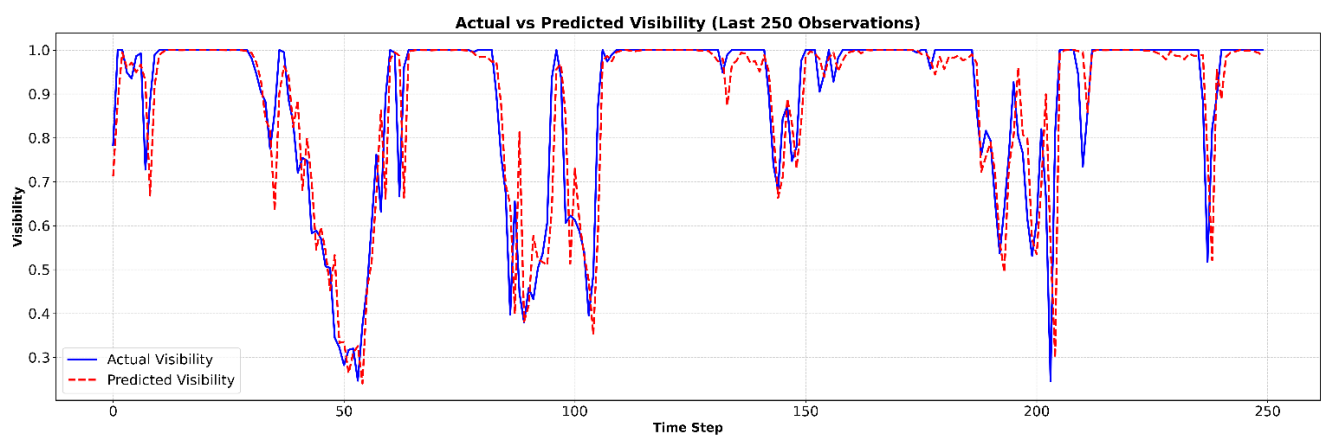
Jika dibandingkan dengan penelitian terdahulu, performa model ini menunjukkan nilai kesalahan yang relatif lebih rendah. Simanjuntak et al [11] melaporkan nilai RMSE sebesar 0,78 untuk suhu dan 0,65 untuk tekanan udara, sementara Supriadi [14] memperoleh RMSE 0,576 untuk suhu dan 1,0647 untuk tekanan udara. Penelitian Khumaidi et al [12] serta Sitompul et al [13] menghasilkan nilai RMSE yang lebih tinggi, berada pada rentang 1,76 hingga 6,97. Bahkan Akbar et al [15], yang menggunakan konfigurasi optimal pada prediksi temperatur Kota Semarang, masih mendapatkan nilai RMSE sebesar 0,725. Dibandingkan dengan hasil tersebut, model LSTM dalam penelitian ini jauh lebih presisi dengan RMSE hanya 0.1305 pada data validasi dan 0.1225 pada data uji. Meskipun demikian, perbandingan ini perlu diinterpretasikan secara hati-hati karena penelitian terdahulu memprediksi parameter meteorologi yang berbeda, seperti suhu dan tekanan udara, sedangkan penelitian ini berfokus pada prediksi jarak pandang. Perbedaan karakteristik variabel target tersebut dapat mempengaruhi tingkat kesulitan prediksi serta skala error yang dihasilkan, sehingga tidak sepenuhnya bersifat sebanding secara langsung.

Capaian performa yang lebih rendah pada penelitian ini dapat dijelaskan dari beberapa aspek metodologis. Pertama, penggunaan data AWOS dengan resolusi temporal tinggi memungkinkan model menangkap dinamika perubahan atmosfer secara lebih detail dibandingkan data dengan resolusi lebih rendah. Kedua, pendekatan multivariat yang mengintegrasikan beberapa parameter meteorologi dari berbagai titik pengamatan memberikan informasi yang lebih kaya dibandingkan pendekatan univariat, sehingga membantu model dalam memahami hubungan kompleks antar variabel. Ketiga, penerapan PCA berperan dalam mereduksi redundansi antar fitur dan menghasilkan representasi input yang lebih efisien, sehingga proses pembelajaran menjadi lebih stabil. Selain itu, arsitektur LSTM yang digunakan mampu menangkap ketergantungan temporal jangka panjang pada data deret waktu, yang relevan dalam memodelkan fenomena meteorologi yang bersifat dinamis. Dengan demikian, perbedaan performa yang diperoleh tidak hanya disebabkan oleh model yang digunakan, tetapi juga dipengaruhi oleh kualitas data, pendekatan multivariat, serta strategi reduksi dimensi yang diterapkan dalam penelitian ini.

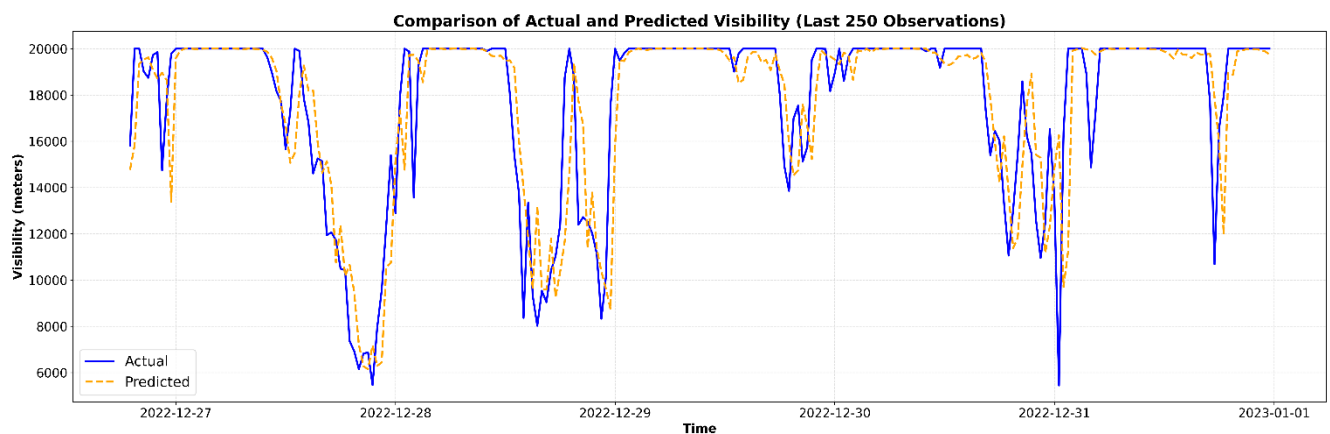
Penggunaan PCA berperan penting dalam mereduksi dimensi data dengan menyoroti fitur-fitur paling relevan sekaligus mengurangi redundansi, sehingga memperkuat kemampuan model dalam menangkap pola kompleks. Pada data meteorologi,

beberapa variabel seperti suhu, kelembaban, tekanan udara, dan lapisan awan sering memiliki korelasi yang cukup tinggi [2]. Kondisi ini dapat menyebabkan reduksi informasi pada fitur input yang berpotensi meningkatkan kompleksitas model dan memperlambat proses konvergensi. Dengan menggunakan PCA, fitur-fitur yang saling berkorelasi ditransformasikan menjadi komponen utama yang saling ortogonal sehingga model dapat mempelajari representasi data yang lebih ringkas namun tetap mempertahankan sebagian besar variansi informasi [19]. Selain itu, dalam konteks pemodelan spatio-temporal, LSTM tidak hanya menangkap dependensi temporal dari data deret waktu, tetapi juga memanfaatkan keterkaitan spasial secara implisit melalui integrasi data dari beberapa titik pengamatan sebagai fitur multivariat. Informasi dari berbagai lokasi diproses dalam urutan waktu, sehingga model mampu mempelajari hubungan antar lokasi dalam konteks temporal. Hal ini memungkinkan model untuk menangkap dinamika perubahan jarak pandang yang dipengaruhi oleh kondisi atmosfer di lokasi sekitar. Kecilnya nilai MAE menunjukkan rata-rata galat absolut prediksi yang rendah, sedangkan nilai MSE yang juga kecil menegaskan bahwa rata-rata kuadrat galat masih dalam batas wajar. Dengan demikian, hasil penelitian ini membuktikan bahwa model LSTM multivariat berbasis data AWOS dengan PCA menunjukkan kemampuan yang baik dalam merepresentasikan pola data serta menghasilkan prediksi jarak pandang yang cukup baik.

Untuk mengevaluasi kemampuan model LSTM secara visual, dilakukan plot terhadap hasil prediksi dibandingkan dengan data aktual jarak pandang pada data uji. Pada Gambar 4 dan 5 menampilkan perbandingan antara nilai aktual dan hasil prediksi model selama rentang waktu pengamatan.



Gambar 4. Perbandingan dan Prediksi vs Aktual (250 data terakhir)



Gambar 5. Grafik Perbandingan dan Prediksi vs Aktual dalam satuan meter (250 data terakhir)

Pada Gambar 4 dan 5 menunjukkan Grafik perbandingan antara nilai aktual dan prediksi jarak pandang menunjukkan bahwa model LSTM mampu mengikuti pola perubahan data dengan cukup baik. Secara umum, garis prediksi (merah) memiliki tren yang sejalan dengan garis aktual (biru), terutama pada periode ketika nilai jarak pandang berada pada kondisi tinggi mendekati nilai maksimum. Hal ini menunjukkan bahwa model berhasil menangkap pola dominan pada data deret waktu, khususnya pada kondisi atmosfer yang relatif stabil. Kedekatan antara kedua kurva pada sebagian besar interval waktu

mengindikasikan bahwa model mampu mempelajari hubungan temporal antar variabel meteorologi yang mempengaruhi perubahan jarak pandang. Namun demikian, pada beberapa titik terlihat bahwa kurva prediksi tidak selalu mengikuti perubahan tajam pada data aktual, terutama ketika terjadi penurunan jarak pandang secara drastis. Pada kondisi tersebut, garis prediksi cenderung lebih halus dan tidak sedalam nilai aktual. Fenomena ini umum terjadi pada model berbasis LSTM karena model cenderung melakukan *temporal smoothing*, yaitu memprediksi nilai berdasarkan pola historis sehingga perubahan ekstrim atau fluktuasi yang sangat cepat tidak sepenuhnya direpresentasikan secara langsung. Meskipun demikian, secara keseluruhan model tetap mampu menangkap arah perubahan (trend) dari peristiwa penurunan jarak pandang tersebut. Hal ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam merepresentasikan dinamika temporal data cuaca, meskipun masih terdapat keterbatasan dalam memprediksi fluktuasi ekstrim secara presisi.

C. Perbandingan Prediksi dan Data Aktual Jarak Pandang

Setelah proses pelatihan dan validasi model LSTM selesai, hasil prediksi yang awalnya dalam skala terstandarisasi dikembalikan ke skala asli (meter) menggunakan metode *inverse transform*. Perbandingan ini berguna untuk menilai seberapa akurat model dalam merepresentasikan kondisi nyata di lapangan. Hasil prediksi model LSTM yang telah dikonversi kembali ke dalam skala aktual (meter) kemudian dibandingkan dengan data jarak pandang aktual pada waktu tertentu. Perbandingan ini menunjukkan seberapa dekat hasil prediksi model dengan nilai sesungguhnya. Hasil perbandingan antara nilai prediksi model dan data aktual ditunjukkan pada Tabel 3.

TABEL 3
PERBANDINGAN DATA PREDIKSI DAN DATA AKTUAL

Waktu	Jarak Pandang Aktual (m)	Jarak Pandang Prediksi (m)
2022-11-28 23:30	6.436	6.757
2022-11-29 00:00	5.628	7.832
2022-11-29 00:30	6.919	7.502
2022-11-29 01:00	8.815	7.453
2022-11-29 01:30	13.133	10.476
2022-11-29 02:00	20.000	12.258
2022-11-29 02:30	20.000	19.605
2022-11-29 03:00	20.000	19.979
2022-11-29 03:30	20.000	19.932
2022-11-29 04:00	20.000	19.999

Berdasarkan Tabel 3, terlihat bahwa model LSTM mampu mengikuti pola kenaikan jarak pandang seiring berjalannya waktu. Pada awal prediksi, model masih menunjukkan perbedaan yang cukup jelas dengan data aktual, misalnya pada pukul 00:00 di mana jarak pandang aktual sebesar 5.628 meter, sedangkan hasil prediksi mencapai 7.832 meter. Namun, seiring bertambahnya waktu, prediksi semakin mendekati nilai aktual, seperti pada pukul 02:30 hingga 03:00 di mana perbedaan antara data aktual dan prediksi relatif kecil. Meskipun demikian, masih terdapat deviasi signifikan pada kondisi ekstrim, misalnya pada pukul 03:30 dan 04:00 di mana nilai prediksi sedikit menyimpang dari data aktual meskipun tetap berada dalam kisaran yang sejalan dengan tren data. Hal ini menunjukkan bahwa model cukup efektif dalam menangkap pola utama pergerakan jarak pandang, meskipun masih memerlukan penyempurnaan agar lebih presisi pada kondisi ekstrim.

Selain dependensi temporal, model juga memanfaatkan informasi spasial dari tiga titik pengamatan AWOS yang terletak di sepanjang landasan pacu. Variasi kondisi meteorologi pada masing-masing titik dapat mencerminkan dinamika atmosfer yang berbeda dalam skala lokal. Dengan menggabungkan data dari beberapa titik pengamatan sebagai fitur multivariat, model LSTM dapat mempelajari keterkaitan spasial antar lokasi secara implisit melalui pola perubahan temporal yang terjadi secara bersamaan.

Meskipun model menunjukkan performa yang baik, terdapat beberapa keterbatasan dalam penelitian ini. Pertama, model hanya dilatih menggunakan data observasi selama satu tahun sehingga belum sepenuhnya merepresentasikan variasi musiman jangka panjang. Kedua, model masih menunjukkan kecenderungan temporal smoothing ketika menghadapi perubahan ekstrim pada nilai jarak pandang, sehingga kedalaman penurunan jarak pandang tidak selalu diprediksi secara presisi. Ketiga, penelitian ini hanya menggunakan arsitektur LSTM dasar sehingga eksplorasi model hibrida seperti CNN-LSTM atau Transformer berpotensi memberikan peningkatan performa pada penelitian selanjutnya.

IV. SIMPULAN

Penelitian ini berhasil membangun model prediksi jarak pandang di Bandara Internasional Juanda menggunakan pendekatan LSTM berbasis data AWOS multivariat yang mempertimbangkan dimensi spasial dan temporal. Proses pemodelan dilakukan melalui tahapan *pre-processing* data, reduksi dimensi menggunakan PCA, pembentukan *sequence data*, serta pelatihan arsitektur LSTM dua lapisan dengan teknik regularisasi *dropout*. Hasil evaluasi menunjukkan bahwa model mampu menghasilkan prediksi jarak pandang yang stabil dengan kesalahan prediksi yang relatif kecil serta kemampuan generalisasi yang baik antara data validasi dan data uji. Kinerja model yang stabil dapat dipengaruhi oleh beberapa faktor utama. Penggunaan data AWOS dengan resolusi temporal tinggi memungkinkan model menangkap perubahan atmosfer secara lebih detail, sementara pendekatan multivariat dari beberapa titik pengamatan memberikan informasi yang lebih kaya dalam merepresentasikan kondisi spasial. Selain itu, penerapan PCA membantu mengurangi redundansi antar fitur sehingga proses pembelajaran menjadi lebih efisien dan stabil. Arsitektur LSTM yang digunakan juga berperan dalam menangkap ketergantungan temporal jangka panjang, sehingga model mampu mengikuti pola fluktuasi jarak pandang, termasuk pada kondisi perubahan yang signifikan. Secara keseluruhan, penelitian ini menegaskan bahwa pendekatan LSTM multivariat berbasis data AWOS dengan dukungan PCA mampu meningkatkan kinerja model dalam mempelajari pola spatio-temporal pada data meteorologi penerbangan, khususnya dalam prediksi jarak pandang. Temuan ini menunjukkan potensi pemanfaatan data AWOS beresolusi tinggi untuk pengembangan metode prediksi cuaca berbasis *deep learning*. Penelitian selanjutnya dapat diarahkan pada model *hybrid* seperti CNN-LSTM atau Transformer, serta eksplorasi integrasi model prediksi dengan sistem pemantauan cuaca yang telah tersedia.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas Islam Negeri Sunan Ampel Surabaya melalui Fakultas Sains dan Teknologi yang telah memberikan dukungan akademik serta fasilitas penelitian. Apresiasi juga diberikan kepada BMKG Juanda, khususnya para staf yang telah menyediakan data AWOS serta memberikan arahan dan masukan berharga selama proses penelitian. Dukungan, arahan, dan kerja sama dari seluruh pihak yang terlibat sangat membantu dalam penyelesaian penelitian

DAFTAR PUSTAKA

- [1] I. Gultepe *et al.*, "A Review of High Impact Weather for Aviation Meteorology," *Pure Appl. Geophys.*, vol. 176, no. 5, pp. 1869–1921, 2019.
- [2] V. Dwi, "Kepala BMKG Pastikan Kesiapan AWOS, LLWAS, dan MAWS, Hadapi Cuaca Ekstrem," BMKG, Surabaya, Indonesia, 2024. Accessed: Dec. 01, 2024. [Online]. Available: <https://www.bmkg.go.id/berita/utama/kepala-bmkg-pastikan-kesiapan-awos-llwas-dan-maws-hadapi-cuaca-ekstrem>
- [3] "Peraturan Badan Meteorologi, Klimatologi, dan Geofisika Nomor 2 Tahun 2015 Tentang Pelaksanaan Pengawasan Pelayanan Informasi Cuaca untuk Penerbangan," BMKG, Jakarta, Indonesia, Regulation, 2015. [Online]. Available: https://web-aviation.bmkg.go.id/storage/files/138/perbanno_2tahun2015.pdf
- [4] M. Lakatos and S. Baran, "Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts," *Meteorological Applications*, vol. 31, no. 6, p. e70015, Nov. 2024, doi: 10.1002/met.70015.
- [5] D. Perić, B. Livada, M. Perić, and S. Vujić, "Thermal Imager Range: Predictions, Expectations, and Reality," *Sensors*, vol. 19, no. 15, p. 3313, 2019.
- [6] C. J. Goodman and J. D. Small Griswold, "Meteorological Impacts on Commercial Aviation Delays and Cancellations in the Continental United States," *Journal of Applied Meteorology and Climatology*, vol. 58, no. 3, pp. 479–494, 2019.
- [7] J. R. Bradley, S. H. Holan, and C. K. Wikle, "Multivariate spatio-temporal models for high-dimensional areal data with application to Longitudinal Employer-Household Dynamics," *Ann. Appl. Stat.*, vol. 9, no. 4, pp. 1761–1791, 2015.
- [8] L. Olivetti and G. Messori, "Advances and prospects of deep learning for medium-range extreme weather forecasting," *Geosci. Model Dev.*, vol. 17, no. 6, pp. 2347–2358, 2024.
- [9] J. R. Bradley, C. K. Wikle, and S. H. Holan, "Spatio-temporal models for big multinomial data using the conditional multivariate logit-beta distribution," *J. Time Series Anal.*, vol. 40, no. 3, pp. 363–382, 2019.
- [10] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [11] H. T. A. Simanjuntak, A. Lumbanraja, G. Samosir, and Regita, "Prediksi Single-Step dan Multi-Step Data Cuaca Menggunakan Model Long Short-Term Memory dan Sarima," *JTIK*, vol. 12, no. 2, pp. 399–410, 2025.
- [12] A. Khumaidi, R. Raafi'udin, and I. P. Solihin, "Pengujian Algoritma Long Short Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung," *J. Telematika*, vol. 15, no. 1, pp. 13–18, 2020.
- [13] P. A. U. Sitompul, H. Sujaini, R. R. Yacoub, and N. Ananda, "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*, vol. 25, no. 1, pp. 33–42, 2025.
- [14] E. Supriyadi, "Prediksi Parameter Cuaca Menggunakan Deep Learning Long-Short Term Memory (LSTM)," *JMG*, vol. 21, no. 2, p. 55, 2021.
- [15] R. Akbar, R. Santoso, and B. Warsito, "Prediksi Tingkat Temperatur Kota Semarang Menggunakan Metode Long Short-Term Memory (LSTM)," *J. Gauss*, vol. 11, no. 4, pp. 572–579, 2023.
- [16] M. Lepot, J.-B. Aubin, and F. Clemens, "Interpolation in Time Series: An Introductory Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment," *Water*, vol. 9, no. 10, p. 796, 2017.
- [17] L. Chaliqi Taufiq, A. Putri, and F. Apriandy, "Simplified Spatial Wind Vector Interpolation Method for Airport Runway Orientation Analysis," *E3S Web Conf.*, vol. 476, p. 01057, 2024.
- [18] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Phil. Trans. R. Soc. A.*, vol. 374, no. 2065, p. 20150202, 2016.
- [19] D. Geng, H. Zhang, and H. Wu, "Short-Term Wind Speed Prediction Based on Principal Component Analysis and LSTM," *Applied Sciences*, vol. 10, no. 13, p. 4416, 2020.

- [20] Shimaila and S. Siddiqi, "PCA-SWF: A Principal Component Analysis-Based Stacked Model Approach for Weather Forecasting," *JISEM*, vol. 10, no. 2s, pp. 203–214, 2025.
- [21] H. B. Kibria, M. Nahiduzzaman, Md. O. F. Goni, M. Ahsan, and J. Haider, "An Ensemble Approach for the Prediction of Diabetes Mellitus Using a Soft Voting Classifier with an Explainable AI," *Sensors*, vol. 22, no. 19, p. 7268, 2022.
- [22] A. Kabori, C. Zargane, A. Guezaz, S. Benkirane, and M. Azrour, "AI-Driven Climate Analysis in a Semi-arid Region: Uncovering Warming Trends and Meteorological Shifts in Southeastern Morocco," *JDSIS*, 2026.
- [23] M. R. Pahlevi, K. Kusrini, and T. Hidayat, "Comparison of LSTM and GRU Models for Forex Prediction," *Sinkron*, vol. 8, no. 4, pp. 2254–2263, 2023.
- [24] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [25] R. Teixeira, A. Cerveira, E. J. S. Pires, and J. Baptista, "Enhancing Weather Forecasting Integrating LSTM and GA," *Applied Sciences*, vol. 14, no. 13, p. 5769, 2024.
- [26] J. Yuan, H. Dengxin, W. Yufeng, Y. Xueting, D. Huige, and Y. rokaya, "Attention mechanism based CNN-LSTM hybrid deep learning model for atmospheric ozone concentration prediction," *Sci Rep*, vol. 15, no. 1, p. 21260, 2025.
- [27] K. Zhou, C. Zhang, B. Xu, J. Huang, C. Li, and Y. Pei, "TE-LSTM: A Prediction Model for Temperature Based on Multivariate Time Series Data," *Remote Sensing*, vol. 16, no. 19, p. 3666, 2024.