

Komparasi *Long Short-Term Memory* dan *Polynomial Regression* untuk Prediksi Tinggi Muka Air Sungai Berbasis Data IoT 92 Hari

Kafka Kautsari¹, Joko Priambodo, S.T., M.T.²

^{1,2} Departemen Teknik Elektro Otomasi, Institut Teknologi Sepuluh Nopember
Email: ¹kautsarikafka@gmail.com, ²joko.priambodo@its.ac.id

*Joko Priambodo, S.T., M.T.

Abstrak

Prediksi dini tinggi muka air (TMA) sungai merupakan komponen kunci sistem peringatan banjir untuk mendukung mitigasi tepat waktu. Penelitian ini melakukan komparasi performa model Long Short-Term Memory (LSTM) terhadap baseline Polynomial Regression dan persistence pada tugas prediksi TMA satu jam ke depan, dengan memanfaatkan data IoT hasil deployment 92 hari di Sungai Kalibokor, Surabaya (1 November 2025 – 31 Januari 2026). Total 264.540 sample tinggi muka air pada resolusi 30 detik di-resampling menjadi 2.208 sample per jam, lalu di-split temporal 70/15/15 untuk train, validation, dan test. Model LSTM dua lapis (64 dan 32 unit) dilatih menggunakan sequence input 48 jam dengan optimasi Adam dan regularisasi dropout 0,2. Hasil evaluasi pada data test (332 sample, 18–31 Januari 2026) menunjukkan LSTM mencapai MAE 4,35 cm, RMSE 7,26 cm, dan R^2 0,953, kompetitif terhadap Polynomial Regression Orde 1 (MAE 3,67 cm, RMSE 7,12 cm, R^2 0,955) namun jauh lebih baik dari Polynomial Orde 2 yang mengalami overfitting (MAE 11,71 cm). Persistence baseline menghasilkan MAE 4,13 cm. Temuan ini mengindikasikan bahwa pada tugas prediksi univariate jangka pendek, polynomial regression orde rendah sudah cukup kuat, sementara LSTM memberikan fondasi arsitektural untuk pengembangan model multivariate dengan integrasi data curah hujan pada penelitian lanjutan.

Kata kunci: Deep Learning, Long Short-Term Memory, Polynomial Regression, Prediksi Banjir, Tinggi Muka Air.

Abstract

Early prediction of river water level (TMA) is a key component of flood warning systems supporting timely mitigation. This study compares the performance of Long Short-Term Memory (LSTM) against Polynomial Regression and persistence baselines for one-hour-ahead TMA prediction, leveraging IoT data from a 92-day deployment at the Kalibokor River, Surabaya (1 November 2025 – 31 January 2026). A total of 264,540 water level samples at 30-second resolution were resampled to 2,208 hourly samples and split temporally 70/15/15 for training, validation, and testing. A two-layer LSTM (64 and 32 units) was trained using 48-hour input sequences with Adam optimization and 0.2 dropout regularization. Evaluation on the test set (332 samples, 18–31 January 2026) shows LSTM achieved MAE 4.35 cm, RMSE 7.26 cm, and R^2 0.953, competitive with Polynomial Regression Order 1 (MAE 3.67 cm, RMSE 7.12 cm, R^2 0.955) but substantially better than Order 2 which overfit (MAE 11.71 cm). Persistence baseline yielded MAE 4.13 cm. These findings indicate that for short-term univariate prediction, low-order polynomial regression suffices, while LSTM provides an architectural foundation for multivariate model development with rainfall integration in subsequent research.

Keywords: Deep Learning, Flood Prediction, Long Short-Term Memory, Polynomial Regression, Water Level.

I. PENDAHULUAN

Banjir merupakan salah satu bencana hidrometeorologis yang sering terjadi di kawasan perkotaan dengan sistem drainase terbatas. Sungai Kalibokor di Surabaya Timur menjadi contoh saluran primer yang rutin mengalami luapan air pada musim hujan, dengan dampak banjir di wilayah hilir seperti Pakuwon City [1]. Sistem peringatan dini berbasis prediksi tinggi muka air (TMA) menjadi instrumen penting untuk memberikan waktu reaksi yang cukup bagi masyarakat dan otoritas dalam mengambil tindakan mitigasi sebelum kondisi mencapai ambang banjir.

Sebagai prasyarat sistem prediksi, telah dikembangkan platform monitoring multisensor berbasis Internet of Things (IoT) dengan catu daya surya off-grid yang menghasilkan data TMA 92 hari di Sungai Kalibokor [2]. Platform tersebut menyajikan TMA dengan datum dasar saluran sebagai variabel utama, namun belum dilengkapi model prediktif yang memanfaatkan pola historis untuk memperkirakan kondisi

ke depan. Penelitian ini menjadi tahap lanjutan dari pengembangan platform tersebut dengan fokus pada implementasi dan validasi model prediksi TMA jangka pendek.

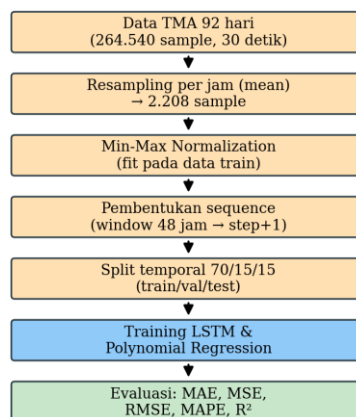
Berbagai pendekatan telah diterapkan untuk prediksi banjir. Mienye dan Swart [3] mengulas perkembangan arsitektur Recurrent Neural Network termasuk LSTM yang dirancang untuk mengatasi vanishing gradient pada data sekuensial. Rafael dkk. [4] mengembangkan sistem IoT terenkripsi dengan kombinasi LSTM, fuzzy logic, dan genetic algorithm untuk prediksi banjir di Filipina, dengan akurasi di atas 92%. Naufaldhianto dkk. [5] mendemonstrasikan rekonstruksi data sensor PLTMG menggunakan RNN dengan hasil yang superior dibandingkan interpolasi konvensional. Di sisi lain, Rezki dkk. [6] menunjukkan bahwa polynomial regression tetap menjadi baseline yang relevan untuk pemodelan hubungan non-linear pada data berdimensi rendah.

Pada konteks univariate forecasting dengan data resolusi menengah (hourly), masih terdapat pertanyaan apakah keunggulan LSTM yang dilaporkan pada studi multivariate juga muncul ketika hanya tersedia history variabel target tanpa predictor eksternal. Penelitian ini berkontribusi pada tiga aspek: (1) implementasi model LSTM dua lapis untuk prediksi TMA satu jam ke depan menggunakan data deployment riil 92 hari; (2) komparasi kuantitatif terhadap baseline persistence dan polynomial regression orde 1–2 dengan metrik MAE, RMSE, MAPE, dan R^2 ; serta (3) analisis kondisi di mana LSTM unggul dan tidak unggul, sebagai landasan pengembangan model multivariate pada penelitian lanjutan.

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental dengan tahapan akuisisi data, pra-pemrosesan, pembentukan sequence, pelatihan model, dan evaluasi terhadap data test yang belum dilihat selama pelatihan. Gambar 1 menampilkan keseluruhan tahapan metodologi.

Tahapan Metodologi Prediksi TMA



Gambar 1. Tahapan Metodologi Prediksi TMA

2.1 Sumber Data dan Pra-pemrosesan

Data primer berasal dari sensor ultrasonik A01NYUB pada platform IoT di Sungai Kalibokor, Surabaya, yang merekam jarak sensor ke permukaan air pada interval nominal 30 detik selama 92 hari (1 November 2025 – 31 Januari 2026), menghasilkan 264.540 sampel. Konversi ke TMA dilakukan dengan datum dasar saluran 458 cm sehingga $TMA = 458 - \text{distance}$, dengan clipping nilai negatif ke 0 untuk konsistensi fisik [2]. Pembacaan anomali ($\text{distance} = 0$ cm) sebanyak 4 sample dibuang.

Data 30 detik di-resampling menjadi resolusi per jam menggunakan operasi rata-rata, menghasilkan 2.208 sample. Resolusi hourly dipilih untuk dua alasan: (1) konsisten dengan skala waktu peringatan dini banjir (jeda 1–1,5 jam antara puncak hujan dan puncak TMA pada studi kasus 5 November 2025), dan (2) efisiensi komputasi untuk pelatihan model. Gap akibat data loss diisi dengan interpolasi linear.

2.2 Normalisasi dan Pembentukan Sequence

Min-max normalization diterapkan untuk menskalakan TMA ke rentang [0, 1] menggunakan Persamaan (1), dengan parameter Xmin dan Xmax dihitung hanya dari data train untuk menghindari data leakage.

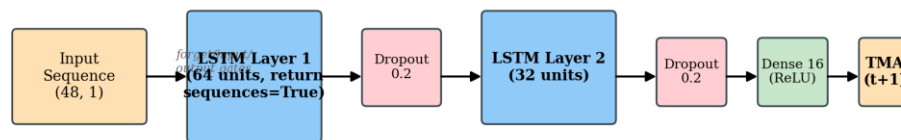
$$X' = (X - Xmin) / (Xmax - Xmin) \quad [1]$$

Sequence input dibentuk dengan sliding window 48 jam ke belakang sebagai input, dan TMA satu jam ke depan sebagai target prediksi. Pemilihan panjang sequence 48 jam mengakomodasi siklus diurnal lengkap dua hari. Total sequence yang dihasilkan adalah 2.184, di-split temporal 70/15/15 menjadi 1.521 train, 331 validation, dan 332 test, sehingga data test (18–31 Januari 2026) sepenuhnya unseen selama pelatihan.

2.3 Arsitektur Model LSTM

Model LSTM dua lapis digunakan sebagai model utama, dengan arsitektur seperti pada Gambar 2. Lapis pertama berisi 64 unit dengan return_sequences=True untuk meneruskan output sequence ke lapis kedua yang berisi 32 unit. Tiap lapis LSTM diikuti oleh dropout layer dengan rate 0,2 sebagai regularisasi untuk mengurangi overfitting. Output dari lapis kedua diteruskan ke dense layer 16 unit dengan aktivasi ReLU, lalu ke dense layer akhir 1 unit untuk menghasilkan prediksi TMA pada t+1.

Arsitektur Model LSTM Dua Lapis untuk Prediksi TMA



Loss: MSE | Optimizer: Adam (lr=0.001) | Batch size: 32 | Early stopping (patience=20)

Total parameter terlatih: ~21.000

Gambar 2. Arsitektur Model LSTM Dua Lapis untuk Prediksi TMA

Tiap unit LSTM beroperasi berdasarkan tiga gerbang utama forget gate, input gate, dan output gate yang mengontrol aliran informasi melalui cell state, sehingga model dapat menangkap dependensi temporal jangka panjang tanpa mengalami vanishing gradient [3]. Total parameter terlatih sekitar 21.000.

2.4 Model Polynomial Regression sebagai Baseline

Sebagai pembandingan, model polynomial regression dilatih dengan fitur 48 lag TMA sebagai input dan TMA pada t+1 sebagai target. Penting dicatat bahwa istilah polynomial regression orde 1 di sini mengacu pada regresi linear multivariate terhadap parameter, bukan terhadap waktu — output prediksi tetap dinamis karena input lag berubah sepanjang waktu. Persamaan polynomial regression orde 1 disajikan pada Persamaan (2), sedangkan orde 2 pada Persamaan (3).

$$\hat{y}(t+1) = \beta_0 + \sum_{i=1}^{48} \beta_i \cdot x(t-i+1) \quad [2]$$

$$\hat{y}(t+1) = \beta_0 + \sum_i \beta_i \cdot x_i + \sum_{ii} \beta_{ii} \cdot x_i^2 + \sum_{ij} \beta_{ij} \cdot x_i \cdot x_j \quad [3]$$

Dengan $\hat{y}(t+1)$ adalah prediksi TMA pada t+1, x_i adalah nilai TMA pada lag ke-i (dalam range $i = 1$ hingga 48 sehingga $x_1 = \text{TMA}(t)$, $x_2 = \text{TMA}(t-1)$, dan seterusnya), β_0 adalah intercept, dan β_i serta β_{ij} adalah koefisien regresi yang dipelajari dengan ordinary least squares. Jumlah parameter untuk orde 1 adalah 49 (48 koefisien linear + 1 intercept). Untuk orde 2, jumlah parameter melonjak menjadi 1.225 yang terdiri atas 48 koefisien linear, 48 koefisien kuadratik, 1.128 koefisien interaksi pasangan, dan 1 intercept — rasio parameter terhadap sampel train ($1.225 / 1.521 \approx 0,81$) sudah sangat dekat dengan jumlah sampel sehingga

rentan overfitting. Pemilihan orde 1–2 berdasarkan kompromi ini; orde 3 dengan 48 fitur menghasilkan lebih dari 20.000 parameter interaksi yang tidak praktis. Persistence baseline (prediksi = nilai terakhir) juga disertakan sebagai referensi paling sederhana.

2.5 Konfigurasi Pelatihan dan Metrik Evaluasi

Konfigurasi hyperparameter LSTM disajikan pada Tabel 1. Pelatihan menggunakan optimizer Adam dengan learning rate 0,001 dan loss function Mean Squared Error (MSE). Early stopping dengan patience 20 epoch dan restoration of best weights mencegah overfitting.

Tabel 1. Konfigurasi Hyperparameter Model LSTM

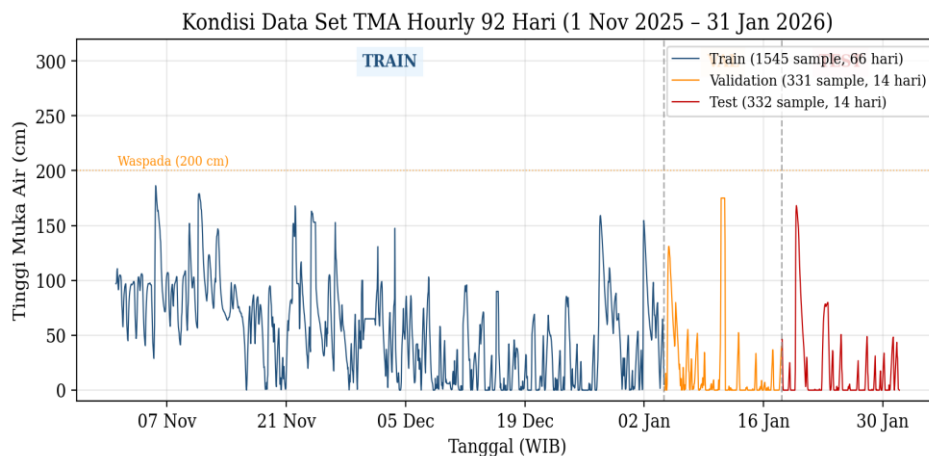
Parameter	Nilai
Sequence length (window)	48 jam
LSTM layer 1 / layer 2	64 / 32 units
Dropout rate	0,2
Dense layer	16 (ReLU) → 1 (linear)
Optimizer / Learning rate	Adam / 0,001
Loss function	Mean Squared Error
Batch size	32
Maksimum epoch / Patience	150 / 20

Empat metrik evaluasi digunakan untuk membandingkan performa model: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) dengan filter TMA > 5 cm untuk menghindari division by near-zero, dan koefisien determinasi (R^2). Semua metrik dihitung pada data test yang sudah didenormalisasi ke skala asli (cm).

III. HASIL DAN PEMBAHASAN

3.1 Karakteristik Data Set untuk Pelatihan dan Pengujian

Kondisi data set TMA hourly yang digunakan untuk pelatihan dan pengujian model ditampilkan pada Gambar 3, dengan pembagian temporal train, validation, dan test ditandai dengan warna berbeda. Pembagian temporal 70/15/15 menghasilkan partisi train (1 November – 4 Januari, 1.545 sample) yang mencakup periode awal musim hujan dengan banyak event TMA tinggi, partisi validation (5–18 Januari, 331 sample) sebagai periode tuning hyperparameter, dan partisi test (18–31 Januari, 332 sample) sebagai data unseen untuk evaluasi akhir.



Gambar 3. Kondisi Data Set TMA Hourly 92 Hari dengan Pembagian Train/Validation/Test

Statistik deskriptif tiap partisi disajikan pada Tabel 3. Terlihat bahwa partisi train memiliki rata-rata TMA yang lebih tinggi (54,0 cm) dibanding val (20,0 cm) dan test (18,0 cm), mencerminkan transisi alami

dari periode puncak musim hujan November ke fase yang lebih kering di Januari. Walaupun terjadi distribution shift antar partisi, peak TMA pada test (168 cm) masih berada dalam range yang ter-cover oleh train (max 186 cm), sehingga model tidak diminta melakukan ekstrapolasi di luar rentang training.

Tabel 2. Statistik Deskriptif TMA per Partisi Data Set

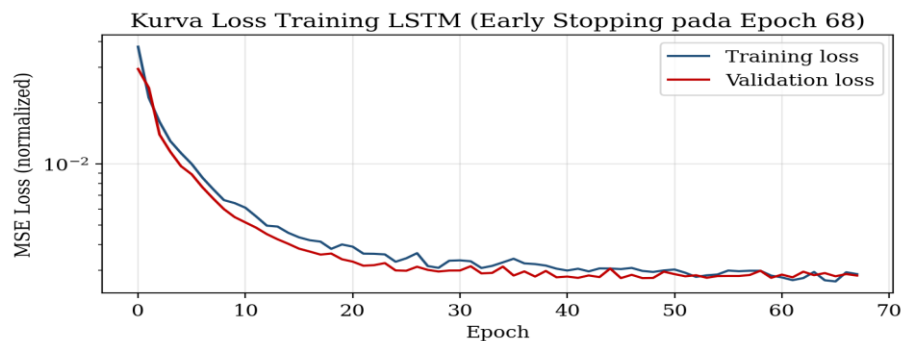
Partisi	Periode	Sample	Mean (cm)	Std (cm)	Min (cm)	Max (cm)
Train	1 Nov – 4 Jan	1.545	54,0	42,0	0	186
Validation	5 – 18 Jan	331	20,0	39,3	0	175
Test	18 – 31 Jan	332	18,0	33,5	0	168
Total	1 Nov – 31 Jan	2.208	37,8	42,9	0	186

Karakteristik dataset menunjukkan tiga properti penting untuk model prediksi: (1) TMA bersifat highly skewed dengan median jauh di bawah mean, banyak nilai mendekati 0 saat saluran kering, dan ekor distribusi panjang pada event hujan; (2) variabilitas tinggi (std hingga 42 cm) menunjukkan dinamika hidrologis yang tidak homogen; serta (3) tidak ada event yang mencapai ambang banjir 315 cm selama periode pengamatan, sehingga model dilatih dan dievaluasi pada kondisi non-banjir.

Seluruh partisi tidak mengandung event banjir aktual, sehingga klaim performa model terbatas pada prediksi TMA dalam kondisi normal hingga waspada ($TMA < 200$ cm). Generalisasi terhadap kondisi banjir nyata memerlukan validasi pada data dengan event ekstrim yang menjadi agenda penelitian lanjutan.

3.2 Hasil Pelatihan LSTM

Kurva loss selama pelatihan ditunjukkan pada Gambar 4. Training loss dan validation loss menurun secara konsisten dari nilai awal $\sim 0,05$ hingga konvergen pada $\sim 0,003$ pada epoch 30, kemudian datar hingga early stopping mengaktifkan restorasi bobot terbaik pada epoch 68. Tidak teramati gap signifikan antara training dan validation loss, menunjukkan model tidak mengalami overfitting selama pelatihan.



Gambar 4. Kurva Loss Training LSTM

3.3 Perbandingan Metrik Performa

Hasil evaluasi keempat model pada data test (332 sample, 18–31 Januari 2026) disajikan pada Tabel 3. LSTM dua lapis mencapai MAE 4,35 cm, RMSE 7,26 cm, MAPE 26,70%, dan R^2 0,953, kompetitif dengan Polynomial Regression Orde 1 yang memberikan hasil terbaik sedikit (MAE 3,67 cm, R^2 0,955). Polynomial Orde 2 mengalami overfitting parah dengan MAE 11,71 cm, sedangkan persistence baseline (prediksi = nilai sekarang) menghasilkan MAE 4,13 cm.

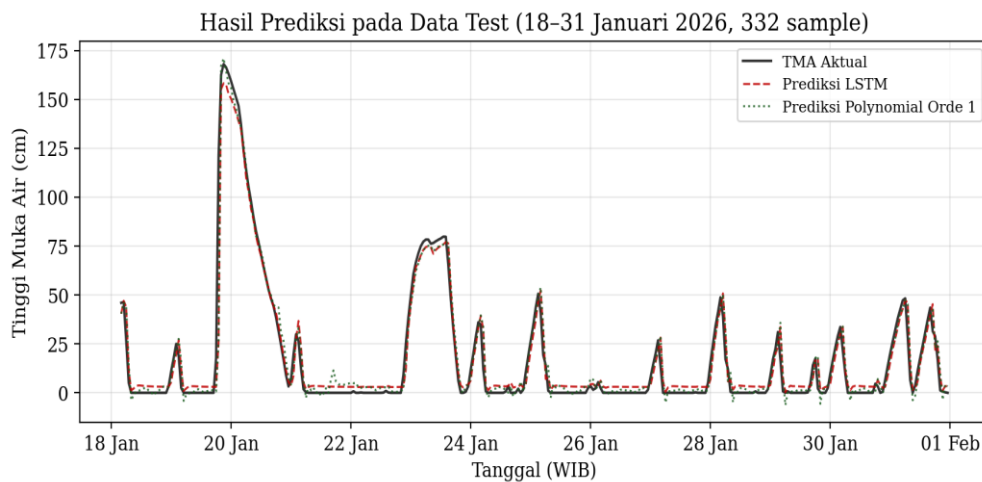
Tabel 3. Perbandingan Metrik Evaluasi pada Data Test

Model	MAE (cm)	RMSE (cm)	MAPE (%)	R^2
Persistence (naive)	4,13	8,75	41,62	—
Polynomial Orde 1	3,67	7,12	26,11	0,955
Polynomial Orde 2	11,71	19,82	56,99	—
LSTM 2-layer	4,35	7,26	26,70	0,953

Polynomial Orde 1 dan LSTM menempati kelompok performa atas dengan selisih MAE hanya 0,68 cm dan selisih R^2 0,002 — secara praktis dapat dikatakan setara. Polynomial Orde 2 menunjukkan anti-pattern klasik overfitting: dengan 48 fitur lag dan ekspansi orde 2 menghasilkan jumlah fitur interaksi yang besar, model menjadi terlalu fleksibel sehingga memorisasi noise pada data train tanpa generalisasi ke test. Persistence baseline cukup kompetitif untuk MAE namun jauh lebih buruk pada MAPE dan RMSE, menandakan kelemahannya pada event dengan perubahan TMA cepat.

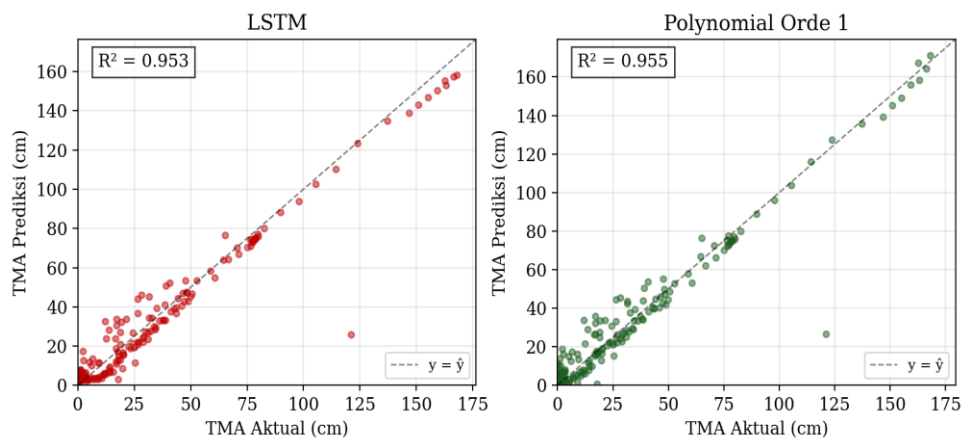
3.4 Visualisasi Prediksi pada Data Test

Gambar 5 menampilkan prediksi LSTM dan Polynomial Orde 1 berdampingan dengan TMA aktual untuk seluruh periode test 14 hari. Kedua model berhasil menangkap dinamika utama termasuk event puncak pada 20 Januari (TMA ~170 cm) serta event sekunder pada 23 dan 25 Januari. Pada periode TMA rendah (di bawah 10 cm), prediksi Polynomial Orde 1 menunjukkan beberapa nilai negatif kecil yang merupakan artefak ekstrapolasi linear, sedangkan LSTM cenderung menjaga prediksi tetap positif berkat aktivasi pada lapis dense.



Gambar 5. Prediksi TMA vs Aktual pada Data Test (18–31 Januari 2026)

Gambar 6 menyajikan scatter plot prediksi vs aktual untuk kedua model. Kedua scatter menunjukkan pola yang sangat dekat dengan garis identity ($y = \hat{y}$), dengan dispersi sedikit lebih tinggi pada rentang TMA rendah (0–50 cm). Satu outlier ekstrim terlihat pada kedua model di sekitar TMA aktual 122 cm yang diprediksi sekitar 25 cm — kemungkinan merupakan event cepat yang tidak dapat di-anticipate dari history 48 jam.



Gambar 6. Scatter Plot Prediksi vs TMA Aktual untuk LSTM dan Polynomial Orde 1

3.5 Diskusi: Mengapa LSTM Tidak Unggul Signifikan?

Hasil komparasi memberikan temuan yang berlawanan dengan ekspektasi umum bahwa LSTM selalu superior pada data sekuensial. Beberapa faktor menjelaskan ini: (1) tugas adalah univariate prediksi 1 jam ke depan, di mana sebagian besar variansi TMA($t+1$) dapat dijelaskan oleh kombinasi linear lag TMA($t-i$) sehingga model linear sudah cukup; (2) data resolusi hourly tidak mengandung dinamika yang membutuhkan memori jangka panjang; (3) ukuran dataset (1.521 sample train) relatif kecil untuk fully leverage kapasitas LSTM yang memiliki 21.000 parameter.

Polynomial Orde 1 yang setara dengan regresi linear pada 48 fitur lag hanya memiliki 49 parameter sesuai Persamaan (2), sehingga sample efficiency-nya jauh lebih tinggi dibanding Polynomial Orde 2 (1.225 parameter) maupun LSTM (~21.000 parameter). Hasil ini sejalan dengan prinsip Occam's razor: pada problem yang well-modeled oleh hubungan linear, model linear menghasilkan generalisasi yang lebih baik daripada model non-linear yang lebih kompleks.

3.6 Implikasi untuk Pengembangan Multivariate

Walaupun belum unggul pada setup univariate ini, LSTM tetap menjadi pilihan arsitektural yang strategis untuk pengembangan berikutnya karena tiga alasan. Pertama, LSTM secara natural mendukung input multivariate (TMA + curah hujan + suhu + kelembapan) tanpa explosion of features seperti pada polynomial regression. Kedua, LSTM dapat di-extend ke multi-step ahead prediction (misal prediksi 6 jam ke depan) yang penting untuk window mitigasi lebih panjang. Ketiga, mekanisme gating LSTM cocok untuk menangkap interaksi non-linear antara variabel meteorologis dan respons TMA yang tidak dapat di-model oleh polynomial orde rendah.

Pengembangan berikutnya akan mengintegrasikan data curah hujan dari rain gauge SEN0186 yang juga berada pada platform IoT, sehingga model menjadi multivariate dengan input TMA + rain1h + rain24h. Diharapkan keunggulan LSTM atas baseline linear akan menjadi lebih signifikan dalam setup ini.

IV. KESIMPULAN

Komparasi model Long Short-Term Memory dan Polynomial Regression untuk prediksi TMA satu jam ke depan berbasis data IoT 92 hari di Sungai Kalibokor telah berhasil dilakukan. LSTM dua lapis (MAE 4,35 cm, RMSE 7,26 cm, R^2 0,953) menunjukkan performa kompetitif dengan Polynomial Regression Orde 1 (MAE 3,67 cm, RMSE 7,12 cm, R^2 0,955), sementara Polynomial Orde 2 mengalami overfitting (MAE 11,71 cm). Pada tugas univariate jangka pendek dengan resolusi hourly, model linear sudah cukup untuk menangkap variansi TMA, dan kapasitas tambahan LSTM tidak memberikan keuntungan signifikan.

LSTM tetap menjadi pilihan arsitektural strategis untuk pengembangan model multivariate dengan integrasi curah hujan dan parameter cuaca lain, serta untuk multi-step ahead prediction yang menjadi agenda penelitian lanjutan. Validasi sistem terhadap event banjir aktual juga memerlukan data dengan kondisi banjir nyata, yang tidak terjadi selama periode deployment 92 hari ini.

V. UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Departemen Teknik Elektro Otomasi, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember Surabaya atas dukungan fasilitas dan pembiayaan penelitian ini.

REFERENSI

- [1] D. Anugrananto, "Perencanaan Jaringan Drainase Sub Sistem Kali Bokor Surabaya," Tugas Akhir, Departemen Teknik Sipil, Institut Teknologi Sepuluh Nopember, Surabaya, 2016.
- [2] K. Kautsari dan J. Priambodo, "Sistem Monitoring Multisensor Sungai Berbasis IoT dengan Catu Daya Surya untuk Mitigasi Dini Banjir di Surabaya," Prosiding Seminar Nasional Karsa Nusantara, vol. 3, 2026 (dalam proses publikasi).
- [3] I. D. Mienye dan T. G. Swart, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," Information, vol. 15, 2024.

-
- [4] H. Rafael, C. Arante, E. Sybingco, M. A. Roque, L. Ambata, A. Chua, dan A. N. Gutierrez, “Development of a Secured IoT-Based Flood Monitoring and Forecasting System Using Genetic-Algorithm-Based Neuro-Fuzzy Network,” *IEEE Access*, vol. 13, hal. 1–19, 2025.
 - [5] E. N. F. Naufaldhianto, B. Al Kindhi, dan I. A. Wicaksono, “Rekonstruksi Data Suhu dan Tekanan pada Pembangkit Listrik Tenaga Mesin Gas (PLTMG) Menggunakan Recurrent Neural Network,” *Prosiding Seminar Nasional Karsa Nusantara*, vol. 2, hal. 342–350, 2025.
 - [6] H. Rezkian, A. Dama, A. A. Supianto, dan N. Y. Setiawan, “Analisis Penggunaan Model Regresi untuk Prediksi Penjualan Spare Part,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 12, hal. 5591–5603, 2021.
 - [7] S. Zhu, J. Miao, W. Chen, C. Liu, C. Weng, dan Y. Luo, “Reconstructing Missing Data Using a Bi-LSTM Model Based on VMD and SSA for Structural Health Monitoring,” *Sensors*, vol. 24, 2024.
 - [8] W. Anbiya dan F. C. Garini, “Application of GARCH Forecasting Method in Predicting The Number of Rail Passengers,” *Jurnal Matematika Statistika dan Komputasi*, vol. 18, no. 2, hal. 198–223, 2022.
 - [9] F. I. Maulana dan D. K. Hakim, “Implementasi Algoritma Prediksi Time Series untuk Konsumsi Energi Listrik,” *Jurnal Teknik Informatika*, 2023.
 - [10] H. Zhao, Z. Lai, H. Leung, dan X. Zhang, *Feature Learning and Understanding: Algorithms and Applications*. Springer International Publishing, 2020.