

PENINGKATAN AKURASI PREDIKSI CURAH HUJAN HARIAN MENGGUNAKAN LSTM-PSO BERBASIS DATA BMKG STASIUN GEOFISIKA PASURUAN

Muhammad Reyhan Dwi Fisena, Eva Yulia Puspaningrum, Ardhon Rakhmadi

Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur

Jalan Rungkut Madya Surabaya, Indonesia

evapuspaningrum.if@upnjatim.ac.id

ABSTRAK

Kabupaten Pasuruan merupakan salah satu daerah dengan risiko tinggi bencana hidrometeorologi di Jawa Timur. Keterbatasan model prediksi konvensional dalam menangkap pola curah hujan yang kompleks dan non-linear, serta proses *hyperparameter tuning* yang tidak efisien, mendorong pengembangan model yang lebih adaptif. Penelitian ini bertujuan mengembangkan model prediksi curah hujan harian berbasis integrasi *Long Short-Term Memory* (LSTM) dan *Particle Swarm Optimization* (PSO) menggunakan data meteorologi BMKG Stasiun Geofisika Pasuruan periode 2021–2025 dengan total 1.815 data pengamatan. Variabel yang digunakan meliputi curah hujan, suhu udara rata-rata, dan kelembapan udara rata-rata. Prapemrosesan data mencakup deteksi *outlier* menggunakan *Interquartile Range* (IQR), imputasi linear, normalisasi Min-Max, dan transformasi *sliding window* 7 hari. PSO dikonfigurasi dengan 20 partikel dan 30 iterasi untuk mengoptimasi *hyperparameter* utama LSTM. Hasil evaluasi menunjukkan model LSTM-PSO mengungguli LSTM *baseline* pada seluruh metrik evaluasi, dengan peningkatan MAE sebesar 7,54%, RMSE sebesar 3,62%, serta NSE dan R^2 sebesar 76,85%. Optimasi *hyperparameter* berbasis PSO terbukti mampu meningkatkan akurasi dan stabilitas model LSTM dalam memprediksi curah hujan harian di Kabupaten Pasuruan.

Kata kunci : Curah Hujan, Deep Learning, Hyperparameter Optimization, LSTM, PSO, Prediksi

1. PENDAHULUAN

Kabupaten Pasuruan termasuk salah satu daerah dengan label zona merah terhadap bencana hidrometeorologi di Jawa Timur[1]. Peningkatan kerentanan bencana hidrometeorologi disebabkan oleh pola curah hujan yang semakin sulit diprediksi dalam beberapa tahun terakhir[2]. Kondisi tersebut juga dipengaruhi oleh kondisi topografi yang beragam dan kompleks di Kabupaten Pasuruan. Kondisi ini menegaskan bahwa ketersediaan model prediksi curah hujan yang akurat dan andal menjadi sangat penting sebagai dasar mitigasi bencana dan pengambilan keputusan.

Berbagai pendekatan prediksi curah hujan telah dikembangkan. Mulai dari metode statistik konvensional seperti ARIMA telah digunakan dalam peramalan deret waktu[3] Namun, metode tersebut memiliki keterbatasan dalam menangkap pola non-linear dan hubungan temporal yang kompleks pada data meteorologi. Untuk mengatasi permasalahan tersebut pendekatan *deep learning* menjadi solusi yang sudah banyak diterapkan khususnya metode *Long Short-Term Memory* (LSTM)[4]. LSTM merupakan pengembangan dari arsitektur *Recurrent Neural Network* (RNN) yang dirancang untuk mempelajari hubungan jangka panjang pada data *time series*[5]. Terbukti oleh penelitian dari Badriyah et al. (2022) LSTM mampu memodelkan data curah hujan multiatribut di Indonesia.

Namun, keterbatasan utama LSTM konvensional terletak pada proses *hyperparameter tuning* atau penentuan *hyperparameter* pada LSTM yang masih bersifat manual atau berbasis *grid search*. Kedua pendekatan tersebut cenderung kurang efisien dan

tidak menjamin solusi optimal secara global. *Particle Swarm Optimization* sebagai algoritma optimasi mampu mengatasi permasalahan tersebut. PSO terinspirasi dari perilaku kawanan burung dan ikan, dan terbukti mampu mencari solusi optimal secara adaptif tanpa memerlukan informasi turunan fungsi objektif[6]. Penelitian oleh Umamaheswari (2025) membuktikan integrasi LSTM-PSO mampu menghasilkan nilai metrik evaluasi MAE=16,43 dan RMSE=20,78, melampaui performa LSTM standar[7].

Penelitian ini mengembangkan model LSTM-PSO untuk meningkatkan akurasi prediksi curah hujan di Kabupaten Pasuruan menggunakan data meteorologi BMKG Stasiun Geofisika Pasuruan periode 2021 hingga 2025. Tujuan spesifik dari penelitian ini meliputi, membangun model LSTM *baseline* sebagai pembandingan, menerapkan PSO untuk mengoptimasi *hyperparameter* LSTM secara sistematis, dan mengevaluasi performa model menggunakan metrik evaluasi MAE, RMSE, NSE dan R^2 . Untuk mengatasi keterbatasan tersebut, penelitian ini mengintegrasikan PSO sebagai optimasi *hyperparameter* pada model LSTM sehingga kontribusi utama penelitian ini meliputi optimasi PSO *hyperparameter* LSTM pada data lokal, evaluasi komprehensif menggunakan metrik evaluasi, serta analisis perbandingan antara model *baseline* dengan model LSTM-PSO dengan *hyperparameter* teroptimasi.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Berbagai penelitian terdahulu telah mengkaji penggunaan metode *deep learning* dan optimasi untuk

prediksi curah hujan. Badriyah menerapkan LSTM multiatribut untuk memprediksi curah hujan di Kota Surabaya menggunakan data periode 2015-2020. Penelitian tersebut memperoleh nilai MSE sebesar 0,489, MAE sebesar 0,537, dan R2 sebesar 0,497, yang menunjukkan bahwa LSTM mampu memodelkan data deret waktu curah hujan meskipun akurasi masih dapat ditingkatkan[4].

Patro dan Bartakke mengembangkan model LSTM untuk prediksi curah hujan harian berbasis data *Automatic Weather Station* (AWS). Model yang dikembangkan menggunakan optimizer Adam terbukti mampu menangkap pola temporal dengan baik, namun belum menerapkan optimasi *hyperparameter* berbasis *swarm intelligence*[5].

Terkait integrasi LSTM dan PSO, Umamaheswari mengembangkan model hibrida LSTM-PSO untuk prediksi curah hujan di wilayah Cuddalore, India, menggunakan data selama 41 tahun dari *NASA Power ACCESS*. PSO digunakan untuk mengoptimasi *hyperparameter* seperti jumlah neuron, *learning rate*, dan *batch size*. Hasil evaluasi menunjukkan PSO-LSTM unggul dibandingkan dengan model lain dengan nilai MAE sebesar 16,43, RMSE sebesar 20,78 dan akurasi 99,78%, membuktikan bahwa integrasi PSO dalam prediksi curah hujan secara signifikan meningkatkan performa model LSTM[7].

Berdasarkan penelitian terdahulu, belum ditemukan penelitian yang secara khusus menerapkan optimasi PSO pada model LSTM untuk memprediksi curah hujan harian menggunakan data observasi resmi BMKG di Kabupaten Pasuruan. Celah penelitian ini mendorong pengembangan model hibrida LSTM-PSO yang disesuaikan dengan karakteristik iklim lokal Kabupaten Pasuruan, yang diharapkan dapat menghasilkan prediksi curah hujan yang lebih akurat dan mendukung upaya mitigasi bencana hidrometeorologi.

2.2. Data Meteorologi

Data meteorologi adalah data yang merepresentasikan kondisi atmosfer dan berperan sebagai dasar dalam penyusunan model prediksi hujan[8]. Dalam penelitian ini data diperoleh melalui Badan Meteorologi Klimatologi dan Geofisika (BMKG) Stasiun Geofisika Pasuruan, yang menyediakan data pengamatan harian yang terstandarisasi sesuai prosedur *World Meteorological Organization* (WMO)[9].

Data yang digunakan mencakup tiga variabel utama, curah hujan (mm) sebagai target prediksi, suhu udara rata-rata (C), dan kelembapan udara rata-rata (%) sebagai variabel input. Ketiga variabel tersebut dipilih karena memiliki hubungan dalam pembentukan hujan, variabel suhu memengaruhi proses penguapan, sedangkan kelembapan menunjukkan tingkat kejenuhan uap air di atmosfer yang berperan dalam kondensasi[10]. Data meteorologi dalam penelitian ini termasuk ke dalam kategori data deret waktu.

Dengan karakteristik tersebut, data meteorologi yang digunakan tidak hanya merepresentasikan kondisi atmosfer secara historis namun, menjadi dasar dalam membangun model prediksi curah hujan yang mampu menangkap pola non-linear dan dependensi jangka panjang melalui integrasi LSTM-PSO.

2.3. Long Short-Term Memory (LSTM)

Long Short-Term Memory diperkenalkan oleh Hochreiter dan Schmidhuber (1997) sebagai solusi permasalahan *vanishing gradient* pada *Recurrent Neural Network* (RNN) konvensional[11]. LSTM memiliki kemampuan dalam menangkap dependensi jangka panjang melalui struktur *cell state* yang berfungsi sebagai media penyimpanan informasi.

Arsitektur LSTM terdiri dari tiga komponen utama berupa *gating mechanism* yaitu, *forget gate*, *input gate*, *output gate* yang bekerja secara terintegrasi untuk mengatur aliran informasi[12]. *Forget gate* berfungsi menentukan informasi yang dipertahankan atau dibuang dari *cell state*, dengan persamaan matematis sebagai berikut:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Input gate mengontrol informasi baru yang akan ditambahkan ke dalam *cell state*, dengan persamaan:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\bar{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

Selanjutnya, *output gate* menentukan informasi yang akan dikeluarkan sebagai *hidden state*:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (5)$$

Melalui mekanisme tersebut LSTM mampu mempertahankan dan memodelkan pola non-linear pada data deret waktu.

2.4. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) merupakan metode optimasi yang bekerja menggunakan pendekatan populasi partikel, diperkenalkan oleh Kennedy dan Eberhart (1995), terinspirasi dari perilaku sosial kawanan burung atau ikan dalam mencari solusi optimal[6]. Dalam PSO, setiap partikel menjadi kandidat solusi yang bergerak dalam ruang pencarian dengan mempertimbangkan pengalaman individu (*personal best/pBest*) dan pengalaman global (*global best/gBest*).

Proses pembaruan posisi partikel dilakukan secara iteratif melalui pembaruan kecepatan dan posisi yang dirumuskan secara matematis:

$$v_i^{(t+1)} = w \cdot v_i^{(t)} + c_1 \cdot r_1 \cdot (pbest_i - x_i^{(t)}) \quad (6)$$

$$+ c_2 \cdot r_2 \cdot (gbest - x_i^{(t)})$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \quad (7)$$

Dengan persamaan (6) merupakan pembaruan kecepatan partikel dan (7) sebagai pembaruan posisi partikel. Di mana w merupakan bobot inersia yang mengontrol keseimbangan antara eksplorasi dan

eksploitasi, c_1 adalah koefisien kognitif yang merepresentasikan pengaruh pengalaman individu, dan c_2 adalah koefisien sosial yang merepresentasikan pengaruh pengalaman kolektif. Nilai r_1 dan r_2 merupakan bilangan acak rentang $[0,1]$ [13]

2.5. Metrik Evaluasi

Evaluasi performa model dilakukan menggunakan empat metrik yaitu, *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *NASH-Sutcliffe Efficiency* (NSE), dan koefisien determinasi (R^2). Keempat metrik ini digunakan untuk mengukur tingkat kesalahan prediksi serta kemampuan model dalam merepresentasikan data aktual [14].

Mean Absolute Error (MAE) digunakan untuk mengukur rata-rata kesalahan absolut antara nilai aktual dan prediksi, yang dinyatakan sebagai:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (8)$$

Root Mean Squared Error (RMSE) memberikan penalti lebih besar terhadap kesalahan yang besar, sehingga sensitif terhadap *outlier*:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (9)$$

Nash-Sutcliffe Efficiency (NSE) digunakan untuk mengevaluasi tingkat kesesuaian model terhadap data aktual, dengan nilai 1 menunjukkan performa terbaik:

$$NSE = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (10)$$

Koefisien determinasi (R^2) digunakan untuk mengukur proporsi variansi data yang dapat dijelaskan model. Penggunaan kombinasi metrik ini bertujuan untuk memberikan evaluasi yang komprehensif terhadap akurasi dan stabilitas model prediksi curah hujan yang dihasilkan [15]

3. METODE PENELITIAN

3.1. Data Penelitian

Data yang digunakan pada penelitian ini berasal dari situs resmi Badan Meteorologi Klimatologi dan Geofisika (BMKG) sebagai penyedia data meteorologi di Indonesia. Data meteorologi dapat diakses secara publik melalui *website* dataonline.bmkg.go.id.

1		ID WMO : 96945		
2		NAMA STA : Stasiun Geofisika Pasuruan		
3		LINTANG : -7.70456		
4		BUJUR : 112.63533		
5		ELEVASI : 832 Meter		
6				
7				
8		TANGGAL	TAVG	RH_AVG
9		01-12-2025	23.1	9999
10		02-12-2025	21.8	9999
11		03-12-2025	21.5	9999
12		04-12-2025	21.3	9999
13		05-12-2025	20.7	9999
14		06-12-2025	21.4	9999
15		07-12-2025	21.6	9999

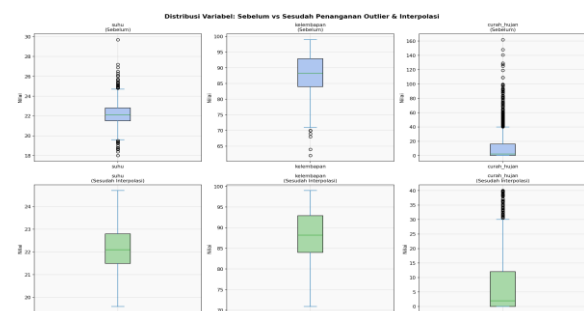
Gambar 1. Data Mentah BMKG

Pada gambar 1 terdapat contoh data mentah yang telah diambil namun, masih memiliki banyak notasi khusus yang perlu dilakukan penanganan agar model dapat memprediksi dengan baik. Pada penelitian ini menggunakan data yang berasal dari UPT Stasiun Geofisika Pasuruan dengan periode pengamatan 1 Januari 2021 hingga 31 Desember 2025, dengan total pengamatan data sebanyak 1.815 hari. Terdapat tiga variabel meteorologi yang digunakan curah hujan (RR,mm), kelembapan udara rata-rata harian (RH_AVG,%), dan suhu rata-rata harian (TAVG,C).

3.2. Prapemrosesan data

Data mentah dari BMKG masih mengandung nilai hilang dengan notasi khusus 8888 dan 9999 yang akan dikonversi ke dalam format *missing value* (NaN). Selanjutnya dilakukan proses untuk meningkatkan kualitas data.

Langkah awal pada prapemrosesan data dengan mendeteksi dan penanganan *outlier* menggunakan metode *Interquartile Range* (IQR), di mana data di luar rentang diubah menjadi NaN. Proses ini menghasilkan deteksi *outlier* pada variabel suhu sebanyak 61 data (3,36%), kelembapan sebanyak 11 data (0,16%), dan curah hujan 163 data (8,98%).



Gambar 2. Distribusi Variabel

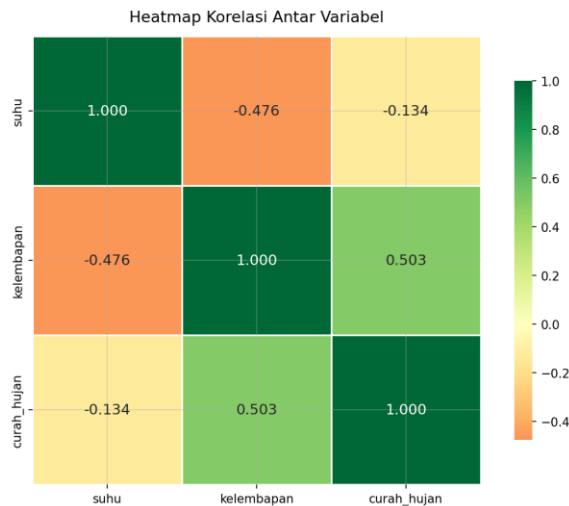
Kedua, dilakukan imputasi untuk menangani *missing value* sehingga dataset menjadi lengkap. Ketiga, normalisasi data menggunakan metode *Min-Max* ke rentang $[0-1]$ untuk menstabilkan proses pelatihan model.

Proses terakhir, data ditransformasikan menggunakan metode *sliding window* dengan *timestep* 7 hari untuk mengubah data deret waktu menjadi format *supervised learning*. Proses ini menghasilkan sebanyak 1808 sampel dari total data 1815.

3.3. Analisis Data Eksploratif

Analisis data eksploratif dilakukan untuk memahami karakteristik dan distribusi data sebelum pemodelan. Analisis meliputi statistik deskriptif, visualisasi *time series*, analisis korelasi antar variabel, serta uji stasioneritas.

Berdasarkan gambar 3, kelembapan memiliki korelasi positif signifikan dengan curah hujan. Hasil uji stasioneritas *Augmented Dickey-Fuller* (ADF) menunjukkan seluruh variabel bersifat stasioner.



Gambar 3. Heatmap Korelasi antar Variabel

3.4. Pembagian Data

Data dibagi secara kronologis (tanpa *shuffling*) dengan rasio 80:20 untuk menjaga dependensi temporal. Data latih menggunakan 1446 sampel (08-01-2021 hingga 03-01-2025) dan data uji 362 sampel (04-01-2025 hingga 31-12-2025). Pembagian secara kronologis ini penting karena model LSTM belajar dari pola temporal yang berurutan, dan pengacakan data akan merusak struktur temporal yang dibutuhkan model.

3.5. Arsitektur Model LSTM

Model yang digunakan dalam penelitian ini berupa *stacked LSTM* yang terdiri dari dua lapisan LSTM dengan penambahan *dropout layer* untuk mencegah *overfitting* serta, *dense layer* sebagai lapisan output. Model menerima input dalam bentuk (*sample, timestep, fitur*).

Dua skenario model dikembangkan yaitu, LSTM *baseline* dengan *hyperparameter* yang ditentukan secara manual berdasarkan penelitian terdahulu dan LSTM-PSO dengan *hyperparameter* hasil optimasi menggunakan *Particle Swarm Optimization*. Perbandingan konfigurasi kedua model ditunjukkan pada tabel 1.

Tabel 1. Perbandingan Konfigurasi Model LSTM

Hyperparameter	LSTM	LSTM-PSO
Unit layer1	200	148
Unit layer2	100	119
Dropout rate	0,1000	0,2727
Learning rate	0,0001	0,000643
Batch size	32	16
Optimizer	Adam	Adam
Loss function	MSE	MSE
Timestep	7	7
Max Epoch	100	100

Hasil optimasi menunjukkan bahwa PSO menghasilkan konfigurasi yang lebih adaptif, model LSTM-PSO membutuhkan waktu pelatihan lebih lama akibat proses optimasi yang dilakukan secara iteratif.

Arsitektur ini dirancang untuk mampu menangkap pola temporal dan hubungan non-linear pada data curah hujan, sehingga diharapkan dapat meningkatkan akurasi prediksi dibandingkan model LSTM konvensional.

3.6. Konfigurasi PSO

Particle Swarm Optimization (PSO) digunakan untuk mengoptimasi *hyperparameter* pada model LSTM. Pada penelitian ini PSO dikonfigurasi dengan jumlah partikel sebanyak 20 ($n_particles=20$), iterasi maksimum 30 ($max_iter=30$), bobot inersia ($w=0,7$), koefisien kognitif ($c1=1,5$), dan koefisien sosial ($c2=1,5$).

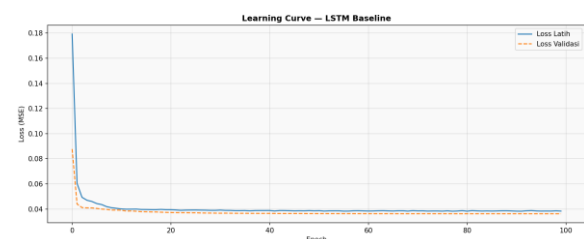
Fungsi objektif yang digunakan adalah *Root Mean Squared Error* (RMSE) pada data validasi, dengan tujuan meminimalkan kesalahan prediksi curah hujan. Setiap partikel merepresentasikan kombinasi *hyperparameter* dalam ruang pencarian, yang meliputi jumlah unit pada dua layer LSTM, *dropout rate*, *learning rate*, dan *batch size*.

Rentang pencarian *hyperparameter* ditentukan secara terbatas untuk menjaga efisiensi komputasi yaitu, jumlah unit layer pertama (32-256), layer kedua (16-128), *dropout rate* (0-0,5), *learning rate* (0,0001-0,01), dan *batch size* (16,32,64,128). Hasil optimasi menunjukkan hasil optimasi menunjukkan konfigurasi terbaik dengan 148 unit pada layer pertama, 119 unit pada layer kedua, *dropout rate* sebesar 0,2727, *learning rate* sebesar 0,000643, dan *batch size* 16

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pelatihan Model

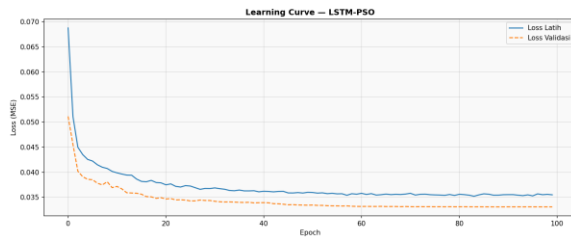
Proses pelatihan model LSTM *baseline* dilakukan selama 100 epoch penuh dengan *batch size* 32 dan *learning rate* awal 0,0001 menggunakan optimizer Adam. Proses pelatihan berjalan stabil tanpa indikasi *overfitting* yang signifikan.



Gambar 4. Learning Curve LSTM Baseline

Model LSTM-PSO dilatih dengan *hyperparameter* hasil optimasi PSO selama 100 epoch penuh dengan *batch size* 26 dan *learning rate* awal 0,000643. Dengan *batch size* yang lebih kecil, model melakukan lebih banyak pembaruan bobot per epoch sehingga proses konvergensi pada loss validasi berlangsung lebih responsif.

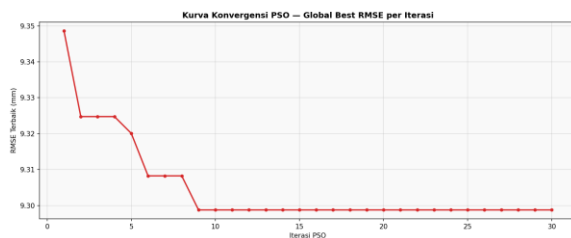
Pada gambar 4 dan gambar 5 terlihat bahwa LSTM-PSO lebih unggul dengan penurunan *learning curve* mencapai angka loss = 0,035 sedangkan LSTM baseline pada 0,04.



Gambar 5. Learning Curve LSTM-PSO

4.2. Proses dan Konvergensi PSO

Proses optimasi PSO dijalankan selama 30 iterasi dengan 20 partikel. Setiap partikel mengevaluasi fungsi *fitness* berupa RMSE pada data validasi curah hujan dalam skala asli (mm). Proses pencarian *hyperparameter* PSO membutuhkan waktu komputasi 774 menit 34,3 detik akibat setiap evaluasi *fitness* membutuhkan pelatihan model LSTM dari awal.



Gambar 6. Kurva Konvergensi PSO

Berdasarkan gambar 6, proses konvergensi PSO menunjukkan penurunan RMSE *global best* yang signifikan pada iterasi awal kemudian stabil hingga iterasi ke-30. Kestabilan ini mengindikasikan bahwa PSO telah menemukan solusi yang mendekati optimal dalam ruang pencarian yang ditetapkan.

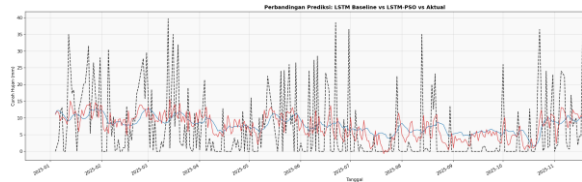
4.3. Evaluasi Performa pada Data Uji

Evaluasi dilakukan pada 362 data uji (periode 4 Januari 2025 hingga 31 Desember 2025) dalam skala nilai asli (mm) setelah proses *inverse transform* normalisasi. Tabel 2 merangkum hasil evaluasi kedua model.

Tabel 2. Perbandingan Metrik Evaluasi Model

Model	MAE	RMSE	NSE	R^2
LSTM Baseline	8,0257	9,7386	0,0847	0,0847
LSTM-PSO	7,4208	9,3857	0,1499	0,1499
Peningkatan	+7,54%	+3,62%	+76,85%	76,85%

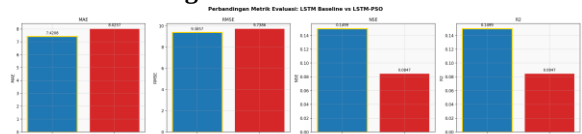
Berdasarkan Tabel 2, model LSTM-PSO unggul pada seluruh metrik evaluasi dibandingkan dengan LSTM baseline. Nilai MAE LSTM-PSO sebesar 7,4208mm menunjukkan rata-rata kesalahan absolut prediksi curah hujan lebih rendah dibandingkan LSTM baseline (8,0257mm), dengan selisih peningkatan sebesar 7,54%. Dengan hasil pada tabel 3 mengindikasikan bahwa model LSTM-PSO lebih robust terhadap kesalahan prediksi yang ekstrem.



Gambar 7. Perbandingan Model vs Data Aktual

Pada Gambar 7 terlihat bahwa garis dengan warna merah sebagai model LSTM-PSO, biru LSTM baseline dan hitam data aktual, menunjukkan peningkatan kualitas prediksi pada model LSTM-PSO dibandingkan LSTM baseline terhadap data aktual.

4.4. Perbandingan Model



Gambar 8. Perbandingan Metrik Evaluasi

Gambar 8 menunjukkan perbandingan kinerja model LSTM *baseline* dan LSTM-PSO berdasarkan metrik evaluasi MAE, RMSE, NSE, dan R^2 . Bar berwarna biru merepresentasikan model LSTM-PSO, sedangkan bar merah menunjukkan model LSTM *baseline*. Secara visual terlihat bahwa model LSTM-PSO menghasilkan nilai kesalahan yang lebih rendah pada metrik MAE dan RMSE dibandingkan LSTM *baseline*. Hal ini menunjukkan bahwa optimasi *hyperparameter* menggunakan PSO mampu meningkatkan akurasi prediksi dengan mengurangi deviasi antara nilai aktual dan hasil prediksi.

Selain itu, pada metrik NSE dan R^2 , model LSTM-PSO juga menunjukkan nilai yang lebih tinggi dibandingkan LSTM *baseline*. Peningkatan nilai NSE mengindikasikan bahwa model LSTM-PSO memiliki kemampuan yang lebih baik dalam merepresentasikan variabilitas data curah hujan, sedangkan peningkatan nilai R^2 menunjukkan bahwa proporsi variansi data yang dapat dijelaskan oleh model menjadi lebih besar. Dengan demikian, model hasil optimasi tidak hanya lebih akurat, tetapi juga lebih stabil dalam menangkap pola data.

Peningkatan performa ini disebabkan oleh kemampuan PSO dalam mengeksplorasi ruang *hyperparameter* secara adaptif untuk menemukan kombinasi parameter optimal, seperti jumlah neuron, *learning rate*, *dropout rate*, dan *batch size*. Berbeda dengan pendekatan manual pada LSTM *baseline*, PSO mampu menghindari pemilihan parameter yang suboptimal sehingga model lebih konvergen dan memiliki generalisasi yang lebih baik. Secara keseluruhan, hasil ini menunjukkan bahwa integrasi LSTM dengan *Particle Swarm Optimization* (PSO) efektif dalam meningkatkan performa prediksi curah hujan harian. Temuan ini memperkuat bahwa pendekatan optimasi berbasis *swarm intelligence* dapat menjadi solusi untuk mengatasi keterbatasan penentuan *hyperparameter* pada model deep learning,

khususnya pada data deret waktu yang bersifat non-linear dan kompleks.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan dan mengevaluasi model prediksi curah hujan harian berbasis integrasi LSTM dan PSO menggunakan data meteorologi BMKG Stasiun Geofisika Pasuruan periode 2021 hingga 2025. Berdasarkan hasil penelitian dapat disimpulkan sebagai berikut. Pertama, model LSTM-PSO menunjukkan performa yang lebih baik dibandingkan LSTM baseline pada seluruh metrik evaluasi. Kedua, optimasi *hyperparameter* berbasis PSO dengan 20 partikel dan 30 iterasi berhasil menemukan konfigurasi optimal yang tidak dapat dicapai melalui pendekatan manual. Ketiga, penggunaan *sliding window* 7 hari terbukti efektif dalam mengubah data meteorologi deret waktu menjadi format *supervised learning* yang sesuai untuk arsitektur LSTM *two layer stacked*. Adapun beberapa saran untuk penelitian selanjutnya. Pertama, memperluas rentang pencarian *hyperparameter* PSO dan meningkatkan jumlah partikel serta iterasi untuk memperoleh konfigurasi yang lebih optimal. Kedua, mengeksplorasi algoritma optimasi lain seperti *Generic Algorithm* (GA) atau *Bayesian Optimization* sebagai pembanding PSO. Ketiga, menambahkan variabel meteorologi lain untuk meningkatkan representasi model. Keempat, mengimplementasikan model pada *website* sederhana untuk digunakan sebagai mitigasi bencana.

DAFTAR PUSTAKA

- [1] BNPB, "Dokumen Kajian Risiko Bencana Nasional Provinsi Jawa Timur 2022–2026," 2021. Accessed: Dec. 20, 2025. [Online]. Available: https://inarisk.bnpb.go.id/pdf/Jawa%20Timur/Dokumen%20KRB%20Prov.%20Jawa%20Timur_final%20draft.pdf
- [2] I. Hanzahri, J. Z. Pagar Alam No, and K. Bandar Lampung, *Dinamika Teknik Sipil Majalah Ilmiah Teknik Sipil*, "Sistem Informasi Geografis dalam Mitigasi Hidrometeorologi: Studi Kasus Intensitas Hujan Mononobe di Pulau Jawa-Bali," *Dinamika Teknik Sipil Majalah Ilmiah Teknik Sipil*, vol. 18, pp. 79–89, Dec. 2025, [Online]. Available: <https://journals.ums.ac.id/index.php/DTS/index>
- [3] M. Mahmudi, A. R. Hidayat, and M. Y. Wijaya, "The Effect of Ensemble Averaging Method on Rainfall Forecasting in Jakarta Using ARIMA and ARIMAX," *Mathline: Jurnal Matematika dan Pendidikan Matematika*, vol. 9, no. 2, pp. 501–512, Apr. 2024, doi: 10.31943/mathline.v9i2.608.
- [4] J. Badriyah, A. Fariza, and T. Harsono, "Prediksi Curah Hujan Menggunakan Long Short Term Memory," *BUDIDARMA*, vol. 6, no. 3, p. 1297, Jul. 2022, doi: 10.30865/mib.v6i3.4008.
- [5] B. S. Patro and P. P. Bartakke, "Daily rainfall prediction using long short-term memory (LSTM) algorithm," *Journal of Agrometeorology*, vol. 26, no. 4, Dec. 2024, doi: 10.54386/jam.v26i4.2745.
- [6] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie, and A. Baroutaji, "Review of Metaheuristic Optimization Algorithms for Power Systems Problems," Jun. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/su15129434.
- [7] P. Umamaheswari, "Enhancing Rainfall Forecasting Using a Hybrid Long Short Term-Particle Swarm Optimization (LSTM-PSO) Method," Jul. 24, 2025. doi: 10.21203/rs.3.rs-6988218/v1.
- [8] Benny Hartanto, Ningrum Astriawati, Supartini, and Damar Kuncoro Yekti, "Pencarian dan Pemanfaatan Informasi Data Badan Meteorologi, Klimatologi, dan Geofisika (BMKG)," *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 5, pp. 553–564, Oct. 2022, doi: 10.55123/insologi.v1i5.906.
- [9] BMKG, "Tentang BMKG," Badan Meteorologi, Klimatologi, dan Geofisika. Accessed: Feb. 20, 2026. [Online]. Available: <https://www.bmkg.go.id/Profil>
- [10] Y. S. Swarinoto, "Pemanfaatan Suhu Udara dan Kelembapan Udara dalam Persamaan Regresi untuk Simulasi Prediksi Total Hujan Bulanan di Bandar Lampung (The Use of Air Temperature and Relative Humidity into Regression Equation to Provide Simulation Predictions of Monthly Rainfall Total at Bandar Lampung)," Bandar Lampung, Dec. 2011.
- [11] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," 1997.
- [12] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.
- [13] E. H. Houssein, A. G. Gad, K. Hussain, and P. N. Suganthan, "Major Advances in Particle Swarm Optimization: Theory, Analysis, and Application," *Swarm Evol. Comput.*, vol. 63, Jun. 2021, doi: 10.1016/j.swevo.2021.100868.
- [14] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," Jul. 19, 2022, *Copernicus GmbH*. doi: 10.5194/gmd-15-5481-2022.
- [15] 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, Jun. 2014, doi: 10.5194/gmd-7-1247-2014