

Pemodelan Spasial-Temporal Kejadian Banjir Menggunakan Random Forest Regression dan Long Short-Term Memory untuk Mitigasi

Spatial-Temporal Modeling of Flood Events Using Random Forest Regression and Long Short-Term Memory for Mitigation

Mochamad Amry Assiva^{1,*}, Hadi Tanuji¹, Eka Renny Noviati Wahyuni², Muhammad Fa'iq Robbani¹

¹Sains Data, Institut Teknologi dan Bisnis Muhammadiyah Grobogan, Grobogan, Indonesia

Jl. Solo - Purwodadi, Sukoharjo, Krangganharjo, Kec. Toroh, Kabupaten Grobogan, Jawa Tengah, Indonesia

²Manajemen Ritel, Institut Teknologi dan Bisnis Muhammadiyah Grobogan, Grobogan, Indonesia

Jl. Solo - Purwodadi, Sukoharjo, Krangganharjo, Kec. Toroh, Kabupaten Grobogan, Jawa Tengah, Indonesia

Email: ^{1,*}amryassiva@itbmg.ac.id, ²haditanuji@itbmg.ac.id, ³ekarennny@itbmg.ac.id, ⁴faiqrobhani@itbmg.ac.id

Email Penulis Korespondensi: amryassiva@itbmg.ac.id

Received: 04/10/2026, Revised: 19/10/2026, Accepted: 22/10/2026, Published: 23/10/2026

Abstrak—Penelitian ini mengevaluasi kerangka pemodelan spasial-temporal kejadian banjir menggunakan Random Forest (RF) dan Long Short-Term Memory (LSTM). RF digunakan sebagai regresor untuk mengestimasi jumlah kejadian banjir tahun 2025 pada 19 kecamatan berdasarkan lima indikator kejadian bencana tahun 2024 dan elevasi rata-rata. Fitur debit intake musim kemarau tidak digunakan karena inventaris bendung yang tersedia tidak memuat deret debit aktual yang konsisten. Evaluasi RF dilakukan dengan leave-one-out cross-validation (LOOCV). Reproduksi model BPBD + elevasi menghasilkan MAE 2.122, RMSE 2.650, dan R^2 -0.531. LSTM digunakan untuk memodelkan curah hujan bulanan dari dua stasiun BMKG selama Agustus 2024–Juli 2026. Konfigurasi look-back tiga bulan menghasilkan MAE 117,34 mm, RMSE 157,09 mm, dan R^2 -0,313, dibandingkan persistence baseline dengan MAE 153,85 mm dan RMSE 216,14 mm. Dengan 19 unit spasial dan 24 observasi hujan, keluaran RF dan LSTM diposisikan sebagai kerangka decision-support, bukan sistem prediksi operasional yang tervalidasi. Hasil ini menegaskan pentingnya evaluasi out-of-sample, audit ketersediaan data, dan pelaporan keterbatasan secara transparan. Kontribusi penelitian ini adalah mengevaluasi model RF melalui validasi LOOCV dan perbandingan fitur topografi, memetakan kesalahan prediksi pada tingkat kecamatan, serta membandingkan LSTM dengan persistence baseline sambil mengaudit keterbatasan data hidrologis.

Kata Kunci: Banjir; Random Forest; Long Short-Term Memory; Curah Hujan; Data Spasial

Abstract—This study evaluates a spatial-temporal flood-event modeling framework using Random Forest (RF) and Long Short-Term Memory (LSTM). RF estimates the number of flood events in 2025 across 19 districts using five disaster indicators from 2024 and mean elevation. Dry-season intake discharge was excluded because the available dam inventory does not provide a consistent series of observed discharge values. RF was evaluated using leave-one-out cross-validation (LOOCV). Reproduction of the BPBD + elevation model yielded MAE 2.122, RMSE 2.650, and R^2 -0.531. LSTM modeled monthly rainfall from two BMKG stations for August 2024–July 2026. The three-month look-back configuration achieved MAE 117.34 mm, RMSE 157.09 mm, and R^2 -0.313, compared with a persistence baseline MAE of 153.85 mm and RMSE of 216.14 mm. Given only 19 spatial units and 24 monthly rainfall observations, RF and LSTM are framed as a decision-support framework rather than a validated operational prediction system. The findings highlight the importance of out-of-sample evaluation, data-availability audits, and transparent reporting of limitations. The contribution of this study is to evaluate RF using LOOCV and topographic-feature comparisons, report prediction errors at the district level, compare LSTM with a persistence baseline, and audit hydrological data limitations.

Keywords: Flood; Random Forest; Long Short-Term Memory; Rainfall; Spatial Data

1. PENDAHULUAN

Banjir merupakan bencana hidrometeorologi dengan variasi spasial dan temporal yang kuat. Data kejadian bencana per kecamatan dapat dimanfaatkan untuk mendukung penentuan prioritas mitigasi, tetapi jumlah kejadian yang tercatat tidak secara otomatis mewakili probabilitas atau kelas risiko. Penelitian ini menggunakan target numerik berupa jumlah kejadian banjir tahun 2025 pada 19 kecamatan. Karena targetnya berupa hitungan yang memiliki jarak numerik dan dapat berbeda lebih dari satu kejadian, regresi Random Forest (RandomForestRegressor) memungkinkan model menghasilkan estimasi kuantitatif dan dievaluasi melalui MAE, RMSE, serta R^2 . Klasifikasi biner memerlukan penetapan ambang batas “berisiko/tidak berisiko” yang akan menghilangkan variasi jumlah kejadian dan membuat hasil sensitif terhadap ambang yang dipilih. Dengan demikian, regresi dipilih untuk menjawab pertanyaan estimasi jumlah kejadian, bukan untuk menyatakan probabilitas banjir.

Penelitian pemetaan banjir banyak menggunakan algoritma ensemble untuk menangkap hubungan nonlinier antara faktor lingkungan dan kejadian banjir. Pham et al. membandingkan algoritma boosting [1], Abu El-Magd mengevaluasi Random Forest dan Naive Bayes [2], Singha et al. menggabungkan pembelajaran mesin dengan analisis geospasial [3], sedangkan Abedi et al. membandingkan Random Forest, XGBoost, dan boosted regression trees [4]. Dalam konteks Indonesia, Purwati dan Pristyanto menggunakan Random Forest dan SVM untuk klasifikasi banjir [5], sementara Perdana Putra et al. mengintegrasikan pembelajaran mesin dengan GIS [6].

Studi tersebut mendukung relevansi metode ensemble dan fitur spasial, tetapi sebagian besar berfokus pada kelas atau kerawanan wilayah, bukan estimasi jumlah kejadian per kecamatan.

Sebagian penelitian kerawanan banjir berfokus pada klasifikasi atau susceptibility mapping, sedangkan penelitian ini menargetkan nilai numerik jumlah kejadian banjir per kecamatan. Regresi tidak berarti bahwa data hitungan telah memenuhi seluruh asumsi distribusi khusus untuk count data; pendekatan ini dipakai sebagai model prediktif berbasis Random Forest dan dievaluasi secara empiris. Dengan hanya 19 unit spasial, validasi perlu dilakukan secara ketat karena satu observasi dapat mengubah metrik secara berarti. LOOCV dipilih agar setiap kecamatan bergiliran menjadi data uji dan tidak ikut dalam pelatihan pada iterasi yang sama. Hasilnya tetap harus dipahami sebagai estimasi awal pada wilayah yang diteliti, bukan jaminan kinerja ketika diterapkan pada daerah atau periode lain.

Komponen temporal menggunakan Long Short-Term Memory (LSTM). Hadi et al. merangkum penggunaan pembelajaran mesin untuk prakiraan banjir dan menyoroti tantangan kualitas data [7]. Zhou et al. membandingkan model untuk prakiraan banjir di sungai pasang surut [8], sedangkan Kumar et al. mengulas peluang dan keterbatasan deep learning dalam manajemen banjir [9]. Zhang et al. menguji generalisasi encoder-decoder LSTM pada DAS tanpa pengukuran lengkap [10], dan Cui et al. mengevaluasi attention untuk prakiraan multi-langkah [11]. Secara umum, studi-studi ini menunjukkan potensi model temporal, tetapi kinerjanya tetap bergantung pada cakupan data, struktur masukan, dan rancangan validasi.

Kesenjangan penelitian ini mencakup tiga hal: penggunaan jumlah kejadian banjir per kecamatan sebagai target numerik dengan evaluasi out-of-sample; pengujian kontribusi topografi melalui perbandingan fitur; serta pemisahan model spasial dan temporal ketika label banjir bulanan per kecamatan belum tersedia. Karena itu, RF digunakan untuk estimasi jumlah kejadian antar-kecamatan, sedangkan LSTM memodelkan curah hujan bulanan. Kedua keluaran tidak digabungkan menjadi satu prediktor hibrida yang tervalidasi.

Kontribusi penelitian adalah evaluasi RF pada data spasial berukuran kecil, pengujian ablation fitur topografi, pelaporan kesalahan setiap kecamatan, dan audit ketersediaan data hidrologis. Inventaris BBWS Pemali Juana menyediakan lokasi, sungai terkait, fungsi, dan luas layanan, tetapi tidak memberikan deret debit intake musim kemarau yang konsisten [12]. Oleh karena itu, debit tersebut tidak digunakan sebagai fitur, nilai kosong tidak diubah menjadi nol, dan peningkatan kinerja akibat debit tidak diklaim.

Penelitian bertujuan mengevaluasi RF untuk mengestimasi jumlah kejadian banjir berdasarkan indikator bencana historis dan elevasi, mengevaluasi LSTM untuk memodelkan curah hujan bulanan, serta menyusun kerangka decision-support yang memisahkan informasi spasial (“di mana”) dan temporal (“kapan”). Kontribusi yang dievaluasi meliputi validasi LOOCV, perbandingan fitur, analisis kesalahan 19 kecamatan, dan pelaporan keterbatasan data.

Prapemrosesan, validasi, pemilihan representasi, dan pengendalian overfitting merupakan bagian penting dari pemodelan prediktif [13]. Pada jaringan saraf untuk deret waktu, panjang sequence dan prosedur pelatihan juga memengaruhi interpretasi hasil [14]. Penelitian sebelumnya oleh Assiva mengenai analisis sentimen pariwisata Grobogan menggunakan konteks lokal dalam analisis data [15], tetapi domain dan targetnya berbeda dari penelitian banjir ini.

Studi komparatif lain memperlihatkan bahwa kinerja pemetaan banjir dipengaruhi metode, resolusi spasial, dan karakteristik wilayah. Saleh et al. menerapkan metode ensemble untuk pemetaan kerawanan banjir bandang [16], Wahba et al. menggunakan RF regression dan GIS dengan variabel lingkungan [17]. Studi LSTM juga mencakup prediksi banjir beresolusi waktu tinggi [18] dan perbandingan dengan model hidrologis berbasis fisik [19]. Penelitian terbaru mengeksplorasi fitur geomorfologi untuk RF [20], sementara kajian mitigasi di Purwodadi menekankan kesiapsiagaan masyarakat [21]. Penelitian ini berbeda karena berfokus pada jumlah kejadian per kecamatan dan secara eksplisit melaporkan keterbatasan validasi pada sampel kecil. Kontribusi penelitian ini meliputi evaluasi estimasi kejadian banjir menggunakan LOOCV, pengujian fitur elevasi dan slope, analisis galat pada 19 kecamatan, serta evaluasi LSTM terhadap persistence baseline dan audit data hidrologis. Kontribusi tersebut mendukung evaluasi metodologis untuk mitigasi banjir, bukan merupakan klaim bahwa sistem prediksi operasional telah tervalidasi.

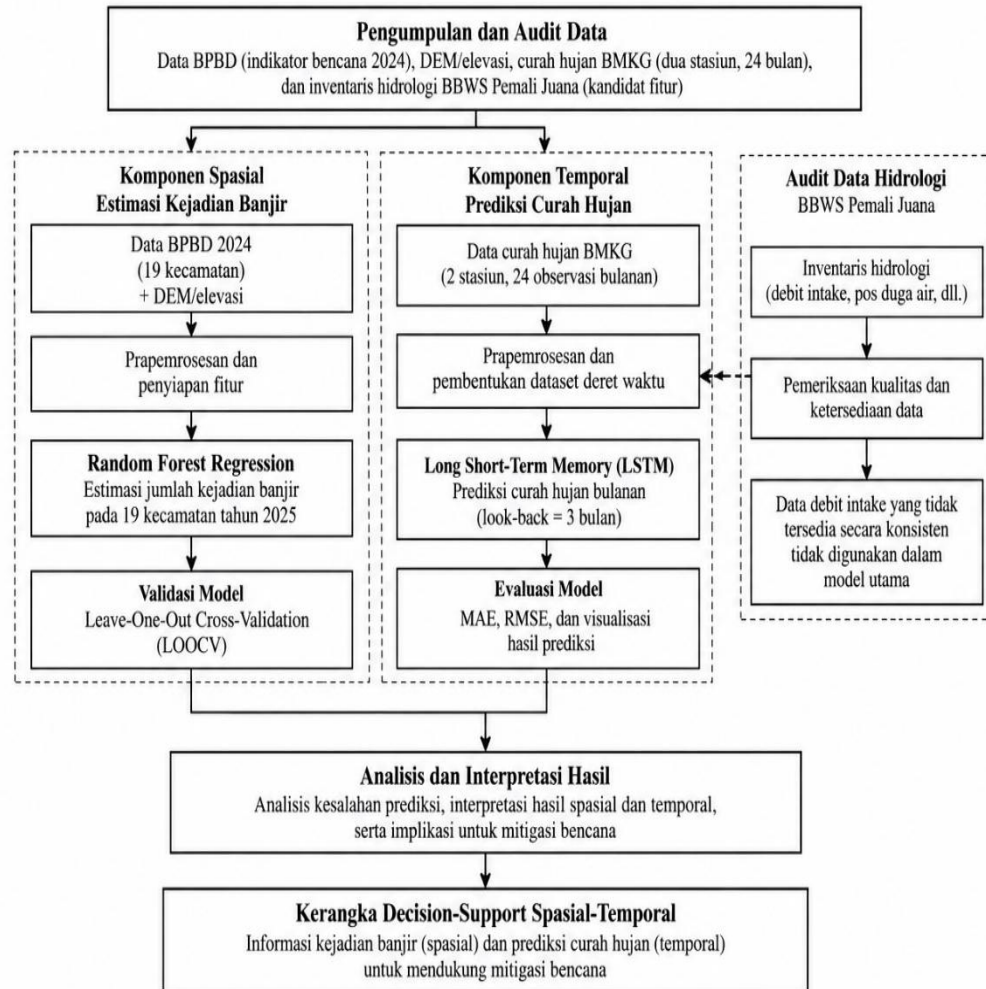
2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif berbasis data dengan dua komponen pemodelan, yaitu pemodelan spasial kejadian banjir menggunakan Random Forest Regression dan pemodelan temporal curah hujan menggunakan Long Short-Term Memory (LSTM). Komponen Random Forest Regression digunakan untuk mengestimasi jumlah kejadian banjir tahun 2025 pada 19 kecamatan berdasarkan indikator bencana tahun 2024 dan elevasi rata-rata wilayah. Evaluasi model dilakukan menggunakan Leave-One-Out Cross-Validation (LOOCV). Sementara itu, komponen LSTM digunakan untuk memprediksi curah hujan bulanan berdasarkan 24 observasi dari dua stasiun BMKG dengan panjang urutan masukan (look-back) tiga bulan.

Tahapan penelitian meliputi pengumpulan dan audit data, prapemrosesan, penyiapan fitur, pemodelan, evaluasi kinerja, analisis kesalahan prediksi, dan interpretasi hasil. Gambar 1 menunjukkan alur pemodelan yang

terdiri atas dua komponen terpisah, yaitu Random Forest Regression untuk estimasi kejadian banjir secara spasial dan LSTM untuk prediksi curah hujan secara temporal. Inventaris data hidrologi dari BBWS Pemali Juana turut diaudit sebagai kandidat fitur, tetapi data debit intake yang tidak tersedia secara konsisten tidak digunakan dalam model utama. Hasil kedua komponen dianalisis untuk memperoleh informasi yang saling melengkapi dalam mendukung interpretasi mitigasi bencana, tanpa mengasumsikan bahwa kedua model telah membentuk satu model prediksi hibrida.



Gambar 1. Tahapan penelitian

2.2 Data dan Preprocessing

Data BPBD mencakup jumlah kejadian banjir, longsor, angin puting beliung, kekeringan, dan kebakaran lahan per kecamatan. Lima indikator bencana tahun 2024 menjadi fitur awal untuk target jumlah banjir tahun 2025. Elevasi rata-rata diekstraksi dari DEMNAS, sedangkan slope diuji dalam analisis ablation. Metadata pos duga air, sungai, bendung, dan bendungan BBWS diaudit sebagai kandidat pengayaan. Karena inventaris tidak memuat debit intake aktual yang konsisten untuk seluruh unit dan pemetaan Q25 belum lengkap, kedua variabel tersebut tidak dimasukkan ke model.

Curah hujan dari dua stasiun BMKG dirata-ratakan per bulan untuk Agustus 2024–Juli 2026, menghasilkan 24 observasi. Nilai tertinggi adalah 602,80 mm pada Januari 2026 dan terendah 3,50 mm pada Juli 2026. Data dibentuk menjadi sliding window dengan look-back utama tiga bulan; jumlah sequence yang terbatas membatasi kemampuan model mempelajari fluktuasi ekstrem. Penskalaan ditentukan dari data pelatihan agar data uji tidak memengaruhi preprocessing. Look-back dua dan enam bulan digunakan sebagai analisis sensitivitas, bukan untuk mengganti konfigurasi utama setelah hasil uji diketahui. Data meteorologis bersumber dari layanan Data Online BMKG [22]. Ringkasan sumber, periode, dan fungsi setiap dataset disajikan pada Tabel 1.

Tabel 1. Data penelitian yang digunakan

No.	Data	Sumber	Periode/Informasi	Pemanfaatan
1	Curah hujan	BMKG WMO 96835	Agustus 2024–Juli 2026; 24 observasi	Input temporal LSTM

No.	Data	Sumber	Periode/Informasi	Pemanfaatan
2	Curah hujan	BMKG WMO 96839	Agustus 2024–Juli 2026; 24 observasi	Input temporal LSTM
3	Kejadian bencana	BPBD Kabupaten Grobogan	2024–2025; 19 kecamatan	Fitur dan target RF
4	Digital Elevation Model	DEMNAS – BIG	Wilayah Kabupaten Grobogan	Variabel spasial RF
5	Sungai, bendung, dan bendungan	BBWS Pemali Juana	Inventaris lokasi, fungsi, sungai, luas layanan; debit aktual tidak tersedia konsisten	Audit kandidat fitur; bukan input debit model
6	Metadata pos duga air dan atribut sungai	BBWS Pemali Juana; inventaris sumber daya air	Inventaris sumber daya air; atribut sungai, Q25, bendung, debit intake, dan infrastruktur	Pengayaan hidrologis dan audit kandidat fitur

Tabel 1 menunjukkan bahwa pemodelan menggunakan dua kelompok data dengan resolusi dan tujuan berbeda: data kejadian bencana BPBD tahun 2024 menjadi fitur untuk mengestimasi jumlah kejadian banjir tahun 2025, sedangkan 24 observasi curah hujan bulanan dari dua stasiun BMKG digunakan untuk pemodelan temporal LSTM. DEMNAS menyediakan fitur elevasi, sementara inventaris BBWS dan metadata pos duga air hanya digunakan untuk audit kandidat fitur hidrologis. Karena deret debit aktual tidak tersedia secara konsisten dan atribut Q25 belum dapat dipetakan secara lengkap ke seluruh kecamatan, kedua informasi tersebut tidak dimasukkan ke model utama.

2.3 Random Forest Regression dan Validasi

RF digunakan karena target berupa jumlah kejadian, bukan kelas kategorikal. RandomForestRegressor menggunakan 100 pohon, random_state 42, dan n_jobs -1. Empat konfigurasi dibandingkan melalui ablation: BPBD; BPBD + elevasi; BPBD + slope; dan BPBD + elevasi + slope. Fitur jarak pos, jumlah pos dalam radius 25 km, debit intake, dan Q25 tidak masuk model utama karena kelengkapan atau pemetaan untuk seluruh kecamatan belum memadai. LOOCV menahan satu kecamatan sebagai data uji dan melatih model pada 18 kecamatan lainnya. MAE mengukur kesalahan absolut rata-rata, RMSE memberi penalti lebih besar pada kesalahan besar, dan R² mengukur kemampuan menjelaskan variasi target. Feature importance dianalisis melalui mean decrease in impurity dan diperiksa menggunakan permutation importance [13].

2.4 Long Short-Term Memory

LSTM menggunakan satu lapisan 32 unit, dropout 0,1, dense 16 unit beraktivasi ReLU, dan satu neuron keluaran. Model dilatih dengan Adam, mean squared error, 200 epoch, dan batch size 4. Look-back tiga bulan ditetapkan sebagai konfigurasi utama; look-back dua dan enam bulan digunakan untuk analisis sensitivitas. Tabel 5 melaporkan mean ± simpangan baku dari beberapa random seed. Karena hasil look-back enam bulan diamati setelah pengujian, hasilnya diperlakukan eksploratif dan memerlukan validasi deret waktu independen sebelum dipilih sebagai konfigurasi utama. Persistence baseline memprediksi curah hujan bulan berikutnya sama dengan nilai aktual bulan sebelumnya. Pemodelan mengikuti praktik umum deep learning dan deret waktu [14].

2.5 Kerangka Integrasi RF–LSTM

Prediksi RF dari konfigurasi BPBD + elevasi dinormalisasi dengan min–max untuk membentuk indeks spasial relatif. LSTM menghasilkan indikator temporal curah hujan. Keduanya dapat divisualisasikan bersama untuk mendukung interpretasi, tetapi tidak digabungkan menjadi satu model prediksi karena label banjir bulanan per kecamatan tidak tersedia. Indeks spasial bukan probabilitas banjir, dan indikator curah hujan yang sama pada bulan tertentu tidak mengubah urutan spasial.

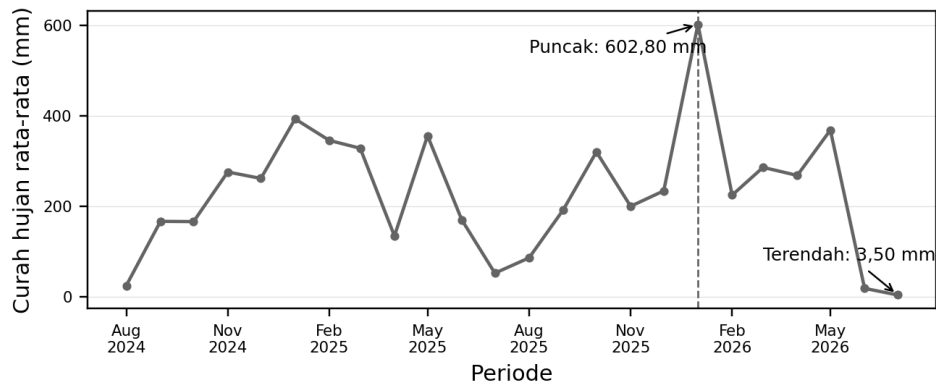
2.6 Analisis Mitigasi

Interpretasi mitigasi mempertimbangkan indeks spasial, kesalahan LOOCV per kecamatan, dan konteks hidrologis. Nilai prediksi yang relatif tinggi dapat menjadi dasar pemeriksaan awal, sedangkan kesalahan besar menunjukkan kebutuhan fitur tambahan. Hasil tidak diperlakukan sebagai penetapan risiko final tanpa validasi lapangan dan data hidrologi yang memadai.

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Data Temporal dan Spasial

Data dua stasiun BMKG menunjukkan variasi bulanan yang besar. Total curah hujan 2025 sebesar 2.814,70 mm (rata-rata 234,56 mm/bulan), sedangkan Januari–Juli 2026 mencapai 1.772,50 mm (rata-rata 253,21 mm/bulan). Puncak 602,80 mm terjadi pada Januari 2026 dan nilai terendah 3,50 mm pada Juli 2026. Rentang tersebut penting untuk menilai keterbatasan model temporal yang dilatih hanya dengan 24 observasi.



Gambar 2. Curah hujan rata-rata bulanan Agustus 2024–Juli 2026

Gambar 2 memperlihatkan perubahan curah hujan rata-rata bulanan dan membantu menjelaskan mengapa deret pendek menyulitkan model recurrent neural network untuk mempelajari kejadian ekstrem secara andal. Elevasi rata-rata 19 kecamatan berada pada rentang sekitar 10,04–109,62 m. Tegowanu, Gubug, dan Purwodadi berada pada rentang elevasi lebih rendah, sedangkan Geyer dan beberapa kecamatan selatan relatif lebih tinggi. Elevasi merupakan ringkasan topografi dan tidak merepresentasikan secara langsung kondisi drainase, kapasitas sungai, tanggul, atau paparan permukaan.

3.2 Penyempurnaan Random Forest dengan Variabel Spasial

Kinerja empat konfigurasi RF berdasarkan LOOCV dirangkum pada Tabel 2. Pada reproduksi dari dataset yang tersedia, model BPBD saja menghasilkan MAE 2.250, RMSE 2.631, dan R^2 -0.509. Penambahan elevasi menghasilkan MAE 2.122, RMSE 2.650, dan R^2 -0.531. Model BPBD + slope menghasilkan MAE 2.192, RMSE 2.693, dan R^2 -0.581, sedangkan kombinasi elevasi dan slope menghasilkan MAE 2.171, RMSE 2.645, dan R^2 -0.525. Penambahan elevasi menurunkan MAE dibandingkan BPBD saja, tetapi RMSE sedikit meningkat dan R^2 tetap negatif. Dengan demikian, peningkatan kinerja tidak seragam pada semua metrik; tidak ada dasar untuk menyatakan bahwa konfigurasi tertentu telah mencapai kemampuan prediksi operasional.

Tabel 2 menunjukkan bahwa semua konfigurasi memiliki R^2 negatif, sehingga kinerjanya belum melampaui prediksi pembanding berbasis rerata pada evaluasi LOOCV. Penambahan elevasi menurunkan MAE dibandingkan fitur BPBD saja, tetapi RMSE sedikit meningkat. Karena itu, hasil tidak mendukung klaim bahwa topografi meningkatkan seluruh aspek kinerja; model masih memiliki keterbatasan dalam menjelaskan variasi kejadian.

Tabel 2. Perbandingan konfigurasi Random Forest berdasarkan LOOCV

Konfigurasi	MAE	RMSE	R^2
BPBD saja	2,250	2,631	-0,509
BPBD + elevasi	2,122	2,650	-0,531
BPBD + slope	2,192	2,693	-0,581
BPBD + elevasi + slope	2,171	2,645	-0,525

Tabel 3 menyajikan importance fitur pada konfigurasi RF + elevasi. Ukuran mean decrease in impurity menunjukkan kontribusi relatif fitur ketika pemisahan pohon dibangun, sedangkan permutation importance menunjukkan perubahan MAE ketika informasi suatu fitur diacak. Elevasi rata-rata dan jumlah kejadian banjir 2024 merupakan fitur penting dalam model. Namun, importance tidak membuktikan hubungan kausal dan nilainya dapat berubah pada dataset kecil, khususnya ketika fitur berkorelasi atau beberapa observasi memiliki pengaruh tinggi. Karena itu, hasil pada Tabel 3 digunakan sebagai indikasi untuk eksplorasi fitur lanjutan, bukan sebagai bukti bahwa elevasi secara tunggal menyebabkan kejadian banjir.

Tabel 3. Feature importance pada RF + elevasi

Fitur	MDI	Permutation importance (Δ MAE)
Elevasi rata-rata	0,4270	0,8125
Banjir 2024	0,3215	0,5270
Angin puting beliung 2024	0,1408	0,2337
Tanah longsor 2024	0,0632	0,1858
Kebakaran lahan 2024	0,0251	0,0381
Kekeringan 2024	0,0224	0,0203

Temuan tersebut sejalan dengan literatur yang menunjukkan pentingnya variabel geospasial dalam pemetaan kerawanan banjir [1], [2], [3], [4], [17], [20]. Penggunaan fitur harus ditentukan oleh kelengkapan dan skala data, bukan hanya ketersediaan teknis.

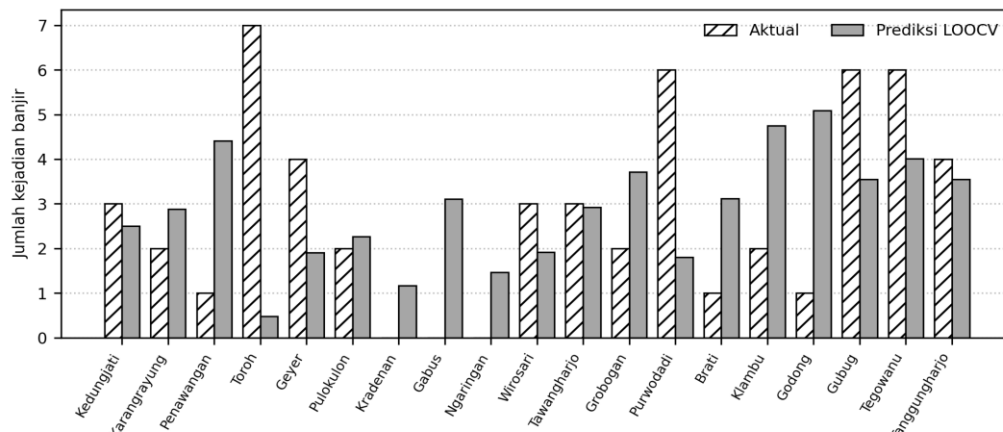
3.3 Evaluasi Out-of-Sample dan Kesalahan Antarkecamatan

Kinerja prediksi out-of-sample dievaluasi menggunakan Leave-One-Out Cross-Validation (LOOCV) pada 19 kecamatan. Pada konfigurasi Random Forest Regression dengan fitur indikator bencana tahun 2024 dan elevasi rata-rata (BPBD + elevasi), model menghasilkan Mean Absolute Error (MAE) sebesar 2,122 kejadian, Root Mean Square Error (RMSE) sebesar 2,650 kejadian, dan koefisien determinasi (R^2) sebesar $-0,531$. Nilai R^2 negatif menunjukkan bahwa prediksi model pada skema validasi tersebut lebih buruk daripada penggunaan rata-rata target sebagai prediktor pembanding. Hasil ini mengindikasikan bahwa kedua fitur yang digunakan belum cukup menjelaskan variasi jumlah kejadian banjir antarkecamatan. Tabel 4 menyajikan jumlah kejadian aktual, prediksi LOOCV, dan kesalahan absolut untuk seluruh 19 kecamatan. Evaluasi menggunakan konfigurasi BPBD + elevasi dipilih karena data debit intake tidak tersedia secara konsisten sehingga tidak dapat digunakan sebagai fitur dalam model utama. Prediksi LOOCV merepresentasikan estimasi untuk setiap kecamatan ketika data kecamatan tersebut dikeluarkan dari proses pelatihan; oleh karena itu, hasilnya tidak dapat langsung ditafsirkan sebagai prediksi operasional untuk periode mendatang. Tabel 4 memperlihatkan variasi kesalahan antarkecamatan.

Tabel 4. Aktual, prediksi LOOCV, dan kesalahan absolut pada 19 kecamatan

Kecamatan	Aktual	Prediksi LOOCV	Error absolut
Kedungjati	3	2,50	0,50
Karangrayung	2	2,88	0,88
Penawangan	1	4,41	3,41
Toroh	7	0,48	6,52
Geyer	4	1,91	2,09
Pulokulon	2	2,26	0,26
Kradenan	0	1,17	1,17
Gabus	0	3,11	3,11
Ngaringan	0	1,46	1,46
Wirosari	3	1,92	1,08
Tawangharjo	3	2,92	0,08
Grobogan	2	3,71	1,71
Purwodadi	6	1,80	4,20
Brati	1	3,12	2,12
Klambu	2	4,75	2,75
Godong	1	5,09	4,09
Gubug	6	3,55	2,45
Tegowanu	6	4,01	1,99
Tanggungharjo	4	3,55	0,45

Toroh memiliki tujuh kejadian aktual, tetapi prediksi LOOCV 0,48, sehingga galat absolut 6,52. Purwodadi memiliki enam kejadian aktual dengan prediksi 1,80 dan galat 4,20. Sebaliknya, Tawangharjo memiliki galat 0,08. Perbedaan ini menunjukkan bahwa kinerja agregat tidak menggambarkan ketepatan yang seragam di setiap kecamatan. Gambar 3 membandingkan jumlah kejadian aktual dan prediksi LOOCV pada seluruh 19 kecamatan untuk memperlihatkan pola underestimation dan overestimation.



Gambar 3. Perbandingan jumlah kejadian banjir aktual dan prediksi LOOCV pada 19 kecamatan

Kasus Toroh dan Purwodadi perlu diperhatikan karena model menghasilkan prediksi yang lebih rendah daripada jumlah kejadian aktual. Berdasarkan hasil LOOCV pada Tabel 4, Toroh memiliki tujuh kejadian aktual dan prediksi 0,48, sedangkan Purwodadi memiliki enam kejadian aktual dan prediksi 1,80. Nilai tersebut menunjukkan bahwa kombinasi indikator bencana tahun 2024 dan elevasi rata-rata belum menangkap seluruh variasi kejadian yang tercatat pada kedua kecamatan. Penelitian ini belum memiliki data kejadian tergeoreferensi, inventaris tanggul jebol yang tervalidasi, maupun data rinci permukiman yang dapat menguji penjelasan tersebut. Karena itu, faktor-faktor tersebut diperlakukan sebagai hipotesis untuk verifikasi lanjutan, bukan sebagai penyebab yang telah dibuktikan. Secara umum, elevasi rata-rata dapat menyamarkan variasi mikro-topografi dan tidak merepresentasikan kapasitas penampang sungai, kondisi tanggul dan saluran, limpasan dari hulu, perubahan penggunaan lahan, maupun paparan penduduk. Pengujian kontribusi setiap faktor memerlukan pencocokan catatan kejadian BPBD dengan lokasi kejadian, data hidrologi BBWS, peta sungai dan permukiman, serta pemeriksaan lapangan bila tersedia. Literatur mengenai flood forecasting dan flood susceptibility juga menekankan bahwa model machine learning sangat bergantung pada kelengkapan dan representasi variabel input [7], [9], [17], [19], [20].

Secara geografis, selisih pada Toroh dan Purwodadi lebih tepat dibaca sebagai keterbatasan representasi skala, bukan bukti bahwa faktor tertentu telah menyebabkan banjir. Di Toroh, tujuh kejadian aktual berhadapan dengan prediksi LOOCV 0,48; di Purwodadi, enam kejadian berhadapan dengan prediksi 1,80. Nilai rata-rata tingkat kecamatan dapat menyamarkan heterogenitas antardesa, ruas sungai, dan kantong permukiman. Konfluensi sungai, saluran drainase, titik penyempitan aliran, kondisi perkotaan, serta tanggul merupakan kandidat faktor yang patut diperiksa, tetapi keterkaitannya dengan catatan kejadian belum diuji dalam penelitian ini. Karena koordinat kejadian, inventaris tanggul jebol tervalidasi, dan data rinci permukiman belum tersedia, faktor tersebut harus diperlakukan sebagai hipotesis, bukan penyebab yang telah terbukti. Verifikasi memerlukan pencocokan catatan BPBD dengan jaringan sungai, data hidrologi BBWS, peta permukiman, dan pemeriksaan lapangan.

3.4 Audit Data Hidrologis dan Dampaknya terhadap Model

Pengayaan hidrologis diawali dengan audit ketersediaan data, bukan asumsi bahwa penambahan fitur pasti meningkatkan kinerja. Inventaris BBWS mencakup lokasi dan fungsi infrastruktur, tetapi tidak menyediakan deret debit intake musim kemarau yang seragam untuk seluruh kecamatan. Karena itu, debit tidak dimasukkan dan nilai hilang tidak diimputasi nol. Keterbatasan ini mencegah perbandingan yang adil jika sebagian kecamatan memiliki nilai debit dan sebagian lainnya tidak.

Atribut Q25 juga tidak digunakan karena belum dapat dipasangkan secara konsisten dengan seluruh kecamatan. Untuk penelitian lanjutan, fitur hidrologis perlu disusun dengan definisi temporal dan spasial yang seragam, misalnya jarak ke sungai utama, kepadatan jaringan sungai, riwayat genangan, dan kapasitas drainase, disertai dokumentasi sumber serta penanganan nilai hilang.

3.5 Evaluasi Long Short-Term Memory

LSTM digunakan untuk memodelkan curah hujan bulanan, bukan secara langsung untuk memprediksi jumlah kejadian banjir. Pemisahan target ini diperlukan karena data banjir yang tersedia tidak memiliki label bulanan pada setiap kecamatan yang sepadan dengan deret curah hujan. Dari 24 observasi bulanan, konfigurasi look-back tiga bulan menghasilkan 21 sequence, yang terdiri atas 15 sequence untuk pelatihan dan 6 sequence untuk pengujian. Proses penskalaan data ditentukan berdasarkan data pelatihan agar informasi dari periode pengujian tidak memengaruhi transformasi masukan model.

Evaluasi sensitivitas dilakukan untuk membandingkan pengaruh panjang look-back terhadap kinerja LSTM. Tabel 5 menyajikan nilai MAE, RMSE, dan R^2 untuk konfigurasi look-back 2, 3, dan 6 bulan, termasuk peran masing-masing konfigurasi dalam analisis. Perbandingan tersebut digunakan untuk menilai perubahan kesalahan prediksi sekaligus mengidentifikasi keterbatasan pemilihan konfigurasi berdasarkan data pengujian.

Tabel 5. Sensitivitas konfigurasi look-back LSTM

Look-back	MAE (mm)	RMSE (mm)	R^2	Peran
2	114,55 ± 0,34	153,37 ± 0,36	-0,252	Sensitivitas
3	117,34 ± 0,55	157,09 ± 0,43	-0,313	Utama
6	112,09 ± 5,31	117,88 ± 4,91	0,260	Sensitivitas; performa terbaik teramati, belum konfirmasi

Berdasarkan Tabel 5, konfigurasi look-back 2 menghasilkan MAE 114,55 mm, RMSE 153,37 mm, dan R^2 -0,252. Konfigurasi look-back 3 sebagai konfigurasi utama menghasilkan MAE 117,34 mm, RMSE 157,09 mm, dan R^2 -0,313. Sementara itu, konfigurasi look-back 6 menghasilkan MAE 112,09 mm, RMSE 117,88 mm, dan R^2 0,260. Dengan demikian, look-back 6 menunjukkan kinerja pengujian terbaik di antara konfigurasi yang dibandingkan. Dibandingkan dengan look-back 3, MAE menurun sekitar 4,47% dan RMSE menurun sekitar 24,96%, sedangkan R^2 meningkat dari -0,313 menjadi 0,260.

Meskipun demikian, hasil tersebut belum dapat dijadikan bukti bahwa look-back 6 merupakan konfigurasi terbaik secara umum. Konfigurasi tersebut menunjukkan kinerja terbaik setelah hasil pengujian diamati sehingga pemilihannya berisiko menimbulkan bias seleksi terhadap test set. Oleh karena itu, look-back 3 tetap dipertahankan sebagai konfigurasi utama yang dilaporkan, sedangkan look-back 2 dan 6 digunakan sebagai analisis sensitivitas. Penetapan look-back 6 sebagai model final memerlukan evaluasi ulang menggunakan data validasi yang terpisah atau prosedur time-series cross-validation yang mencegah penggunaan data pengujian untuk memilih konfigurasi model.

Untuk menilai manfaat LSTM dibandingkan metode prediksi sederhana, dilakukan perbandingan dengan persistence baseline, yaitu metode yang menggunakan curah hujan aktual pada periode sebelumnya sebagai prediksi untuk periode berikutnya. Tabel 6 menyajikan perbandingan kinerja konfigurasi utama LSTM dengan persistence baseline berdasarkan MAE, RMSE, dan R^2 pada enam periode pengujian.

Tabel 6. Perbandingan LSTM utama dengan persistence baseline

Model	MAE	RMSE	R^2
LSTM look-back 3	117,34 mm	157,09 mm	-0,313
Persistence	153,85 mm	216,14 mm	-1,486

Tabel 6 menunjukkan bahwa LSTM dengan look-back 3 menghasilkan MAE dan RMSE yang lebih rendah daripada persistence baseline. MAE berkurang sekitar 23,73%, sedangkan RMSE berkurang sekitar 27,32%. Hasil tersebut menunjukkan bahwa pada periode pengujian yang digunakan, LSTM menghasilkan kesalahan absolut dan kesalahan kuadrat rata-rata yang lebih rendah dibandingkan prediksi naif. Namun, nilai R^2 LSTM masih negatif, yaitu -0,313. Dengan demikian, hasil perbandingan ini belum cukup untuk menyatakan bahwa LSTM memiliki kemampuan generalisasi yang kuat. Jumlah titik uji yang hanya enam juga membuat hasil evaluasi sensitif terhadap karakteristik curah hujan pada setiap periode pengujian.

Kesalahan prediksi LSTM perlu ditafsirkan dengan mempertimbangkan perubahan curah hujan yang sangat tajam. Pada Juni 2026, curah hujan aktual turun menjadi 18,10 mm, sedangkan prediksi cenderung masih lebih tinggi. Pada Juli 2026, curah hujan aktual hanya 3,50 mm setelah mencapai puncak 602,80 mm pada Januari 2026. Kondisi tersebut menggambarkan keterbatasan pembelajaran ketika jumlah observasi dan keragaman pola yang tersedia tidak memadai, yang dalam pembahasan ini disebut sebagai data starvation. Dengan hanya 24 observasi bulanan, konfigurasi look-back tiga bulan menghasilkan 21 sequence, dengan 15 sequence untuk pelatihan dan enam untuk pengujian. Jumlah sampel efektif tersebut sangat terbatas bagi jaringan recurrent neural network seperti LSTM yang perlu mempelajari dependensi temporal, pola musiman, dan perubahan ekstrem melalui sejumlah parameter model.

Keterbatasan jumlah sampel dapat meningkatkan risiko overfitting, membatasi kemampuan model dalam merepresentasikan kejadian ekstrem yang jarang, dan menyebabkan prediksi kurang responsif terhadap perubahan mendadak. Namun, keterbatasan ini tidak berarti bahwa LSTM selalu tidak sesuai untuk pemodelan hidrologis. Kelayakan penggunaannya perlu dinilai berdasarkan panjang deret waktu, kompleksitas arsitektur, karakteristik data masukan, dan rancangan validasi. Kajian flood forecasting menunjukkan bahwa kinerja model temporal dapat berubah menurut panjang data, karakteristik masukan, dan horizon prediksi [7], [8], [9], [10], [11], [18]. Kajian kritis Waqas dan Humphries meninjau variasi RNN/LSTM untuk deret waktu hidrologis dan menyoroti pentingnya panjang data, pemilihan arsitektur, serta validasi yang sesuai [23]. Tinjauan scoping Byaruhanga et al. juga menunjukkan bahwa pemilihan model prediksi banjir perlu mempertimbangkan tujuan peramalan, karakteristik data, dan konteks sistem peringatan dini [24].

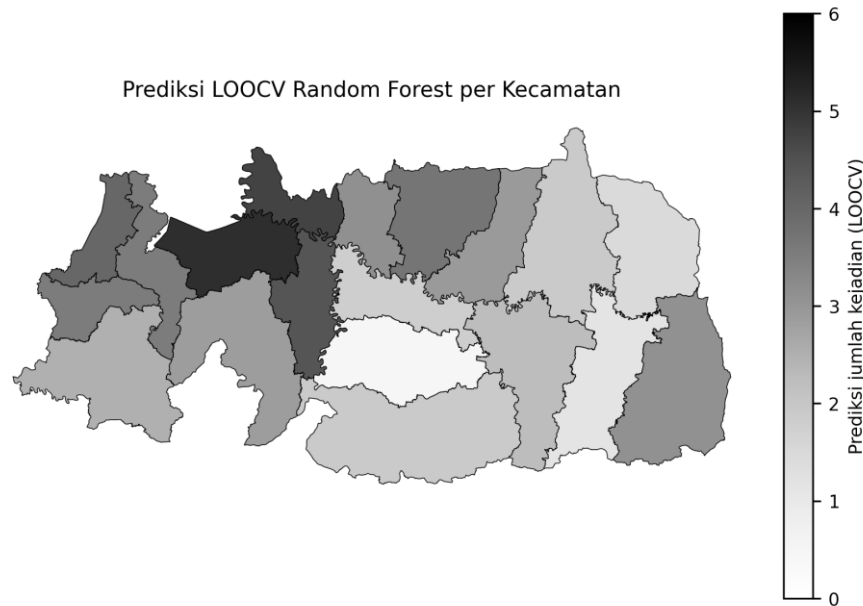
Istilah data starvation pada penelitian ini digunakan secara deskriptif untuk menyebut keterbatasan jumlah sampel efektif, bukan sebagai istilah diagnostik formal atau bukti bahwa arsitektur LSTM secara intrinsik gagal. Dengan enam titik uji yang saling berurutan dan berasal dari deret yang sama, hasil MAE, RMSE, dan R^2 memiliki ketidakpastian tinggi serta tidak cukup untuk menyimpulkan kinerja pada musim atau tahun lain. Oleh karena itu, hasil model perlu diperlakukan sebagai eksplorasi awal sampai tersedia deret yang lebih panjang dan evaluasi rolling-origin atau time-series cross-validation yang independen.

3.6 Kerangka Integrasi RF–LSTM

RF dan LSTM memiliki target berbeda: RF mengestimasi jumlah kejadian banjir per kecamatan untuk 2025, sedangkan LSTM memprediksi curah hujan bulanan dari 24 observasi. Oleh sebab itu, keduanya belum menjadi model hibrida terlatih. RF memberikan informasi relatif mengenai variasi spasial, sementara LSTM menyediakan konteks temporal hujan. Penggabungan keduanya hanya bersifat interpretatif sampai tersedia label banjir bulanan per kecamatan yang tersinkronisasi.

Gambar 4 menyajikan distribusi spasial prediksi Leave-One-Out Cross-Validation (LOOCV) dari model RF dengan konfigurasi fitur indikator bencana tahun 2024 dan elevasi rata-rata (BPBD + elevasi) pada 19 kecamatan. Peta tersebut memperlihatkan variasi estimasi jumlah kejadian banjir yang dihasilkan ketika setiap kecamatan dikeluarkan secara bergantian dari proses pelatihan. Nilai pada peta merupakan prediksi hasil validasi silang, bukan probabilitas banjir atau indeks bahaya yang telah divalidasi untuk penggunaan operasional. Oleh

karena itu, hasil visualisasi perlu diinterpretasikan dengan mempertimbangkan metrik evaluasi model dan keterbatasan fitur yang digunakan.



Gambar 4. Peta prediksi LOOCV Random Forest per kecamatan

Dalam mitigasi, prediksi RF dapat menjadi masukan awal untuk memilih lokasi pemeriksaan, bukan satu-satunya dasar penetapan risiko. Kecamatan dengan prediksi relatif tinggi atau galat besar dapat diprioritaskan untuk verifikasi catatan kejadian, pemeriksaan drainase dan saluran, serta pengumpulan data hidrologi lokal. Keputusan tetap memerlukan validasi lapangan dan informasi paparan penduduk maupun aset.

Keluaran LSTM dapat memberi konteks perubahan curah hujan bulanan, tetapi tidak menunjukkan risiko banjir khusus untuk setiap kecamatan. Penggunaan operasional membutuhkan data hujan yang lebih panjang, pengujian lintas periode, serta hubungan yang terukur antara hujan, debit atau tinggi muka air, dan kejadian banjir.

3.7 Pembahasan

Hasil RF menunjukkan bahwa estimasi jumlah kejadian banjir dengan 19 unit spasial masih memiliki keterbatasan. Seluruh konfigurasi menghasilkan R^2 negatif pada LOOCV, dan kesalahan berbeda tajam antarkecamatan. Hal ini menguatkan pentingnya melaporkan hasil out-of-sample dan galat per unit, bukan hanya satu metrik agregat. Literatur pemetaan banjir menggunakan ensemble dan fitur geospasial menunjukkan potensi metode tersebut [1], [2], [3], [4], [16], [17], tetapi hasil penelitian ini menegaskan bahwa kinerja bergantung pada kesesuaian fitur dengan target serta resolusi data.

Untuk LSTM, konfigurasi look-back tiga bulan memiliki kesalahan lebih rendah daripada persistence baseline pada pengujian yang dilakukan, tetapi jumlah sampel dan test set sangat kecil. Studi forecasting banjir dengan LSTM dan attention menunjukkan potensi model temporal [10], [11], [18], [20], sementara kajian terbaru juga menyoroti ketergantungan kinerja pada data dan rancangan validasi [23], [24], [25]. Oleh karena itu, hasil ini belum dapat digeneralisasi sebagai bukti keunggulan LSTM untuk kondisi ekstrem.

Penelitian terdahulu di Indonesia telah mengintegrasikan pembelajaran mesin dengan GIS untuk mendukung prediksi banjir [6], sedangkan kegiatan mitigasi di Purwodadi menekankan kesiapsiagaan masyarakat [21]. Penelitian ini melengkapi arah tersebut melalui analisis jumlah kejadian per kecamatan dan audit keterbatasan data. Namun, keluaran model tetap merupakan informasi pendukung; penilaian operasional memerlukan verifikasi geospasial, hidrologis, dan lapangan.

Implikasi mitigasi yang paling dapat dipertanggungjawabkan adalah penggunaan hasil sebagai alat penyaringan awal. Wilayah dengan kesalahan prediksi tinggi menunjukkan lokasi yang membutuhkan audit data lebih lanjut, sedangkan perubahan curah hujan dapat digunakan sebagai konteks temporal. Keluaran tidak boleh diterjemahkan menjadi probabilitas banjir atau peringatan dini tanpa validasi dan integrasi data real-time.

Keterbatasan penelitian mencakup 19 unit spasial, 24 observasi hujan bulanan, tidak tersedianya deret debit numerik yang seragam, dan belum adanya label banjir bulanan per kecamatan yang tersinkronisasi dengan curah hujan. Penelitian berikutnya perlu memperpanjang periode data, menambahkan fitur hidrologis yang terukur, dan menggunakan validasi temporal independen.

Kontribusi utama penelitian adalah proses evaluasi yang transparan: perbandingan fitur topografi, pelaporan kesalahan LOOCV seluruh kecamatan, audit data hidrologis, perbandingan LSTM dengan baseline, dan

penjelasan keterbatasan sampel. Hasil belum mendukung klaim model prediksi operasional, tetapi memberikan dasar terukur untuk memperbaiki data dan rancangan validasi pada penelitian selanjutnya.

4. KESIMPULAN

Penelitian ini mengevaluasi kerangka pemodelan spasial-temporal banjir menggunakan Random Forest (RF) dan Long Short-Term Memory (LSTM) untuk mendukung analisis mitigasi. Pada komponen spasial, RF mengestimasi jumlah kejadian banjir tahun 2025 pada 19 kecamatan berdasarkan lima indikator bencana tahun 2024 dan elevasi rata-rata. Model BPBD + elevasi menghasilkan MAE 2,122, RMSE 2,650, dan $R^2 -0,531$ pada LOOCV; penambahan elevasi menurunkan MAE dibandingkan model BPBD saja, tetapi RMSE dan R^2 tidak membaik sehingga peningkatan kinerja tidak konsisten. Debit intake musim kemarau dan Q25 tidak digunakan karena data tidak tersedia atau belum dapat dipetakan secara konsisten. Galat terbesar terdapat pada Toroh dan Purwodadi, sedangkan kemungkinan peran konfluensi sungai, tanggul, drainase, dan pola permukiman masih harus diverifikasi, bukan diasumsikan sebagai penyebab. Pada komponen temporal, LSTM dengan look-back tiga bulan menghasilkan MAE 117,34 mm, RMSE 157,09 mm, dan $R^2 -0,313$, dibandingkan persistence baseline dengan MAE 153,85 mm dan RMSE 216,14 mm. Namun, 24 observasi bulanan dan enam titik uji membatasi kesimpulan generalisasi, terutama saat terjadi fluktuasi ekstrem. Dengan demikian, RF dan LSTM hanya diposisikan sebagai kerangka decision-support yang menyajikan informasi spasial dan temporal secara berdampingan, bukan prediktor hibrida atau sistem peringatan dini tervalidasi. Penelitian selanjutnya memerlukan deret hujan lebih panjang, label banjir bulanan per kecamatan, data debit dan tinggi muka air aktual, fitur penggunaan lahan dan jaringan sungai, serta validasi temporal-spasial independen. Kontribusi penelitian ini terletak pada evaluasi metodologis melalui perbandingan fitur spasial, analisis galat LOOCV pada 19 kecamatan, evaluasi LSTM terhadap persistence baseline, serta audit keterbatasan data hidrologis. Hasil penelitian menjadi dasar untuk perbaikan pengumpulan data dan rancangan validasi pada studi lanjutan, tanpa mengklaim kesiapan operasional model.

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: Mochamad Amry Assiva, Hadi Tanuji, Eka Renny Noviati Wahyuni.

Metodologi: Mochamad Amry Assiva, Hadi Tanuji.

Perangkat Lunak: Mochamad Amry Assiva, Hadi Tanuji.

Validasi: Mochamad Amry Assiva, Hadi Tanuji, Muhammad Fa'iq Robbani.

Analisis Formal: Mochamad Amry Assiva, Hadi Tanuji, Eka Renny Noviati Wahyuni.

Investigasi: Seluruh penulis.

Sumber Daya: Seluruh penulis.

Kurasi Data: Hadi Tanuji, Eka Renny Noviati Wahyuni, Muhammad Fa'iq Robbani.

Penulisan Draf Awal: Mochamad Amry Assiva, Hadi Tanuji.

Penulisan, Peninjauan, dan Penyuntingan: Seluruh penulis.

Visualisasi: Mochamad Amry Assiva, Hadi Tanuji.

Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data yang disajikan dalam penelitian ini tersedia berdasarkan permintaan kepada penulis korespondensi.

5.3 Pendanaan

Penelitian ini didanai melalui skema Penelitian Dosen Pemula (PDP) tahun 2026 sesuai proposal penelitian.

5.4 Pernyataan Dewan Peninjau Institusional

Tidak Berlaku (Not applicable).

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Tidak Berlaku (Not applicable).

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan



Para penulis menggunakan AI Generatif untuk meningkatkan kejelasan penulisan dalam artikel ini. Para penulis telah meninjau dan menyunting konten yang dihasilkan dengan bantuan AI serta bertanggung jawab penuh atas versi akhir publikasi.

REFERENCES

- [1] Q. B. Pham et al., “Evaluation of various boosting ensemble algorithms for predicting flood hazard susceptibility areas,” *Geomatics, Nat. Hazards Risk*, vol. 12, no. 1, pp. 2607–2628, Jan. 2021, doi: 10.1080/19475705.2021.1968510.
- [2] S. A. Abu El-Magd, “Random forest and naïve Bayes approaches as tools for flash flood hazard susceptibility prediction, South Ras El-Zait, Gulf of Suez Coast, Egypt,” *Arab. J. Geosci.*, vol. 15, no. 3, p. 217, Feb. 2022, doi: 10.1007/s12517-022-09531-3.
- [3] C. Singha, V. K. Rana, Q. B. Pham, D. C. Nguyen, and E. Łupikasza, “Integrating machine learning and geospatial data analysis for comprehensive flood hazard assessment,” *Environ. Sci. Pollut. Res.*, vol. 31, no. 35, pp. 48497–48522, Jul. 2024, doi: 10.1007/s11356-024-34286-7.
- [4] R. Abedi, R. Costache, H. Shafizadeh-Moghadam, and Q. B. Pham, “Flash-flood susceptibility mapping based on XGBoost, random forest and boosted regression trees,” *Geocarto Int.*, vol. 37, no. 19, pp. 5479–5496, Oct. 2022, doi: 10.1080/10106049.2021.1920636.
- [5] S. E. Purwati and Y. Pristyanto, “Model Random Forest and Support Vector Machine for Flood Classification in Indonesia,” *sinkron*, vol. 8, no. 4, pp. 2261–2268, Oct. 2024, doi: 10.33395/sinkron.v8i4.13973.
- [6] M. R. Perdana Putra, Rama Ashari, Muhirin, Azib Widad Zuhaily Imam, and Kusriani, “Flood Prediction Using Machine Learning Model Integrated with Geographical Information System,” *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 10, no. 2, pp. 121–126, Mar. 2025, doi: 10.23917/khif.v10i2.3723.
- [7] F. A. A. Hadi et al., “Machine learning techniques for flood forecasting,” *J. Hydroinformatics*, vol. 26, no. 4, pp. 779–799, Apr. 2024, doi: 10.2166/hydro.2024.208.
- [8] J. Zhou, L. Chen, T. Hu, H. Lu, Y. Shi, and L. Chen, “The comparative study of machine learning agent models in flood forecasting for tidal river reaches,” *Sci. Rep.*, vol. 15, no. 1, p. 19130, May 2025, doi: 10.1038/s41598-025-04633-w.
- [9] V. Kumar, H. M. Azamathulla, K. V. Sharma, D. J. Mehta, and K. T. Maharaj, “The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management,” *Sustainability*, vol. 15, no. 13, p. 10543, Jul. 2023, doi: 10.3390/su151310543.
- [10] Y. Zhang, S. Ragetti, P. Molnar, O. Fink, and N. Peleg, “Generalization of an Encoder-Decoder LSTM model for flood prediction in ungauged catchments,” *J. Hydrol.*, vol. 614, p. 128577, Nov. 2022, doi: 10.1016/j.jhydrol.2022.128577.
- [11] Z. Cui, S. Guo, Y. Zhou, and J. Wang, “Exploration of dual-attention mechanism-based deep learning for multi-step-ahead flood probabilistic forecasting,” *J. Hydrol.*, vol. 622, p. 129688, Jul. 2023, doi: 10.1016/j.jhydrol.2023.129688.
- [12] Balai Besar Wilayah Sungai Pemali Juana, “Data Sumber Daya Air,” Kementerian Pekerjaan Umum. Accessed: Aug. 04, 2026. [Online]. Available: <https://sda.pu.go.id/balai/bbwspemalijuana/pages/informasisetiapaat>
- [13] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
- [14] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
- [15] M. A. Assiva, “Analisis sentimen terhadap pariwisata di Kabupaten Grobogan berbasis Orange menggunakan Naive Bayes,” *Innov. J. Soc. Sci. Res.*, vol. 4, no. 6, pp. 2351–2359, 2024, doi: 10.31004/innovative.v4i6.16469.
- [16] A. Saleh, A. Yuzir, N. Sabtu, S. K. M. Abujayyab, M. R. Bunmi, and Q. B. Pham, “Flash flood susceptibility mapping in urban area using genetic algorithm and ensemble method,” *Geocarto Int.*, vol. 37, no. 25, pp. 10199–10228, Dec. 2022, doi: 10.1080/10106049.2022.2032394.
- [17] M. Wahba, R. Essam, M. El-Rawy, N. Al-Arif, F. Abdalla, and W. M. Elsadek, “Forecasting of flash flood susceptibility mapping using random forest regression model and geographic information systems,” *Heliyon*, vol. 10, no. 13, p. e33982, Jul. 2024, doi: 10.1016/j.heliyon.2024.e33982.
- [18] L. Zhang, H. Qin, J. Mao, X. Cao, and G. Fu, “High temporal resolution urban flood prediction using attention-based LSTM models,” *J. Hydrol.*, vol. 620, p. 129499, May 2023, doi: 10.1016/j.jhydrol.2023.129499.
- [19] B. Sabzipour et al., “Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment,” *J. Hydrol.*, vol. 627, p. 130380, Dec. 2023, doi: 10.1016/j.jhydrol.2023.130380.
- [20] J. S. Navarro, R. Zhuang, C. Albertini, and S. Manfreda, “Mapping flood susceptibility using Random Forest exploiting satellite observations and geomorphic features,” *Sci. Total Environ.*, vol. 1002, p. 180592, Nov. 2025, doi: 10.1016/j.scitotenv.2025.180592.
- [21] B. L. Al Zakina, Z. Zainuddin, T. B. Santoso, and L. Safitri, “Penyuluhan Mitigasi Bencana Banjir Sebagai Upaya Kesiapsiagaan Bencana Di Kota Purwodadi,” *J. Hum. Educ.*, vol. 4, no. 4, pp. 80–84, Jul. 2024, doi: 10.31004/jh.v4i4.1147.
- [22] Badan Meteorologi Klimatologi dan Geofisika, “Data Online Meteorologi, Klimatologi, dan Geofisika,” Badan Meteorologi, Klimatologi, dan Geofisika. Accessed: Aug. 04, 2026. [Online]. Available: <https://dataonline.bmkg.go.id/dataonline-home>
- [23] M. Waqas and U. W. Humphries, “A critical review of RNN and LSTM variants in hydrological time series predictions,” *MethodsX*, vol. 13, p. 102946, 2024, doi: 10.1016/j.mex.2024.102946.
- [24] N. Byaruhanga, D. Kibirige, S. Gokool, and G. Mkhonta, “Evolution of flood prediction and forecasting models for flood early warning systems: A scoping review,” *Water*, vol. 16, no. 13, p. 1763, 2024, doi: 10.3390/w16131763.
- [25] M. Asif, M. M. Kuglitsch, I. Pelivan, and R. Albano, “Review and Intercomparison of Machine Learning Applications for Short-term Flood Forecasting,” *Water Resour. Manag.*, vol. 39, no. 5, pp. 1971–1991, Mar. 2025, doi: 10.1007/s11269-025-04093-x.