

OPTIMASI PREDIKSI KETINGGIAN MUKA AIR SUNGAI CILIWUNG DI PINTU AIR POS DEPOK MENGGUNAKAN ENSEMBLE LEARNING

Ririn Andriyani^{1*}, Wulan Kusuma Wardani²

¹Sains Atmosfer dan Keplanetan, Fakultas Sains, Institut Teknologi Sumatera; Jl. Terusan Ryacudu, Way Huwi, Kec. Jati Agung, Kabupaten Lampung Selatan, Lampung, 35365, (0721) 8030188

²Teknik Sistem Energi, Fakultas Teknologi Industri, Institut Teknologi Sumatera; Jl. Terusan Ryacudu, Way Huwi, Kec. Jati Agung, Kabupaten Lampung Selatan, Lampung, 35365, (0721) 8030188

Keywords:

Tinggi Muka Air;
Machine Learning;
Hidrologi.

Correspondent Email:

ririn.andriyani@sap.itera.ac.id

Abstrak. Prediksi ketinggian muka air menjadi aspek penting dalam sistem peringatan dini banjir, terutama di wilayah hilir Sungai Ciliwung. Penelitian ini dilakukan pada Pintu Air Depok, titik pengamatan strategis aliran air dari hulu menuju DKI Jakarta, dengan tujuan membandingkan dan mengoptimalkan prediksi menggunakan ensemble machine learning. Dataset berupa data deret waktu resolusi 1 jam mencakup ketinggian muka air sebagai target, serta curah hujan, suhu, dan ketinggian muka air sebelumnya sebagai variabel prediktor, diperoleh dari BPBD DKI Jakarta dan ERA5 ECMWF. Metode yang digunakan meliputi XGBoost, Random Forest, dan Stacked Learning. Hasil menunjukkan seluruh model mampu memprediksi ketinggian muka air dengan baik, ditunjukkan oleh kombinasi $R^2 > 0,74$, MAE rendah, RMSE terkendali, dan MAPE kecil pada data pengujian. Stacked Learning memberikan keseimbangan terbaik antara akurasi dan generalisasi dengan error terendah, XGBoost unggul pada data training namun cenderung overfit ringan, sedangkan Random Forest menunjukkan performa yang stabil. Temuan ini menegaskan bahwa pendekatan ensemble mampu menghasilkan prediksi yang akurat, stabil, dan dapat mendukung sistem prediksi real-time untuk mitigasi banjir.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. *Water level prediction is a critical component of flood early warning systems, particularly in the downstream area of the Ciliwung River. This study was conducted at Depok Water Gate, a strategic monitoring point for water flow from the upstream regions to Jakarta, aiming to compare and optimize predictions using ensemble machine learning approaches. The dataset consists of hourly time-series data, including water level as the target variable, and rainfall, temperature, and previous water levels as predictors, obtained from BPBD DKI Jakarta and ERA5 ECMWF. The models applied in this study include XGBoost, Random Forest, and Stacked Learning. Results show that all models achieved satisfactory predictive performance, indicated by R^2 above 0.74, low MAE, controlled RMSE, and minimal MAPE on the testing dataset. Stacked Learning provided the best balance between accuracy and generalization, with the lowest prediction errors, while XGBoost excelled on training data but exhibited slight overfitting. Random Forest demonstrated stable performance across training and testing datasets. These findings highlight that ensemble approaches can deliver accurate and robust predictions, making them suitable for real-time water level forecasting and flood mitigation in downstream regions of the Ciliwung River.*

1. PENDAHULUAN

Kejadian banjir di Indonesia menunjukkan tren peningkatan dalam frekuensi dan dampak yang ditimbulkan [1], [2]. Data Badan Nasional Penanggulangan Bencana (BNPB) mencatat terjadi sebanyak 1.584 kejadian banjir sepanjang tahun 2025 [3]. Selain itu, sekitar 49.776.300 hektar wilayah di Indonesia termasuk daerah rawan banjir, dengan risiko 109.618.626 jiwa terpapar [4], menunjukkan bahwa ancaman banjir merupakan masalah nasional yang berdampak luas pada populasi dan infrastruktur. Peningkatan muka air sungai atau waduk sering menjadi indikator awal terjadinya banjir, karena luapan air terjadi ketika kapasitas aliran tidak lagi mampu menampung volume air [5]. Oleh karena itu, prediksi ketinggian muka air menjadi aspek penting dalam upaya mitigasi banjir serta perencanaan pengelolaan sumber daya air.

Karakteristik hidrologi dalam penelitian ini difokuskan pada wilayah Daerah Aliran Sungai (DAS) Ciliwung, khususnya pada titik pengamatan Pintu Air Depok yang berperan sebagai indikator penting dalam sistem pengendalian banjir. Kenaikan ketinggian muka air pada titik ini berdampak terhadap wilayah hilir yang sebagian besar berada di Jakarta Timur, Jakarta Selatan, dan sebagian Jakarta Pusat [6]. Secara geografis, DAS Ciliwung melintasi wilayah dengan kepadatan penduduk tinggi serta penggunaan lahan yang beragam, mulai dari kawasan hulu yang didominasi daerah pegunungan hingga wilayah hilir berupa dataran rendah perkotaan [7]. Kondisi ini menyebabkan dinamika aliran sungai sangat dipengaruhi oleh interaksi antara faktor topografi, curah hujan, serta perubahan tata guna lahan.

Berbagai penelitian telah dilakukan untuk memprediksi ketinggian muka air menggunakan metode numerik dan machine learning. Model hidrologi fisik, seperti HEC-RAS dan SWAT, mengandalkan persamaan hidrodinamika dan input curah hujan, topografi, serta penggunaan lahan [8], [9], [10], [11]. Meskipun model ini banyak digunakan dalam analisis hidrologi, pendekatan tersebut memiliki keterbatasan dalam menangani ketidakpastian data serta pola hubungan non-linear yang kompleks. Sebagai alternatif, machine learning mulai banyak diterapkan untuk prediksi ketinggian muka air berbasis

data time series. Penelitian oleh Hasyimi, dkk menunjukkan bahwa model LSTM berhasil memprediksi ketinggian muka air sungai ciliwung, namun belum cukup akurat karena hanya menggunakan data historisnya [7]. Selain itu, studi Irawan, dkk menggunakan Random Forest untuk memprediksi tinggi muka air berbasis data iklim dan menunjukkan kinerja model yang cukup baik dengan curah hujan menjadi faktor paling berpengaruh dalam dinamikanya [12]. XGBoost dan Random Forest juga dilaporkan efektif dalam memprediksi tinggi gelombang signifikan dan kenaikan muka air laut berbasis data meteorologi–oseanografi dengan tingkat akurasi yang sangat tinggi [13]. Meskipun berbagai penelitian telah menunjukkan potensi metode machine learning dalam prediksi ketinggian muka air, beberapa keterbatasan masih ditemukan. Pertama, banyak penelitian hanya berfokus pada satu atau dua model sehingga belum terdapat perbandingan yang komprehensif antar algoritma machine learning pada dataset lokal di Indonesia. Kedua, evaluasi terhadap performa prediksi jangka pendek maupun jangka menengah masih terbatas sehingga sulit menentukan model yang paling optimal untuk kebutuhan sistem peringatan dini. Ketiga, sebagian besar penelitian belum melakukan optimasi hyperparameter secara sistematis, sehingga potensi peningkatan akurasi model belum dimanfaatkan secara maksimal.

Berdasarkan kondisi tersebut, penelitian ini dilakukan untuk mengisi kesenjangan penelitian dengan membandingkan beberapa metode machine learning, yaitu Random Forest, XGBoost, dan Stacked Ensemble Model dalam memprediksi ketinggian muka air di Pintu Air Depok, yang merupakan salah satu titik pengamatan penting pada DAS Ciliwung. Penelitian ini juga melakukan optimasi hyperparameter serta evaluasi kinerja model secara komprehensif. Kebaruan penelitian ini terletak pada perbandingan sistematis beberapa algoritma machine learning pada dataset lokal ketinggian muka air serta penerapan optimasi model untuk meningkatkan akurasi prediksi, yang selanjutnya dapat digunakan sebagai dasar pengembangan sistem peringatan dini banjir bagi wilayah hilir, khususnya di Jakarta Timur, Jakarta Selatan, dan sebagian Jakarta Pusat.

2. TINJAUAN PUSTAKA

2.1 Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan yang berfokus pada pengembangan sistem komputer yang mampu meningkatkan kinerjanya secara otomatis melalui pengalaman dan proses pembelajaran dari data [14]. Pendekatan ini banyak digunakan pada berbagai bidang seperti keuangan, kesehatan, hingga lingkungan karena kemampuannya dalam mengidentifikasi hubungan non-linear yang kompleks dalam data [14], [15], [16], [17], [18], [19]. Dalam analisis berbasis data time series, machine learning mampu memanfaatkan informasi historis untuk memodelkan pola perubahan suatu variabel dari waktu ke waktu sehingga dapat digunakan untuk melakukan prediksi pada periode selanjutnya [19].

Salah satu pendekatan dalam machine learning yang banyak digunakan untuk meningkatkan performa model adalah ensemble learning. Ensemble learning merupakan teknik yang menggabungkan beberapa model pembelajaran untuk menghasilkan prediksi yang lebih akurat dan stabil dibandingkan menggunakan satu model tunggal [20]. Prinsip dasar metode ini adalah bahwa kombinasi beberapa model dapat meningkatkan akurasi prediksi model [21]. Beberapa pendekatan utama dalam ensemble learning antara lain bagging, boosting, dan stacking, yang masing-masing memiliki mekanisme berbeda dalam membangun dan menggabungkan model [22]. Pendekatan bagging (Bootstrap Aggregating) bekerja dengan membangun beberapa model menggunakan sampel data yang berbeda yang dihasilkan melalui proses bootstrap. Setiap model dilatih secara independen, kemudian hasil prediksi dari seluruh model digabungkan, biasanya dengan melakukan rata-rata untuk prediksi numerik atau pemungutan suara untuk klasifikasi untuk menghasilkan prediksi akhir yang lebih stabil serta mampu mengurangi variansi model [23]. Berbagai algoritma machine learning dapat digunakan sebagai model dasar dalam bagging, seperti Decision Tree, Logistic Regression, K-Nearest Neighbors (KNN), Gaussian Naive Bayes, dan Support Vector Machine (SVM). Salah satu algoritma yang menerapkan teknik bagging adalah Random Forest, yang membangun banyak decision tree dari berbagai subset data

dan menggabungkan hasil prediksinya untuk meningkatkan performa model [24]. Pendekatan boosting bekerja dengan menggabungkan beberapa model yang dibangun secara bertahap. Setiap model baru dilatih dengan memberikan perhatian lebih pada data yang sebelumnya sulit diprediksi oleh model sebelumnya, sehingga kesalahan prediksi dapat diperbaiki secara iteratif. Hasil prediksi dari seluruh model kemudian digabungkan untuk menghasilkan model akhir yang memiliki performa lebih baik serta kemampuan generalisasi yang lebih tinggi [25]. Salah satu algoritma boosting yang banyak digunakan adalah Extreme gradient boost (XGBoost), categorical boost (CatBoost), adaptive boost (AdaBoost), and light gradient boosting machine (LGBM) [26]. Pendekatan stacking merupakan metode ensemble learning yang menggabungkan beberapa model dasar (base learners) dengan menggunakan model lain sebagai meta-learner. Pada metode ini, prediksi dari model-model dasar digunakan sebagai input bagi meta-learner untuk menghasilkan prediksi akhir yang lebih akurat. Algoritma yang sering digunakan sebagai meta-learner antara lain Logistic Regression atau Linear Regression [27].

2.2 Ketinggian Muka Air

Ketinggian muka air (water level) merupakan salah satu parameter penting dalam kajian hidrologi yang menggambarkan posisi permukaan air pada suatu badan air seperti sungai, danau, atau waduk pada waktu tertentu ([28]. Parameter ini berperan penting dalam memahami dinamika aliran, ketersediaan sumber daya air, serta potensi terjadinya bencana hidrologi seperti banjir dan kekeringan.

Ketinggian muka air dipengaruhi oleh berbagai faktor, antara lain curah hujan, suhu, karakteristik daerah aliran sungai (DAS), serta aktivitas antropogenik seperti pengoperasian bendungan dan waduk [29]. Kompleksitas faktor-faktor tersebut menyebabkan pola perubahan ketinggian muka air bersifat dinamis dan non-linear, sehingga memerlukan pendekatan analisis yang mampu menangkap hubungan antar variabel secara lebih komprehensif.

3. METODE PENELITIAN

3.1 Lokasi dan Data Penelitian

Penelitian ini dilakukan pada Pintu Air Pos Depok yang terletak pada aliran Sungai Ciliwung, yang merupakan salah satu sungai utama yang mengalir dari wilayah hulu di Bogor melalui Depok hingga ke hilir di DKI Jakarta [7]. Secara hidrologis, wilayah ini memiliki peran penting dalam sistem pengendalian banjir Jakarta karena berfungsi sebagai titik pemantauan awal aliran air dari hulu. Kondisi tersebut menjadikan Pintu Air Pos Depok sebagai lokasi strategis dalam memantau dinamika ketinggian muka air serta sebagai bagian dari sistem peringatan dini banjir di wilayah hilir.

Dataset yang digunakan dalam penelitian ini terdiri dari dua komponen utama, yaitu data ketinggian muka air sebagai variabel target dan data curah hujan, suhu, serta nilai lag ketinggian muka air sebagai variabel prediktor.

Data ketinggian muka air diperoleh dari BPBD DKI Jakarta melalui sistem pemantauan daring pada stasiun Pintu Air Pos Depok [30]. Data ini merupakan data deret waktu (time series) dengan interval pencatatan setiap 1 jam, sehingga mampu merepresentasikan dinamika perubahan ketinggian muka air secara detail. Sementara itu, data curah hujan dan suhu diperoleh dari ERA5 yang merupakan produk reanalysis dari European Centre for Medium-Range Weather Forecasts (ECMWF). Data ERA5 menyediakan informasi meteorologi dengan resolusi temporal yang tinggi dan cakupan global, sehingga banyak digunakan dalam penelitian hidrologi sebagai sumber data alternatif pada wilayah yang memiliki keterbatasan data observasi lapangan [31]. Penggunaan variabel lag bertujuan untuk menangkap ketergantungan temporal dalam data deret waktu

3.2 Implementasi Ensemble Machine Learning

3.2.1 Preprocessing dan Optimasi Data

Data preprocessing merupakan tahap fundamental dalam penelitian ini karena berperan langsung terhadap kualitas dan performa model prediksi [32]. Selain itu, penelitian ini juga menerapkan pendekatan data-level optimization, yaitu optimasi yang dilakukan melalui peningkatan kualitas data sebelum proses pemodelan [33].

Tahapan yang dilakukan meliputi:

- a) Outlier Removal (IQR Method)
Outlier diidentifikasi dan dihapus menggunakan metode Interquartile Range (IQR), di mana suatu data dikategorikan sebagai outlier apabila berada di bawah $Q1 - 1,5 \times IQR$ atau di atas $Q3 + 1,5 \times IQR$, sebagaimana diperkenalkan dalam pendekatan Tukey [34]. Proses ini dilakukan pada semua variabel untuk mengurangi noise dan meningkatkan stabilitas model.
- b) Feature Engineering (Lag Variable)
Untuk menangkap pola temporal pada data deret waktu, dilakukan pembentukan variabel lag, yaitu representasi nilai pada periode sebelumnya sebagai prediktor. Penggunaan variabel ini memungkinkan model memahami ketergantungan antar waktu (autokorelasi) sehingga meningkatkan akurasi prediksi [35]. Pada penelitian ini, lag diterapkan pada variabel ketinggian muka air karena nilainya sangat dipengaruhi oleh kondisi sebelumnya, sehingga membantu model menangkap pola perubahan secara lebih akurat dan stabil.
- c) Handling Missing Values
Missing values ditangani menggunakan metode mean imputation dengan SimpleImputer. Pendekatan ini menjaga distribusi data tetap stabil tanpa mengurangi jumlah observasi [36].
- d) Feature Scaling (Standardization)
Normalisasi dilakukan menggunakan StandardScaler untuk menyamakan skala antar variabel. Transformasi dilakukan menggunakan persamaan [37]:

$$z = \frac{(x_i - \bar{x})}{\sigma} \quad (1)$$

Langkah ini sangat penting terutama untuk model seperti SVR yang sensitif terhadap skala data [38]. Dengan standarisasi, proses pembelajaran menjadi lebih stabil dan konvergen lebih cepat.

e) Data Splitting (Time Series Approach)

Dataset dibagi menjadi data training (70%) dan testing (30%) tanpa proses pengacakan (`shuffle = False`). Pendekatan ini mempertahankan urutan waktu sehingga model dapat mengevaluasi kemampuan prediksi pada data masa depan secara realistis.

3.2.2 Ensemble Learning

Penelitian ini melakukan perbandingan tiga jenis metode ensemble learning, yaitu boosting, bagging, dan stacking, untuk memprediksi ketinggian muka air. Model yang digunakan mencakup XGBoost sebagai representasi boosting, Random Forest untuk bagging, serta Stacked Ensemble dengan meta-learner untuk stacking. Pendekatan ini memungkinkan evaluasi komprehensif terhadap performa berbagai tipe ensemble pada dataset yang sama.

a) Implementasi Boosting

Metode boosting diterapkan melalui algoritma XGBoost menggunakan kelas `XGBRegressor` untuk membangun model prediksi ketinggian muka air. Model ini memanfaatkan variabel curah hujan, suhu, serta lag dari ketinggian muka air sebelumnya sebagai input.

Fungsi objektif yang digunakan adalah `reg:squarederror` untuk meminimalkan error kuadrat. Optimasi hiperparameter dilakukan menggunakan metode `GridSearchCV` dengan skema 3-fold cross-validation. Parameter yang dioptimasi meliputi `n_estimators`, `max_depth`, `learning_rate`, `subsample`, serta `colsample_bytree`. Dataset yang telah melalui tahap preprocessing dan pembagian data sebelumnya digunakan sebagai input model, dengan urutan temporal yang dipertahankan.

b) Implementasi Bagging

Metode Bagging dibangun melalui algoritma Random Forest Regressor. Random Forest terdiri dari sekumpulan

pohon keputusan yang dibangun secara paralel menggunakan bootstrap sampling, di mana setiap pohon dilatih pada subset data acak untuk meningkatkan stabilitas dan kemampuan generalisasi [39]. Prediksi akhir dihasilkan dari rata-rata prediksi seluruh pohon (ensemble averaging).

Berbeda dengan pendekatan pada XGBoost, proses optimasi model ini menggunakan `TimeSeriesSplit` dalam skema validasi silang untuk menjaga urutan temporal data selama proses pelatihan. Optimasi hiperparameter dilakukan menggunakan `GridSearchCV`, dengan parameter yang meliputi `n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf`.

c) Stacking Implementation

Metode stacking diterapkan menggunakan pendekatan `StackingRegressor`, yang mengkombinasikan model dasar dari `XGBRegressor`, `SVM`, dan `LinearRegression`. Ketiga model tersebut dipilih untuk merepresentasikan karakteristik pembelajaran yang berbeda, yaitu boosting, kernel-based learning, dan model linear. Prediksi dari masing-masing model dasar kemudian digunakan sebagai input bagi meta-learner, yaitu Linear Regression, untuk menghasilkan prediksi akhir. Proses pelatihan dilakukan dengan skema cross-validation sebanyak 3 fold (`cv=3`) untuk menghasilkan prediksi antar model yang lebih robust. Selain itu, parameter `passthrough=True` digunakan untuk menyertakan fitur asli bersama output model dasar ke dalam meta-learner, sehingga informasi awal tetap dimanfaatkan dalam proses pembelajaran akhir.

3.3 Metrik Evaluasi

Evaluasi performa model dilakukan untuk mengukur kemampuan prediksi dari masing-masing metode yang digunakan. Pengukuran dilakukan pada data training

dan testing menggunakan beberapa metrik evaluasi yang umum digunakan dalam permasalahan regresi. Metrik evaluasi yang digunakan meliputi koefisien determinasi (R^2), *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). R^2 digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data, sedangkan MAE, MSE, dan RMSE digunakan untuk mengukur besar kesalahan prediksi. MAPE digunakan untuk mengukur kesalahan dalam bentuk persentase.

Perhitungan metrik dilakukan dengan membandingkan nilai aktual dan prediksi menggunakan fungsi dari library **scikit-learn**. Nilai RMSE diperoleh dari akar kuadrat MSE, sedangkan MAPE dihitung sebagai rata-rata persentase kesalahan absolut.

4. HASIL DAN PEMBAHASAN

4.1 Kinerja Model Prediksi

Kinerja prediksi pada dataset pelatihan dan pengujian berdasarkan hasil evaluasi metrik dari model XGBoost, Random Forest (RF), dan Stacked Learning disajikan pada Tabel 1. Hasil pengujian menunjukkan bahwa seluruh model memiliki performa yang baik dalam memodelkan hubungan antara variabel input dan output, dengan nilai R^2 pada data testing berada di atas 0,74. Hal ini mengindikasikan bahwa model mampu menjelaskan lebih dari 74% variasi ketinggian muka air.

Model XGBoost menunjukkan performa terbaik pada data training dengan nilai R^2 sebesar 0,8570, diikuti oleh Stacked Learning sebesar 0,8035 dan Random Forest sebesar 0,7973. Nilai error pada XGBoost juga relatif rendah (MAE = 2,07; RMSE = 4,06; MAPE = 2,03%), yang menunjukkan kemampuan model dalam menangkap pola kompleks pada data pelatihan. Namun, pada data testing, model Stacked Learning menunjukkan performa paling unggul dengan nilai R^2 sebesar 0,7815, serta nilai error terendah (MAE = 1,81; RMSE = 3,57; MAPE = 1,68%). Model Random Forest dan XGBoost masing-masing menghasilkan R^2 sebesar 0,7437 dan 0,7547, dengan nilai error yang sedikit lebih tinggi dibandingkan model

stacking. Perbandingan antara performa data training dan testing menunjukkan adanya perbedaan karakteristik antar model dalam hal generalisasi. Model XGBoost memiliki selisih performa yang cukup signifikan antara training dan testing, yang mengindikasikan adanya kecenderungan overfitting ringan.

Sebaliknya, Random Forest menunjukkan performa yang lebih stabil antara data training dan testing. Pendekatan bagging yang digunakan mampu mengurangi variansi model, meskipun dengan konsekuensi akurasi yang sedikit lebih rendah dalam menangkap pola non-linear kompleks.

Model Stacked Learning menunjukkan keseimbangan terbaik antara akurasi dan generalisasi. Meskipun nilai R^2 pada data training tidak setinggi XGBoost, performa pada data testing justru lebih baik, menunjukkan kemampuan stacking dalam meningkatkan generalisasi dan menurunkan error [40].

Meskipun demikian, perbedaan kinerja antar ketiga model secara keseluruhan relatif tidak signifikan. Nilai evaluasi seperti R^2 , MAE, RMSE, dan MAPE berada dalam rentang yang berdekatan, yang menunjukkan bahwa seluruh model telah mampu bekerja secara optimal dalam menangkap pola hubungan antara variabel input dan output. Dengan kata lain, tidak terdapat satu metode yang secara mutlak mendominasi performa prediksi.

Kondisi ini mengindikasikan bahwa karakteristik data yang digunakan, khususnya jumlah variabel yang terbatas namun relevan, yaitu ketinggian muka air sebelumnya, curah hujan sebelumnya, dan suhu saat ini, telah cukup representatif untuk dimodelkan oleh berbagai algoritma machine learning. Oleh karena itu, peningkatan performa yang dihasilkan oleh model stacking lebih bersifat *incremental*, terutama dalam hal peningkatan generalisasi dan penurunan error, dibandingkan sebagai lompatan performa yang signifikan.

Temuan ini menunjukkan bahwa dalam konteks prediksi ketinggian muka air berbasis data time-series sederhana, pemilihan algoritma bukan menjadi satu-satunya faktor penentu kinerja model. Sebaliknya, kualitas dan representasi data, termasuk pemilihan variabel dan pendekatan *time-lag*, memiliki peran yang sama pentingnya dalam menentukan akurasi prediksi.

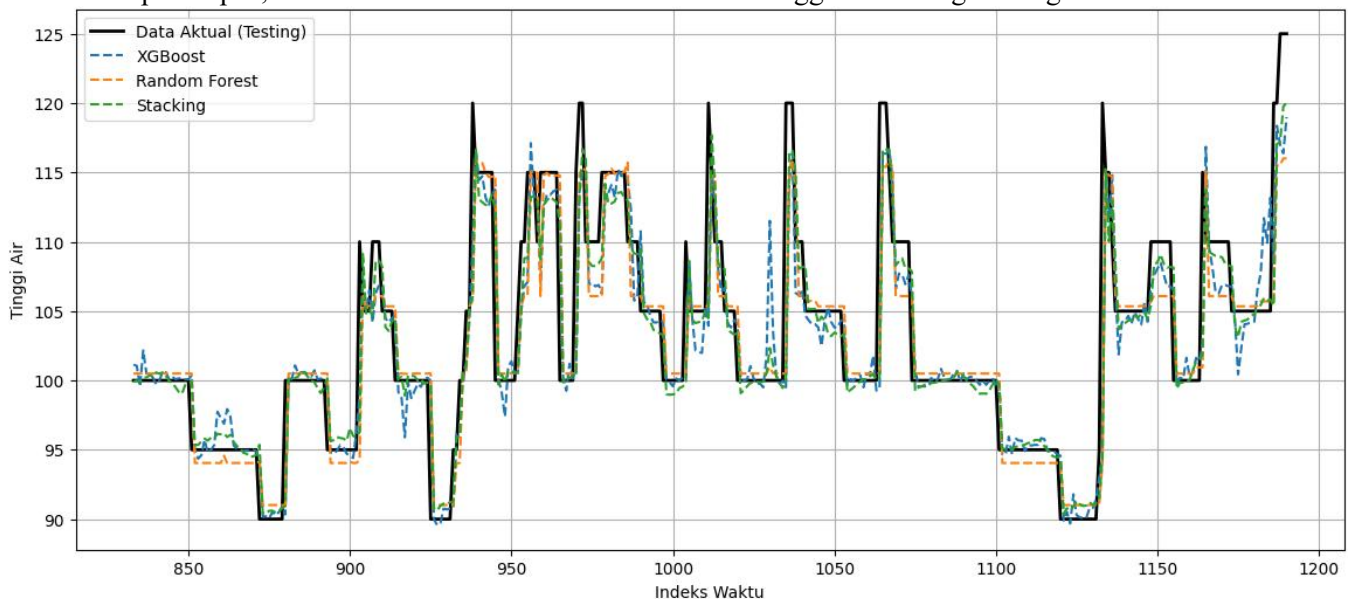
Tabel 1. Hasil Evaluasi Metrik Model Ensemble Learning

Methods	Training Dataset					Testing Dataset				
	R ²	MAE	MSE	RMSE	MAPE	R ²	MAE	MSE	RMSE	MAPE
XGBoost	0.8570	2.07	16.47	4.06	2.03%	0.7547	2.04	14.33	3.79	1.89%
RF	0.7973	2.41	23.34	4.83	2.37%	0.7437	2.02	14.97	3.87	1.88%
Stacking	0.8035	2.29	22.64	4.76	2.26%	0.7815	1.81	12.76	3.57	1.68%

Gambar 1 menunjukkan perbandingan antara nilai tinggi air aktual (garis hitam) dengan hasil prediksi dari masing-masing model, yaitu XGBoost (garis biru putus-putus), Random Forest (garis oranye putus-putus), dan Stacking (garis hijau putus-putus). Secara umum, ketiga model mampu mengikuti pola utama data, terutama dalam menangkap tren naik dan turun tinggi air. Namun demikian, terdapat perbedaan yang cukup jelas dalam tingkat akurasi dan kemampuan masing-masing model dalam merespons perubahan yang terjadi secara cepat.

Model XGBoost (garis biru putus-putus) menunjukkan kemampuan yang cukup baik dalam mengikuti pola data, khususnya pada perubahan yang relatif dinamis. Model ini cenderung lebih responsif dibandingkan model lainnya ketika terjadi fluktuasi. Namun, pada beberapa titik, terutama saat terjadi lonjakan atau penurunan yang tajam, prediksi XGBoost menunjukkan deviasi terhadap nilai aktual. Hal ini mengindikasikan bahwa meskipun XGBoost cukup adaptif, model ini masih memiliki

keterbatasan dalam menangkap pola yang sangat kompleks atau ekstrem secara konsisten. Sementara itu, model Random Forest (garis oranye putus-putus) menghasilkan prediksi yang lebih stabil dan cenderung halus. Model ini mampu merepresentasikan pola umum data dengan baik, tetapi terlihat kurang sensitif terhadap perubahan yang bersifat tiba-tiba. Pada kondisi di mana terjadi lonjakan atau penurunan yang signifikan, Random Forest cenderung tidak mampu mengikuti perubahan tersebut secara akurat. Hal ini menunjukkan bahwa Random Forest lebih unggul dalam generalisasi pola, namun memiliki keterbatasan dalam menangkap dinamika data yang bersifat fluktuatif. Di sisi lain, model Stacking (garis hijau putus-putus) menunjukkan performa yang paling baik di antara ketiga model. Prediksi yang dihasilkan cenderung lebih dekat dengan nilai aktual secara konsisten, baik pada kondisi stabil maupun saat terjadi perubahan yang cukup tajam. Hal ini menunjukkan bahwa pendekatan Stacking, yang mengombinasikan beberapa model, mampu memanfaatkan keunggulan masing-masing model dasar



Gambar 1. Perbandingan antara nilai tinggi air aktual (garis hitam) dengan hasil prediksi dari masing-masing model, yaitu XGBoost (garis biru putus-putus), Random Forest (garis oranye putus-putus), dan Stacking (garis hijau putus-putus)

sehingga menghasilkan prediksi yang lebih akurat. Model ini juga terlihat lebih seimbang dalam hal stabilitas dan sensitivitas terhadap perubahan.

Secara keseluruhan, model Stacking dapat dikatakan unggul dalam memprediksi tinggi air karena mampu menangkap pola data secara lebih komprehensif dibandingkan model tunggal seperti XGBoost dan Random Forest. Meskipun demikian, pemilihan model tetap

perlu disesuaikan dengan kebutuhan. XGBoost dapat menjadi pilihan yang baik ketika diperlukan model yang lebih responsif terhadap perubahan, sedangkan Random Forest lebih sesuai untuk kondisi yang mengutamakan stabilitas prediksi. Sementara itu, Stacking lebih direkomendasikan untuk kasus yang membutuhkan akurasi tinggi dengan mempertimbangkan kompleksitas pola data yang ada.

5. KESIMPULAN

Berdasarkan hasil penelitian sebelumnya, dapat disimpulkan sebagai berikut.

- Semua model (XGBoost, Random Forest, dan Stacked Learning) mampu memprediksi ketinggian muka air dengan baik, ditunjukkan oleh nilai $R^2 > 0,74$ pada data testing, serta nilai error MAE, RMSE, dan MAPE yang relatif rendah.
- Stacked Learning memberikan performa terbaik pada data pengujian dengan error terendah, sementara XGBoost unggul pada data pelatihan dan Random Forest lebih stabil.
- Perbedaan kinerja antar model relatif kecil, sehingga tidak terdapat metode yang secara signifikan lebih unggul.
- Pendekatan ensemble, khususnya Stacked Learning, meningkatkan stabilitas dan kemampuan generalisasi, meskipun peningkatan performa bersifat bertahap.

Keterbatasan penelitian terletak pada jumlah variabel dan pendekatan *time-lag* yang masih sederhana. Penelitian selanjutnya disarankan menambah variabel dan mengeksplorasi *time-lag* yang lebih panjang atau metode yang lebih kompleks untuk meningkatkan akurasi prediksi jangka panjang dan kemampuan generalisasi model.

DAFTAR PUSTAKA

- [1] V. S. A. Hendrawan, A. P. Rahardjo, H. G. Mawandha, E. Aldrian, A. Muhari, and D. Komori, "Review article: Past and future climate-related hazards in Indonesia," Jun. 10, 2025. doi: 10.5194/egusphere-2025-584.
- [2] A. Kawasaki and N. Shimomura, "Accelerated widening of economic disparity due to recurrent floods," *International Journal of Disaster Risk Reduction*, vol. 102, p. 104273, Feb. 2024, doi: 10.1016/j.ijdr.2024.104273.
- [3] Badan Nasional Penanggulangan Bencana, "Rangkuman Bencana tahun 2025," Geoport Data Bencana Indonesia. Accessed: Dec. 25, 2025. [Online]. Available: <https://gis.bnpb.go.id/>
- [4] Badan Nasional Penanggulangan Bencana (BNPB), "Indeks Risiko Bencana Indonesia." Accessed: Dec. 26, 2025. [Online]. Available: <https://inarisk.bnpb.go.id/irbi>
- [5] V. Kumar, K. V. Sharma, N. K. Mangukiya, D. K. Tiwari, P. V. Ramkar, and U. Rathnayake, "Machine learning applications in flood forecasting and predictions, challenges, and way-out in the perspective of changing environment," *AIMS Environ. Sci.*, vol. 12, no. 1, pp. 72–105, 2025, doi: 10.3934/environsci.2025004.
- [6] Unit Pengelola Data dan Informasi Ketataairan Dinas Sumber Daya Air Provinsi DKI Jakarta, "Tinggi Muka Air Online." Accessed: Apr. 08, 2026. [Online]. Available: <https://poskobanjir.dsdadki.web.id/>
- [7] M. Hasyimi, L. Nada, and F. Zain, "Model Prediksi Ketinggian Air Disemua Pintu Muka Air Sungai Ciliwung Dengan Menggunakan Metode Long Short-Term Memory," *Journal of Innovative and Creativity (Joecy)*, vol. 5, no. 2, pp. 21964–21973, Aug. 2025, doi: 10.31004/joecy.v5i2.3317.
- [8] M. I. Ullah, K. S. Qureshi, A. ur Rauf, and L. A. Shah, "Advanced floodplain mapping: HEC-RAS and ArcGIS pro application on Swat River," *Journal of Umm Al-Qura University for Engineering and Architecture*, vol. 15, no. 3, pp. 245–258, Sep. 2024, doi: 10.1007/s43995-024-00054-4.
- [9] M. B. Ansori, U. Lasminto, and A. A. G. Kartika, "Runoff Hydrograph Analysis of HEC-RAS 2D Flow Hydrodynamics Meteorological Rain-on-Grid on Observed Watershed: A Case Study of Wiroko Sub-Watershed," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 14, no. 2, pp. 575–581, Apr. 2024, doi: 10.18517/ijaseit.14.2.19813.

- [10] M. F. Ikhwal, M. I. Rau, S. Nur, T. Ferijal, W. Prayogo, and S. F. D. Saputra, "Application of Soil and Water Assessment Tool in Indonesia – a review and challenges," *Desalination Water Treat.*, vol. 277, pp. 105–119, Nov. 2022, doi: 10.5004/dwt.2022.29018.
- [11] E. Koltsida, N. Mamassis, and A. Kallioras, "Hydrological modeling using the Soil and Water Assessment Tool in urban and peri-urban environments: the case of Kifisos experimental subbasin (Athens, Greece)," *Hydrol. Earth Syst. Sci.*, vol. 27, no. 4, pp. 917–931, Feb. 2023, doi: 10.5194/hess-27-917-2023.
- [12] R. Irawan, G. Kholijah, and C. Multahadah, "Prediksi Tinggi Muka Air di Lahan Gambut Provinsi Jambi Berdasarkan Curah Hujan, Suhu dan Kelembapan Udara Menggunakan Metode Random Forest Regression," *JISTech (Journal of Islamic Science and Technology)* *JISTech*, vol. 10, no. 1, pp. 46–55, 2025, doi: 10.30829/jistech.v10i1.24569.
- [13] G. Erutjahjo and A. Supriyanto, "Prediksi Tinggi Gelombang Laut di Perairan Semarang – Demak dengan Menggunakan Random Forest dan XGBoost," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 4, pp. 869–881, Sep. 2025, doi: 10.30591/jpit.v10i4.9315.
- [14] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science (1979)*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.
- [15] J. Baehr *et al.*, "Predicting product life cycle environmental impacts with machine learning: Uncertainties and implications for future reporting requirements," *Sustain. Prod. Consum.*, vol. 52, pp. 511–526, Dec. 2024, doi: 10.1016/j.spc.2024.11.005.
- [16] R. Nopour, "Machine learning models in enhancing prediction of health-related indices among older adults: A scoping review," *Heliyon*, vol. 11, no. 12, p. e43510, Jul. 2025, doi: 10.1016/j.heliyon.2025.e43510.
- [17] F. Furizal, A. Ma'arif, and D. Rifaldi, "Application of Machine Learning in Healthcare and Medicine: A Review," *Journal of Robotics and Control (JRC)*, vol. 4, no. 5, pp. 621–631, Sep. 2023, doi: 10.18196/jrc.v4i5.19640.
- [18] I. Evdokimov, M. Kampouridis, and T. Papastilianou, "Application Of Machine Learning Algorithms to Free Cash Flows Growth Rate Estimation," *Procedia Comput. Sci.*, vol. 222, pp. 529–538, 2023, doi: 10.1016/j.procs.2023.08.191.
- [19] S. Ahmed, M. M. Alshater, A. El Ammari, and H. Hammami, "Artificial intelligence and machine learning in finance: A bibliometric review," *Res. Int. Bus. Finance*, vol. 61, p. 101646, Oct. 2022, doi: 10.1016/j.ribaf.2022.101646.
- [20] Syawaludin and A. B. Sunu W.A, "Perbandingan Model Ensemble Learning dalam Memprediksi Harga Sewa Indekos di Jakarta," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.7737.
- [21] T. N. Rincy and R. Gupta, "Ensemble Learning Techniques and its Efficiency in Machine Learning: A Survey," in *2nd International Conference on Data, Engineering and Applications (IDEA)*, IEEE, Feb. 2020, pp. 1–6. doi: 10.1109/IDEA49133.2020.9170675.
- [22] K. A. Nguyen, W. Chen, B.-S. Lin, and U. Seeboonruang, "Comparison of Ensemble Machine Learning Methods for Soil Erosion Pin Measurements," *ISPRS Int. J. Geoinf.*, vol. 10, no. 1, p. 42, Jan. 2021, doi: 10.3390/ijgi10010042.
- [23] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/BF00058655.
- [24] R. Samantaray and H. Das, "Performance Analysis of Machine Learning Algorithms Using Bagging Ensemble Technique for Software Fault Prediction," in *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, IEEE, Mar. 2023, pp. 1–7. doi: 10.1109/ISCON57294.2023.10111952.
- [25] R. E. Schapire, "The Boosting Approach to Machine Learning: An Overview," 2003, pp. 149–171. doi: 10.1007/978-0-387-21579-2_9.
- [26] A. M. Khan, A. BinZiad, and A. Al Subaii, "Boosting Algorithm Choice in Predictive Machine Learning Models for Fracturing Applications," in *SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition*, SPE, Oct. 2021. doi: 10.2118/205642-MS.
- [27] Q. Wang and H. Lu, "A novel stacking ensemble learner for predicting residual strength of corroded pipelines," *Npj Mater. Degrad.*, vol. 8, no. 1, p. 87, Aug. 2024, doi: 10.1038/s41529-024-00508-z.
- [28] C. Ven Te Chow, D. R. Maidment, and L. W. Mays, *Applied Hydrology*. Wellington: Technical Correspondence School, 1963.
- [29] Q. Wang and H. Wang, "Estimating river water level using satellite altimetry and water surface temperature data in a regulated subtropical river," *J. Hydrol. Reg. Stud.*, vol. 59, p. 102376, Jun. 2025, doi: 10.1016/j.ejrh.2025.102376.

- [30] Dinas SDA Prov. DKI Jakarta, “Pemantauan Tinggi Muka Air (TMA).” Accessed: Apr. 08, 2026. [Online]. Available: <https://bpbd.jakarta.go.id/waterlevel>
- [31] H. Hersbach *et al.*, “The ERA5 global reanalysis,” *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, Jul. 2020, doi: 10.1002/qj.3803.
- [32] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd Edition. Burlington: Morgan Kaufmann Publishers, 2011.
- [33] W. Mu, R. Cardelli, and S. Ferrari, “Data Preprocessing Techniques for Machine Learning Towards Improving Building Energy Performance: A Systematic Review,” *Energies (Basel)*, vol. 19, no. 6, p. 1561, Mar. 2026, doi: 10.3390/en19061561.
- [34] *Statistical Methods*. Elsevier, 2022. doi: 10.1016/C2019-0-02521-6.
- [35] M. Shafiuzzaman *et al.*, “Enhanced very short-term load forecasting with multi-lag feature engineering and prophet-XGBoost-CatBoost architecture,” *Energy*, vol. 335, p. 137981, Oct. 2025, doi: 10.1016/j.energy.2025.137981.
- [36] S. Caton, S. Malisetty, and C. Haas, “Impact of Imputation Strategies on Fairness in Machine Learning,” *Journal of Artificial Intelligence Research*, vol. 74, pp. 1011–1035, Jun. 2022, doi: 10.1613/jair.1.13197.
- [37] C. Yan, “A review on spectral data preprocessing techniques for machine learning and quantitative analysis,” *iScience*, vol. 28, no. 7, p. 112759, Jul. 2025, doi: 10.1016/j.isci.2025.112759.
- [38] Y. Yu, X. Wang, and B. Liu, “Combining feature scaling estimation with SVM classifier design using GA approach,” *Journal of Electronics (China)*, vol. 22, no. 5, pp. 550–557, Sep. 2005, doi: 10.1007/s11767-005-0017-x.
- [39] D. Horn, T. M. Krael, T. N. T. Tran, A. Groll, and C. Jentsch, “The impact of random tree depth—a novel randomization process for ensemble methods,” *Comput. Stat.*, vol. 41, no. 1, p. 25, Jan. 2026, doi: 10.1007/s00180-025-01697-0.
- [40] D. R. K. Saputra, Y. V. Via, and A. N. Sihananto, “Deteksi Anomali Menggunakan Ensemble Learning dan Random Oversampling pada Penipuan Transaksi Keuangan,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4910.