

VALIDASI SILANG REGRESI RIDGE DAN STEPWISE UNTUK PERSAMAAN TERGENERALISASI ANTARA DEBIT BENDUNGAN SUTAMI DAN CURAH HUJAN WILAYAH

Ahmad Luthfi^{1,2}, Eko Yuli Handoko^{1*}, Muhammad Aldila Syariz¹,
Ardhasena Sopaheluwakan², Andang Kurniawan²

¹Departemen Teknik Geomatika, Fakultas Teknik Sipil Perencanaan dan Kebumihan-
Institut Teknologi Sepuluh Nopember Surabaya

²Badan Meteorologi Klimatologi dan Geofisika
e-mail : ekoyh@its.ac.id

Abstrak. Informasi debit masuk bendungan sangat krusial untuk pengelolaan sumber daya air di Bendungan Sutami, Daerah Aliran Sungai (DAS) Brantas Hulu, Jawa Timur. Tantangan utama dalam penelitian ini adalah konversi data curah hujan dasarian menjadi informasi debit yang mampu menghadapi kendala terkait konsistensi sehingga memerlukan pendekatan yang lebih robust melalui validasi silang. Penelitian ini mengevaluasi persamaan regresi ridge dan stepwise dengan pendekatan jeda untuk menyusun persamaan curah hujan-debit menggunakan data curah hujan wilayah menggunakan enam raster untuk periode Januari I 2021–Desember III 2023 (108 dasarian). Pemilihan jeda dilakukan melalui analisis korelasi, dengan korelasi negatif sebagai batas penentuan. Regresi stepwise dioptimalkan secara bertahap hingga jeda maksimum, sedangkan regresi ridge diuji dengan rentang λ 0–50. Persamaan yang dihasilkan divalidasi menggunakan metode validasi silang dengan rasio pembagian data 50:50. Hasil menunjukkan bahwa metode interpolasi *Inverse Distance Weighting* (IDW) menghasilkan persamaan terbaik, dengan regresi stepwise optimal pada lima variabel prediktor dan regresi ridge pada λ 0,40. Nilai *Mean Absolute Error* (MAE) dan koefisien korelasi menunjukkan bahwa persamaan tergeneralisasi dapat menghasilkan galat rendah. Penelitian ini merekomendasikan penggunaan informasi debit permukaan yang dikonversi dari data curah hujan sebagai pengganti parameter sifat hujan dalam perencanaan operasional bendungan, dengan evaluasi bulanan untuk memastikan akurasi model. Kolaborasi antarinstansi disarankan untuk mendukung analisis lanjutan. Hasil penelitian menunjukkan bahwa informasi debit yang dihasilkan mampu mendukung pengelolaan operasional Bendungan Sutami dengan tingkat akurasi yang memadai untuk aplikasi praktis.

Kata Kunci: Persamaan Debit, Validasi Silang, Regresi Ridge, Regresi Stepwise

Abstract. Inflow discharge information is critical for water resource management at the Sutami Dam, located in the Upper Brantas River Basin, East Java. The primary challenge in this study is converting 10-days rainfall data into inflow information capable of addressing consistency and necessitating a more robust approach through cross-validation. This study evaluates ridge regression and stepwise regression with lag approach to formulate rainfall-inflow equations using zonal rainfall data derived from six raster for period of January I 2021–December III 2023 (108 10-days periods). Lag selection is performed through correlation analysis, with negative correlation as the threshold. Stepwise regression was incrementally optimized up to maximum lag, while ridge regression was tested with a λ range of 0–50. The equations were validated using a cross-validation method with a 50:50 split ratio. The results indicate that the *Inverse Distance Weighting* (IDW) interpolation method produced the best equation, with stepwise regression optimized at five predictor variables and ridge regression at a λ value of 0.40. The *Mean Absolute Error* (MAE) and correlation coefficient values demonstrate that the generalized equations yield low error rates. This study recommends the use of surface inflow information converted from rainfall data as a substitute for rainfall characteristic parameters in the operational planning, with monthly evaluations to ensure model accuracy. Inter-institutional collaboration is recommended to support further analysis. The findings demonstrate that generating inflow information effectively supports the operational management of the Sutami Dam with a sufficient accuracy for practical applications.

Keywords: Inflow Equation, Cross-Validation, Ridge Regression, Stepwise Regression

PENDAHULUAN

Prediksi debit masuk bendungan yang akurat merupakan kebutuhan krusial untuk pengelolaan dan perencanaan sumber daya air (Sanikhani dkk., 2025). Debit masuk menjadi faktor utama dalam operasional Bendungan Sutami di Daerah Aliran Sungai (DAS) Brantas, Jawa Timur, yang berfungsi sebagai infrastruktur pengendali banjir, penyedia air irigasi hilir, dan pembangkit listrik (Nathania dkk., 2021). Sebagai infrastruktur vital di wilayah padat penduduk dan terindustrialisasi, ketepatan prediksi debit masuk sangat penting untuk

mengoptimalkan pengelolaan bendungan dan menjamin ketersediaan air pada musim kemarau. Namun, kompleksitas proses hidrologi, yang dipengaruhi oleh variabilitas curah hujan, karakteristik DAS, dan faktor lingkungan lainnya, menyulitkan upaya prediksi ini.

Pengelolaan Bendungan Sutami juga menghadapi kendala dalam perencanaan tahunan, di mana rencana operasional disusun mulai Juli tahun sebelumnya dan difinalisasi pada November tahun sebelumnya. Sebagai contoh, rencana tahun 2023 harus diselesaikan pada November 2022. Prakiraan yang digunakan secara status quo menggunakan prakiraan September yang hanya mencakup informasi untuk periode hingga Juni tahun rencana. Perencanaan bendungan juga masih mengandalkan informasi sifat musim sebagai acuan (Bestari, 2023). Hal ini tidak sesuai dengan perhitungan kategori debit yang digunakan oleh pihak pengelola bendungan. Prakiraan curah hujan dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), yang digunakan dalam perencanaan, umumnya berupa *percent normal index* atau sifat hujan berdasarkan rata-rata periode 1991–2020. Pengelola bendungan menggunakan dengan konsep persentil basah-kering. Ketidaksesuaian ini menambah tantangan dalam pemanfaatan data iklim untuk prediksi debit.

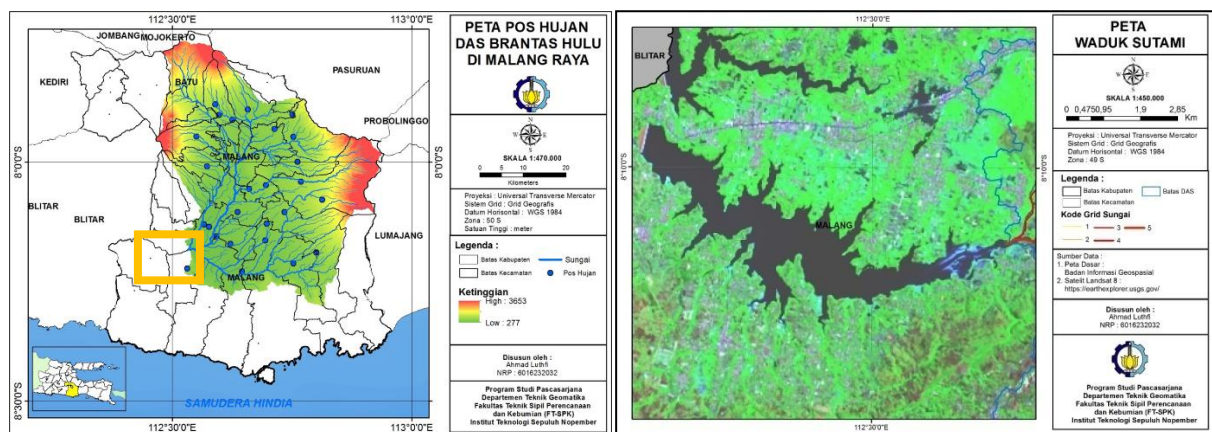
Studi yang dilakukan Mahmudiah dkk., (2024) menunjukkan verifikasi prakiraan curah hujan yang diolah menjadi debit menjanjikan peran yang lebih baik dibandingkan menggunakan pendekatan sifat hujan namun hubungan curah-debit secara spesifik masih perlu dicari. Curah hujan wilayah merupakan prediktor utama untuk debit masuk bendungan karena hubungannya yang langsung dengan aliran air. Persamaan hujan-debit menjadi kompleks karena adanya respons tertunda (*lag*) akibat faktor seperti infiltrasi tanah, topografi, dan vegetasi. Pendekatan regresi berbasis jeda yang menggunakan data curah hujan dari waktu sebelumnya sebagai prediktor, telah terbukti mampu menangkap dinamika temporal secara efektif (García-Feal dkk., 2022). Studi Han dkk., 2023 mengindikasikan bahwa regresi multivariat berbasis jeda dapat merepresentasikan pola debit dengan akurat. Dibandingkan dengan model hidrologi berbasis fisika yang memerlukan data luas dan komputasi berat, analisis regresi menawarkan solusi yang lebih hemat sumber daya dan mudah diterapkan. Regresi ridge efektif dalam menangani multikolinearitas antar prediktor iklim (Arisandi dkk., 2021), sedangkan regresi stepwise memfasilitasi seleksi prediktor secara sistematis. Pemilihan prediktor yang optimal sangat penting dalam penyusunan persamaan debit (Tran dkk., 2021), sehingga perbandingan regresi ridge dan stepwise menjadi pendekatan yang relevan untuk menghubungkan informasi curah hujan dengan informasi debit.

Pertukaran bias-varians (*bias-variance trade-off*) merupakan tantangan utama dalam pemodelan statistik modern (Rashidi dkk., 2019). Baik regresi ridge maupun stepwise memiliki keterbatasan yang perlu dievaluasi terkait masalah ini. Penggunaan regresi rentan terhadap *overfitting* karena proses pengolahan variabelnya bergantung pada karakteristik data sehingga dapat menghasilkan persamaan yang kurang stabil. Praktik terbaik (*best practice*) atau pendekatan yang lazim dipercaya konsisten menghasilkan hasil yang optimal menunjukkan bahwa validasi silang diperlukan untuk memastikan kemampuan generalisasi model (Burkov, 2020). Validasi silang membantu mencegah *overfitting* dengan memberikan estimasi performa yang lebih akurat pada data uji, dibandingkan hanya meminimalkan galat pada data pelatihan (Chollet dkk., 2022; Géron, 2023). Oleh karena itu diperlukan studi khusus untuk memastikan bahwa transformasi informasi dari prakiraan curah hujan menjadi prakiraan debit menggunakan persamaan yang tergeneralisasi dengan baik. Studi ini juga dapat menetapkan standar yang lebih *modern* terkait pentingnya validasi silang dalam pengembangan persamaan prakiraan iklim. Dengan semakin terjangkau biaya komputasi, validasi silang kini menjadi standar baru dalam pengembangan model di berbagai disiplin ilmu (Wilimitis & Walsh, 2023). Hal ini menunjukkan perlu adanya studi terkait validasi silang dalam pemodelan curah hujan dengan debit.

Studi tentang debit masuk bendungan di Indonesia masih jarang dilakukan, terutama yang memanfaatkan pendekatan validasi silang sebagai solusi untuk konsistensi persamaan. Beberapa studi hidrologi lokal lebih mengutamakan pendekatan berbasis fisika atau analisis kualitatif, sedangkan model berbasis data masih kurang mendapat perhatian. Dalam penelitian ini dilakukan pengembangan persamaan regresi ridge dan stepwise sebagai fungsi untuk mengubah data curah hujan menjadi debit dengan pendekatan jeda serta mengevaluasi performa kedua persamaan tersebut menggunakan metode validasi silang.

METODOLOGI

Objek penelitian ini adalah Bendungan Sutami, yang berfungsi sebagai infrastruktur pengelolaan Waduk Sutami. Gambar 1 (kiri) menggambarkan lokasi Bendungan Sutami beserta Daerah Aliran Sungai (DAS) yang mengisi waduk. DAS ditunjukkan melalui wilayah yang diarsir berdasarkan gradasi ketinggian. DAS tersebut mencakup seluruh wilayah Kota Batu, Kota Malang, dan sebagian Kabupaten Malang. Sebanyak 33 stasiun pengukur curah hujan (ombrometer) digunakan untuk menghasilkan data raster curah hujan yang menjadi dasar analisis zonal (perata-rataan spasial) dalam penelitian ini. Gambar 1 (kanan) memperlihatkan perbesaran waduk Sutami, yang diambil dari area yang ditandai dengan kotak oranye pada Gambar 1 (kiri), untuk memberikan gambaran detail tentang wujud waduk dari atas.



Gambar 1. Koordinat 33 Pos Hujan dan Foto Atas Waduk Sutami

Studi ini memanfaatkan data curah hujan wilayah hasil zonal perata-rataan dan debit masuk Bendungan Sutami di Daerah Aliran Sungai (DAS) Brantas, Jawa Timur, untuk periode dasarian (10 harian) dari Januari I 2021 hingga Desember III 2023, mencakup 108 dasarian. Data curah hujan wilayah diperoleh melalui analisis statistika zona (*zonal statistics*) dari data raster menggunakan perangkat lunak *R Statistics* (R Core Team, 2023). Data raster dikonversi agar memiliki resolusi spasial yang seragam yaitu 0,01 x 0,01 derajat. Terdapat enam jenis data raster yang digunakan, meliputi data *Climate Hazards Group InfraRed Precipitation* (CHIRP) (Funk dkk., 2015) dan lima metode berbasis interpolasi, yaitu *Inverse Distance Weighting* (IDW), *Ordinary Kriging* (OK), *Universal Kriging* (UK), dan *Kriging with External Drift* (KED).

Tabel 1. Konfigurasi Data Raster

Metode Interpolasi Raster	Input	Parameter	Studi Terdahulu
IDW	curah hujan 33 pos	nn=12, power=5	(Kurniawan dkk., 2021)
OK	curah hujan 33 pos		
UK	curah hujan 33 pos		
KED1	curah hujan 33 pos, CHIRP	variogram=exponential (nugget=0; sill=1; range=0,1)	(Utomo dkk., 2025)
KED2	curah hujan 33 pos, CHIRP, DEM		

Semua metode interpolasi menggunakan input yang sama, yaitu data dari 33 pos hujan di DAS Hulu Brantas (Gambar 1-kiri) kecuali KED1 yang mengintegrasikan CHIRP dan KED2 yang menggabungkan CHIRP serta *Digital Elevation Model* (DEM). Keterangan lebih lanjut dapat dilihat pada Tabel 1. Penentuan parameter optimal untuk interpolasi di luar lingkup bahasan studi ini. Data debit masuk bendungan yang digunakan adalah periode yang sama dengan curah hujan sehingga mengakibatkan pemotongan data akibat

jeda (*lag*). Penelitian ini mengevaluasi perbedaan performa persamaan hujan-debit dari hasil zonal perataan enam raster sehingga dapat juga digunakan sebagai justifikasi metode interpolasi terbaik,

Untuk menangkap dinamika respons tertunda antara curah hujan dan debit masuk Bendungan Sutami, pendekatan jeda digunakan dengan memasukkan data curah hujan dari dasarian sebelumnya sebagai prediktor. Seleksi periode jeda dilakukan dengan menganalisis korelasi antara data curah hujan wilayah dasarian sesuai DAS hulu (Gambar 1) dan debit masuk, sesuai dengan konfigurasi data raster pada Tabel 1. Korelasi Pearson dihitung untuk setiap jeda, dari jeda -15 hingga jeda 15 dengan asumsi hubungan positif maksimum 15 dasarian atau 5 bulan. Curah hujan yang menunjukkan korelasi negatif dijadikan batas untuk pelibatan jeda, karena korelasi negatif mencerminkan hubungan yang tidak konsisten dengan prinsip hidrologi. Analisis korelasi ini bertujuan untuk menentukan jeda optimal dalam prediksi debit masuk bendungan, memastikan hanya variabel yang relevan secara hidrologis yang digunakan dalam persamaan regresi.

Studi ini menetapkan 50% dari total data yang tersedia sebagai data pelatihan, sesuai dengan metodologi penelitian Kim dkk. (2022). Pertimbangan yang digunakan adalah panjang data yang terbatas, yaitu 108 dasarian. Validasi silang dilakukan dengan rasio 50:50 untuk pelatihan dan pengujian. Dataset dibagi secara berurutan. Urutan temporal data dijaga agar konsisten dengan analisis jeda curah hujan dan debit. Metode validasi silang ini dipilih untuk mengoptimalkan jumlah data pelatihan sekaligus memastikan evaluasi persamaan yang andal merata. Prosedur ini diulang beberapa kali untuk memastikan konsistensi dan stabilitas hasil validasi.

Regresi stepwise diterapkan dengan menyusun persamaan prediksi dengan memilih variabel jeda secara bertahap dan terkoordinasi untuk menjaga kesinambungan dengan analisis korelasi curah hujan dan debit. Proses ini dimulai dengan memasukkan variabel jeda 1 (curah hujan dari dasarian sebelumnya), kemudian menambahkan variabel jeda berikutnya (jeda 2, jeda 3, dan seterusnya) hingga batas jeda yang ditetapkan berdasarkan analisis korelasi. Pada setiap tahap, persamaan dinilai menggunakan validasi silang untuk mengukur galat pelatihan dan pengujian. Kriteria pemilihan variabel didasarkan pada kontribusinya terhadap pengurangan galat, dengan menerapkan metode *forward selection* untuk meminimalkan risiko *overfitting*. Pendekatan ini selaras dengan praktik yang diuraikan oleh Tran dkk., (2021), yang menekankan pentingnya seleksi variabel bertahap untuk meningkatkan kemampuan generalisasi persamaan regresi.

Regresi ridge digunakan dengan menangani multikolinearitas yang dapat muncul dari variabel jeda yang berkorelasi tinggi dalam analisis curah hujan dan debit. Studi sebelumnya (Nugrahani dkk., 2025) memberikan contoh aplikasi serupa pada data iklim dan menjelaskan mekanisme regresi ridge. Persamaan ini mengintegrasikan semua variabel jeda hingga batas yang ditentukan melalui analisis korelasi tanpa proses seleksi bertahap seperti yang diterapkan pada regresi stepwise. Parameter penalti (*lambda*) diatur dalam rentang 0 hingga 50, dengan pendekatan iteratif untuk menentukan nilai *lambda* yang optimal. Untuk setiap nilai *lambda*, performa persamaan dinilai menggunakan validasi silang, dengan fokus pada galat pelatihan dan pengujian untuk mengevaluasi efektivitas model. Eksperimen diawali dengan melakukan uji *lambda* dari 0 hingga 50 dengan pertambahan 1. Nilai *lambda* optimal kemudian akan dipilih berdasarkan nilai minimum dari galat pengujian. Optimasi ini dilakukan secara iteratif, dengan membandingkan hasil validasi silang untuk setiap konfigurasi persamaan. Pendekatan ini menjamin bahwa persamaan yang dipilih tidak hanya akurat tetapi juga dapat digeneralisasikan dengan baik.

Kinerja persamaan regresi stepwise dan ridge dievaluasi berdasarkan galat pelatihan dan galat pengujian yang diperoleh dari hasil validasi silang. Validasi silang menjadi pendekatan efektif untuk memilih model, terutama ketika literatur terkait masih terbatas dalam menjelaskan alasan pemilihan model tertentu (Burkov, 2020). Evaluasi pada tahap pelatihan-pengujian menggunakan *Mean Square Error* (MSE) sebagai metrik utama karena mampu menunjukkan perbedaan yang jelas antar konfigurasi model (Kumar dkk., 2023). Persamaan MSE adalah sebagai berikut.

$$MSE = \frac{\sum (x - y)^2}{n} \quad (1)$$

Dengan

x = debit estimasi dari persamaan regresi

y = debit observasi

n = jumlah observasi

Persamaan dianggap tergeneralisasi dengan baik jika total dari galat pengujian dan galat pelatihan merupakan terkecil dibandingkan konfigurasi lainnya. Total galat mencerminkan kemampuan persamaan untuk berfungsi optimal pada data latih dan data baru, sebagaimana direkomendasikan oleh Wilimitis dan Walsh, (2023).

Pengujian validasi silang dilakukan dengan perangkat lunak *R-Statistics* (Chollet dkk., 2022) Galat dari setiap iterasi pelatihan dan pengujian dirata-ratakan untuk divisualisasikan dalam plot. Total galat digunakan untuk menentukan tiga persamaan terbaik dari regresi ridge dan stepwise. Persamaan regresi ridge terbaik menghasilkan konfigurasi sumber raster dan nilai lambda yang paling optimal, sedangkan persamaan regresi stepwise terbaik menghasilkan konfigurasi sumber raster dan jumlah variabel prediktor yang ideal. Untuk memfasilitasi interpretasi konvensional, konfigurasi persamaan terbaik dihitung ulang menggunakan seluruh data yang tersedia. *Mean Absolute Error* (MAE) dan korelasi digunakan sebagai parameter evaluasi tambahan (Enayati dkk., 2024). Persamaan koefisien korelasi dan MAE berturut-turut adalah sebagai berikut.

$$Korelasi = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

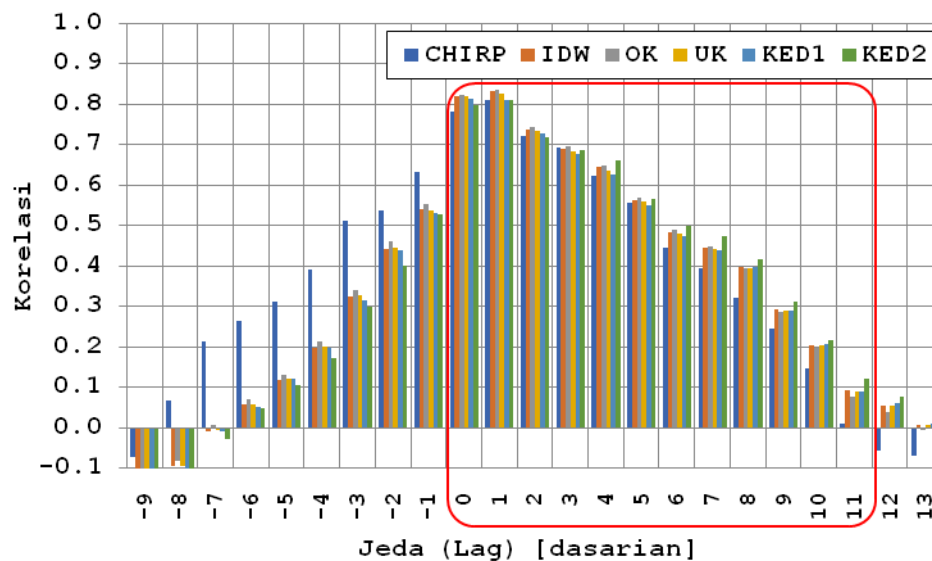
$$MAE = \frac{\sum |x - y|}{n} \quad (3)$$

Korelasi antara nilai debit hasil estimasi dan debit observasi dihitung untuk menilai hubungan linier antara prediksi dan data aktual. Selain itu, MAE memberikan interpretasi yang mudah dipahami tentang magnitudo kesalahan prediksi dalam satuan debit.

HASIL DAN PEMBAHASAN

Analisis Penentuan Batas Jeda (*Lag*)

Penentuan batas jeda (*lag*) dilakukan melalui analisis periode curah hujan yang relevan secara historis sebagai prediktor debit masuk Bendungan Sutami. Data curah hujan wilayah dari sumber CHIRP, IDW, OK, UK, KED1, dan KED2 dirata-ratakan menggunakan paket raster (Hijmans, 2025) dalam perangkat lunak *R-Statistics* untuk dibandingkan dengan rata-rata debit masuk bendungan. Curah hujan dasarian diukur dalam milimeter per dasarian, sedangkan rata-rata debit masuk diukur dalam meter kubik per detik. Analisis korelasi mencakup periode Januari I 2021 hingga Desember III 2023, yang terdiri dari 108 dasarian. Korelasi dihitung untuk jeda dari -15 hingga +15 dasarian, tetapi Grafik Analisis Korelasi Curah Hujan dan Debit Dasarian dengan Jeda hanya menampilkan rentang dari -9 hingga 13. Hasilnya menunjukkan bahwa semua metode interpolasi menghasilkan korelasi positif yang signifikan pada jeda 0 hingga 11, dengan nilai tertinggi pada jeda 1 (0,832 untuk IDW; 0,836 untuk OK; 0,826 untuk UK). Sebaliknya, korelasi negatif teramati pada sampai dengan jeda -7 dan mulai muncul kembali pada jeda 12 hingga 15 seperti terlihat pada Gambar 2.



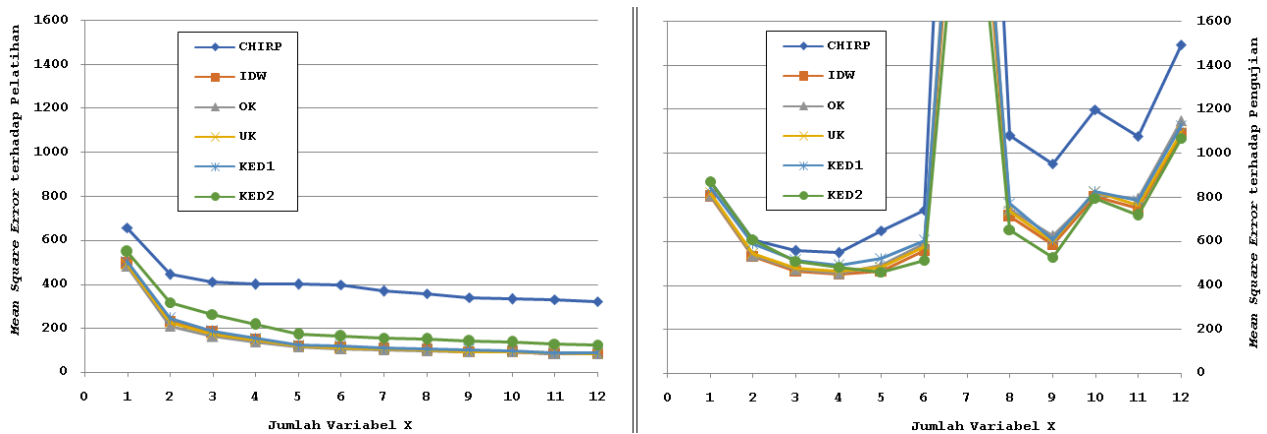
Gambar 2. Grafik Analisis Korelasi Curah Hujan dan Debit Dasarian dengan Jeda

Berdasarkan analisis korelasi, rentang jeda dari 0 hingga 11 dasarian dipilih untuk digunakan dalam pengujian persamaan regresi. Jeda negatif dianggap tidak relevan secara hidrologis, karena memprediksi debit masa lalu dengan curah hujan masa depan bertentangan dengan prinsip kausalitas fisis. Demikian pula, jeda di atas 11 dasarian tidak dimasukkan karena munculnya korelasi negatif, seperti -0,056 untuk CHIRP pada jeda 12, yang mengindikasikan hubungan yang tidak lagi signifikan secara hidrologis dan hanya mencerminkan korelasi statistik acak. Temuan ini sejalan dengan studi sebelumnya yang menyimpulkan bahwa diperlukan jeda hingga 7 dasarian (Kumar dkk., 2023). Pemilihan batas jeda ini memastikan bahwa hanya variabel prediktor yang memiliki keterkaitan fisis dengan dinamika hidrologi Bendungan Sutami yang dimasukkan dalam persamaan regresi ridge dan stepwise.

Penetapan batas jeda pada rentang 0 hingga 11 dasarian diperkuat oleh karakteristik DAS Brantas Hulu yang memiliki waktu respons hidrologi yang bervariasi akibat pengaruh infiltrasi tanah dan aliran permukaan. Tidak semua curah hujan langsung berkontribusi pada debit masuk, karena sebagian diserap oleh sistem hidrologi yang menghasilkan mekanisme jeda (Sulistyo dan Fauzi, 2023). Metode IDW menunjukkan kinerja yang andal dengan korelasi tertinggi pada jeda 1 (0,832), menegaskan keunggulannya dalam interpolasi data curah hujan lokal, sebagaimana dilaporkan oleh Kurniawan dkk., (2021) Analisis ini menyediakan landasan yang kuat untuk penerapan pendekatan jeda dalam pemodelan dan memastikan bahwa persamaan yang dihasilkan mencerminkan dinamika temporal curah hujan dan debit masuk secara realistis

Galat (*Error*) Pelatihan dan Pengujian

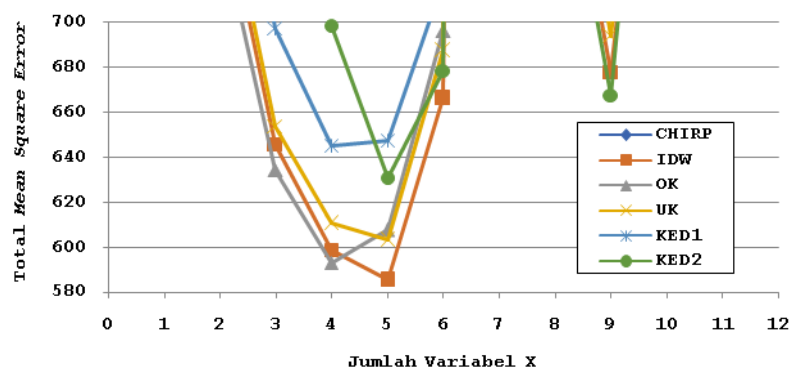
Untuk mengevaluasi performa persamaan regresi stepwise dalam memprediksi debit masuk Bendungan Sutami, metode validasi silang 50:50 diterapkan pada dataset yang mencakup 97 pasang data curah hujan dan debit, dengan batas jeda dari 0 (simultan) hingga 11 dasarian. Pasangan data curah hujan dan debit dibagi secara berurutan dan bergantian, dengan 48 pasang sebagai set pelatihan dan 49 pasang sebagai set pengujian. Total 49 iterasi validasi silang dilakukan untuk setiap kombinasi metode interpolasi (CHIRPS, IDW, OK, UK, KED1, KED2), rentang lambda 0–50, dan variasi jumlah prediktor. Galat pelatihan dan galat pengujian dihitung menggunakan MSE, yang mewakili bias (selisih dengan data latih, Gambar 3 kiri) dan varians (selisih dengan data uji, Gambar 3 kanan) persamaan. MSE dihitung untuk setiap iterasi, menghasilkan 49 nilai per metrik, yang dirata-ratakan untuk menilai kinerja persamaan. Pendekatan ini, sesuai rekomendasi Wilimitis dan Walsh, (2023), menjamin evaluasi yang robust terhadap generalisasi persamaan pada data yang jumlahnya terbatas.



Gambar 3. MSE dari Persamaan Regresi Stepwise terhadap Data Latih (kiri) dan terhadap Data Uji (kanan)

Regresi stepwise diterapkan melalui pendekatan pemilihan maju yang bersambung (*continued forward selection*), dengan prediktor ditambahkan secara berurutan dari jeda 0 (simultan) hingga jeda 11. Sebagai contoh, persamaan dengan satu prediktor hanya menggunakan jeda 0, persamaan dengan dua prediktor mencakup jeda 0 dan jeda 1, dan seterusnya hingga 12 prediktor yang mencakup jeda 0 hingga jeda 11. Pendekatan ini didukung oleh pertimbangan fisis bahwa curah hujan pada jeda tertentu harus disertai oleh curah hujan pada jeda sebelumnya yang lebih dekat secara temporal, sesuai dengan kontinuitas respons hidrologi DAS Brantas. Seleksi prediktor berurutan ini konsisten dengan prinsip hidrologi yang dijelaskan oleh Kim dkk., (2022), yang menegaskan perlunya mempertimbangkan hubungan temporal yang logis dalam pemodelan aliran untuk menghasilkan prediksi yang akurat.

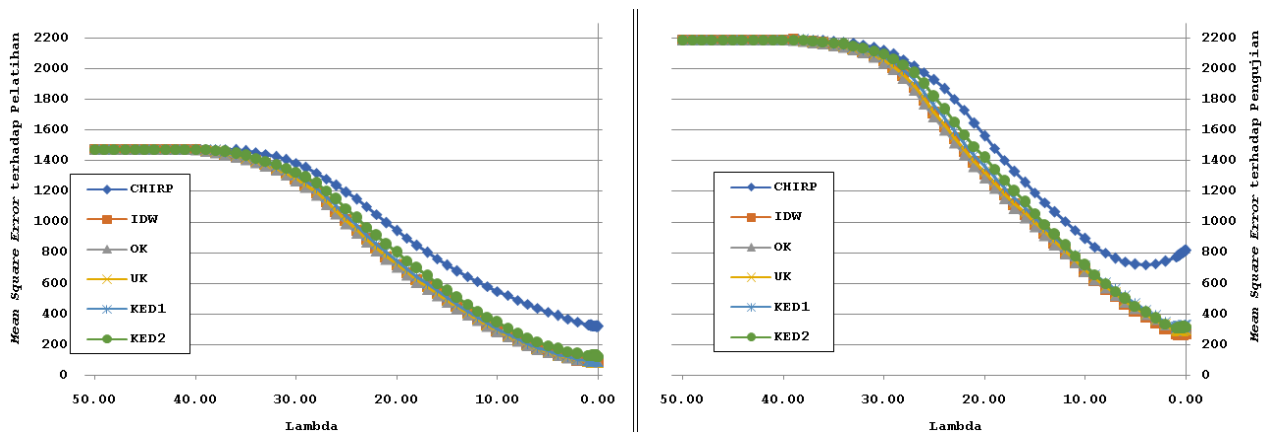
Hasil evaluasi galat pelatihan (bias) dan galat pengujian (varians), sebagaimana ditampilkan dalam Gambar 3, mengindikasikan bahwa metode interpolasi IDW menghasilkan galat terendah dibandingkan metode interpolasi lainnya. Kurva memprioritaskan tampilan dengan skala-Y yang sama antara gambar kanan dan kiri sehingga dapat diperbandingkan dengan mudah. Bagian terpotong mengindikasikan bahwa konfigurasi tersebut tidak baik untuk digunakan. Pada regresi stepwise, galat pelatihan berkurang seiring bertambahnya jumlah prediktor, mencapai nilai terendah pada 12 prediktor (misalnya, 82,96 untuk IDW), menunjukkan bias yang lebih kecil. Namun, galat pengujian menunjukkan tren berbeda, dengan nilai terendah pada 4 variabel untuk IDW (448,43), tetapi meningkat tajam pada jumlah prediktor yang lebih banyak (misalnya, 1084,41 untuk 12 prediktor), menandakan adanya *overfitting*. Fenomena ini, sebagaimana dijelaskan oleh Bradshaw dkk., (2023), mengkonfirmasi bahwa penambahan prediktor berlebihan dapat menghasilkan *overfitting* atau meningkatkan varians dan mengurangi kemampuan generalisasi persamaan. Dengan demikian, persamaan dengan 4 variabel menggunakan IDW dianggap optimal, mencapai keseimbangan bias dan varians untuk prediksi debit masuk Bendungan Sutami.



Gambar 4. Total MSE Terendah Persamaan Regresi Stepwise

Untuk memastikan persamaan regresi stepwise yang paling tergeneralisasi dalam memprediksi debit masuk Bendungan Sutami, galat total dihitung sebagai penjumlahan galat pelatihan dan galat pengujian (yang masing-masing mencerminkan bias dan varians persamaan). Hal ini divisualisasikan dalam Gambar 4. Galat total dianalisis untuk setiap metode interpolasi curah hujan (CHIRPS, IDW, OK, UK, KED1, dan KED2) dengan jumlah variabel bervariasi dari 1 hingga 12, menggunakan 97 pasang curah-debit (jeda 0 hingga 11) untuk periode debit April III 2021 hingga Desember III 2023. Hasil tabulasi galat total menunjukkan bahwa persamaan dengan interpolasi IDW dengan 5 variabel prediktor menghasilkan galat terkecil (585,37), diikuti oleh OK dengan 4 variabel (592,96) dan IDW dengan 4 variabel (598,99). Nilai galat total yang rendah pada kombinasi ini mengindikasikan keseimbangan optimal antara bias dan varians sehingga mencerminkan generalisasi yang baik. Sebaliknya, galat total meningkat tajam pada jumlah prediktor yang lebih tinggi (misalnya, 3165,24 untuk IDW dengan 7 variabel prediktor), menunjukkan risiko *overfitting* akibat kompleksitas persamaan yang berlebihan.

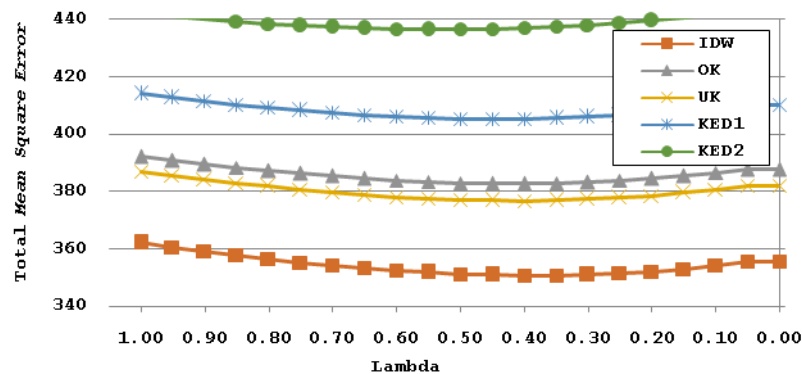
Penilaian performa regresi ridge dalam memprediksi debit masuk Bendungan Sutami juga dilakukan melalui analisis galat pelatihan dan galat pengujian, diukur dengan MSE menggunakan validasi silang pada 97 pasang data curah hujan dan debit dengan jeda 0 hingga 11 dasarian. Sebagaimana ditampilkan dalam Gambar 5, pola melenting baru muncul pada lambda kurang dari 10. Tidak seperti regresi stepwise, yang kompleksitas persamaannya bertambah dengan jumlah prediktor, pada regresi ridge kompleksitas persamaan meningkat seiring penurunan parameter penalti lambda. Itulah mengapa sumbu X pada Gambar 5 dibalik. Pengujian awal menggunakan rentang lambda dari 0 hingga 50 dengan kenaikan 1, mengindikasikan bahwa galat terendah terjadi pada rentang 0 hingga 10. Untuk analisis yang lebih mendetail, pengujian lanjutan dilakukan dengan lambda berkisar dari 0 hingga 1 dengan interval 0,05 (0; 0,05; 0,10; hingga 0,95). Hasil tabulasi menunjukkan bahwa galat pelatihan meningkat seiring kenaikan lambda untuk semua metode interpolasi (CHIRPS, IDW, OK, UK, KED1, dan KED2). Hal ini mencerminkan penalti yang lebih ketat pada koefisien regresi. Galat pengujian mencapai nilai minimum pada lambda rendah, dengan IDW menghasilkan galat pengujian terkecil (267,41) pada lambda 0,25. Pendekatan ini menjamin evaluasi yang *robust* terhadap generalisasi persamaan dalam konteks data hidrologi.



Gambar 5. MSE dari Persamaan Regresi Ridge terhadap Data Latih (Bias, kiri) dan terhadap Data Uji (Varians, kanan)

Hasil analisis menunjukkan bahwa regresi ridge dengan lambda 0,25 menggunakan data curah hujan dari metode IDW menghasilkan galat pelatihan terkecil sebesar 83,97 dan galat pengujian sebesar 267,41, hal ini sekali lagi menunjukkan superioritas metode interpolasi ini dibandingkan dengan metode lainnya. Nilai lambda yang rendah ini mengindikasikan bahwa multikolinearitas antar variabel curah hujan dasarian yang berjeda dapat dikendalikan, memungkinkan persamaan untuk mempertahankan daya prediktif tanpa penalti yang berlebihan. Peningkatan galat pengujian pada nilai lambda yang lebih tinggi (misalnya, 272,46 untuk IDW pada lambda 1,00) menunjukkan bahwa penalti yang terlalu ketat dapat melemahkan akurasi pada data

pengujian. Tapi tidak serta merta persamaan regresi ridge IDW dengan lambda 0,25 dipilih untuk analisis lanjutan. Sekali lagi, persamaan tergeneralisasi ditentukan berdasarkan total galat sehingga mewakili kondisi optimal berdasarkan bias dan varians.



Gambar 6. Total MSE Terendah Persamaan Regresi Ridge

Gambar 6 memvisualisasikan identifikasi persamaan regresi ridge yang paling tergeneralisasi untuk membuat persamaan debit masuk Bendungan Sutami menggunakan curah hujan dasarian berdasarkan galat total. Galat total diperoleh dari penjumlahan galat pelatihan dan galat pengujian untuk setiap metode interpolasi curah hujan (CHIRPS, IDW, OK, UK, KED1, dan KED2) pada rentang lambda dari 0,0 hingga 1,0. Hasil tabulasi mengungkapkan bahwa tiga konfigurasi dengan galat total terendah adalah IDW dengan lambda 0,40 (350,62), lambda 0,35 (350,65), dan lambda 0,45 (350,80). Temuan ini menunjukkan keseimbangan ideal antara bias dan varians. Dengan demikian, ketiga konfigurasi ini dipilih untuk analisis lanjutan. Konfigurasi berbasis IDW ini menunjukkan kenaikan galat total pada lambda yang lebih tinggi (misalnya, 362,09 untuk IDW pada lambda 1,00) maupun pada lambda yang lebih rendah. Hal ini mencerminkan penalti berlebihan yang mengurangi ketepatan persamaan dan adanya titik optimal dalam validasi silang.

Persamaan Curah-Debit



Gambar 7. Grafik Perbandingan Debit Observasi dengan Debit Hasil Persamaan Tergeneralisasi

Grafik perbandingan di atas membandingkan hasil dari persamaan curah hujan-debit dengan debit observasi untuk mengevaluasi akurasi enam persamaan secara umum. Gambar 7 menunjukkan hasil yang tidak terlalu berbeda dari 6 (enam) persamaan yang diidentifikasi sebagai persamaan tergeneralisasi, yaitu 3 (tiga) persamaan regresi stepwise (IDW dengan 5 variabel prediktor, OK dengan 4 variabel prediktor, IDW dengan 4 variabel prediktor) dan 3 (tiga) persamaan regresi ridge (IDW dengan lambda 0,40, 0,35, dan 0,45). Analisis

ulang dilakukan menggunakan seluruh pasangan data tersedia (97 pasang data, jeda 0 hingga 11) dari periode April III 2021 hingga Desember III 2023. Secara visual umum, hasil persamaan curah hujan dengan debit benar-benar memiliki kesamaan pola dengan data debit observasi. Bahkan tanpa menyediakan perhitungan lain seperti MAE atau korelasi, pola yang terbentuk sudah sangat identik. OK dengan 4 variabel sebagai satu-satunya metode non-IDW tidak menunjukkan hasil yang lebih baik ketika persamaan berbasis IDW menjauhi pola garis debit observasi. Hal ini mendukung IDW sebagai metode interpolasi yang kemampuannya mampu menangkap variabilitas curah hujan lokal, sebagaimana didukung oleh Kurniawan dkk., (2021). Secara visual sederhana Gambar 7 memastikan bahwa persamaan yang dipilih tidak hanya tergeneralisasi berdasarkan validasi silang, tetapi juga akurat ketika diterapkan pada seluruh dataset dan disandingkan dengan debit observasi.

$$\text{Estimasi Debit} = \text{Konstanta} + \sum_{i=1}^{12} \text{Curah Hujan Dasarian jeda } (i - 1) \times \text{Koefisien } (i) \quad (4)$$

Bentuk umum persamaan linier yang dihasilkan dalam penelitian ini disajikan dalam Persamaan 4 di atas. Enam persamaan terbaik yang diidentifikasi sebagai tergeneralisasi dianalisis lebih lanjut untuk mengevaluasi koefisien dan konstanta persamaan linier yang terbentuk. Persamaan dibentuk berdasarkan data curah hujan zonal wilayah (dalam milimeter per dasarian) dari periode April III 2021 hingga Desember III 2023. Rincian konstanta dan koefisien persamaan terdapat dalam Tabel 2.

Tabel 2. Koefisien dan Konstanta Persamaan Tergeneralisasi

[Stepwise/Ridge]	Stepwise			Ridge		
[Interpolasi]	IDW, nvar=5	OK, nvar=4	IDW, nvar=4	IDW, lambda=0.40	IDW, lambda=0.35	IDW, lambda=0.45
Total MSE Validasi Silang	585,37	592,96	598,99	350,62	350,65	350,80
Konstanta	44,410	46,868	47,141	36,045	35,878	36,212
CHD_jeda 0	0,289	0,269	0,267	0,304	0,305	0,304
CHD_jeda 1	0,215	0,270	0,258	0,228	0,228	0,228
CHD_jeda 2	0,088	0,093	0,083	0,068	0,068	0,068
CHD_jeda 3	0,054	0,148	0,155	0,042	0,042	0,042
CHD_jeda 4	0,156	x	x	0,066	0,067	0,066
CHD_jeda 5	x	x	x	0,056	0,056	0,056
CHD_jeda 6	x	x	x	0,007	0,007	0,007
CHD_jeda 7	x	x	x	0,025	0,025	0,025
CHD_jeda 8	x	x	x	0,047	0,047	0,047
CHD_jeda 9	x	x	x	0,010	0,010	0,010
CHD_jeda 10	x	x	x	0,046	0,046	0,046
CHD_jeda 11	x	x	x	0,016	0,016	0,015

Istilah “nvar” pada Tabel 2 menunjukkan jumlah variabel prediktor sesuai hasil analisa sebelumnya. Koefisien untuk curah hujan simultan (jeda 0) secara konsisten menunjukkan nilai terbesar di hampir semua persamaan (misalnya, 0,289 untuk regresi stepwise IDW dengan 5 prediktor dan 0,304 untuk regresi ridge IDW dengan lambda 0,40). Pengecualian terjadi pada persamaan regresi stepwise OK dengan 4 prediktor, di mana koefisien jeda 1 lebih besar (0,270 dibandingkan 0,269 untuk jeda 0). Hal ini mengindikasikan bahwa curah hujan pada waktu bersamaan memiliki pengaruh dominan terhadap debit pada dasarian tersebut, sejalan dengan respons hidrologi cepat di DAS Brantas Hulu yang didominasi sawah. Tutupan lahan sawah menyebabkan air hujan cenderung menjadi limpasan (*run-off*) daripada terserap tanah. Namun, ketergantungan

lebih besar pada jeda 1 dalam regresi stepwise OK dengan 4 variabel prediktor mungkin mencerminkan variasi dalam metode interpolasi atau karakteristik data lokal yang memerlukan eksplorasi pada studi selanjutnya.

Koefisien untuk jeda berikutnya (jeda 1 hingga jeda 11) menunjukkan kontribusi yang semakin kecil, dengan nilai yang menurun seiring bertambahnya jeda, misalnya dari 0,215 (jeda 1) hingga 0,054 (jeda 3) pada regresi stepwise IDW dengan 5 prediktor, dan lebih kecil lagi pada regresi ridge (misalnya, 0,016 untuk jeda 11 pada λ 0,45). Meskipun dihasilkan melalui pendekatan statistik, persamaan ini mampu menangkap karakteristik fisis yang relevan, di mana curah hujan pada jeda yang lebih lama memiliki pengaruh yang semakin rendah terhadap debit. Hal ini juga konsisten dengan waktu respons hidrologi DAS yang pasti dipengaruhi oleh infiltrasi tanah dan retensi air. Temuan ini sesuai studi Sulistyو dan Fauzi, (2023) yang menegaskan adanya jeda temporal dalam persamaan hidrologi. Konsistensi koefisien yang lebih besar pada jeda 0 hingga jeda 1 dibandingkan jeda yang lebih panjang menunjukkan bahwa persamaan ini sesuai dengan prinsip kausalitas hidrologi. Namun, keterbatasan dataset (97 pasang data) mengindikasikan perlunya validasi dengan data yang lebih ekstensif untuk memastikan stabilitas koefisien, terutama pada jeda yang lebih jauh.

Diskusi

Untuk memberikan penilaian yang lebih intuitif terhadap kinerja enam persamaan terbaik—tiga persamaan regresi stepwise (IDW dengan 5 prediktor, OK dengan 4 prediktor, IDW dengan 4 prediktor) dan tiga persamaan regresi ridge (IDW dengan λ 0,40, 0,35, dan 0,45)— analisis korelasi, MAE, persentil ke-80 selisih absolut (P80), median selisih absolut (MED), dan maksimum selisih absolut (MAX) dilakukan setelah persamaan dihitung ulang menggunakan seluruh dataset (97 pasang data, jeda 0 hingga 11) dari periode April III 2021 hingga Desember III 2023. Hasil analisis, sebagaimana disajikan dalam Tabel 3, mengindikasikan bahwa persamaan regresi ridge dengan IDW secara konsisten menunjukkan performa terbaik, dengan MAE terendah pada λ 0,45 (8,9 m³/detik), diikuti oleh λ 0,35 dan 0,40 (masing-masing 9,0 m³/detik). Sebaliknya, persamaan regresi stepwise menghasilkan MAE yang lebih besar, dengan IDW dengan 5 variabel prediktor mencatat MAE 10,6 m³/detik, diikuti oleh OK dengan 4 variabel prediktor (11,9 m³/detik) dan IDW dengan 4 variabel prediktor (12,0 m³/detik). Koefisien korelasi tertinggi diperoleh oleh persamaan regresi ridge (0,959 untuk semua λ). Hal ini mengkonfirmasi bahwa persamaan tergeneralisasi ini mampu menunjukkan hubungan linier yang kuat antara curah hujan dan debit masuk Bendungan Sutami pada skala dasarian. Hasil persamaan ini lebih sederhana dibandingkan pendekatan FJ Mock, yang melibatkan banyak variabel dan uji coba *trial-and-error* kelembapan tanah (A. Baskoro dkk., 2024).

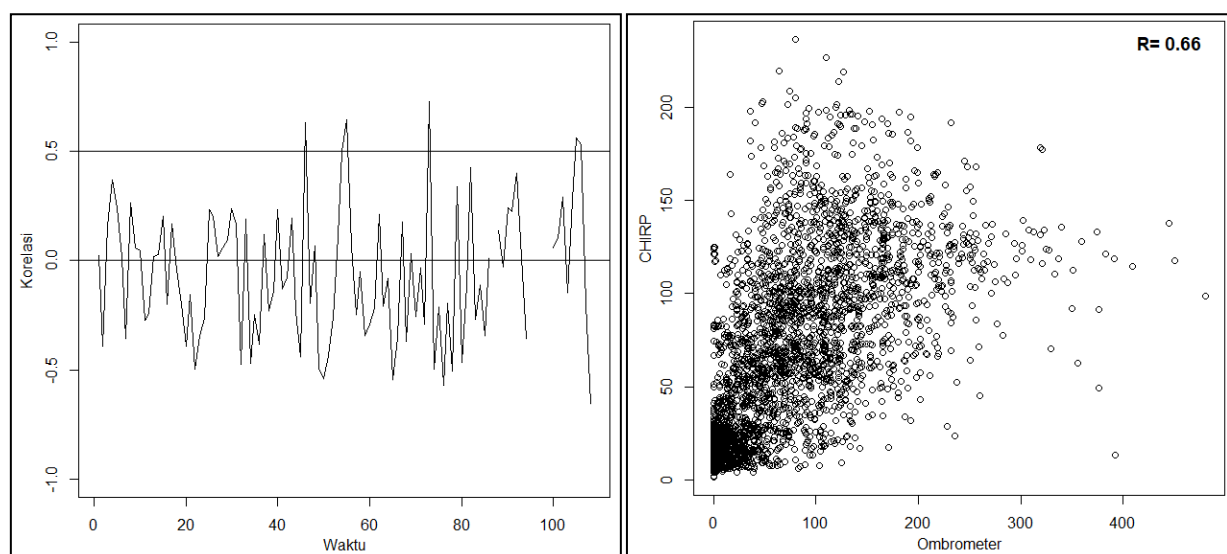
Tabel 3. Korelasi dan Selisih Absolut

[Stepwise/Ridge]	Stepwise			Ridge		
[Interpolasi]	IDW, nvar=5	OK, nvar=4	IDW, nvar=4	IDW, λ =0.40	IDW, λ =0.35	IDW, λ =0.45
MAE	10,6	11,9	12,0	9,0	9,0	8,9
P80 selisih abs	17,1	18,2	18,6	15,2	15,1	15,1
MED selisih abs	8,2	9,6	10,0	7,3	7,2	7,4
MAX selisih abs	40,4	46,3	47,3	31,2	31,0	31,1

Metrik tambahan seperti median selisih absolut (MED), persentil ke-80 selisih absolut (P80), dan maksimum selisih absolut (MAX) memberikan wawasan mendalam tentang distribusi selisih absolut. Persamaan regresi ridge dengan IDW pada λ 0,35 mencatat MED terendah (7,2 m³/detik), mengindikasikan bahwa sebagian besar prediksi memiliki deviasi kecil dari data observasi. P80, yang mencerminkan galat pada 80% data menunjukkan nilai rendah pada persamaan regresi ridge (15,1–15,2 m³/detik) dengan selisih yang tidak terlalu besar dibandingkan regresi stepwise (17,1–18,6 m³/detik), menunjukkan konsistensi prediksi yang baik. MAX mengungkapkan bahwa persamaan regresi stepwise

memiliki galat maksimum lebih tinggi (hingga 47,3 m³/detik untuk IDW dengan 4 prediktor) dibandingkan regresi ridge (31,0–31,2 m³/detik), menandakan kemampuan regresi ridge yang lebih baik dalam menangani kasus ekstrem. Penelitian ini mampu menegaskan keunggulan regresi ridge dibandingkan studi sebelumnya (Arisandi dkk., 2021) melalui penyampaian bukti bahwa penalti lambda yang optimal menghasilkan prediksi yang lebih andal. Rendahnya MED dan P80 pada persamaan regresi ridge mengkonfirmasi keandalan untuk aplikasi operasional. Galat maksimum yang masih signifikan menyoroti perlunya pengujian pada kondisi hidrologi ekstrem, seperti curah hujan tinggi pada skala dasarian. Performa unggul regresi ridge menunjukkan bahwa penalti lambda yang tepat menyeimbangkan kompleksitas dan akurasi persamaan tetapi keterbatasan dataset (97 pasang data) tetap menekankan pentingnya validasi dengan data lebih ekstensif untuk memastikan robustitas dalam pengelolaan sumber daya air.

Tidak munculnya konfigurasi persamaan tergeneralisasi dari metode interpolasi yang lebih rumit seperti KED (*Kriging with External Drift*) dapat dijelaskan dari tidak seragamnya nilai korelasi spasial dari *external drift* seperti data CHIRP dengan curah hujan observasi pada area yang relatif sempit. Gambar 8 menunjukkan bahwa dalam skala tiap dasarian, korelasi negatif dan korelasi positif lemah ($<0,5$) sering terjadi pada pasangan data curah hujan (ombrometer) dengan data estimasi curah hujan berbasis infra merah (CHIRP). Hal ini sangat bertentangan dengan syarat korelasi tinggi agar sebuah variabel dapat digunakan sebagai masukan interpolasi dalam interpolasi seperti KED. Selain itu, kegagalan metode yang lebih rumit seperti KED dapat ditelusuri ke fakta dasar bahwa penginderaan jarak jauh cukup sulit menggambarkan hujan di Indonesia (Firdaus dkk., 2018). Di sisi lain, ketika data dijadikan satu plot, pola dengan relasi cukup kuat justru muncul. Kondisi ini menunjukkan bahwa pengolahan seperti KED atau *Geographically Weighted Regression* (GWR) perlu mempertimbangkan relasi secara spasio-temporal yang tidak sekadar mengandalkan hubungan spasial satu waktu saja.



Gambar 8. Korelasi Tiap Dasarian (kiri) dan Korelasi Seluruh Waktu (kanan)

Hasil analisis menunjukkan bahwa persamaan tergeneralisasi, yang dipilih berdasarkan validasi silang, mampu memenuhi harapan baik dari segi metrik intuitif seperti MAE yang rendah dan koefisien korelasi yang tinggi, maupun dari segi keandalan yang diukur melalui performa konsisten pada data pelatihan dan pengujian, sehingga menghindari jebakan *overfitting* yang sering muncul saat hanya mengejar galat kecil pada data pelatihan saja. Dalam konteks hidrologi Bendungan Sutami, di mana variabilitas data dari Daerah Aliran Sungai (DAS) Brantas sangat kompleks, pendekatan ini menegaskan bahwa persamaan tidak hanya akurat secara numerik tetapi juga memiliki kemampuan generalisasi yang kuat untuk menangani data baru—sesuatu yang krusial untuk prediksi debit masuk yang dapat diandalkan. Studi ini melengkapi studi spesifik sebelumnya (Mahmudiah dkk., 2024) terkait Bendungan Sutami untuk mendapatkan mekanisme perencanaan yang lebih baik. Dengan demikian, persamaan tergeneralisasi ini membuktikan keunggulannya dibandingkan

pendekatan yang semata-mata fokus pada galat kecil, karena menggabungkan akurasi jangka pendek dengan robustitas jangka panjang yang didukung oleh validasi silang. Namun hal yang tetap perlu dicatat adalah sesuai studi Sweet dkk., (2023) bahwa pemilihan terhadap validasi silang memiliki dampak yang signifikan terhadap hasil studi sehingga pada era modern ini justru suatu studi tidak dapat dibandingkan langsung dengan studi lain. Kolaborasi yang terbuka dan komprehensif masih tetap perlu dilakukan secara rutin.

PENUTUP

Simpulan dan Saran

Studi ini berhasil menunjukkan bahwa data curah hujan wilayah dalam skala waktu dasarian yang dihasilkan melalui metode interpolasi *Inverse Distance Weighting* (IDW) merupakan prediktor yang efektif untuk memprediksi debit masuk Bendungan Sutami. Validasi silang 50:50 menunjukkan regresi stepwise optimal digunakan dengan melibatkan 4 sampai 5 variabel jeda yang tersambung. Demikian juga regresi ridge optimal pada λ 0,35-0,45. Kedua persamaan ini menunjukkan kemampuan untuk menghasilkan prediksi yang tergeneralisasi (*well-generalized*) sekaligus memiliki galat yang kecil. Hal ini sekaligus menunjukkan bahwa IDW memberikan representasi spasial curah hujan yang lebih baik dibandingkan metode interpolasi lain seperti Ordinary Kriging, Universal Kriging, Kriging with External Drift, dan CHIRP dalam kasus Bendungan Sutami. Temuan ini memberikan wawasan penting tentang potensi pendekatan statistik dalam mendukung pengelolaan sumber daya air, meskipun keterbatasan seperti asumsi linearitas dan panjang data yang relatif pendek (108 dasarian) masih perlu mendapat perhatian. Studi ini juga menyoroti sebuah jembatan atas tantangan dalam mengintegrasikan prakiraan curah hujan, yang sebelumnya terbatas pada sifat hujan BMKG menjadi prakiraan yang langsung berbentuk informasi debit. Dengan demikian, studi ini menjadi langkah awal untuk memperpendek kesenjangan antara prakiraan curah hujan dan prediksi debit untuk bendungan lainnya di DAS Brantas.

Berdasarkan hasil studi, terdapat beberapa saran yang dapat dilakukan untuk meningkatkan kualitas perencanaan dan pengelolaan operasional Bendungan Sutami. Pertama, penggunaan informasi berupa prakiraan sifat hujan musim dalam perencanaan bendungan harus segera digantikan karena prakiraan debit berbasis curah hujan, seperti yang ditunjukkan dalam studi ini, menawarkan alternatif yang lebih intuitif dan relevan untuk operasional bendungan. Pendekatan ini memungkinkan prediksi yang lebih terukur dibandingkan sifat hujan yang cenderung bersifat kualitatif. Kedua, rapat koordinasi terkait operasional bendungan harus mempertimbangkan kesesuaian antara keterbatasan informasi prakiraan curah hujan (seperti cakupan pada suatu waktu ke depan dan verifikasinya) dengan prakiraan debit. Hal ini sangat perlu dipertimbangkan untuk memastikan keputusan operasional didasarkan pada data yang konsisten dan realistis. Ketiga, evaluasi terhadap operasional bendungan perlu dilakukan lebih sering, setidaknya setiap bulan, untuk mengakomodasi pembaruan informasi prakiraan curah hujan dan debit. Pendekatan ini akan memungkinkan penyesuaian rencana operasional secara dinamis, terutama mengingat variabilitas hidrologi yang tinggi di DAS Brantas. Keempat, hasil studi ini tidak seharusnya langsung diterapkan secara operasional, tetapi perlu diuji lebih lanjut dengan data yang lebih panjang. Pengujian perlu menggunakan metode yang sama melalui kolaborasi dengan seluruh pemangku kepentingan. Selain itu, eksplorasi metode nonlinear seperti *Long Short-Term Memory* (LSTM), sebagaimana disarankan oleh Kim dkk., (2022) dapat dipertimbangkan untuk menangkap dinamika hidrologi yang lebih kompleks, terutama dalam kondisi ekstrem seperti banjir atau kekeringan. Kolaborasi lintas sektor sangat diperlukan untuk mengintegrasikan variabel tambahan seperti penggunaan lahan dan air di hulu, yang dapat meningkatkan akurasi informasi di masa depan.

Ucapan Terima Kasih

Kami mengungkapkan rasa terima kasih yang tulus kepada Pusat Pengembangan Sumber Daya Manusia Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) atas bantuan beasiswa yang diberikan, TKPSDA (Tim Koordinasi Pengelolaan Sumber Daya Air) Brantas yang telah memberikan wawasan nyata serta keterbukaan kolaboratif mengenai pengelolaan alokasi air, serta PT. PLN Nusantara Power yang menyediakan

data debit yang sangat berharga. Semangat kolaborasi ini memainkan peran krusial dalam keberhasilan penyelesaian studi ini.

DAFTAR PUSTAKA

- Arisandi, R., Ruhiat, D., & Marlina, D. E. (2021). Implementasi Ridge Regression untuk Mengatasi Gejala Multikolinearitas pada Pemodelan Curah Hujan Berbasis Data Time Series Klimatologi. *Jurnal Riset Matematika dan Sains Terapan*, 1(1), 1–11.
- Baskoro, A., Suripin, & Suprpto. (2024). *Analisis Evapotranspirasi Metode Penman Modifikasi Dan Thornthwaite Terhadap Pemodelan Debit Fj. Mock: Analysis Of Evapotranspiration Using Modified Penman And Thornthwaite Methods On Fj. Mock's Run-Off Modelling* (Vol. 12, Nomor 1). <https://doi.org/10.33084/mits.v12i1.6134>
- Bestari, A. (2023). *Laporan Akhir: Penyusunan Alokasi Air WS Brantas*.
- Bradshaw, T. J., Huemann, Z., Hu, J., & Rahmim, A. (2023). A Guide to Cross-Validation for Artificial Intelligence in Medical Imaging. Dalam *Radiology: Artificial Intelligence* (Vol. 5, Nomor 4). Radiological Society of North America Inc. <https://doi.org/10.1148/ryai.220232>
- Burkov, A. (2020). *Machine Learning Engineering*. True Positive Incorporated.
- Chollet, F., Kalinowski, T., & Allaire, J. J. (2022). *Deep learning with R*. (Second edition). Manning Publications Co.
- Enayati, S. M., Najarchi, M., Mohammadpour, O., & Mirhosseini, S. M. (2024). Evaluating machine learning models in predicting dam inflow and hydroelectric power production in multi-purpose dams (case study: Mahabad Dam, Iran). *Applied Water Science*, 14(9). <https://doi.org/10.1007/s13201-024-02260-w>
- Firdaus, P., Kurniawan, A., Fachruddin Lubis, & Soetamto. (2018). *Informasi Penginderaan Jarak Jauh sebagai Pelengkap Analisis Hujan 3.0 mm pada Banjir Tana Toraja*. <https://www.researchgate.net/publication/344180748>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(1), 150066. <https://doi.org/10.1038/sdata.2015.66>
- García-Feal, O., González-Cao, J., Fernández-Nóvoa, D., Astray Dopazo, G., & Gómez-Gesteira, M. (2022). Comparison of machine learning techniques for reservoir outflow forecasting. *Natural Hazards and Earth System Sciences*, 22(12), 3859–3874. <https://doi.org/10.5194/nhess-22-3859-2022>
- Géron, A. (2023). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow THIRD EDITION Concepts, Tools, and Techniques to Build Intelligent Systems*. <https://oreilly.com>
- Han, H., Kim, D., Wang, W., & Kim, H. S. (2023). Dam inflow prediction using large-scale climate variability and deep learning approach: a case study in South Korea. *Water Supply*, 23(2), 934–947. <https://doi.org/10.2166/WS.2023.012>
- Hijmans, R. J. (2025). *Geographic Data Analysis and Modeling*. <https://doi.org/10.32614/CRAN.package.raster>
- Kim, B. J., Lee, Y. T., & Kim, B. H. (2022). A Study on the Optimal Deep Learning Model for Dam Inflow Prediction. *Water (Switzerland)*, 14(17). <https://doi.org/10.3390/w14172766>
- Kumar, V., Kedam, N., Sharma, K. V., Mehta, D. J., & Caloiero, T. (2023). Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. *Water (Switzerland)*, 15(14). <https://doi.org/10.3390/w15142572>
- Kurniawan, A., Makmur, E., & Supari, S. (2021). MENENTUKAN METODE INTERPOLASI SPASIAL CURAH HUJAN BULANAN TERBAIK DI JAWA TIMUR. *Seminar Nasional Geomatika*, 263. <https://doi.org/10.24895/sng.2020.0-0.1142>
- Mahmudiah, R. R., Crysdiyan, C., Hariyadi, M. A., & Kurniawan, A. (2024). AUTOMATED MULTI-MODEL PREDICTION AND EVALUATION FOR CONNECTING RAINFALL PREDICTION INFORMATION AND SINGLE-YEAR OPERATIONAL PLAN OF LAHOR-SUTAMI DAM. *EnviroSciencieae*, 20(4), 433. <https://doi.org/10.20527/es.v20i4.21054>

- Nathania, N. A., Sayekti, R. W., & Sholichin Moch. (2021). Nathania_2021. *Jurnal Teknologi dan Rekayasa Sumber Daya Air* 1(2), 890–903. <https://doi.org/10.21776/ub.jtresda.2021.001.02.45>
- Nugrahani, E., Nurdianti, S., Sopaheluwakan, A., Septiawan, P., & Pratama, W. (2025). Analysis of Evaporation Variables in Open Water and Penman Monteith Evaporation Variables Using Ridge Regression [Case Study Forest Fire Model]. *Journal of Atmospheric Science Research*, 51–65. <https://doi.org/10.30564/jasr.v7i4.6916>
- Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P., & Green, R. (2019). Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. Dalam *Academic Pathology* (Vol. 6). SAGE Publications Ltd. <https://doi.org/10.1177/2374289519873088>
- Sanikhani, H., Nikpour, M. R., & Jamshidi, F. (2025). Advanced Framework for Predicting Rainfall-Runoff: Comparative Evaluation of AI Models for Enhanced Forecasting Accuracy. *Water Resources Management* 39(7), 3273–3294. <https://doi.org/10.1007/s11269-025-04106-9>
- Sulistyo, T., & Fauzi, R. (2023). Soil Infiltration Rate Prediction using Machine Learning Regression Model: A Case Study on Sepinggan River Basin, Balikpapan, Indonesia. *Indonesian Journal on Geoscience*, 10(3), 335–347. <https://doi.org/10.17014/ijog.10.3.335-347>
- Sweet, L., Müller, C., Anand, M., & Zscheischler, J. (2023). Cross-Validation Strategy Impacts the Performance and Interpretation of Machine Learning Models. *American Meteorological Society*. <https://doi.org/DOI:10.1175/AIES-D-23-0026.1>
- Tran, T. D., Tran, V. N., & Kim, J. (2021). Improving the accuracy of dam inflow predictions using a long short-term memory network coupled with wavelet transform and predictor selection. *Mathematics*, 9(5), 1–21. <https://doi.org/10.3390/math9050551>
- Wilimitis, D., & Walsh, C. G. (2023). Practical Considerations and Applied Examples of Cross-Validation for Model Development and Evaluation in Health Care: Tutorial. *JMIR AI*, 2, e49023. <https://doi.org/10.2196/49023>