

Prediksi Data Inflow Waduk Das Citarum Untuk Perencanaan Pengelolaan Air Menggunakan Metode SARIMA-ANN

Sudarno, Tukiyyat, Ferhat Aziz

e-mail: dosen00119@unpam.ac.id, dosen02711@unpam.ac.id, dosen02728@unpam.ac.id

Abstrak--Prediksi inflow yang akurat sangat krusial untuk perencanaan pengelolaan air waduk secara efektif, termasuk optimalisasi operasi Pembangkit Listrik Tenaga Air (PLTA), mitigasi banjir dan kekeringan, serta pemenuhan kebutuhan air masyarakat. Penelitian ini bertujuan untuk mengembangkan model prediksi debit aliran masuk (inflow) air di Daerah Aliran Sungai (DAS) Citarum dengan menggunakan metode hibrida SARIMA-Artificial Neural Network (SARIMAANN). Model SARIMA, yang merupakan model linier, memiliki kemampuan untuk menangkap pola musiman dan tren dalam data deret waktu, namun memiliki keterbatasan dalam memodelkan pola non-linier yang kompleks. Untuk mengatasi kelemahan tersebut, model SARIMA digabungkan dengan ANN, yang dikenal efektif dalam mengidentifikasi pola non-linier dari residual atau kesalahan prediksi SARIMA. Hasil evaluasi model menunjukkan bahwa SARIMA-ANN secara signifikan lebih unggul dibandingkan model SARIMA tunggal, dengan nilai Mean Absolute Error (MAE), Mean Squared Error (MSE), dan Root Mean Squared Error (RMSE) yang jauh lebih rendah. Peningkatan akurasi ini menegaskan efektivitas pendekatan hibrida dalam memprediksi data inflow yang kompleks.

Kata Kunci— Peramalan, Inflow Waduk, DAS Citarum, SARIMA, ANN, SARIMA-ANN.

I. PENDAHULUAN

Daerah Aliran Sungai (DAS) merupakan ekosistem yang tersusun atas tanah, vegetasi, iklim, dan manusia, di mana gangguan pada salah satu komponen akan berdampak pada komponen lainnya. Vegetasi/hutan merupakan elemen paling rentan yang memengaruhi fungsi lingkungan, termasuk jasa air bagi PLTA, PDAM, dan masyarakat. Sungai Citarum sebagai sungai terpanjang di Jawa Barat berperan vital untuk kehidupan sehari-hari, namun mengalami kerusakan akibat pencemaran dan pemanfaatan lahan yang tidak terkontrol hingga menjadikannya salah satu sungai paling tercemar di dunia. DAS Citarum memiliki peran strategis karena mendukung kebutuhan hampir 60% penduduk Jawa Barat dan memasok 20% listrik Jawa-Bali melalui tiga waduk besar (Saguling, Cirata, dan Jatiluhur). Oleh karena itu, pengelolaan inflow waduk sangat penting untuk menjamin keberlanjutan pasokan air dan operasi PLTA. Prediksi inflow diperlukan agar pengelolaan air waduk lebih optimal dalam mencegah banjir, kekeringan, serta menjaga kebutuhan air masyarakat. Metode SARIMA digunakan untuk menangkap pola musiman dan linier, sedangkan ANN dapat memprediksi pola non-linear. Kombinasi SARIMA-ANN dipilih karena mampu meningkatkan akurasi, fleksibilitas, dan penanganan pola kompleks dibanding penggunaan metode tunggal. Studi terdahulu menunjukkan pendekatan hybrid ini efektif dalam memprediksi inflow waduk dan mendukung pengambilan keputusan operasional waduk di tengah ketidakpastian perubahan iklim dan variabilitas aliran air.

II. METODE PENELITIAN

Metode SARIMA-ANN merupakan pendekatan hibrida yang menggabungkan kemampuan model SARIMA dalam menangkap pola linier dan musiman dengan kemampuan ANN dalam mengenali pola non-linier.

1. Pengumpulan dan Prakondisi Data
Data inflow waduk serta variabel iklim seperti curah hujan, suhu, dan kelembaban tanah dikumpulkan dalam bentuk deret waktu. Data diperiksa untuk mengatasi nilai hilang atau anomali agar siap dipakai dalam pemodelan.
2. Pemodelan SARIMA
Tahap ini meliputi identifikasi pola linier dan musiman menggunakan analisis ACF dan PACF. Parameter SARIMA ditentukan kemudian dilakukan estimasi model. Hasil model diuji melalui analisis residual. Model SARIMA menghasilkan prediksi inflow berdasarkan pola musiman dan tren linier.
3. Pengolahan Residu dengan ANN

Residu hasil prediksi SARIMA digunakan sebagai input untuk ANN, ditambah variabel eksogen jika tersedia. ANN dilatih untuk menangkap pola non-linier yang tidak ditangkap SARIMA. Proses ini melibatkan pembagian data menjadi data latih dan data uji serta penyesuaian arsitektur jaringan.

4. Kombinasi Prediksi SARIMA dan ANN

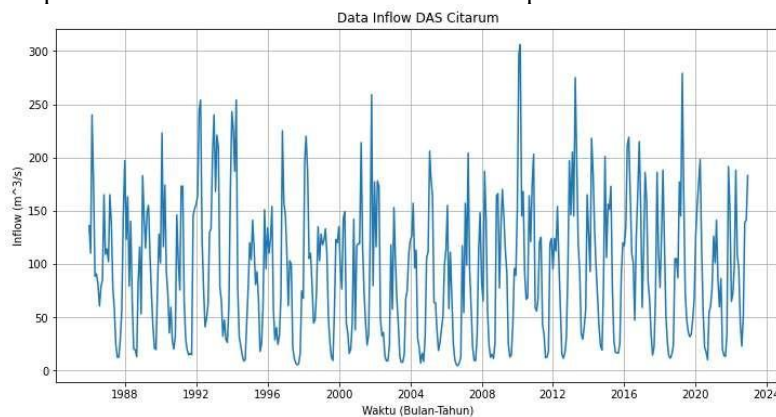
Hasil prediksi dari SARIMA dan ANN digabungkan untuk menghasilkan prediksi akhir. Kombinasi dilakukan dengan memberikan bobot sesuai performa masing-masing model. Evaluasi kinerja dilakukan dengan ukuran kesalahan seperti MAE, MAPE, dan RMSE.

5. Implementasi dalam Pengelolaan Waduk

Model SARIMA-ANN yang telah divalidasi dapat digunakan secara operasional untuk memprediksi inflow jangka pendek maupun jangka panjang. Model ini mendukung pengambilan keputusan pengelolaan waduk, termasuk penyimpanan air, pelepasan air, dan pengendalian banjir.

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini data penelitian berupa data sekunder yang diperoleh dari Perum Jasa Tirta II Bandung. Data yang dikumpulkan adalah data inflow DAS Citarum dari tahun 1986 sampai tahun 2022. Data yang terkumpul dianalisis dengan model SARIMA-ANN. Pola perubahan debit DAS Citarum tersebut ditampilkan dalam Gambar 1.



Gambar 1. Data debit Inflow DAS Citarum

Analisis statistik data inflow DAS Sungai Citarum menunjukkan pola yang menarik selama periode dari tahun 1986 hingga 2022. Data ini mencakup total 396 observasi bulanan, dengan nilai minimum inflow sebesar 12.3 m³/detik dan nilai maksimum yang mencapai 518 m³/detik. Rentang yang luas ini mengindikasikan variasi yang signifikan dalam volume air yang mengalir ke sungai dari waktu ke waktu. Secara keseluruhan, data menunjukkan kecenderungan yang sedikit bervariasi dengan nilai rata-rata (mean) sebesar 134.4 m³/detik. Distribusi data inflow menunjukkan kecenderungan yang positif (positive skewness), dengan nilai skewness sebesar 0.94, yang berarti sebagian besar data inflow cenderung berada pada nilai yang lebih rendah dari rata-rata. Hal ini juga didukung oleh nilai median sebesar 116 m³/detik, yang lebih kecil dari nilai rata-rata (mean). Kondisi ini mengisyaratkan bahwa meskipun ada beberapa bulan dengan inflow yang sangat tinggi (nilai ekstrem), sebagian besar waktu, aliran air sungai Citarum berada pada volume yang relatif lebih rendah.

Untuk melihat seberapa tersebar data, kita dapat melihat deviasi standar (standard deviation) yang mencapai 79.7 m³/detik. Nilai deviasi standar yang cukup tinggi ini menunjukkan bahwa data inflow memiliki variabilitas yang besar. Selain itu, nilai kurtosis sebesar 0.58 menunjukkan bahwa distribusi data inflow memiliki puncak yang relatif lebih datar dibandingkan dengan distribusi normal. Hal ini menunjukkan bahwa frekuensi nilai-nilai ekstrem (baik sangat rendah maupun sangat tinggi) tidak terlalu dominan. Secara ringkas, statistik deskriptif dari data inflow Citarum ini mengungkapkan gambaran yang jelas mengenai karakteristik aliran sungai. Data ini memiliki rentang yang lebar, dengan kecenderungan nilai yang lebih rendah, namun sesekali dipengaruhi oleh lonjakan aliran yang signifikan. Variabilitas yang tinggi dan distribusi yang tidak normal ini penting untuk dipertimbangkan dalam studi hidrologi atau manajemen sumber daya air, karena menunjukkan perlunya strategi pengelolaan yang adaptif untuk menghadapi fluktuasi volume air yang besar sepanjang tahun.

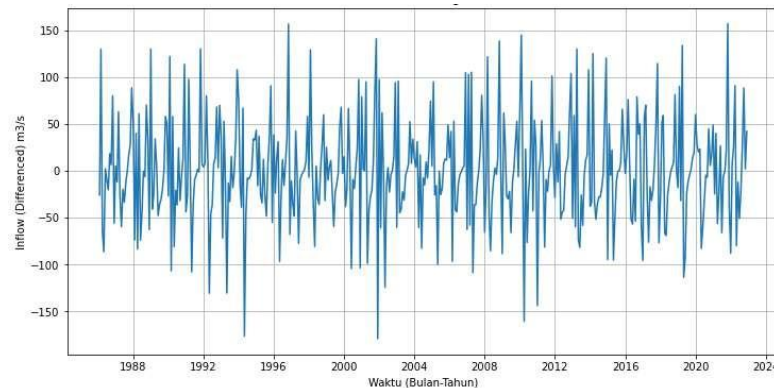
4.1. Pemodelan SARIMA

Proses perancangan model SARIMA bertujuan untuk mendapatkan parameter non musiman (p,d,q) dan parameter musiman (P,D,Q,s) yang paling mendekati untuk data Inflow DAS Citarum. Implementasi pemodelan SARIMA-ANN dilakukan dengan perangkat pemrograman Python. Langkah pertama dalam pemodelan deret waktu, khususnya dengan metode SARIMA, adalah memeriksa apakah data tersebut bersifat **stasioner**. Tanpa stasioneritas, model SARIMA tidak

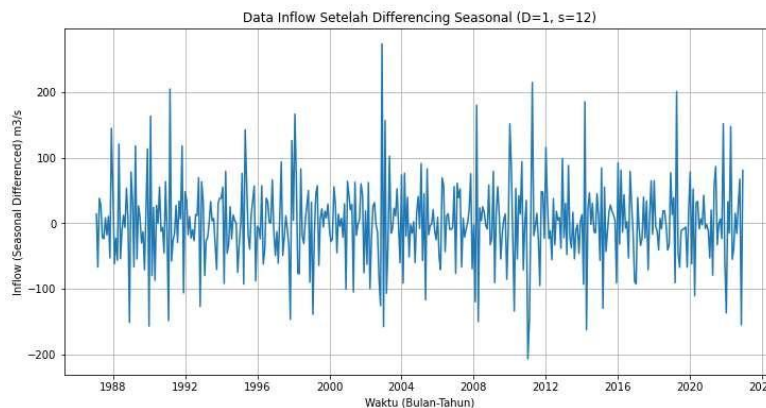
dapat memberikan hasil yang valid atau andal. Oleh karena itu, sebelum membangun model peramalan untuk data inflow Sungai Citarum, sangat penting untuk menguji properti ini. Jika data inflow tidak stasioner, kita perlu melakukan transformasi data seperti differencing untuk memenuhi asumsi model.

Untuk mengkonfirmasi stasioneritas secara objektif, kita dapat menggunakan uji statistik formal seperti Augmented Dickey-Fuller (ADF) Test. Uji ini memiliki hipotesis nol (H_0) bahwa deret waktu memiliki unit root, yang berarti data tidak stasioner. Sebaliknya, hipotesis alternatif (H_1) menyatakan bahwa data stasioner. Dengan menjalankan uji ADF pada data inflow Citarum, kita memperoleh nilai uji (ADF statistic) dan **p-value**. Dari hasil pengujian ADF nilai p-value yang dihasilkan lebih besar dari tingkat signifikansi yang ditentukan (> 0.05), sehingga dapat disimpulkan bahwa data inflow tersebut tidak stasioner.

Langkah berikutnya adalah melakukan differencing data. Proses ini biasanya dilakukan secara berulang hingga data menjadi stasioner. Proses differencing data dilakukan pada bagian data musiman (Gambar 2) untuk menentukan nilai parameter d dan bagian data non musiman (Gambar 3) untuk menentukan nilai parameter D . Untuk data Inflow DAS ini periode musiman diambil nilai $s = 12$, yaitu sifat musiman per 12 bulan atau 1 tahun.



Gambar 2. Data Inflow setelah differencing non-seasonal ($d=1$).



Gambar 3. Data Inflow setelah differencing Seasonal ($D = 1, s = 12$)

Untuk mengetahui apakah hasil differencing non-seasonal dan seasonal tersebut sudah stasioner, perlu dilakukan uji statistic ADF. Hasil uji ADF adalah sebagai berikut :

--- Uji Stasioneritas (ADF) untuk Data Setelah Differencing Non-Seasonal dan Seasonal (periode 12) ---

Statistik ADF: -10.03

Nilai p-value: 0.00 Nilai Kritis:

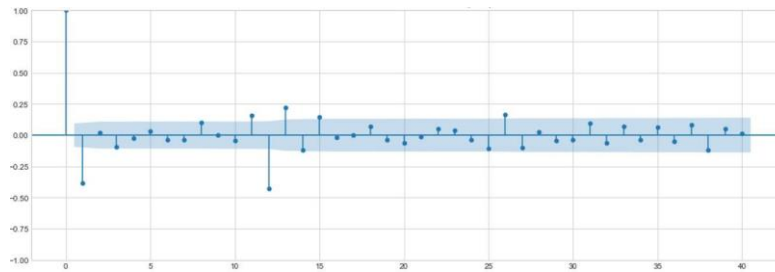
1%: -3.45

5%: -2.87

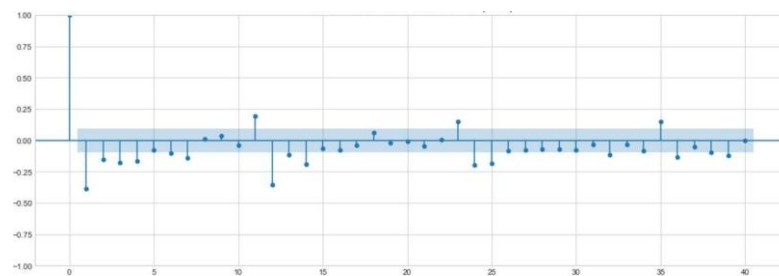
10%: -2.57

Data setelah differencing Non-Seasonal dan Seasonal (periode 12) mempunyai nilai p-value = 0.00, sehingga dapat disimpulkan bahwa data tersebut stasioner ($p\text{-value} \leq 0.05$). Dengan demikian didapatkan nilai parameter SARIMA $d = 1$ dan $D = 1$.

Langkah berikutnya adalah menentukan nilai parameter p , q untuk non musiman dan P , Q untuk musiman. Tahap ini dilakukan dengan pendekatan ACF (Auto Correlation Function) dan PACF (Partial ACF).



Gambar 4. Grafik ACF

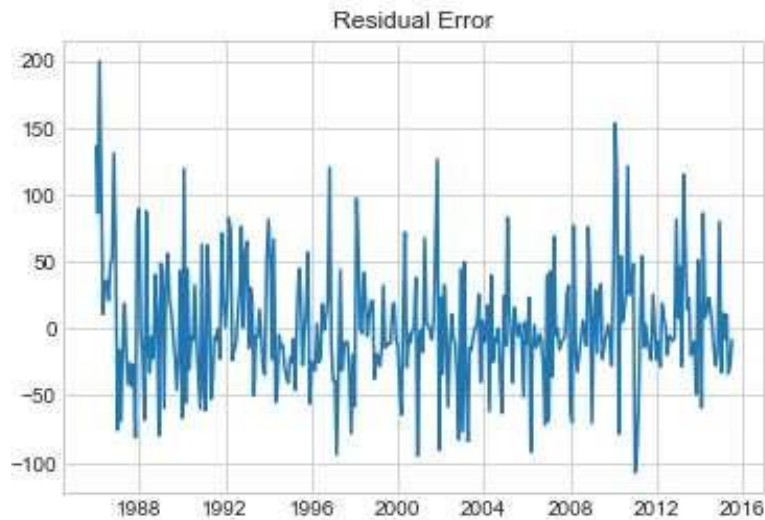


Gambar 5. Grafik PACF

Berdasarkan analisis grafik ACF dan PACF dari data inflow Citarum yang telah di-differencing, kita dapat mengidentifikasi parameter model SARIMA non-musiman (p, q) dan musiman (P, Q). Untuk komponen non-musiman model SARIMA, grafik PACF menunjukkan adanya lonjakan signifikan pada lag 1 dan kemudian menurun drastis. Fenomena ini mengindikasikan adanya komponen autoregresif (AR) orde 1, sehingga nilai $p = 1$. Sebaliknya, pada grafik ACF, tidak ada lonjakan yang jelas pada lag 1, namun ada pola penurunan bertahap. Hal ini tidak mengindikasikan adanya komponen moving average (MA), sehingga nilai $q = 0$. Selanjutnya, untuk komponen musiman, kita fokus pada lag musiman yang terjadi setiap 12 bulan (karena data bulanan). Pada grafik ACF, terlihat ada lonjakan signifikan yang melampaui batas kepercayaan pada lag 12. Hal ini menunjukkan adanya ketergantungan musiman. Pola ini mengindikasikan komponen moving average musiman (SMA) orde 1, sehingga nilai $Q = 1$. Sementara itu, pada grafik PACF, tidak ada lonjakan yang signifikan pada lag 12 yang melampaui batas kepercayaan. Oleh karena itu, kita dapat menyimpulkan bahwa tidak ada komponen autoregresif musiman (SAR), sehingga nilai $P = 0$. Dengan demikian, model awal yang diidentifikasi dari grafik adalah SARIMA(1, 1, 0)(0, 1, 1)12.

4.2. Pemodelan ANN

Setelah model SARIMA berhasil diidentifikasi dan dilatih, tahapan selanjutnya adalah menganalisis residual atau sisaan dari model tersebut. Proses ini sangat penting karena residual mewakili informasi dalam deret waktu yang tidak berhasil ditangkap oleh model SARIMA. Idealnya, residual harus bersifat acak (white noise), artinya tidak ada lagi pola atau struktur yang dapat diprediksi. Namun, seringkali residual masih mengandung pola non-linear yang tidak dapat dijelaskan oleh model SARIMA linier. Residual inilah yang akan menjadi input kunci untuk model Artificial Neural Network (ANN), yang dikenal sangat efektif dalam menangkap pola-pola non-linear yang kompleks. Dengan demikian, model ANN akan berfungsi sebagai "pembersih" residual, meningkatkan akurasi peramalan secara keseluruhan.

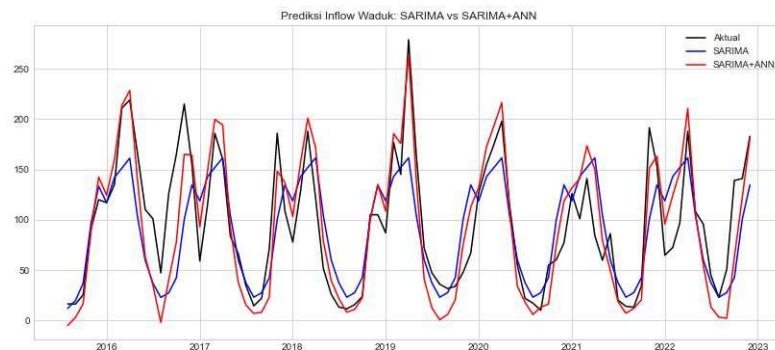


Gambar 6. Residual error dari pelatihan SARIMA

Data residual error dari hasil pelatihan SARIMA ditampilkan dalam Gambar 6. Langkah berikutnya adalah mempersiapkannya untuk melatih model ANN. Residual ini akan diubah menjadi input untuk model ANN, di mana nilai residual pada waktu t digunakan untuk memprediksi nilai residual pada waktu $t+1$. Struktur ANN yang digunakan adalah tipe *Multi Layer Perceptron Regressor* (MLP Regressor). Model ini akan dilatih dengan menggunakan data residual historis untuk menemukan hubungan non-linear. Dalam penelitian ini digunakan model MLP Regressor dengan 2 lapisan tersembunyi, aktivasi RELU dan solver optimasi ADAM. MLP Regressor paling cocok untuk melengkapi model SARIMA karena arsitekturnya yang sederhana namun kuat untuk tugas regresi non-linear. Residual dari model SARIMA dapat dianggap sebagai data deret waktu white noise yang idealnya tidak memiliki korelasi. Namun, seringkali masih ada pola non-linear yang tersembunyi. MLP, dengan arsitektur multi-lapisan dan fungsi aktivasi non-linier, sangat efektif dalam menangkap pola-pola tersebut tanpa perlu mempertimbangkan urutan temporal data. Hal ini membedakannya dari RNN yang secara inheren dirancang untuk data deret waktu yang memiliki dependensi sekuensial yang kuat. Residual SARIMA diasumsikan tidak memiliki dependensi linier lagi, sehingga MLP yang fokus pada pemodelan hubungan input-output non-linier menjadi pilihan yang lebih efisien dan tidak terlalu kompleks.

4.3. Integrasi SARIMA-ANN

Proses terakhir adalah mengintegrasikan hasil prediksi dari kedua model. Output dari model ANN adalah nilai residual yang diprediksi untuk periode waktu mendatang. Nilai ini kemudian ditambahkan kembali ke hasil prediksi dari model SARIMA. Dengan kata lain, peramalan akhir adalah penjumlahan dari prediksi SARIMA dan prediksi ANN. Untuk mengetahui kinerja dari model SARIMA-ANN, dilakukan pengujian dengan data uji. Dimana data awal dibagi 80 % data training dan 20% data testing. Hasil pengujian SARIMA-ANN ditampilkan dalam Gambar 7.



Gambar 7. Hasil prediksi Inflow DAS Citarum dengan SARIMA dan SARIMA- ANN.

Berdasarkan grafik perbandingan antara nilai aktual, prediksi SARIMA, dan prediksi SARIMA-ANN, dapat disimpulkan bahwa model hibrida SARIMA-ANN menunjukkan kinerja yang superior dalam memprediksi data deret waktu dibandingkan model SARIMA tunggal. Model SARIMA, meskipun telah berhasil menangkap pola tren dan musiman, masih menunjukkan deviasi yang signifikan dari nilai aktual pada beberapa titik data. Prediksi SARIMA cenderung mengikuti pola

umum data, namun seringkali gagal untuk memprediksi puncak (peak) dan lembah (trough) dengan akurat. Sebaliknya, prediksi yang dihasilkan oleh model SARIMA-ANN, yang menggunakan MLP Regressor untuk mengoreksi residual dari SARIMA, menunjukkan kecocokan yang jauh lebih baik dengan nilai aktual. Garis prediksi SARIMA-ANN terlihat lebih dinamis dan mampu mengikuti fluktuasi data yang tidak dapat dijelaskan oleh model linier SARIMA. Perbaikan ini membuktikan bahwa MLP Regressor berhasil menangkap pola non-linear yang tersembunyi dalam residual SARIMA, sehingga secara efektif mengurangi kesalahan prediksi. Oleh karena itu, model hibrida SARIMA-ANN lebih akurat dalam mereplikasi perilaku data deret waktu yang kompleks, menjadikannya pilihan yang lebih unggul untuk peramalan yang presisi. Evaluasi kinerja antara SARIMA dan SARIMA-ANN juga bisa dilihat dari perbandingan nilai metrik evaluasi MAPE, RMSE.

Tabel 1. Metrik Prediksi Kinerja SARIMA dan SARIMA-ANN

METRI K	SARIM A	SARIMA-ANN
MAE	32.21	25.51
MSE	1829.16	1028.65
RMSE	42.77	32.07

Berdasarkan hasil evaluasi pada Tabel 1, model hibrida SARIMA-ANN menunjukkan performa yang secara signifikan lebih baik dibandingkan dengan model SARIMA tunggal. Hal ini terbukti dari nilai-nilai metrik evaluasi yang jauh lebih rendah pada model SARIMA-ANN. Mean Absolute Error (MAE) SARIMA- ANN sebesar 25.51, yang lebih rendah dari MAE SARIMA sebesar 32.21. Penurunan ini mengindikasikan bahwa rata-rata selisih absolut antara nilai prediksi dan nilai aktual pada SARIMA-ANN lebih kecil. Artinya, secara umum, prediksi model hibrida lebih dekat dengan data sesungguhnya.

IV. KESIMPULAN

Berdasarkan hasil analisis dan evaluasi, penelitian ini menyimpulkan bahwa model hibrida SARIMA-ANN sangat efektif dalam memprediksi debit aliran masuk (inflow) DAS Citarum. Penggabungan metode linier SARIMA, yang mampu menangkap pola musiman dan tren, dengan metode non-linier ANN, yang efektif dalam mengidentifikasi pola-pola tersembunyi dalam residual, dapat menghasilkan peningkatan akurasi peramalan yang signifikan. Hal ini dibuktikan dengan nilai MAE, MSE, dan RMSE yang jauh lebih rendah pada model SARIMA-ANN dibandingkan dengan model SARIMA tunggal. Keunggulan model hibrida ini menjadikannya alat yang andal untuk perencanaan dan pengelolaan sumber daya air. Implementasi model ini memberikan manfaat besar bagi Perum Jasa Tirta dalam mengoptimalkan operasi waduk, mengendalikan risiko banjir dan kekeringan, serta memastikan ketersediaan air yang berkelanjutan.

DAFTAR PUSTAKA

- [1] Azad, A.S.; Sokkalingam, R.; Daud, H.; Adhikary, S.K.; Khurshid, H.; Mazlan, S.N.A.;Rabbani, M.B.A. (2022) Water Level Prediction through Hybrid SARIMA and ANN Models Based on Time Series Analysis: Red Hills Reservoir Case Study. *Sustainability* **2022**, 14, 1843. <https://doi.org/10.3390/su14031843>
- [2] Birhanu, K., Abcdef, T., Olumana, M., & Abcdef, D. (2017). Application of SARIMA model to forecasting monthly flows in Waterval River , South Africa. <https://doi.org/10.1515/jwld-2017-0088>
- [3] Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018). Time Series Forecasting of Temperatures using SARIMA : An Example from Nanjing <https://doi.org/10.1088/1757-899X/394/5/052024>
- [4] Cong, J., Ren, M., Xie, S., & Wang, P. (2019). Predicting seasonal influenza based on SARIMA model, in mainland China from 2005 to 2018. *International Journal of Environmental Research and Public Health*, 16(23). <https://doi.org/10.3390/ijerph16234760>
- [5] Dimri, T., Ahmad, S., & Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. *Journal of Earth System Science*, 129(1). <https://doi.org/10.1007/s12040-020-01408-x>
- [6] Etuk, E. H., Aboko, S., Harcourt, P., & Harcourt, P. (2014). An additive seasonal Box-Jenkins model for Nigerian monthly savings deposit rates. 2(March), 54–59.
- [7] Farajzadeh, J., Fakheri Fard, A., & Lotfi, S. (2014). Modeling of monthly rainfall and runoff of Urmia lake basin using “feed forward neural network” and “time series analysis” model. *Water Resources and Industry*, 7–8, 38–48. <https://doi.org/10.1016/j.wri.2014.10.003>
- [8] Hamidi, K., Sedghi, H., & Telvari, A. (2017). Forecasting by Stochastic Models to Inflow of Karkheh Dam at. 3(5), 340–350.
- [9] Kasus, S., Hujan, C., Ambon, K., Maluku, P., Kafara, Z., Rumlawang, F. Y., & Sinay, L. J. (2017). Peramalan Curauh hujan

- dengan Pendekatan Seasonal Autoregressive Integrated Moving Average(SARIMA) Rainfall Forecasting Using Seasonal Autoregressive Integrated Moving Average Model (SARIMA). 11, 63–74.
- [10] Liu, H., Li, C., Shao, Y., Zhang, X., Zhai, Z., Wang, X., & Qi, X. (2020). Journal of Infection and Public Health Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011 – 2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. *Journal of Infection and Public Health*, 13 (2), 287–294. <https://doi.org/10.1016/j.jiph.2019.12.008>
- [11] Maulana, H. A., Sampe, M. Z., & Hanifah, F. (2017). Pemodelan dan Peramalan Deret Waktu Studi kasus : Suhu Permukaan Laut di Selatan Jawa Timur. 1(2), 187–199.
- [12] Nastasia F. Margini a,* , Nadjadji Anwar a, Wasis Wardoyo a, D. D. Prastyo b, Zulkifli Yusop (2022), Monthly Inflow Forecasting of Three Multi-Purpose Reservoirs, *International Journal on Advance Science Engineering Information Technology (IJESIT)* Vol.12 (2022) No. 6 ISSN: 2088-5334
- [13] Prawaka, F., Zakaria, A., & Tugiono, S. (2016). Analisis Data Curah Hujan yang Hilang Dengan Menggunakan Metode Normal Ratio , Inversed Square Distance , dan Rata-Rata Aljabar (Studi Kasus Curah Hujan Beberapa Stasiun Hujan Daerah Bandar Lampung). 4(3), 397–406.
- [14] Ruhiat, D. dan A. E. (2018). Pengaruh faktor musiman pada pemodelan deret waktu untuk peramalan debit sungai dengan metode sarima. 2(2), 117–128.