

PEMODELAN VARIABILITAS MUSIMAN DALAM CURAH HUJAN: PENDEKATAN DERET WAKTU

Ali Mustopa¹⁾, Joko Dwi Santoso²⁾, Muhammad Sachib Farhan Nauvaldhi³⁾, Devi Wulandari⁴⁾

^{1,4)}Sistem Informasi Universitas Amikom Yogyakarta

²⁾Teknik Komputer Universitas Amikom Yogyakarta

³⁾Informatika Universitas Amikom Yogyakarta

email : ali.m@amikom.ac.id¹⁾, jds@amikom.ac.id²⁾, farhan.nauvaldi2014@students.amikom.ac.id³⁾, devi@amikom.ac.id⁴⁾

Abstraksi

Presipitasi merupakan komponen fundamental dalam siklus hidrologi Bumi yang memiliki dampak signifikan terhadap ekosistem, produktivitas pertanian, dan pengelolaan sumber daya air. Variabilitas presipitasi, terutama pola musiman, sangat penting untuk dipahami dalam konteks perencanaan yang efektif dan strategi adaptasi terhadap perubahan iklim yang semakin nyata. Penelitian ini bertujuan untuk mengeksplorasi variabilitas musiman curah hujan di Bangladesh melalui analisis deret waktu yang komprehensif, dengan memanfaatkan teknik statistik canggih, termasuk model Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), dan Exponential Smoothing State Space (ETS). Data yang digunakan dalam penelitian ini mencakup catatan curah hujan dan suhu rata-rata bulanan dari tahun 1901 hingga 2023, yang diambil dari Humanitarian Data Exchange (HDX). Analisis dilakukan untuk mengidentifikasi pola dan tren musiman yang ada dalam data tersebut. Hasil analisis menunjukkan bahwa model SARIMA $(2, 0, 1) \times (0, 1, 1, 12)$ memberikan kinerja terbaik dibandingkan dengan model lainnya, dengan nilai Root Mean Squared Error (RMSE) dan Mean Absolute Percentage Error (MAPE) yang rendah, yang menunjukkan akurasi prediksi yang tinggi. Sementara itu, model ETS juga menunjukkan hasil yang baik dalam hal akurasi, tetapi dengan nilai MAPE yang lebih tinggi, yang mengindikasikan bahwa meskipun model ini mampu menangkap pola musiman, proporsi kesalahan relatif terhadap data aktual lebih besar. Penelitian ini menekankan pentingnya mempertimbangkan komponen musiman dalam pemodelan deret waktu, karena fluktuasi musiman yang terjadi secara periodik dapat ditangkap dengan lebih baik, sehingga menghasilkan prediksi yang lebih konsisten dan akurat. Dengan demikian, hasil penelitian ini memberikan wawasan berharga untuk perencanaan dan strategi adaptasi di masa depan, terutama dalam konteks perubahan iklim yang dapat mempengaruhi pola curah hujan.

Kata Kunci :

ARIMA, SARIMA, ETS, Analisis Deret Waktu, Analisis Kinerja, Model Terbaik

Abstract

Precipitation is recognized as a fundamental component of the Earth's hydrological cycle, significantly impacting ecosystems, agricultural productivity, and water resource management. The variability of precipitation, particularly seasonal patterns, is deemed essential for understanding effective planning and adaptation strategies in response to increasingly evident climate change. This study aims to explore the seasonal variability of rainfall in Bangladesh through comprehensive time series analysis, utilizing advanced statistical techniques, including Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Exponential Smoothing State Space (ETS). The data used in this research includes monthly average rainfall and temperature records from 1901 to 2023, sourced from the Humanitarian Data Exchange (HDX). The analysis is conducted to identify existing seasonal patterns and trends within the data. The results indicate that the SARIMA $(2, 0, 1) \times (0, 1, 1, 12)$ model is found to provide the best performance compared to other models, with low values of Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE), demonstrating high prediction accuracy. Meanwhile, the ETS model is also shown to yield good results in terms of accuracy, but with higher MAPE values, indicating that although this model is capable of capturing seasonal patterns, the proportion of error relative to actual data is greater. This research emphasizes the importance of considering seasonal components in time series modeling, as periodic seasonal fluctuations can be captured more effectively, resulting in more consistent and accurate predictions. Thus, the findings of this study provide valuable insights for future planning and adaptation strategies, particularly in the context of climate change that may affect rainfall patterns.

Keywords :

Seasonal Variability, Precipitation, Time Series Analysis, SARIMA, ETS

Pendahuluan

Presipitasi [1] merupakan komponen fundamental dari siklus hidrologi Bumi, memainkan peran penting dalam membentuk ekosistem, produktivitas pertanian, dan pengelolaan sumber daya air. Memahami variabilitas presipitasi, khususnya pola musiman, sangat penting untuk perencanaan yang efektif dan strategi adaptasi dalam menghadapi perubahan iklim serta kemunculan hujan [2]. Variabilitas musiman dalam hujan dapat berdampak signifikan terhadap ketersediaan air, hasil panen, dan frekuensi peristiwa cuaca ekstrem, sehingga sangat penting untuk mengembangkan model yang kuat yang dapat menangkap dinamika ini secara akurat [3]. Penelitian ini bertujuan untuk mengeksplorasi variabilitas musim hujan melalui analisis deret waktu yang komprehensif, dengan memanfaatkan teknik statistik canggih untuk memahami pola curah hujan historis dan meningkatkan kemampuan peramalan di masa depan. Signifikansi pemodelan variabilitas musiman semakin penting mengingat meningkatnya frekuensi dan intensitas peristiwa terkait iklim, seperti kekeringan dan banjir. Metode tradisional sering kali mengabaikan pola musiman yang kompleks, sehingga penelitian ini akan menerapkan pendekatan deret waktu untuk mengidentifikasi dan mengukur tren musiman dengan menggunakan data curah hujan historis yang mencakup beberapa dekade. Artikel ini akan disusun dalam beberapa bagian, dimulai dengan tinjauan pustaka tentang pemodelan presipitasi dan analisis musiman. Selanjutnya, metodologi yang digunakan, termasuk pengumpulan data dan model deret waktu seperti ARIMA, SARIMA dan ETS. Penggunaan SARIMA dan ETS untuk melihat variasi pola dan dapat digunakan untuk melihat variasi musim. Hasil analisis deret waktu akan menampilkan pola dan tren musiman, diikuti dengan validasi model untuk menilai akurasi. Bagian terakhir akan membahas perbandingan model untuk mencari yang terbaik dari masing model prediksi.

Tinjauan Pustaka

ARIMA adalah model statistik yang digunakan untuk analisis deret waktu yang menggabungkan tiga komponen utama: autoregressive (AR), differencing (I), dan moving average (MA). Model ini sangat efektif untuk data yang stasioner, di mana pola tidak berubah seiring waktu. Box dan Jenkins [4] menunjukkan bahwa ARIMA dapat digunakan untuk memprediksi data deret waktu dengan baik, terutama ketika data menunjukkan pola yang jelas. Namun, model ini memerlukan data yang stasioner, sehingga sering kali perlu dilakukan transformasi untuk mencapai stasionaritas. ARIMA juga dapat diadaptasi untuk data non-stasioner dengan menggunakan differencing, yang membantu menghilangkan tren dari data. Selain itu, Wei [5] menekankan pentingnya pemilihan parameter yang tepat dalam model ARIMA untuk meningkatkan akurasi peramalan.

SARIMA adalah pengembangan dari model ARIMA yang dirancang untuk menangani data deret waktu

dengan komponen musiman. Model ini menambahkan parameter musiman ke dalam model ARIMA, memungkinkan analisis yang lebih mendalam terhadap pola musiman dalam data. Menurut penelitian oleh Hyndman dan Athanasopoulos [6], SARIMA dapat secara efektif menangkap fluktuasi musiman dan memberikan prediksi yang lebih akurat dibandingkan dengan model ARIMA standar ketika diterapkan pada data yang memiliki pola musiman yang jelas. Penelitian oleh Clements dan Hendry [7] juga menunjukkan bahwa pemilihan parameter musiman yang tepat sangat penting untuk meningkatkan akurasi model, serta menekankan pentingnya analisis residual untuk memastikan model yang baik.

Exponential Smoothing State Space Models (ETS) adalah pendekatan lain untuk analisis deret waktu yang berfokus pada peramalan. Model ini menggunakan metode peluruhan eksponensial untuk memberikan bobot yang lebih besar pada data terbaru, sehingga lebih responsif terhadap perubahan dalam pola data. Penelitian oleh Hyndman dan Khandakar [8] menunjukkan bahwa ETS dapat menangani data dengan tren dan musiman, dan sering kali lebih mudah diinterpretasikan dibandingkan dengan model ARIMA. Model ini juga menunjukkan kinerja yang kompetitif dalam hal akurasi peramalan, terutama untuk data yang memiliki pola musiman yang kuat. Selain itu, penelitian oleh Makridakis et al. [9] menekankan bahwa ETS sering kali lebih sederhana dan lebih cepat dalam hal komputasi dibandingkan dengan model ARIMA, menjadikannya pilihan yang menarik untuk peramalan waktu nyata.

Ritu, dkk [10] menyajikan studi tentang prediksi curah hujan musim panas di Bangladesh untuk tahun 2021 menggunakan Climate Predictability Tool (CPT), yang menggunakan suhu permukaan laut (SST) sebagai prediktor dan curah hujan yang diamati sebagai prediktor berdasarkan data dari tahun 1990 hingga 2020. Studi ini menemukan korelasi yang kuat (0,7360) antara curah hujan yang diamati dan yang dihasilkan CPT, meskipun CPT cenderung melebih-lebihkan curah hujan maksimum. Penelitian ini menyoroti variasi yang signifikan dalam pola curah hujan, dengan curah hujan tertinggi pada tahun 2021 terjadi di Cox's Bazar, menyimpang dari tren historis curah hujan maksimum di Sylhet dan Srimangal. Studi ini menekankan pentingnya prediksi curah hujan yang akurat untuk pertanian dan manajemen bencana di Bangladesh, dan memvalidasi efektivitas CPT melalui berbagai metode statistik, yang pada akhirnya berkontribusi pada peningkatan kemampuan peramalan dalam konteks perubahan iklim.

Ashwini dkk [11] membahas penerapan model ARIMA Musiman (SARIMA) untuk memperkirakan curah hujan musim hujan di Tamilnadu. Penelitian ini menggunakan data curah hujan dari Januari 1990 hingga Desember 2017, yang diperoleh dari Departemen Meteorologi India. Model SARIMA menunjukkan stasioneritas deret waktu dan

dievaluasi menggunakan Mean Squared Error (MSE) dan Root Mean Squared Error (RMSE). Hasilnya menunjukkan bahwa model ARIMA secara akurat memperkirakan curah hujan dengan kesalahan minimal, menjadikannya alat yang andal untuk memperkirakan curah hujan musim hujan di masa mendatang dan berpotensi mengurangi dampak bencana alam seperti banjir dan kekeringan.

Prapcoyo dkk. [12] membahas tentang memprediksi inflasi bulanan di Yogyakarta dengan menggunakan model Exponential Smoothing-State Space (ETS), yaitu model ETS (A, N, A) yang merupakan gabungan dari additive error, no trend, dan additive seasonality. Penelitian ini menggunakan data sekunder dari Januari 2015 sampai dengan Desember 2021 yang diperoleh dari Badan Pusat Statistik (BPS). Model ETS (A, N, A) dinilai paling tepat dengan nilai Akaike Information Criterion (AIC) sebesar 145,1996, Root Mean Square Error (RMSE) sebesar 0,2166014, dan Mean Absolute Percentage Error (MAPE) sebesar 127,1662. Prakiraan untuk tiga periode mendatang menunjukkan inflasi bulanan di Yogyakarta cenderung stabil dengan sedikit fluktuasi namun tetap berada di bawah 10%.

Sujito dkk. [13] membandingkan kinerja model Single Exponential Smoothing (SES) dan Exponential Smoothing-State Space (ETS) untuk memperkirakan harga emas harian. Dengan menggunakan data dari 25 Maret 2019 hingga 24 Maret 2020, studi ini mengevaluasi model berdasarkan Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Mean Absolute Percentage Error (MAPE), dan Root Mean Square Error (RMSE). Hasilnya menunjukkan bahwa model ETS(M, N, N), yang memiliki kesalahan perkalian dan tidak memiliki tren atau musim, mengungguli model SES. Model ETS(M, N, N) mencapai koefisien eksponensial sebesar 0,9999, dengan nilai AIC dan BIC masing-masing sebesar 2902,143 dan 2912,882, serta nilai MAPE dan RMSE masing-masing sebesar 0,6513446 dan 15,01525. Prakiraan untuk tiga periode berikutnya berfluktuasi sekitar \$1600 per troy ons.

Ashwini dkk. [11] membahas perilaku masa depan curah hujan dan suhu rata-rata bulanan di negara-negara Asia Selatan menggunakan model Seasonal Autoregressive Moving Average (SARIMA). Data dari Januari 1961 hingga Desember 2016 dikumpulkan dari Portal Pengetahuan Perubahan Iklim Grup Bank Dunia. Uji tren Mann-Kendall dan Sen's Slope Estimator digunakan untuk menganalisis tren, yang mengungkap perubahan signifikan dalam pola curah hujan dan suhu. Studi ini menemukan bahwa perubahan iklim terlihat jelas dalam curah hujan dan suhu di seluruh wilayah. Model SARIMA yang dipilih divalidasi dan digunakan untuk memperkirakan curah hujan dan suhu bulanan dari Januari 2011 hingga Desember 2021, yang memberikan wawasan berharga untuk penelitian dan pembuatan kebijakan masa depan dalam adaptasi perubahan iklim.

Widdya dan Puspita. [14] menggunakan metode Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk memprediksi jumlah kasus Covid-19 di Kota Padang. Data harian dari 2 Agustus 2020 hingga 6 Januari 2021 digunakan untuk membangun model SARIMA (0, 1, 1)(0, 1, 1)₇, yang terbukti menjadi model terbaik dengan nilai Mean Squared Error sebesar 330.9333. Hasil peramalan menunjukkan bahwa jumlah kasus Covid-19 masih fluktuatif, berkisar antara 20 hingga 66 kasus per hari. Analisis menunjukkan bahwa model ini memenuhi asumsi white noise dengan parameter yang signifikan berbeda dari nol, sehingga layak digunakan untuk peramalan. Upaya pemerintah daerah dalam mendisiplinkan masyarakat sangat diperlukan agar Kota Padang dapat segera berubah menjadi zona hijau dari Covid-19.

Metode Penelitian

Tahap awal dari metode yang diusulkan melibatkan persiapan data. Data Curah Hujan Bangladesh, yang diselenggarakan oleh Humanitarian Data Exchange (HDX), menawarkan kumpulan data lengkap terkait iklim dan cuaca, yang secara khusus berfokus pada statistik curah hujan subnasional. Data ini menggabungkan suhu rata-rata bulanan (dalam derajat Celsius) dan total curah hujan bulanan (dalam milimeter), yang disusun dalam format CSV dengan kolom untuk suhu, bulan, tahun, dan curah hujan. Meliputi rentang waktu dari tahun 1901 hingga 2023, data ini memberikan wawasan historis yang berharga tentang tren iklim, yang memfasilitasi analisis gabungan pola suhu dan curah hujan selama lebih dari satu abad. Data tersebut diambil dari laman kaggle [15]. Setelah persiapan data, tahap berikutnya adalah pembagian kumpulan data menjadi subset pelatihan dan pengujian. Hal ini dicapai melalui pemisahan pelatihan-pengujian, di mana proporsi data yang telah ditentukan sebelumnya, biasanya 80%, dialokasikan untuk melatih model, sementara 20% sisanya dicadangkan untuk menguji kinerja prediktifnya. Pembagian ini penting karena memungkinkan penilaian kemampuan model untuk menggeneralisasi ke data yang tidak terlihat, sehingga memberikan wawasan tentang akurasi prediktifnya.

Setelah data disiapkan dan dibagi, metodologi berlanjut ke pemilihan model, yang merupakan komponen penting dari analisis deret waktu. Dalam studi ini, kami berfokus pada dua teknik pemodelan utama: Autoregressive Integrated Moving Average (ARIMA) dan Seasonal Autoregressive Integrated Moving Average (SARIMA). Model ARIMA didefinisikan oleh tiga parameter: (p), (d), dan (q). Parameter (p) menunjukkan jumlah suku autoregresif, (d) menunjukkan derajat perbedaan yang diperlukan untuk mencapai stasioneritas, dan (q) menandakan jumlah kesalahan prakiraan yang tertinggal dalam persamaan prediksi. Model SARIMA memperluas kerangka kerja ARIMA dengan menggabungkan komponen musiman, yang

dicirikan oleh parameter tambahan yang memperhitungkan efek musiman.

Untuk mengidentifikasi parameter optimal bagi model ARIMA dan SARIMA, pendekatan pencarian grid digunakan. Pendekatan ini melibatkan penjelajahan sistematis rentang nilai yang telah ditetapkan sebelumnya untuk setiap parameter. Untuk ARIMA, parameter (p) , (d) , dan (q) biasanya ditetapkan dalam rentang bilangan bulat kecil, seperti $(p, q \in \{0, 1, 2\})$ dan $(d = 0)$ (dengan asumsi datanya stasioner). Untuk SARIMA, parameter musiman diperkenalkan, dengan (m) mewakili periode musiman, yang dapat mengambil nilai seperti 4, 6, atau 12, tergantung pada sifat data.

Evaluasi model dilakukan menggunakan metrik kinerja yang mengukur keakuratan prediksi. Root Mean Squared Error (RMSE) dan Mean Absolute Percentage Error (MAPE) digunakan sebagai metrik evaluasi utama. RMSE dihitung menggunakan rumus:

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

di mana (y_t) mewakili nilai aktual, (\hat{y}_t) menunjukkan nilai prediksi, dan (n) adalah jumlah observasi. MAPE didefinisikan sebagai:

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

Metrik ini memberikan penilaian komprehensif atas kemampuan prediksi model, yang memungkinkan perbandingan berbagai konfigurasi dan pemilihan model dengan kinerja terbaik.

Setelah mengidentifikasi model ARIMA dan SARIMA yang optimal, metodologi berlanjut ke tahap peramalan. Pada tahap ini, model yang dipilih disesuaikan dengan data pelatihan, dan prediksi dibuat untuk kumpulan data pengujian. Proses penyesuaian melibatkan estimasi parameter model berdasarkan data pelatihan, diikuti dengan pembuatan prakiraan untuk titik waktu mendatang yang ditentukan. Prakiraan tersebut kemudian dibandingkan dengan nilai aktual dalam kumpulan pengujian untuk mengevaluasi kinerja model.

Selain ARIMA dan SARIMA, metode yang diusulkan juga menggabungkan Exponential Smoothing State Space Model (ETS) sebagai teknik peramalan alternatif. Model ETS dicirikan oleh kemampuannya untuk menangkap level, tren, dan komponen musiman dari data deret waktu. Model ini didefinisikan oleh tiga komponen: jenis kesalahan, jenis tren, dan jenis musiman. Pemilihan komponen-komponen ini dipandu oleh karakteristik data, dan

berbagai konfigurasi diuji untuk mengidentifikasi model yang optimal.

Hasil dan Pembahasan

Bab ini menyajikan analisis data curah hujan untuk Bangladesh, dengan fokus pada tren dan hubungan yang diamati selama periode 1901 hingga 2023. Tujuannya adalah untuk memberikan wawasan tentang bagaimana variabel iklim telah berubah dari waktu ke waktu. Bab ini disusun untuk pertama-tama menjelaskan kumpulan data, diikuti oleh metodologi yang digunakan, dan kemudian menyajikan dan membahas hasil analisis. Dataset yang digunakan untuk analisis ini terdiri dari catatan historis curah hujan.

Kumpulan data yang digunakan untuk analisis ini bersumber dari Humanitarian Data Exchange (HDX) dan mencakup data suhu rata-rata bulanan dan total curah hujan bulanan untuk Bangladesh. Kumpulan data tersebut disusun dalam format CSV dengan kolom-kolom utama berikut: suhu rata-rata (tem), bulan, tahun, dan total curah hujan (rain).

Table 1: Ringkasan Dataset

Characteristic	Value
Number of Observations	1,476
Date Range	1901 - 2023
Variables Included	Temperature, Rainfall
Geographic Coverage	Subnational (Bangladesh)

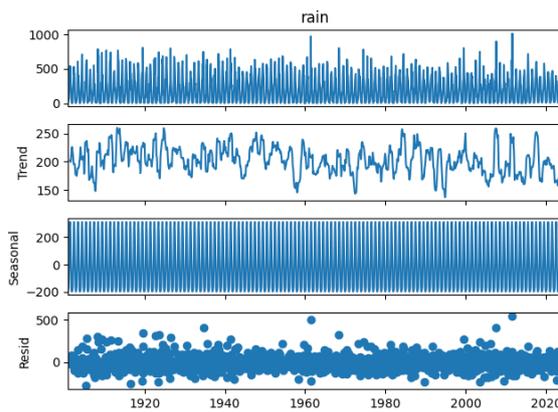
Ringkasan statistik deskriptif untuk variabel-variabel utama disajikan dalam Tabel 1. Tabel 2 berikut ini mencakup ukuran-ukuran seperti nilai rata-rata, median, deviasi standar, nilai minimum, dan nilai maksimum curah hujan.

Tabel 2. Statistik deskriptif

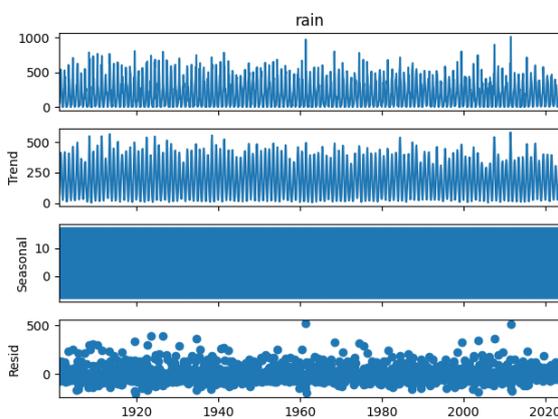
Dataset	tem	Month	Year	rain
Count	1.380	1.380	1.380	1.380
mean	25,0662	6,5000	1.958	203,2754
std	3,6824	3,4533	33,208	202,7309
min	16,8006	1,0000	1.901	0,0000
25%	22,1141	3,7500	1.929	18,4989
50%	27,0130	6,5000	1.958	145,0860
75%	28,0079	9,2500	1.987	347,6395
max	29,5260	12,0000	2.015	1.012,02

Grafik menunjukkan dekomposisi data deret waktu yang disebut 'hujan' menjadi komponen-komponennya: tren, musiman, dan residual. Panel atas menampilkan data deret waktu 'hujan' asli, yang berfluktuasi seiring waktu. Panel kedua menggambarkan komponen tren, yang menggambarkan pola naik dan turun jangka panjang dalam data. Panel ketiga menyajikan komponen musiman, yang menyoroti pola siklus kuat yang berulang setiap tahun. Panel bawah menunjukkan komponen residual, yang mewakili variabilitas yang

tersisa dalam data setelah memperhitungkan tren dan musim.

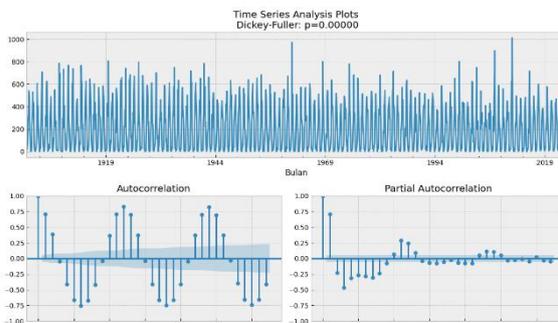


Grafik 1. Dekomposisi Data Periodic 12



Grafik 2. Dekomposisi Data Periodic 6

Grafik 1 menggambarkan data deret waktu asli, yang memperlihatkan pola yang relatif konsisten dengan lonjakan sesekali. Grafik 2, di sisi lain, menggambarkan residual, yang merupakan deviasi data asli dari komponen tren dan musiman. Residu tersebut tampak terdistribusi secara acak tanpa pola atau tren yang jelas. Hal ini menunjukkan bahwa data deret waktu dapat dimodelkan secara efektif menggunakan komponen tren dan musiman, sehingga menghasilkan gangguan acak yang minimal. Tidak adanya komponen musiman yang kuat di grafik tengah semakin memperkuat pengamatan ini.



Gambar 3 Analisis Deret Waktu

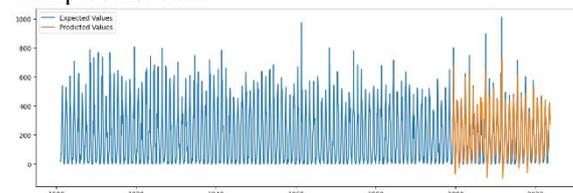
Gambar 3 menggambarkan analisis deret waktu dari suatu kumpulan data. Plot bagian atas menampilkan deret waktu itu sendiri, yang memperlihatkan pola

musiman yang jelas. Plot bagian tengah, fungsi autokorelasi (ACF), menyoroti korelasi positif yang kuat pada jeda waktu 12, yang mengonfirmasi keberadaan pola musiman yang kuat. Plot bagian bawah, fungsi autokorelasi parsial (PACF), semakin menekankan pola musiman ini dengan puncak yang signifikan pada jeda waktu 12. Hasil uji Dickey-Fuller, dengan nilai p sebesar 0,00000, menunjukkan bahwa deret waktu tersebut stasioner, yang berarti tidak memiliki tren atau kemusiman.

Tabel 3. Evaluasi Model Arima

ORDER	RMSE	MAPE	AIC	BIC
(0, 0, 0)	189.672	1205.27	17851.1	17861.4
(0, 0, 1)	151.673	1166.77	17232.2	17247.8
(0, 0, 2)	138.188	951.705	16888.7	16909.5
(1, 0, 0)	137.578	1050.22	16939.6	16955.2
(1, 0, 1)	136.603	1062.28	16910.5	16931.2
(1, 0, 2)	132.308	818.437	16767.3	16793.3
(2, 0, 0)	135.598	1053.67	16871.3	16892.1
(2, 0, 1)	115.602	865.436	16371.5	16397.5

Berdasarkan data hasil evaluasi model ARIMA pada tabel 3, terlihat bahwa berbagai konfigurasi order ARIMA telah diuji menggunakan metrik evaluasi RMSE, MAPE, AIC, dan BIC. Model ARIMA dengan konfigurasi order (2, 0, 1) menghasilkan performa terbaik dibandingkan model lainnya. Hal ini ditunjukkan dengan nilai RMSE (Root Mean Square Error) terendah sebesar 115.602, nilai MAPE (Mean Absolute Percentage Error) yang juga relatif kecil sebesar 865.436, serta nilai AIC (Akaike Information Criterion) dan BIC (Bayesian Information Criterion) yang lebih rendah masing-masing sebesar 16,371.5 dan 16,397.5. Secara umum, nilai AIC dan BIC yang lebih rendah menunjukkan bahwa model lebih sederhana dan memberikan hasil prediksi yang lebih baik. Sebagai perbandingan, model dengan order (0, 0, 0) memiliki kinerja terburuk, dengan nilai RMSE sebesar 189.672 dan MAPE sebesar 1,205.27, menandakan prediksi yang kurang akurat. Oleh karena itu, ARIMA (2, 0, 1) direkomendasikan sebagai model terbaik untuk digunakan dalam memprediksi data.



Gambar 4. Hasil Prediksi

Gambar 4 menunjukan grafik hasil prediksi di atas menunjukkan perbandingan antara nilai aktual (ditampilkan dengan garis biru) dan nilai prediksi (ditampilkan dengan garis oranye). Secara keseluruhan, nilai prediksi mengikuti pola dan tren dari data aktual dengan baik, terutama dalam menangkap fluktuasi musiman yang signifikan. Pada

bagian akhir grafik, di mana nilai prediksi lebih dominan, terlihat bahwa model cukup akurat dalam memprediksi data testing, meskipun terdapat beberapa perbedaan kecil antara prediksi dan nilai aktual di beberapa titik. Fluktuasi musiman yang konsisten pada data aktual juga berhasil ditangkap oleh model, mencerminkan bahwa komponen musiman telah dimodelkan dengan baik. Namun, terdapat penyimpangan di beberapa titik, yang kemungkinan disebabkan oleh noise atau variasi data yang tidak sepenuhnya teridentifikasi oleh model. Secara keseluruhan, model prediktif menunjukkan performa yang baik dalam mengikuti pola data, meskipun akurasi dapat ditingkatkan lebih lanjut melalui pengoptimalan parameter atau penambahan informasi pada data. Kurangnya tren yang kuat dalam rata-rata bergulir menunjukkan bahwa data curah hujan tidak meningkat atau menurun secara signifikan dari waktu ke waktu. Deviasi standar bergulir yang stabil menunjukkan bahwa variabilitas curah hujan relatif konsisten dari waktu ke waktu. Informasi ini dapat berguna untuk memahami pola curah hujan jangka panjang di area tersebut.

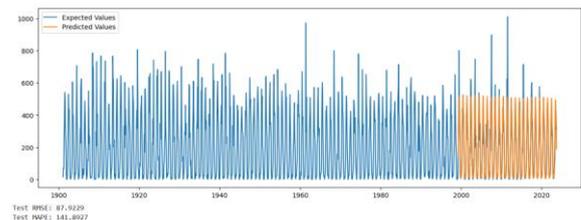
Tabel 4. Evaluasi Model Sarima

SARIMA	RMSE	MAPE	AIC	BIC
(0, 0, 0)x (0, 1, 1, 4)	189,72	1170,81	17821,09	17831,46
(0, 0, 0)x (0, 1, 1, 6)	187,45	1128,37	17749,96	17760,33
(0, 0, 0)x (0, 1, 1, 12)	87,92	141,89	15345,07	15355,42
(0, 0, 1)x (0, 1, 1, 4)	150,93	1083,92	17194,57	17210,13
(0, 0, 1)x (0, 1, 1, 6)	149,25	1098,04	17129,66	17145,21
(0, 0, 1)x (0, 1, 1, 12)	90,19	142,68	15355,33	15370,86
(0, 0, 2)x (0, 1, 1, 4)	138,06	891,03	16854,33	16875,07
(0, 0, 2)x (0, 1, 1, 6)	135,38	888,79	16768,07	16788,81
(0, 0, 2)x (0, 1, 1, 12)	88,30	128,20	15346,78	15367,49
(1, 0, 0)x (0, 1, 1, 4)	137,19	977,14	16901,30	16916,85
(1, 0, 0)x (0, 1, 1, 6)	134,28	1058,12	16793,15	16808,70
(1, 0, 0)x (0, 1, 1, 12)	89,72	143,11	15356,01	15371,55
(1, 0, 1)x (0, 1, 1, 4)	136,06	982,39	16870,55	16891,29
(1, 0, 1)x (0, 1, 1, 6)	133,42	1054,95	16773,60	16794,33
(1, 0, 1)x (0, 1, 1, 12)	90,16	136,33	15356,07	15376,78
(1, 0, 2)x (0, 1, 1, 4)	132,18	737,17	16725,62	16751,54
(1, 0, 2)x (0, 1, 1, 6)	143,29	953,99	16991,34	17017,26
(1, 0, 2)x (0, 1, 1, 12)	89,47	130,38	15352,64	15378,53
(2, 0, 0)x (0, 1, 1, 4)	134,94	966,46	16829,67	16850,41
(2, 0, 0)x (0, 1, 1, 6)	132,49	1036,78	16745,15	16765,89
(2, 0, 0)x (0, 1, 1, 12)	88,41	129,06	15352,24	15372,95

(2, 0, 1)x (0, 1, 1, 4)	114,87	785,39	16329,99	16355,92
(2, 0, 1)x (0, 1, 1, 6)	112,33	859,19	16216,11	16242,03
(2, 0, 1)x (0, 1, 1, 12)	88,94	112,98	15307,58	15333,47
(2, 0, 2)x (0, 1, 1, 4)	6068,02	402,95	15749,09	15780,20
(2, 0, 2)x (0, 1, 1, 6)	92,01	83,07	15500,09	15531,19
(2, 0, 2)x (0, 1, 1, 12)	89,14	138,84	15354,79	15385,86

Tabel 4 mengevaluasi berbagai model SARIMA berdasarkan konfigurasi (p, d, q) x (P, D, Q, s), menggunakan metrik seperti RMSE (besar kesalahan), MAPE (persentase kesalahan), AIC (keseimbangan antara kecocokan model dan kesederhanaan), dan BIC (mirip AIC tetapi dengan penalti untuk kompleksitas). Model dengan periode musiman (s = 12) umumnya mengungguli periode musiman yang lebih pendek, dengan nilai RMSE, MAPE, AIC, dan BIC yang lebih rendah. Secara khusus, model ((2, 0, 1) x (0, 1, 1, 12)) memiliki AIC terendah (15307.58), sementara ((2, 0, 2) x (0, 1, 1, 6)) memiliki MAPE terendah (83.07), dan ((0, 0, 0) x (0, 1, 1, 12)) mencapai RMSE terendah (87.92). Hasil ini menyoroti trade-off antara kompleksitas model dan kinerja, dengan periode musiman yang lebih panjang menunjukkan kemampuan prediksi yang lebih baik.

Model SARIMA terbaik menyeimbangkan akurasi prediktif dan kesederhanaan, berdasarkan RMSE, MAPE, AIC, dan BIC. Dari opsi yang ada, model ((2, 0, 1) x (0, 1, 1, 12)) menonjol dengan AIC terendah (15307.58), RMSE yang kompetitif (88.94), dan MAPE rendah (112.98), menunjukkan tingkat akurasi yang tinggi. Selain itu, model ini memiliki nilai BIC yang wajar (15333.47), mencerminkan keseimbangan yang baik antara kompleksitas dan kecocokan model. Sementara model ((2, 0, 2)x(0, 1, 1, 6)) memiliki MAPE terendah (83.07), nilai AIC dan BIC yang lebih tinggi menunjukkan kompleksitas yang lebih besar. Oleh karena itu, ((2, 0, 1) x (0, 1, 1, 12)) adalah pilihan paling cocok untuk peramalan yang akurat dan efisien.



Gambar 5. Hasil Prediksi

Grafik pada gambar 5 membandingkan nilai ekspektasi (garis biru) dan prediksi (garis oranye) dari sebuah deret waktu. Nilai prediksi ditampilkan pada bagian akhir timeline, yang kemungkinan mewakili periode pengujian atau prediksi. Meskipun nilai prediksi secara umum mengikuti tren nilai ekspektasi, terdapat penyimpangan, terutama dalam menangkap puncak dan lembah data. Metrik kinerja model, dengan Test RMSE sebesar 88.9445 dan Test MAPE sebesar 112.9786.

Tabel 5. Evaluasi Model ETS Sesional 4,6,12

ETS	RMSE	MAPE	MAE
(add, add,4)	262,6464	75,3821	186,1419
(add, add,6)	3277,4207	15258,5965	2878,6792
(add, add,12)	88,4194	141,8832	55,5360

Tabel 5 di atas menyajikan evaluasi kinerja model ETS (*Exponential Smoothing*) dengan konfigurasi komponen *additive trend* dan *additive seasonality*, menggunakan berbagai panjang periode musiman, yaitu 4, 6, dan 12. Model dengan periode musiman 12 memberikan hasil terbaik, ditunjukkan oleh nilai RMSE sebesar 88,4194, MAPE sebesar 141,8832, dan MAE sebesar 55,5360, yang menunjukkan prediksi yang paling akurat. Sebaliknya, model dengan periode musiman 6 memiliki kinerja terburuk, dengan nilai RMSE yang sangat tinggi sebesar 3277,4207, MAPE sebesar 15258,5965, dan MAE sebesar 2878,6792, menunjukkan bahwa panjang periode ini tidak cocok untuk data yang digunakan. Model dengan periode musiman 4 memiliki kinerja yang cukup baik, tetapi masih kalah dibandingkan dengan periode musiman 12, dengan nilai RMSE sebesar 262,6464, MAPE sebesar 75,3821, dan MAE sebesar 186,1419. Oleh karena itu, konfigurasi **ETS (add, add, 12)** disarankan sebagai pilihan terbaik karena lebih mampu menangkap pola musiman pada data dengan akurasi yang lebih tinggi.

Tabel 6. perbandingan Model Terbaik

Best Model	RMSE	MAPE
Arima (2, 0, 1)	115,6020	865,4360
Sarima(0, 0, 0)x(0, 1, 1, 12)	88,9445	112,9786
ETS (add, add,12)	88,4194	141,8832

Tabel 6 membandingkan akurasi tiga model peramalan—ARIMA, SARIMA, dan ETS—menggunakan RMSE (besar kesalahan) dan MAPE (kesalahan relatif). Model ARIMA (2, 0, 1) memiliki performa terburuk dengan RMSE tertinggi (115,6020) dan MAPE (865,4360), menunjukkan prediksi yang kurang akurat. Model ETS (add, add, 12) mencapai RMSE terendah (88,4194), yang mencerminkan pengurangan besar kesalahan yang baik, tetapi MAPE yang lebih tinggi (141,8832) menunjukkan akurasi proporsional yang lebih rendah. Model SARIMA(2, 0, 1) × (0, 1, 1, 12) berhasil menyeimbangkan kedua metrik secara efektif, dengan RMSE rendah (88,9445) dan MAPE terbaik (112,9786), menjadikannya pilihan paling cocok untuk peramalan yang akurat dan konsisten.

Kesimpulan dan Saran

Kesimpulannya, dari ketiga model yang dibandingkan, SARIMA(2, 0, 1) × (0, 1, 1, 12) merupakan model terbaik untuk peramalan. Model ini memiliki keseimbangan yang baik antara RMSE yang rendah (88,9445) dan MAPE terendah (112,9786),

menunjukkan akurasi prediksi yang konsisten baik dalam besar kesalahan maupun kesalahan relatif. Sementara model ETS memiliki RMSE yang sedikit lebih rendah, MAPE yang lebih tinggi menunjukkan bahwa ETS kurang akurat dalam proporsi terhadap data sebenarnya. Dengan demikian, SARIMA lebih unggul dalam memberikan hasil peramalan yang akurat dan andal.

Penggunaan model periodik dalam peramalan, khususnya dalam konteks model SARIMA, menunjukkan bahwa mempertimbangkan faktor musiman dalam analisis waktu sangat penting untuk meningkatkan akurasi prediksi. Model dengan periode musiman yang lebih panjang, seperti (s = 12), cenderung memberikan hasil yang lebih baik, baik dalam hal RMSE maupun MAPE. Hal ini mengindikasikan bahwa fluktuasi yang terjadi secara periodik dalam data dapat ditangkap dengan lebih baik, sehingga menghasilkan prediksi yang lebih konsisten dan akurat. Dengan demikian, memasukkan komponen musiman dalam model peramalan dapat meningkatkan kinerja model dan memberikan wawasan yang lebih baik terhadap pola yang ada dalam data.

Daftar Pustaka

- [1] H. P. Dasari and V. S. Challa, "A Study of Precipitation Climatology and Its Variability over Europe Using an Advanced Regional Model (WRF)," *Am. J. Clim. Chang.*, vol. 04, no. 01, pp. 22–39, 2015, doi: 10.4236/ajcc.2015.41003.
- [2] K. S. Rawat, R. K. Pal, and S. K. Singh, "Rainfall variability analysis using Precipitation Concentration Index: a case study of the western agro-climatic zone of Punjab, India," *Indones. J. Geogr.*, vol. 53, no. 3, Dec. 2021, doi: 10.22146/ijg.64890.
- [3] A. Azwar, E. Meilianda, and M. Masimin, "KAJIAN POLA CURAH HUJAN DURASI PANJANG TERKAIT DENGAN WAKTU KEJADIAN BANJIR DI KABUPATEN ACEH UTARA," *J. Arsip Rekayasa Sipil dan Perenc.*, vol. 4, no. 1, pp. 39–48, Jan. 2022, doi: 10.24815/jarsp.v4i1.16723.
- [4] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [5] W. W. S. Wei, *Multivariate time series analysis and applications*. John Wiley & Sons, 2019.
- [6] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.
- [7] M. Clements and D. F. Hendry, *Forecasting Economic Time Series*. in *Forecasting economic time series*. Cambridge University Press, 1998.
- [8] R. J. Hyndman and Y. Khandakar, "Automatic Time Series Forecasting: The forecast Package for R," *J. Stat. Softw.*, vol. 27, no. 3, 2008, doi: 10.18637/jss.v027.i03.
- [9] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition:

- Results, findings, conclusion and way forward,”*
Int. J. Forecast., vol. 34, no. 4, pp. 802–808,
Oct. 2018, doi:
10.1016/j.ijforecast.2018.06.001.
- [10] A. I. Ritu, M. N. Ahasan, and M. N. Ahasan,
“*Prediction of Summer Monsoon Rainfall over
Bangladesh Using Climate Predictability Tool
(CPT)*,” *DEW-DROP*, vol. 9, no. 1, pp. 63–73,
2023, Accessed: Jan. 08, 2025. [Online].
Available: <https://live6.bmd.gov.bd/p/Dew-Drop-Vol.9-No.1>
- [11] U. Ashwini, K. Kalaivani, K. Ulagapriya, and A.
Saritha, “*Time Series Analysis based Tamilnadu
Monsoon Rainfall Prediction using Seasonal
ARIMA*,” in *2021 6th International Conference
on Inventive Computation Technologies
(ICICT)*, IEEE, Jan. 2021, pp. 1293–1297. doi:
10.1109/ICICT50816.2021.9358615.
- [12] H. Prapcoyo and M. As’ad, “*THE
FORECASTING OF MONTHLY INFLATION
IN YOGYAKARTA CITY USES AN
EXPONENTIAL SMOOTHING-STATE SPACE
MODEL*,” *Int. J. Econ. Bus. Account. Res.*, vol.
6, no. 2, pp. 1144–1152, Jun. 2022, doi:
10.29040/IJEBAR.V6I2.4853.
- [13] M. As *et al.*, “*Forecasting Performance Base on
Exponential Smoothing with SES and ETS
Model for Gold Price Forecasting*,” *JATISI
(Jurnal Tek. Inform. dan Sist. Informasi)*, vol.
11, no. 1, pp. 267–274, Mar. 2024, doi:
10.35957/JATISI.V11I1.3804.
- [14] W. Rahmalina, “*Pemodelan Seasonal
Autoregressive Integrated Moving Average
Untuk Memprediksi Jumlah Kasus Covid-19 di
Padang*,” *J. Mat. Integr.*, vol. 17, no. 1, p. 23,
Aug. 2021, doi: 10.24198/jmi.v17.n1.32024.23-
31.
- [15] Yakin, “*Bangladesh Weather Dataset (1901 -
2023)*.” Accessed: Jan. 08, 2025. [Online].
Available:
[https://www.kaggle.com/datasets/yakinrubaiait/
bangladesh-weather-dataset](https://www.kaggle.com/datasets/yakinrubaiait/bangladesh-weather-dataset)