

Rainfall Forecasting Using Seasonal Autoregressive Integrated Moving Average Method And Decomposition Method in Pangkalan Bun

Grisella Estefania R¹, Agustini Tripena², Triyani³

^{1,2,3}Universitas Jenderal Soedirman, Indonesia

*Corresponding Author: agustinitripena@gmail.com

Article History:

Received: 03 Mei 2025

Revised: 25 Mei 2025

Accepted: 28 Mei 2025

Keywords: forecasting, rainfall, SARIMA, decomposition.

Abstract: : Flash floods and droughts that often occur in Pangkalan Bun, Central Kalimantan are clear evidence of the impact of climate change. Global climate change has caused an increase in the frequency and intensity of extreme weather, including significant fluctuations in rainfall. The purpose of this study is to create the best model for predicting rainfall using the SARIMA method and the Decomposition method in the case of data in Pangkalan Bun. The data used are rainfall data in Pangkalan Bun as many as 180 observations (January 2009 - December 2023). Rainfall data is divided into training data and testing data. Training data is used to determine the forecasting model using the SARIMA method and the Decomposition method. The results, the SARIMA (1,0,1) (1,0,1)¹² model gives the lowest MAPE value compared to other decomposition models, namely a MAPE value of 26.05%. The MAPE value indicates that the SARIMA (1,0,1) (1,0,1)¹² model is suitable for use in predicting future data. The forecast results show that the highest rainfall will occur in November 2024 at 263.97 mm and the lowest in August 2024 at 106.26 mm.

PENDAHULUAN

Perubahan iklim global telah menjadi tantangan besar bagi berbagai negara, termasuk Indonesia. Salah satu dampak yang paling terasa adalah fluktuasi curah hujan yang ekstrem, yang dapat menyebabkan bencana alam seperti banjir dan kekeringan. Curah hujan merupakan salah satu faktor penting dalam menentukan kondisi iklim suatu wilayah, dan ketidakpastiannya dapat berdampak signifikan pada sektor pertanian, ekonomi, dan kehidupan masyarakat. Oleh karena itu, peramalan curah hujan yang akurat sangat diperlukan untuk membantu pemerintah dan masyarakat dalam mengambil langkah-langkah untuk meminimalkan dampak tersebut.

Curah hujan di Indonesia, khususnya di wilayah tropis seperti Pangkalan Bun, memiliki pola musiman yang kuat. Namun, pola ini seringkali mengalami fluktuasi yang tidak menentu, sehingga memerlukan metode peramalan yang mampu menangkap karakteristik musiman tersebut. Beberapa faktor yang memengaruhi curah hujan antara lain suhu, tekanan udara,

kelembapan, kecepatan angin, dan jumlah lapisan awan. Selain itu, perubahan iklim global juga telah menyebabkan peningkatan frekuensi dan intensitas cuaca ekstrem, seperti hujan lebat dan kemarau panjang.

Berdasarkan data dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), Pangkalan Bun merupakan salah satu wilayah di Indonesia yang sering mengalami fluktuasi curah hujan yang ekstrem. Kejadian banjir bandang pada tahun 2022 menjadi salah satu contoh dampak negatif dari curah hujan yang tidak menentu, yang menyebabkan kerugian besar bagi pemerintah daerah dan masyarakat setempat. Oleh karena itu, penting untuk memiliki informasi yang akurat mengenai prakiraan curah hujan agar dapat meminimalkan dampak buruk dari perubahan cuaca ekstrem.

Salah satu metode yang dapat digunakan untuk meramalkan curah hujan adalah dengan menggunakan pendekatan analisis deret waktu. Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA) dan metode dekomposisi merupakan dua pendekatan yang sering digunakan dalam peramalan data deret waktu, terutama untuk data yang memiliki pola musiman. SARIMA secara khusus dirancang untuk menangani data dengan komponen musiman, sementara metode dekomposisi memisahkan data menjadi komponen-komponen seperti tren, musiman, siklus, dan acak.

Penelitian terdahulu telah menunjukkan efektivitas kedua metode ini dalam berbagai konteks. Misalnya, penelitian oleh Yuni dkk. (2015) menggunakan metode dekomposisi untuk meramalkan jumlah pengunjung perpustakaan, dengan hasil model dekomposisi multiplikatif adalah model terbaik dalam meramalkan jumlah pengunjung. Sementara itu, penelitian oleh Fitriastutik dan Anityasari (2020) membandingkan metode SARIMA dan dekomposisi dalam meramalkan timbulan sampah, dan menemukan bahwa model SARIMA menghasilkan *Mean Absolute Percentage Error* (MAPE) yang lebih kecil dibandingkan model dekomposisi. Selain itu, penelitian oleh Sajidul dkk. (2023) menggunakan metode SARIMA untuk meramalkan curah hujan di Kota Bandung, dengan hasil model SARIMA cukup signifikan untuk meramalkan curah hujan di Kota Bandung.

Berdasarkan penelitian-penelitian terdahulu, penelitian ini akan membandingkan kinerja model SARIMA dan metode dekomposisi dalam meramalkan curah hujan di Pangkalan Bun. Data curah hujan bulanan dari tahun 2009 hingga 2021 (data *training*) akan digunakan untuk membangun model peramalan dan data dari tahun 2022 hingga 2023 (data *testing*) akan digunakan untuk evaluasi model. Selanjutnya, model yang dihasilkan akan dievaluasi menggunakan MAPE untuk menentukan model terbaik. Hasil penelitian ini diharapkan dapat memberikan informasi yang akurat mengenai prakiraan curah hujan di Pangkalan Bun, sehingga dapat digunakan sebagai acuan dalam membuat kebijakan untuk meminimalkan dampak buruk akibat cuaca ekstrem.

LANDASAN TEORI

1. Curah Hujan

Curah hujan adalah jumlah air hujan yang turun pada suatu wilayah tertentu dalam jangka waktu tertentu. Besar kecilnya curah hujan juga bergantung pada waktu tertentu atau musim tertentu. Alat yang digunakan untuk mengukur curah hujan adalah fluviometer (Sujalu dkk., 2022:163). Alat ini merupakan alat yang terdiri dari corong dan tabung penampung (Aldrian dkk., 2011:73).

Menurut Susilowati dan Sadad (2015), curah hujan harian adalah jumlah hujan yang terukur dalam satu hari (selama 24 jam) pada stasiun pengamatan curah hujan. Curah hujan bulanan adalah jumlah curah hujan harian yang tercatat selama satu bulan pengamatan pada

stasiun curah hujan tertentu. Menurut BMKG (2024), jumlah curah hujan bulanan dibagi menjadi empat kelompok, antara lain: curah hujan rendah, curah hujan menengah, curah hujan tinggi, curah hujan sangat tinggi.

2. Analisis Deret Waktu

Analisis deret waktu adalah salah satu prosedur statistika yang dilakukan untuk meramalkan struktur probabilitas keadaan yang akan datang dalam rangka pengambilan keputusan (Aswi dan Sukarna, 2006: 5). Prosedur statistika yang dimaksud mencakup analisis terhadap pengamatan, pencatatan, dan penyusunan kejadian dari waktu ke waktu. Data deret waktu merupakan serangkaian data pengamatan yang berurutan berdasarkan indeks dengan interval waktu yang tetap.

Menurut Makridakis dkk. (1992: 10-11), pola deret waktu dapat dibedakan menjadi empat jenis, yakni pola data horizontal, pola data musiman, pola data siklis, pola data tren

3. Stasioneritas Deret Waktu

Stasioneritas deret waktu merupakan asumsi dasar bahwa proses suatu deret pengamatan tidak berubah seiring dengan adanya perubahan waktu (Montgomery, 2008: 25). Deret waktu stasioner berarti deret yang relatif tidak terjadi kenaikan atau penurunan nilai secara tajam pada data. Menurut Aswi dan Sukarna (2006:7), kondisi stasioner terbagi menjadi dua, yaitu:

a. Stasioner dalam rata-rata

Dalam deret waktu, data dapat dikatakan stasioner dalam rata-rata apabila fluktuasi data berada di sekitar nilai rata-rata yang konstan, tidak bergantung pada waktu dan variansi dari fluktuasi tersebut, Pengujian stasioneritas dalam rata-rata dari suatu data deret waktu dapat dilakukan dengan dua cara yaitu dengan menganalisis plot ACF atau dengan uji statistik *augmented* Dickey-Fuller (ADF). Untuk deret waktu yang stasioner dalam rata-rata, ACF akan turun menuju nol relatif cepat, sedangkan ACF dari data yang tidak stasioner menurun perlahan dan ditandai dengan *lag* pertama yang seringkali besar dan bernilai positif (Hyndman dan Athanasopoulos, 2018:293).

b. Stasioner dalam variansi

Data dapat dikatakan stasioner dalam variansi apabila fluktuasi data dari waktu ke waktu menunjukkan bahwa data bersifat konstan atau tidak berubah secara signifikan. Secara visual, hal ini dapat diamati melalui plot deret waktu, dengan melihat pola fluktuasi yang relatif stabil. Untuk pengujian yang lebih akurat, dapat digunakan transformasi pangkat (*power transformation*) seperti transformasi Box-Cox. Secara umum rumus untuk transformasi Box-Cox dapat dituliskan sebagai berikut:

$$C_t = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & \text{jika } \lambda \neq 0, \\ \log Z_t & \text{jika } \lambda = 0, \end{cases} \quad (1)$$

Setelah mentransformasi data, selanjutnya dilakukan pemodelan dan peramalan menggunakan metode yang sesuai dengan pola datanya. Hasil ramalan yang diperoleh, perlu ditransformasikan kembali skala aslinya. Hal ini bertujuan agar hasil ramalan dapat diinterpretasikan dan dibandingkan dengan data historis dengan lebih mudah. Menurut Hyndman dan Athanasopoulos (2018:66) invers transformasi Box-Cox diberikan oleh rumus berikut:

$$Z_t = \begin{cases} ((\lambda C_t + 1)^{1/\lambda}) & \text{jika } \lambda \neq 0, \\ \exp(C_t) & \text{jika } \lambda = 0. \end{cases} \quad (2)$$

4. Differencing

Differencing (pembedaan) digunakan untuk mengatasi data yang tidak stasioner dalam rata-

rata (Makridakis, 1992:383). Proses ini bertujuan untuk menghilangkan pola perubahan rata-rata dalam data sehingga data menjadi stasioner dalam rata-rata. Proses *differencing* dilakukan setelah data stasioner dalam variansi dengan cara dengan mengurangi nilai data saat ini dengan nilai data sebelumnya. Secara matematis operasi *differencing* orde pertama dapat ditulis sebagai berikut.

$$\Delta Z_t = Z_t - Z_{t-1}, \quad (3)$$

Menurut Montgomery dkk.(2008:36), *differencing* juga dapat dinyatakan dalam bentuk operator *backshift* yang dilambangkan dengan notasi B . Operator *backshift* digunakan dalam *differencing* untuk menggeser mundur waktu sebanyak satu satuan waktu untuk membentuk data yang baru. Dengan kata lain, $BZ_t = Z_{t-1}$. Secara umum, notasi B dapat ditulis menjadi $B^d Z_t = Z_{t-d}$ untuk bilangan positif d . Proses *differencing* orde ke- d dari Z_t dapat ditulis sebagai berikut:

$$\Delta^d Z_t = (1 - B)^d Z_t,$$

5. *Autocorrelation Function*

Menurut Cryer dan Chan (2008:109-111) *Autocorrelation Function* (ACF) atau fungsi autokorelasi adalah fungsi yang merepresentasikan hubungan (korelasi) antara nilai pengamatan waktu ke- t (Z_t) dengan pengamatan pada waktu-waktu sebelumnya ($Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$). Fungsi ACF pada *lag* ke- k untuk suatu deret waktu Z_1, Z_2, \dots, Z_n dapat ditulis sebagai berikut:

$$\begin{aligned} r_k &= \text{corr}(Z_t, Z_{t-k}) \\ &= \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{n-k} (Z_t - \bar{Z})^2} \end{aligned} \quad (4)$$

Untuk memeriksa apakah autokorelasi r_k signifikan, dapat digunakan rumus kesalahan baku (*standard errors*) dari r_k , yakni.

$$se_{r_k} = \sqrt{\frac{1 + \sum_{j=1}^{k-1} r_j^2}{n}},$$

statistik uji t untuk $r_k = 0$ atau $r_k \neq 0$ sebagai berikut:

$$t_{r_k} = \frac{r_k}{se_{r_k}}.$$

Menurut Aswi dan Sukarna (2006:13), diagram ACF dapat digunakan untuk mengidentifikasi kestasioneran data. Jika diagram ACF cenderung turun lambat, maka dapat disimpulkan data belum stasioner dalam rata-rata.

6. *Partial Autocorrelation Function*

Menurut Cryer dan Chan (2008: 112-116) *Partial Autocorrelation Function* (PACF) atau fungsi autokorelasi parsial merupakan fungsi yang merepresentasikan hubungan parsial antara pengamatan waktu ke- t dinotasikan Z_t dengan pengamatan pada waktu-waktu sebelumnya $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$. Autokorelasi parsial pada *lag* ke- k dinyatakan sebagai korelasi antara Z_t dan Z_{t+k} setelah dihilangkannya efek dari variabel-variabel $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$. Menurut Aswi dan Sukarna (2006:11), rumus PACF adalah sebagai berikut

$$\phi_{kk} = \text{corr}(Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}), \quad (5)$$

Persamaan 5 dapat ditentukan melalui persamaan Yule-Walker, yakni:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j},$$

7. *ARIMA Box-Jenkins*

ARIMA Box-Jenkins merupakan integrasi dari beberapa model deret waktu yang sudah ada dan telah dikembangkan oleh George Box dan Gwilym Jenkins. Nama mereka sering disinonimkan dengan proses ARIMA yang diterapkan untuk analisis deret waktu, peramalan, dan

Suatu deret waktu dikatakan mengikuti model *Autoregressive Moving Average* (ARMA), jika proses tersebut memiliki model yang sebagian merupakan proses *Autoregressive* dan sebagian yang lain merupakan proses *Moving Average*. Proses ARMA (p, q) dapat dinyatakan secara sistematis dalam bentuk persamaan berikut (Cryer dan Chan, 2008:77):

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}. \quad (11)$$

Pada persamaan (11), Z_t merupakan nilai pengamatan pada waktu ke- t , yang dipengaruhi secara linier oleh nilai-nilai masa lalu Z_{t-1}, \dots, Z_{t-p} serta nilai-nilai *error* masa lalu $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$. Model ini dikenal sebagai proses ARMA dengan orde p untuk komponen *autoregressive* dan orde q untuk komponen *moving average*.

8. Proses *Autoregressive Integrated Moving Average* (ARIMA)

Suatu deret waktu Z_t dikatakan mengikuti model *Autoregressive Integrated Moving Average* (ARIMA) jika setelah dilakukan *differencing* sebanyak orde d , deret tersebut menjadi stasioner dan dapat dimodelkan sebagai proses ARMA. Artinya, deret waktu asli dapat diubah menjadi stasioner melalui proses *differencing*, kemudian dimodelkan menggunakan kombinasi model AR dan MA. Dalam bentuk operator *backshift* model ARIMA dapat ditulis sebagai berikut:

$$\phi(B)(1-B)^d Z_t = \theta(B)\varepsilon_t, \quad (12)$$

9. Model *Seasonal Autoregressive Intergered Moving Average* (SARIMA)

Model *seasonal* ARIMA merupakan bentuk khusus dari model ARIMA jika terdapat unsur musiman yang jelas pada hasil observasi Z_t . Hal ini berarti data memiliki pola berulang-ulang dalam interval waktu yang tetap atau musiman. Pola musiman ini dapat terlihat tidak hanya melalui grafik data, tetapi juga dapat dilihat melalui grafik ACF dan PACF. Untuk mengatasi ketidakstasioneran data akibat unsur musiman, dapat dilakukan proses *differencing* musiman dengan periode musiman s (Cryer dan Chan, 2008:230-233).

Proses *differencing* musiman dari deret waktu Z_t dapat dituliskan sebagai

$$X_t = (1 - B^s)Z_t, \quad (13)$$

Model *seasonal* MA (Q) dengan musiman sepanjang s dinyatakan oleh (Cryer dan Chan, 2008:229):

$$Z_t = \varepsilon_t - \theta_1 \varepsilon_{t-s} - \theta_2 \varepsilon_{t-2s} - \dots - \theta_Q \varepsilon_{t-Qs}, \quad (14)$$

dengan θ : koefisien *moving average* musiman orde ke Q . Model MA musiman dapat ditulis dalam bentuk operator *backshift*, yakni.

$$Z_t = (1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_Q B^{Qs}) \varepsilon_t, \quad (15)$$

$$= \theta_s(B) \varepsilon_t.$$

Sedangkan, untuk model *seasonal* AR (P) dengan musiman sepanjang s dapat dinyatakan oleh (Cryer dan Chan, 2008:229):

$$Z_t = \phi_1 Z_{t-s} + \phi_2 Z_{t-2s} + \dots + \phi_P Z_{t-P} + \varepsilon_t, \quad (16)$$

dengan ϕ : koefisien *autoregressive* musiman orde ke P . Model AR musiman dapat ditulis dalam bentuk operator *backshift* yakni,

$$Z_t - \phi_1 Z_{t-s} - \phi_2 Z_{t-2s} - \dots - \phi_P Z_{t-P} = \varepsilon_t,$$

$$(1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_P B^{Ps}) Z_t = \varepsilon_t,$$

$$\phi_s(B) Z_t = \varepsilon_t. \quad (17)$$

Selanjutnya jika suatu hasil observasi Z_t mengikuti proses yang dibentuk oleh gabungan antara model ARIMA (p, d, q) dan model SARIMA (P, D, Q), maka modelnya dapat dituliskan dalam bentuk operator *backshift* sebagai berikut:

$$\phi(B)\Phi_s(B)\nabla^d\nabla^D Z_t = \theta(B)\Theta_s(B)\varepsilon_t, \quad (18)$$

10. Identifikasi orde AR dan MA

Proses identifikasi orde AR dan MA merupakan langkah yang penting dalam pemodelan ARIMA. Orde-orde ini berperan dalam menentukan struktur dan kompleksitas model ARIMA yang digunakan. Untuk mengidentifikasi orde AR dan MA dapat dilakukan dengan menganalisis pola pada grafik ACF dan PACF. Tabel 1 berikut mengenai indentifikasi orde AR dan MA berdasarkan plot ACF dan PACF (Aswi dan Sukarna, 2006:79).

Tabel 1 Model ARIMA Box-Jenkins berdasarkan ACF dan PACF

No	Model	ACF	PACF
1	AR (p)	Menurun secara perlahan setelah $lag\ p$ (<i>dies down</i>)	Terputus setelah $lag\ p$ (<i>cut off</i>)
2	MA (q)	Terputus setelah $lag\ q$ (<i>cut off</i>)	Menurun secara perlahan setelah $lag\ q$ (<i>dies down</i>)
3	ARMA (p, q)	Menurun secara perlahan setelah $lag\ p, q$ (<i>dies down</i>)	Menurun secara perlahan setelah $lag\ p, q$ (<i>dies down</i>)

11. Proses *White Noise*

Suatu model deret waktu yang baik akan memiliki sifat *white noise*, yaitu memenuhi asumsi *error* yang bersifat acak dan berdistribusi normal (Cryer dan Chan, 2008:175-184).

12. *Error Bersifat Acak*

Keacakan *error* model dapat diperiksa dengan memperhatikan fungsi autokorelasi dari barisan *error* e_1, e_2, \dots, e_n dengan $e_t = F_t - Z_t$. *Error* model dikatakan acak apabila tidak terdapat autokorelasi yang signifikan untuk setiap lag yang ditentukan. Keacakan *error* berdasarkan data residual dari suatu model dapat diuji menggunakan uji statistik Ljung-Box Q dengan hipotesis sebagai berikut:

$$H_0: r_1 = r_2 = \dots = r_k = 0 \text{ (error bersifat acak)}$$

$$H_1: \exists r_j \neq 0 \text{ (error tidak bersifat acak)}$$

Statistik uji yang digunakan:

$$Q^* = n(n+2) \sum_{k=1}^n \frac{r_k^2}{n-k}, \quad (19)$$

13. Metode Dekomposisi

Menurut Makridakis (1992:122-127) suatu pendekatan pada analisis data deret berkala meliputi upaya untuk mengidentifikasi komponen-komponen yang mempengaruhi setiap nilai pada sebuah data deret berkala. Prosedur identifikasi ini disebut dekomposisi. Setiap komponen diidentifikasi secara terpisah dan proyeksi setiap komponen ini kemudian digabung untuk menghasilkan ramalan nilai-nilai masa mendatang dari data deret berkala tersebut.

Metode dekomposisi biasanya mencoba memisahkan ketiga komponen dari pola dasar yaitu musiman, tren dan siklis. Perbedaan antara musiman dan siklis adalah bahwa musiman berulang dengan sendirinya pada interval waktu yang tetap, sedangkan faktor siklis mempunyai jangka waktu yang lebih lama dan panjangnya berbeda dari siklis satu dengan yang lainnya.

Metode dekomposisi mengasumsikan bahwa data tersusun sebagai berikut (Makridakis dkk, 1992:123)

$$\begin{aligned} \text{data} &= \text{pola} + \text{kesalahan} \\ &= f(\text{trend}, \text{siklus}, \text{musiman}) + \text{kesalahan}. \end{aligned}$$

Jadi selain komponen pola, terdapat pula unsur kesalahan yang acak. Empat komponen dalam analisis deret berkala adalah sebagai berikut (Makridakis dkk, 1992:123):

1. Komponen tren, merupakan komponen jangka panjang yang mendasari pertumbuhan atau penurunan dalam suatu data deret berkala.
2. Komponen musiman, menggambarkan pola perubahan yang berulang secara teratur dari waktu ke waktu.
3. Komponen siklis, merupakan fluktuasi gelombang yang mempengaruhi keadaan selama lebih dari semusim.
4. Komponen *random* (kesalahan), merupakan komponen tak beraturan yang terbentuk dari fluktuasi yang disebabkan oleh peristiwa tak terduga.

Metode dekomposisi merupakan salah satu metode peramalan tertua. Terdapat beberapa pendekatan alternatif untuk mendekomposisi suatu deret berkala, yang semuanya bertujuan untuk memisahkan komponen data deret berkala secara teliti dengan kesalahan terkecil. Konsep dasar dalam pemisahan tersebut bersifat empiris dan tetap, pertama memisahkan musiman, lalu tren, dan terakhir siklus.

Menurut Makridakis (1992:124) secara umum penulisan matematis dari model dekomposisi adalah:

$$Z_t = f(I_t, T_t, S_t, E_t), \quad (20)$$

Bentuk fungsional yang pasti dari persamaan (20) bergantung pada metode dekomposisi yang digunakan diantaranya yakni metode dekomposisi rata-rata sederhana atau model dekomposisi aditif. Model dekomposisi aditif adalah model yang terbentuk dengan menambahkan komponen musiman, tren, siklis dan komponen *random*. Persamaan (21) berikut mengasumsikan model dekomposisi aditif.

$$Z_t = I_t + T_t + C_t + E_t. \quad (21)$$

Metode dekomposisi *rasio-trend* atau model dekomposisi multiplikatif. Model dekomposisi multiplikatif adalah model yang terbentuk dengan mengalikan komponen musiman, tren, siklis dan komponen *random*. Model dekomposisi multiplikatif baik digunakan untuk data yang memiliki variasi musim tidak menentu. Persamaan (22) berikut mengasumsikan model dekomposisi multiplikatif:

$$Z_t = I_t \times T_t \times C_t \times E_t. \quad (22)$$

Sebelum dilakukan pengolahan data untuk peramalan periode selanjutnya, perlu diidentifikasi terlebih dahulu model yang akan digunakan dengan melihat plot data deret waktu. Adapun langkah-langkah pengolahan data dengan menggunakan metode dekomposisi adalah sebagai berikut.

1. Memisahkan indeks musiman dengan cara membagi data aktual dengan indeks musiman. Indeks musiman dapat dihitung dengan menggunakan metode rata-rata sederhana. (Suharyadi dan Purwanto, 2003: 167-191)

$$\frac{Z_t}{I_t} = T_t S_t E_t, \quad (23)$$

$$\text{dengan, } I_t = \text{jumlah musim} \times \frac{\text{Rata-rata per musim}}{\text{Rata-rata total}}. \quad (24)$$

Memisahkan komponen tren dari data yang sudah dihilangkan komponen musimannya dengan cara membagikan data yang sudah dipisahkan komponen musimannya dengan nilai tren. Nilai tren dapat dihitung dengan analisis tren. (Suharyadi dan Purwanto, 2003: 167-191).

$$\frac{T_t S_t E_t}{T_t} = S_t E_t, \quad (25)$$

Menurut Makridakris (1992:127-128) rumus untuk menghitung nilai nilai tren linier sebagai berikut:

$$T_t = a + bt, \quad (26)$$

$$b = \frac{n \sum (t \cdot Z_t) - \sum t \cdot \sum Z_t}{n \sum t^2 - (\sum t)^2}, \quad (27)$$

$$a = \frac{\sum Z_t}{n} - b \frac{\sum t}{n}, \quad (28)$$

dengan, a adalah intersep dan b adalah kemiringan.

- Memisahkan hasil data yang sudah dihilangkan komponen musiman dan tren untuk memperoleh komponen siklis. Nilai siklis dapat dihitung menggunakan metode rata-rata bergerak. (Suharyadi dan Purwanto, 2003: 167-191)

$$S_t = \frac{Z_t + Z_{t+1} + Z_{t+2} + \dots + Z_{t+n-1}}{n}. \quad (29)$$

- Memisahkan komponen musiman, tren, siklis dari data asli untuk mendapatkan unsur random yang ada. (Suharyadi dan Purwanto, 2003: 167-191)

$$\frac{S_t E_t}{S_t} = E_t. \quad (30)$$

- Melakukan peramalan dengan masing-masing komponen menggunakan model yang telah terpilih. (Suharyadi dan Purwanto, 2003: 167-191)

14. Ukuran Hasil Peramalan

Model peramalan yang baik dicirikan oleh tingkat keakuratan yang tinggi. Untuk mengukur tingkat keakuratan ini, ada beberapa alat ukur yang dapat digunakan untuk mengevaluasi hasil peramalan model terhadap data observasi. Beberapa alat ukur yang dapat digunakan antara lain (Makridakis, 1992:40-44).

a. Akaike Information Criterion (AIC)

AIC digunakan untuk memilih model terbaik dari beberapa model yang mungkin seperti saat menganalisis dengan menggunakan metode SARIMA. AIC mengukur informasi relatif yang hilang ketika suatu model digunakan untuk merepresentasikan proses yang menghasilkan data. Secara sistematis untuk menghitung nilai AIC dapat dinyatakan sebagai berikut:

$$AIC = n \log \left(\frac{SSE}{n} \right) + 2(k + 2), \quad (31)$$

AIC memberikan penalti untuk model yang lebih kompleks, sehingga mencegah *overfitting*. Dalam konteks model SARIMA, AIC digunakan untuk memilih orde $(p, d, q)(P, D, Q)^s$ yang optimal (Hyndman dan Athanasopoulos, 2018: 164).

b. Mean Absolute Percentage Error (MAPE)

MAPE mengukur rata-rata persentase nilai mutlak dari selisih antara nilai observasi dan nilai peramalan. MAPE mudah diinterpretasikan karena dinyatakan dalam persentase.

$$MAPE = \frac{1}{m} \sum_{t=1}^m \frac{|Z_t - F_t|}{Z_t} 100\%, \quad (32)$$

METODE PENELITIAN

Data yang digunakan adalah data sekunder yang diperoleh dari *website* BPS Kalimantan Tengah. Penelitian ini menggunakan data curah hujan Pangkalan Bun dari tahun 2009 hingga 2023. Kemudian data dibagi menjadi 2 dengan pembagian data *training* (2009-2021) dan data *testing* (2022-2023) untuk evaluasi kinerja model dengan nilai MAPE dan MSE.

Analisis data dilakukan dengan langkah-langkah berikut:

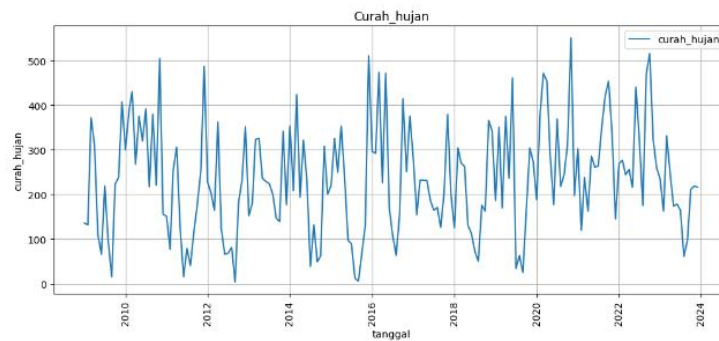
- Mendeskripsikan data dan menganalisis pola data curah hujan di Pangkalan Bun tahun 2009-2023;
- Membagi data curah hujan di Pangkalan Bun menjadi 2 bagian yaitu data *training* (2009-2021) dan data *testing* (2022-2023);
- Memodelkan data dengan metode SARIMA;

4. Memodelkan data dengan metode dekomposisi model aditif dan model multiplikatif;
5. Membandingkan model SARIMA dan dekomposisi multiplikatif berdasarkan nilai MAPE dan MSE;
6. Meramalkan curah hujan Pangkalan Bun tahun 2024 dengan model terpilih.

HASIL DAN PEMBAHASAN

1. Deskripsi dan Analisis Pola Data

Data curah hujan bulanan di Pangkalan Bun (2009–2023) menunjukkan pola musiman yang kuat, dengan puncak curah hujan tertinggi pada Januari dan penurunan signifikan pada Juni–Agustus. Kemudian statistika deskriptif menunjukkan curah hujan rata-rata curah hujan sebesar 233,58 mm/bulan, standar deviasi sebesar 122,07mm menunjukkan variabilitas tinggi, dan rentang data sebesar 4–550,3 mm. Pola data curah hujan di Pangkalan Bun ditampilkan pada Gambar 1 berikut.



Gambar 1

Plot *time series* data

Statistika deskriptif data ditampilkan pada Tabel 1.

Tabel 1 Statistika deskriptif data

	Curah hujan
<i>count</i>	180
<i>mean</i>	233,585
<i>std</i>	122,066
<i>min</i>	4
25%	152,175
50%	225,150
75%	320,350
<i>max</i>	550,300

Curah hujan cenderung tinggi pada bulan November hingga Maret (musim hujan) dan rendah pada bulan Juni hingga September (musim kemarau).

1. Pembagian Data

Data dibagi menjadi dua bagian : Data *training*: Januari 2009 – Desember 2021 (156 bulan) dan Data *testing*: Januari 2022 – Desember 2023 (24 bulan)

Pembagian ini dilakukan untuk memastikan model dapat diuji keakuratannya menggunakan data yang belum pernah digunakan dalam pembuatan model.

2. Pemodelan data dengan metode SARIMA

a. Identifikasi Model

Data diuji stasioneritasnya dalam rata-rata dan variansi. Untuk menguji stasioneritas dalam rata-rata menggunakan Augmented Dickey-Fuller (ADF) dan transformasi Box-Cox untuk menguji stasioneritas dalam variansi. Hasil uji ADF ditampilkan pada Tabel 2 berikut.

Tabel 2 Hasil uji ADF

<i>Critical Values</i> : '5%':	-2,880622899711496
<i>ADF Statistic</i>	: -6,3779436306845465
<i>p-value</i>	: 2,259341345599163E-08

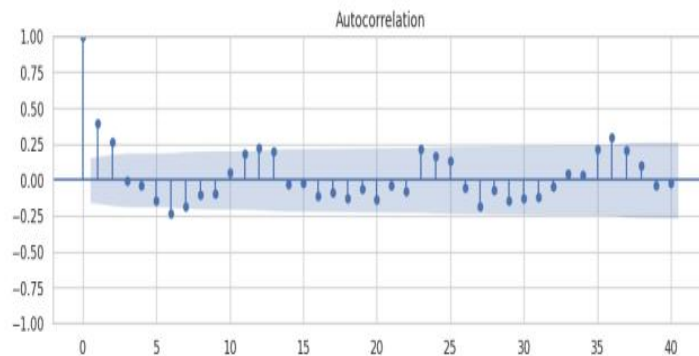
Hipotesis:

H_0 : data stasioner dalam rata-rata;

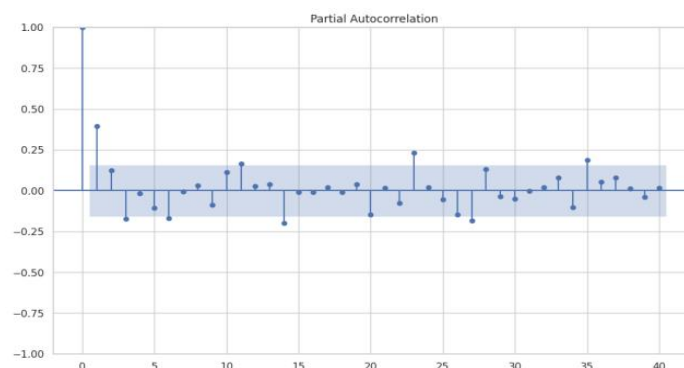
H_1 : data tidak stasioner dalam rata-rata.

Berdasarkan hasil uji ADF diperoleh bahwa nilai *p-value* sebesar $2,25 \times 10^{-8}$ lebih kecil dari $\alpha = 0,05$ sehingga H_0 diterima. Kesimpulannya adalah data stasioner dalam rata-rata artinya, data tidak perlu dilakukan *differencing*.

Transformasi Box-Cox dilakukan dengan *software Google Colaboratory*. Diperoleh hasil nilai $\lambda = 0,69$ sehingga data perlu ditransformasi karena nilai λ memiliki *rounded-value* (nilai pembulatan) tidak sama dengan 1. Begitupun sebaliknya, apabila nilai λ memiliki *rounded-value* sama dengan 1 maka data tidak perlu ditransformasi. Kemudian identifikasi orde AR, MA, dan komponen musiman dengan menganalisis plot ACF dan PACF. Plot ACF dan PACF ditampilkan pada Gambar 2 dan Gambar 3.



Gambar 2 Plot ACF data transformasi



Gambar 3 Plot PACF data transformasi

Berdasarkan Gambar 2 dan Gambar 3 didapatkan model SARIMA yang mungkin yaitu $SARIMA(\{0,1\},0,\{0,1,2\})(\{0,1\},0,\{0,1\})^{12}$. Model-model ini akan dibentuk dan diuji signifikansi parameter-parameter model. Model dengan semua parameternya memiliki *pvalue* $< 0,05$ merupakan model yang signifikan dan dapat dilanjutkan untuk uji diagnostik model.

b. Estimasi Parameter

Estimasi parameter model dilakukan dengan uji diagnostik pada model. Uji yang dilakukan yaitu uji asumsi *white noise* pada residual data. Asumsi *white noise* artinya data memiliki residual yang bersifat acak dan berdistribusi normal. Dari 18 model yang signifikan diperoleh 7 model SARIMA yang memenuhi asumsi *white noise* yaitu SARIMA(0,0,1)(1,0,0)¹², SARIMA(0,0,2)(1,0,0)¹², SARIMA(1,0,1)(0,0,0)¹², SARIMA(1,0,1)(0,0,1)¹², SARIMA(1,0,1)(1,0,0)¹², SARIMA(1,0,1)(1,0,1)¹², SARIMA(1,0,2)(0,0,0)¹².

c. Model SARIMA Terbaik

Hasil estimasi parameter, model SARIMA (1,0,1)(1,0,1)¹² adalah model dengan nilai AIC terkecil yaitu sebesar 1.420,69. Model SARIMA (1,0,1)(1,0,1)¹² juga memenuhi asumsi *white noise* pada uji diagnostik. Hasil uji diagnostik SARIMA (1,0,1)(1,0,1)¹² ditampilkan pada Tabel 3.

Tabel 3 Hasil uji diagnostik model SARIMA (1,0,1)(1,0,1)¹²

Uji yang digunakan	P-value
Ljung Box Q*	0,67
Shapiro-Wilk	0,739

Berdasarkan Tabel 3, pada uji Ljung-Box Q*: $p\text{-value} = 0,67 > 0,05$ artinya, residual model bersifat acak. Pada uji normalitas Shapiro-Wilk: $p\text{-value} = 0,739 > 0,05$ artinya residual bersifat normal. Dengan begitu residual model memenuhi asumsi *white noise* sehingga model SARIMA(1,0,1)(1,0,1)¹² dapat digunakan untuk meramalkan data masa depan. Selanjutnya, dengan mensubstitusi nilai parameter signifikannya yaitu AR(1): $\phi = 0,9389$; MA(1): $\theta = -0,5871$. Untuk orde musiman: SAR(12) $\Phi = 0,9983$; SMA(12) $\Theta = -0,9605$ pada persamaan SARIMA. Diperoleh persamaan untuk model SARIMA (1,0,1)(1,0,1)¹² sebagai berikut

$$Z_t = 0,9389Z_{t-1} + 0,9983Z_{t-12} - 0,9373Z_{t-13} + \varepsilon_t + 0,5871\varepsilon_{t-1} + 0,9605\varepsilon_{t-12} + 0,5639\varepsilon_{t-13} \quad (1)$$

d. Evaluasi dengan Data Testing

Evaluasi model dengan data testing menggunakan ukuran peramalan MAPE. Perhitungan peramalan dilakukan dengan mensubstitusi periode waktu pada persamaan (1) dan diperoleh nilai MAPE sebesar 26,05%. Nilai MAPE tersebut

3. Pemodelan Data dengan Metode Dekomposisi Multiplikatif

a. Pemisahan Komponen Musiman

Pemisahan komponen musiman dilakukan dengan menghitung indeks musiman setiap bulan. Indeks musiman dihitung dengan menggunakan metode rata-rata sederhana. Indeks musiman ditampilkan pada Tabel 4.

Tabel 4 Perhitungan indeks musiman

Bulan	Indeks Musiman ($I_t = \text{jumlah musim} \times \frac{\text{mean per musim}}{\text{mean total}}$)
Januari	0,9724
Februari	1,0392
Maret	1,3585
April	1,2777
Mei	1,0580
Juni	0,7975
Juli	0,7080

Agustus	0,5652
September	0,6221
Oktober	0,9273
November	1,3985
Desember	1,2757

b. Pemisahan Komponen Tren

Pemisahan komponen tren yaitu dengan menghitung rumus persamaan tren. Persamaan tren yang digunakan adalah persamaan tren linear dengan nilai $a = 193,661$ dan nilai $b = 0,471$ sehingga diperoleh persamaan (2)

$$Tt = 193,661 + 0,471t \quad (2)$$

Dengan menggunakan persamaan (2) kemudian dihitung nilai tren untuk periode Januari 2022 – Desember 2023 ($t = 157-180$) yang ditampilkan pada Tabel 5.

Tabel 5 Perhitungan nilai tren

t	Tt	t	Tt
157	267,6124	169	273,2648
158	268,0835	170	273,7358
159	268,5545	171	274,2068
160	269,0255	172	274,6779
161	269,4966	173	275,1489
162	269,9676	174	275,6199
163	270,4386	175	276,0909
164	270,9096	176	276,562
165	271,3807	177	277,033
166	271,8517	178	277,504
167	272,3227	179	277,9751
168	272,7937	180	278,4461

Hasil Tabel 5 akan digunakan untuk meramalkan data curah hujan pada tahun tersebut. Kemudian dapat dilakukan evaluasi menggunakan data *testing*.

c. Evaluasi dengan Data *Testing*

Hasil peramalan tahun 2022-2023 dengan model dekomposisi aditif dan model dekomposisi multiplikatif ditampilkan pada Tabel 6 dan Tabel 7.

Tabel 6 Hasil ramalan tahun 2022-2023 model aditif

Bulan	Ramalan ($Ft = It \times Tt$)	Bulan	Ramalan ($Ft = It \times Tt$)
Jan'22	268,58	Jan'23	274,24
Feb'22	269,12	Feb'23	274,77
Mar'22	269,91	Mar'23	275,57
Apr'22	270,30	Apr'23	275,96
Mei'22	270,55	Mei'23	276,21
Jun'22	270,77	Jun'23	276,42
Jul'22	271,15	Jul'23	276,80
Ags'22	271,47	Ags'23	277,13
Sep'22	272,00	Sep'23	277,66
Okt'22	272,78	Okt'23	278,43
Nov'22	273,72	Nov'23	279,37
Des'22	274,07	Des'23	279,72

Tabel 7 Hasil ramalan tahun 2022-2023 model aditif

Bulan	Ramalan (Ft = It×Tt)	Bulan	Ramalan (Ft = It×Tt)
Jan'22	267,6124	Jan'23	273,2648
Feb'22	268,0835	Feb'23	273,7358
Mar'22	268,5545	Mar'23	274,2068
Apr'22	269,0255	Apr'23	274,6779
Mei'22	269,4966	Mei'23	275,1489
Jun'22	269,9676	Jun'23	275,6199
Jul'22	270,4386	Jul'23	276,0909
Ags'22	270,9096	Ags'23	276,562
Sep'22	271,3807	Sep'23	277,033
Okt'22	271,8517	Okt'23	277,504
Nov'22	272,3227	Nov'23	277,9751
Des'22	272,7937	Des'23	278,4461

Perhitungan nilai MAPE menggunakan data *testing*, adapun perhitungan nilai untuk model dekomposisi aditif dan model dekomposisi multiplikatif adalah sebagai berikut. Dengan mensubstitusikan $m = 24$, nilai MAPE yang diperoleh dari data *testing* dan hasil peramalan adalah:

Nilai MAPE model dekomposisi aditif,

$$MAPE = \frac{1}{24} \left(\frac{|268,9 - 268,58|}{268,9} + \dots + \frac{|216,3 - 279,72|}{216,3} \right) 100\% = \frac{1}{24} (11,89139) 100\% = 49,55\%.$$

selanjutnya, menghitung nilai MAPE model dekomposisi multiplikatif,

$$MAPE = \frac{1}{24} \left(\frac{|268,9 - 260,2|}{268,9} + \dots + \frac{|216,3 - 355,2|}{216,3} \right) 100\% = \frac{1}{24} (10,4070) 100\% = 43,36\%.$$

Berdasarkan hasil perhitungan, dapat disimpulkan bahwa peramalan menggunakan model dekomposisi multiplikatif diperoleh MAPE sebesar 43,36%. Nilai MAPE tersebut berada direntang 20-50%, ini berarti model layak digunakan untuk peramalan curah hujan di Pangkalan Bun.

4. Pemilihan Model Terbaik

Perbandingan antara model SARIMA dan dekomposisi multiplikatif ditampilkan pada Tabel 5.

Tabel 5 Perbandingan Kinerja Model

Model	MAPE (%)
SARIMA (1,0,1)(1,0,1) ¹²	26,05
Dekomposisi Aditif	49,55
Dekomposisi Multiplikatif	43,36

Berdasarkan Tabel 5 nilai MAPE model SARIMA (1,0,1)(1,0,1)¹² lebih kecil daripada model dekomposisi aditif dan model dekomposisi multiplikatif. Hal ini menunjukkan model SARIMA lebih akurat jika digunakan untuk meramalkan curah hujan di Pangkalan Bun dibandingkan model dekomposisi aditif dan model dekomposisi multiplikatif. Perbedaan signifikan pada MAPE menunjukkan SARIMA lebih mampu menangkap pola musiman dan non-linear data.

5. Peramalan Curah Hujan Tahun 2024

Peramalan curah hujan di Pangkalan Bun tahun 2024 menggunakan model SARIMA (1,0,1)(1,0,1)¹². Peramalan dilakukan dengan mensubstitusikan periode waktu pada persamaan (1) sehingga diperoleh hasil *forecast* seperti pada Tabel 6

Tabel 6 Hasil *forecast*

Tanggal	Forecast	Forecast (skala asli)
Jan'24	53,93206	194,3358241
Feb'24	55,72184	203,4790815
Mar'24	66,73482	262,4871774
Apr'24	64,17266	248,3499091
Mei'24	55,85903	204,185244
Jun'24	45,03038	150,8515235
Jul'24	40,77082	131,283026
Ags'24	35,02586	106,263715
Sep'24	35,59892	108,6854929
Okt'24	49,9966	174,6941832
Nov'24	67,00242	263,9775925
Des'24	61,83199	235,6477622

Tabel 6 Hasil *forecast* tersebut juga ditransformasikan kembali ke skala asli (mm) menggunakan persamaan (1). Diperoleh hasil peramalan menggunakan model SARIMA (1,0,1)(1,0,1)¹² seperti pada tabel dengan: curah hujan tertinggi: November 2024 (263,98 mm) dan curah hujan terendah: Agustus 2024 (106,26 mm).

KESIMPULAN

Model terbaik untuk data curah hujan di Pangkalan Bun adalah model SARIMA (1,0,1)(1,0,1)¹² karena memiliki nilai AIC terkecil daripada model SARIMA lainnya yaitu sebesar 1.420,69 Adapun model SARIMA yang diperoleh adalah

$$Z_t = 0,9389Z_{t-1} + 0,9983Z_{t-12} - 0,9373Z_{t-13} + \varepsilon_t + 0,5871\varepsilon_{t-1} + 0,9605\varepsilon_{t-12} + 0,5639\varepsilon_{t-13}$$

Nilai MAPE sebesar 26,05%. Nilai MAPE tersebut berada di rentang 20-50% yang artinya model dikatakan layak untuk meramalkan data curah hujan di Pangkalan Bun. Diperoleh nilai ramalan pada satuan curah hujan (mm) yaitu untuk bulan Januari 2024 sebesar 194,33 mm, Februari 2024 sebesar 203,47 mm, Maret 2024 sebesar 262,48 mm, April 2024 sebesar 248,35 mm, Mei 2024 sebesar 204,18 mm, Juni 2024 sebesar 150,85 mm, Juli 2024 sebesar 131,28 mm, Agustus 2024 sebesar 106,26 mm, September 2024 sebesar 108,68 mm, Oktober 2024 sebesar 174,69 mm, November 2024 sebesar 263,98 mm dan Desember 2024 sebesar 235,65 mm.

DAFTAR REFERENSI

- Aldrian, E., Karmini, M., dan Budiman, B. (2011). *Adaptasi dan Mitigasi Perubahan Iklim di Indonesia*. Jakarta: Pusat Perubahan Iklim dan Kualitas Udara ,Kedeputan Bidang Klimatologi, Badan Meteorologi, Klimatologi, dan Geofisika.
- Aswi, dan Sukarna. (2006). *Analisis Deret Waktu: Teori dan Aplikasi*. Makassar: Andira Publisher.
- BMKG. (2024, Maret 12). *Istilah Iklim*. Retrieved from <https://staklim-ntb.bmkg.go.id/pemahaman-iklim>
- Cryer, J. D., dan Chan, K.-S. (2008). *Time Series Analysis*. Iowa: Springer.

-
- Ervina, M. S. (2018, Oktober). Peramalan Jumlah Penumpang Kereta Api di Indonesia dengan Resilient Back-Propagation (Rprop) Neutral Network. *Jurnal Matematika "MANTIK"*, 4(2), 90-99. doi:<https://doi.org/10.15642/mantik.2018.4.2.90-99>
- Fitriastutik, E., dan Anityasari, M. (2020). Forecasting Timbulan Sampah Kota Surabaya Menggunakan Time Series Analysis. *Jurnal Teknik*, 9(2), 348-354.
- Hyndman, R., dan Athanasopoulos, G. (2014). *Forecasting: Principles and Practice*. Melbourne: OTexts.
- Hyndman, R., dan Athanasopoulos, G. (2018). *Forecasting : Principles and Practice*. Melbourne: OTexts.
- Makridakis, S., dan Wheelwright, S. (1992). *Metode dan Aplikasi Peramalan*. Jakarta: Erlangga.
- Montgomery, D. C. (2008). *Introduction to Time Series Analysis and Forecasting*. New York: John Wiley And Son, Inc.
- Navianti, D. R., dan Widjajati, F. A. (2012). Penerapan fuzzy inference system pada prediksi curah hujan di Surabaya Utara. *Jurnal Sains dan Seni ITS*, 1(1), A23-A28.
- Nurlatifah, A., Hatmaja, R. B., dan Rakhman, A. A. (2023). Analisis Potensi Kejadian Curah Hujan Ekstrem di Masa Mendatang. *Jurnal Ilmu Lingkungan*, 21(4), 980-986.
- Sajidul Fajri, E. K. (2023). Pemodelan Curah Hujan Kota Bandung Menggunakan Seasonal Autoregressive Integrated Moving Average Pada Data Time Series dengan Bantuan Minitab. *Bandung Conference Series: Mathematics*, 7-17.
- Setiawan, A., Sari, R., dan Aprilia, R. (2024). Peramalan Hasil Panen Kelapa Sawit di PTPN IV Berangir Dengan Metode Dekomposisi. *Jurnal Pembelajaran dan Matematika Sigma*, 10(1), 82-92.
- Suharyadi, dan Purwanto, S. (2003). *Statistika untuk Ekonomi dan Keuangan Modern*. Jakarta Selatan: Salemba Empat.
- Sujalu, A. P., Pulihasih, A. Y., dan Biantary, M. P. (2022). *Instrumen Klimatologi dan Meteorologi*. Yogyakarta: Zahir Publishing.
- Susilowati, S., dan Sadad, I. (2015). Analisa karakteristik curah hujan di Kota Bandar Lampung. *Jurnal Konstruksia*, 7(1), 13-26.
- Wei, W. (2006). *Time Series Analysis, Addison Wesley*. Redwood City: CA.
- Yuni, S., Talakua, M. W., dan Lesnussa, Y. A. (2015). Peramalan jumlah pengunjung perpustakaan Universitas Pattimura Ambon menggunakan metode dekomposisi. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 9(1), 41-50.
-