

PREDIKSI RATA-RATA KELEMBAPAN MENGGUNAKAN METODE SARIMAX DENGAN RATA-RATA TEMPERATUR SEBAGAI VARIABEL EXOGENOUS

Predict Of Average Humadity Using The SARIMAX Method With Average Temperature As The Exogenous Variable

Ihsan Fathoni Amri^{1*}, Selvi Ana Windia Sari², Ailsha Syafa Kinanta³, M.Al Haris⁴,
Isnaeni Miftahul Sidqi⁵, Mochamad Fahmi Choirudin⁶

¹ Program Studi S1 Sains Data, Fakultas Sains dan Teknologi Pertanian
Universitas Muhammadiyah Semarang
Jl. Kedungmundu No.18, Semarang, 50273, Jawa Tengah, Indonesia
^{2,3,4,5,6} Program Studi S1 Statistika, Fakultas Sains dan Teknologi Pertanian
Universitas Muhammadiyah Semarang
Jl. Kedungmundu No.18, Semarang, 50273, Jawa Tengah, Indonesia

E-mail Correspondence Author: ihsanfathoni@unimus.ac.id

Abstrak

Rata-rata kelembapan di Indonesia memiliki variansi yang bergantung pada lokasi dan musim. Kelembapan juga bisa bervariasi sepanjang hari, dengan puncak kelembapan biasanya terjadi pada pagi hari dan menurun pada siang hari sebelum meningkat kembali pada malam hari. Penelitian ini bertujuan untuk memprediksi rata-rata kelembapan di Stasiun Klimatologi Jawa Tengah dengan menggunakan metode SARIMAX (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*). Metode SARIMAX dipilih karena memiliki kemampuan dalam menangani data time series yang memiliki komponen musiman dan melibatkan variabel *exogenous*. Rata-rata temperature digunakan sebagai variabel *exogenous* karena adanya korelasi yang signifikan antara rata-rata temperature dan rata-rata kelembapan. Data rata-rata kelembapan dan temperature diambil dari catatan harian pada periode yang digunakan dalam penelitian. Model SARIMAX kemudian dikembangkan dengan parameter yang dioptimalkan melalui proses iteratif untuk mencapai tingkat akurasi prediksi yang maksimal. Hasil penelitian menunjukkan bahwa model SARIMAX (1, 1, 1)(1, 1, 1)₄ dengan nilai AIC sebesar 323,89 dan nilai MAPE sebesar 2,863913 mampu memberikan prediksi yang cukup akurat terhadap rata-rata Kelembapan di Stasiun Klimatologi Jawa Tengah, dengan error prediksi paling rendah. Model ini dapat memprediksi rata-rata kelembapan selama 8 hari kedepan. Temuan ini dapat membantu dalam perencanaan dan pengelolaan berbagai sektor kegiatan di wilayah tersebut.

Kata Kunci: Kelembapan, Temperature, SARIMAX, Peramalan

Abstract

Average humidity in Indonesia varies depending on location and season. Humidity can also vary throughout the day, with peak humidity usually occurring in the morning and decreasing during the day before increasing again at night. This study aims to predict the average humidity at the Central Java Climatology Station using the SARIMAX (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*) method. The SARIMAX method was chosen because it has the ability to handle time series data that has a seasonal component and involves exogenous variables. Average temperature is

Research Article • Open Acces

Email: jurnalparameter@gmail.com

Homepage: <https://ojs3.unpatti.ac.id/index.php/parameter>

used as an exogenous variable because there is a significant correlation between average temperature and average humidity. The average humidity and temperature data were taken from daily records for the period used in the study. The SARIMAX model was then developed with optimized parameters through an iterative process to achieve maximum prediction accuracy. The results showed that the SARIMAX (1, 1, 1)(1, 1, 1)4 model with an AIC value of 323.89 and a MAPE value of 2.863913 was able to provide a fairly accurate prediction of the average humidity at the Central Java Climatology Station, with the lowest prediction error. This model can predict the average humidity for the next 8 days. These findings can help in planning and managing various sectors of activity in the region.

Keywords: Humidity, Temperature, SARIMAX, Forecasting



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

1. PENDAHULUAN

Pengukuran kelembapan di Stasiun Klimatologi Jawa Tengah memiliki implikasi signifikan dalam menentukan kualitas udara dan mengantisipasi potensi bencana alam seperti banjir dan kekeringan [1]. Cuaca adalah kondisi atmosfer di suatu wilayah tertentu dalam jangka waktu yang relatif singkat, yang mencakup kondisi seperti suhu, kelembapan, dan tekanan atmosfer sebagai komponen utamanya. Cuaca adalah fenomena alam yang menunjukkan kondisi fisik yang terjadi di suatu wilayah tertentu selama periode tertentu. Contoh fenomena cuaca meliputi angin kencang, hujan, dan panas yang menyengat [2]. Perubahan iklim dapat menyebabkan erosi, percepatan aliran air permukaan, serta terjadinya kekeringan dan kelebihan air sehingga menyebabkan banjir dan tanah longsor. Perubahan iklim juga menyebabkan hujan lebat dan banjir. Upaya mitigasi dan pencegahan risiko juga harus dilakukan [3].

Kelembapan didefinisikan sebagai jumlah uap air yang terkandung di udara atau atmosfer. Sedangkan hujan merupakan tetesan air atau kristal es yang jatuh atau muncul dari awan atau gugusan awan [4]. Temperature adalah satuan termodinamika yang mengukur energi alkinetik translasi rata-rata molekul pasir atau gas. Termometer digunakan untuk mengukur temperature. Temperature menunjukkan derajat pemanasan suatu benda. Semakin tinggi temperature suatu benda maka semakin panas pula temperature benda tersebut. Secara mikroskopis, temperature menunjukkan energi yang dimiliki suatu benda [5]. Pemantauan suhu dan kelembapan memegang peranan penting dalam berbagai aspek kehidupan sehari-hari. Di bidang pertanian, ruang server industri, pemantauan lingkungan, dan bahkan kehidupan rumah tangga, pemantauan kondisi temperature dan kelembapan memberikan informasi berharga untuk mengambil keputusan yang tepat [6].

Stasiun Iklim di Jawa Tengah merupakan lokasi pengumpulan dan analisis data cuaca, termasuk suhu. Metode SARIMAX (*Seasonal Autoregressive Integrated Moving Average Exogenous Regressors*) digunakan untuk memprediksi rata-rata kelembapan di stasiun tersebut, dengan temperature sebagai variabel *exogenous* yang mempengaruhi kelembapan udara. Stasiun Klimatologi Jawa Tengah sebagai pusat pengumpulan data iklim dan meteorologi memiliki peran penting dalam memantau dan memprediksi perubahan kelembapan di wilayahnya. Dalam beberapa tahun terakhir, kelembapan udara di Jawa Tengah telah menunjukkan pola fluktuatif yang kompleks, dengan variasi yang signifikan antara bulan dan tahun. Dengan demikian, peramalan kelembapan yang

menggunakan metode yang tepat dan akurat sangat diperlukan untuk memprediksi dan mengantisipasi perubahan kelembapan di masa mendatang [7].

Forecasting merupakan meramalkan suatu kejadian atau peristiwa di masa yang akan datang dengan melihat peristiwa di masa lampau [8]. Peramalan deret waktu adalah rencana pengamatan terhadap suatu variabel yang diperoleh setiap periode dan disusun dalam susunan yang sesuai dengan periode penyusunannya [9]. Ada beberapa metode untuk meramalkan data deret waktu, antara lain *Autoregressive Integrated Moving Average* (ARIMA). Namun jika data menunjukkan pola yang berulang pada suatu titik waktu tertentu, maka metode peramalan yang cocok *Seasonal Autoregressive Integrated Moving Average* (SARIMA) [10]. Kekurangan metode SARIMA adalah hanya dapat digunakan untuk prediksi data deret waktu singular dan tidak memperhatikan pengaruh variabel lain. Untuk mengatasi kekurangan dari model SARIMA, Dikembangkan model baru yaitu SARIMAX (*Seasonal Autoregressive Integrated Moving Average Exogenous*) [11].

Penelitian ini mengkaji bagaimana penggunaan metode SARIMAX untuk meramalkan kelembapan rata-rata di Stasiun Klimatologi Jawa Tengah dengan mempertimbangkan temperature sebagai variabel *exogenous* dan Tingkat akurasi yang dihasilkan dibandingkan metode lainnya.

2. METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang di ambil dari BMKG melalui laman https://dataonline.bmkg.go.id/data_iklim, data ini adalah data *time series* dengan periode waktu dari 1 April sampai 14 Juni 2024, dengan total data sebanyak 75 sampel.

2.2 Variabel penelitian

Variabel yang diteliti pada kasus ini yaitu rata-rata kelembapan dan rata-rata temperature. Berikut ini disajikan **Tabel 1.** variabel penelitian.

Notasi	Variabel
Y	Rata-rata kelembapan
X	Rata-rata temperatur

Sumber : bmkg.go.id

Berdasarkan **Tabel 1.** dapat dilihat bahwa rata-rata kelembapan sebagai variabel dependent atau variabel Y, dengan variabel rata-rata temperatur sebagai variabel independent atau variabel X.

2.3 Stasioneritas

Stasioneritas mengacu pada kondisi di mana rata-rata (*mean*) dan varians data tidak berubah seiring waktu, tetap konstan tanpa ada pertumbuhan atau penurunan. Data *stasioner* adalah data yang rata-ratanya tidak berubah seiring waktu, dengan fluktuasi yang terjadi hanya di sekitar nilai rata-rata dan varians yang tetap. Peneliti menganalisis pola pada plot data untuk menentukan apakah data tersebut stasioner atau tidak. Jika plot data menunjukkan konsistensi tanpa ada tren kenaikan atau penurunan, maka data tersebut dianggap stasioner [12].

Untuk mengatasi masalah kestasioneran data, ada beberapa metode yang dapat digunakan. Masalah kestasioneran data dapat diklasifikasikan menjadi dua jenis, yaitu masalah kestasioneran dalam rata-rata dan masalah kestasioneran dalam varians. Untuk mengatasi masalah kestasioneran dalam rata-rata, dapat digunakan metode perbedaan (*differencing*), yaitu dengan mengurangi nilai data saat ini dengan nilai data sebelumnya secara berurutan. Jika *differencing* ordo pertama tidak cukup untuk mencapai stasioneritas, maka proses *differencing* dilanjutkan ke ordo kedua, dan seterusnya, sampai data menjadi stasioner [13].

2.4 Evaluasi Model

Dalam proses pembuatan model forecasting, peneliti membagi data menjadi data train dan data test. Data train digunakan untuk membangun model peramalan, sedangkan data test digunakan untuk mengevaluasi hasil prediksi. Dua parameter yang dapat digunakan untuk mengukur kebaikan model adalah Mean Absolute Percentage Error (MAPE), dengan persamaan sebagai berikut [14] :

$$MAPE = 100\% \frac{\sum_t y_t - \hat{y}_t}{n}$$

(1)Keterangan :

y_t = nilai actual pada waktu ke t

\hat{y}_t = nilai prediksi pada waktu ke t

n = jumlah total data

2.5 SARIMA

Model ARIMA dapat diperluas untuk menangani faktor musiman, yang ditulis dengan notasi ARIMA (p,d,q)(P,D,Q)s. Rumus ini dijelaskan lebih lanjut sebagai berikut [15]:

$$\varphi_p B \varphi_p(B)^s (1 - B)^d (1 - B^s)^D Z_t = \theta_q B \theta_q(B^s) \alpha_t \quad (2)$$

Di mana s adalah periode musiman, dan

$$\varphi_p(B)^s = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps} \quad (3)$$

$$\theta_q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_q B^{qs} \quad (4)$$

Keterangan :

$\varphi_p B$ = komponen AR non-musiman

$\varphi_p(B)^s$ = komponen AR musiman

$(1 - B)^d$ = komponen differencing non-musiman dengan orde d

$(1 - B^s)^D$ = Komponen differencing musiman dengan orde D

Z_t = data actual atau nilai observasi pada waktu t

$\theta_q B$ = Komponen MA non-musiman

$\theta_q(B^s)$ = komponen MA musiman

α_t = nilai error pada waktu t

2.6 SARIMAX

Salah satu pengembangan dari model deret waktu SARIMA adalah model SARIMAX. Model SARIMAX merupakan perluasan dari SARIMA dengan memasukkan variabel tambahan atau variabel *exogenous*. Dalam model ini, faktor-faktor yang mempengaruhi variabel dependen Y pada waktu t tidak hanya merupakan fungsi dari variabel Y itu sendiri dalam berbagai waktu, tetapi juga dipengaruhi oleh variabel-

variabel independen lainnya pada waktu t . Secara umum, bentuk model SARIMAX (p, d, q) (P, D, Q) S adalah sebagai berikut:

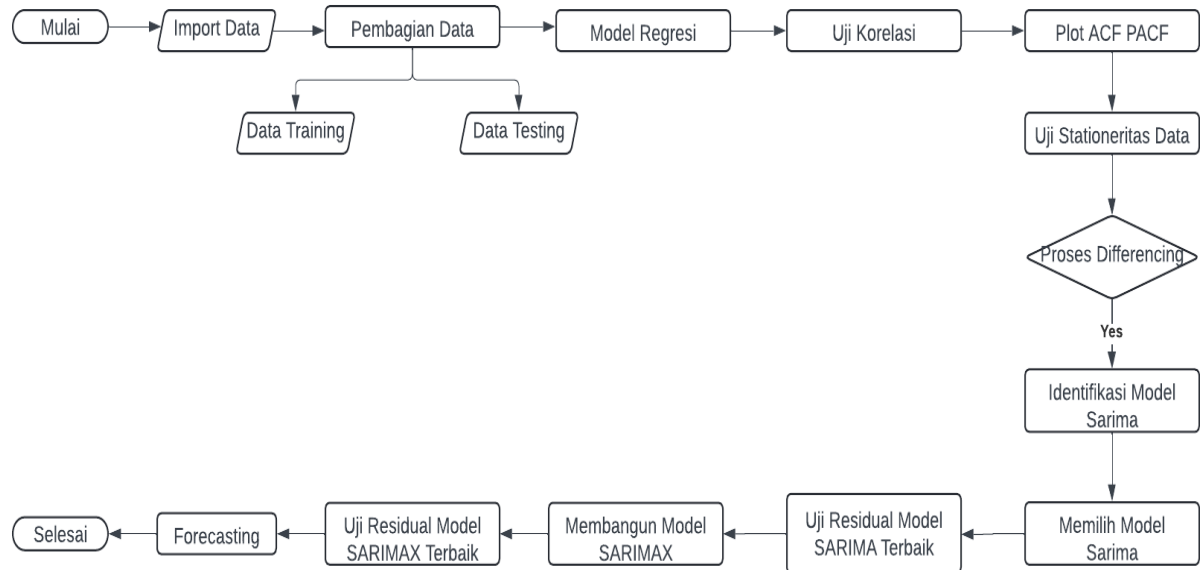
$$(1 - B)^d(1 - B^S)^D\phi_p(B)\phi_P(B^S)Y_t = \theta_q(B)\theta_Q(B^S)e_t + \alpha_1X_{1,t} + \alpha_2X_{2,t} + \dots + \alpha_kX_{k,t} \quad (5)$$

dengan:

$X_{k,t}$ = variabel tambahan atau sebagai variabel *exogenous* ke- k pada waktu t [16].

2.7 Tahap Penelitian

Teknik pengumpulan data dilakukan dengan mengambil data kelembapan dan temperature untuk wilayah Jawa Tengah. Data tersebut diperoleh dari BMKG Provinsi Jawa Tengah. Penelitian ini menggunakan *software R-Studio* dan *Python*. Tahapan Analisis Metode Sarimax adalah sebagai berikut:



Gambar 1. Tahapan Penelitian

Sumber : lucidchart.com

Gambar 1. Menunjukkan tahapan pemodelan, adapun penjelasannya sebagai berikut :

1. Langkah awal yaitu import data atau memasukkan data deret waktu yang akan digunakan untuk pemodelan.
2. Pembagian data :data dibagi menjadi data training digunakan untuk membangun model, dan data testing untuk mengevaluasi keakuratan model.
3. Model regresi: berperan untuk menggabungkan variabel eksogen.
4. Uji korelasi : tahap ini digunakan untuk melihat hubungan antara nilai deret waktu pada periode berbeda.
5. Plot ACF dan PACF: digunakan untuk mengidentifikasi orde model AR dan MA untuk model Sarima.
6. Uji stasioneritas data: digunakan untuk mengetahui apakah data stasioner atau tidak.
7. Proses differencing: jika data tidak stasioner proses ini dilakukan untuk menghilangkan tren atau komponen musiman agar data menjadi stasioner.
8. Identifikasi model SARIMA: Setelah data menjadi stasioner, proses identifikasi model SARIMA dilakukan dengan menetapkan nilai parameter p, d, q , serta parameter musiman P, D , dan Q .

9. Memilih Model SARIMA: Berdasarkan uji ACF, PACF, dan analisis differencing, model SARIMA terbaik dipilih.
10. Uji Residual Model SARIMA Terbaik: Setelah model SARIMA dipilih, residualnya (error) diuji untuk memastikan bahwa model telah menangkap semua pola yang ada dalam data, dan residual bersifat acak.
11. Membangun Model SARIMAX: setelah model sarima didapatkan, kemudian membangun model SARIMA dengan tambahan variabel eksogen yang bisa memengaruhi prediksi.
12. Uji Residual Model SARIMAX Terbaik: Sama seperti langkah sebelumnya, model SARIMAX juga diuji residualnya untuk memastikan tidak ada pola sistematis yang tersisa dalam error.
13. Forecasting: Setelah model terbaik dipilih dan diuji, dilakukan proses peramalan atau forecasting untuk memprediksi nilai-nilai masa depan berdasarkan data historis.
14. Selesai: Proses pemodelan dan forecasting selesai.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Deskriptif

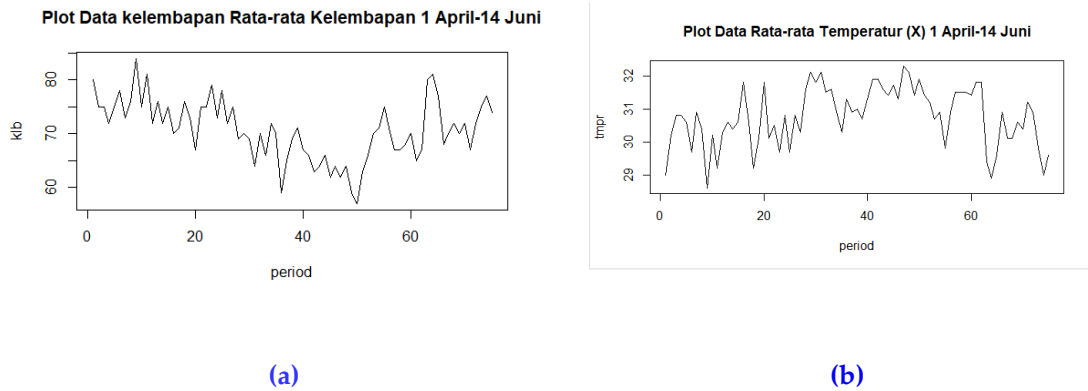
Analisis deskriptif adalah metode statistik yang digunakan untuk menggambarkan atau merangkum data. Tujuan utama dari analisis deskriptif adalah memberikan gambaran umum tentang data yang akan dilakukan pengujian. Berikut ini disajikan beberapa elemen dari analisis deskriptif variabel rata-rata kelembapan (Y) dan variabel rata-rata temperature (X).

Tabel 2. Analisis deskriptif

Analisis	Nilai	
	Rata-rata kelembapan	Rata-rata temperatur
Mean	70,75	30,76
Median	71,00	30,80
Maksimum	84,00	32,30
Minimum	57,00	28,60
Standar deviasi	5,58	0,88

Sumber : *Open Source Software R*

Berdasarkan **Tabel 2**, diketahui bahwa nilai rata-rata dari variabel Y sebesar 70,75, dengan median senilai 71,00, nilai maksimum 84,00, nilai minimum 57,00, dan nilai standar deviasi sebesar 5,58. Selain itu diketahui juga nilai rata-rata dari variabel x sebesar 30,76, dengan median senilai 30,80, nilai maksimum 32,30, nilai minimum 28,60, dan nilai standar deviasi sebesar 50,88. Selanjutnya diberikan grafik dari data actual rata-rata kelembapan dan rata-rata temperature di Stasiun Klimatologi Jawa Tengah.



Gambar 2. (a) Plot data rata-rata kelembapan (b) Plot data rata-rata temperatur

Sumber : Open Source Software R

Berdasarkan **Gambar 2. (a)** dan **(b)** menunjukkan bahwa data tidak stasioner, hal ini dikarenakan nilai *p-value* yang diperoleh dari hasil pengujian *Augmented Dickey-Fuller Test* menunjukkan nilai *p-value* lebih besar dari taraf signifikansi sebesar 0,05 (5%). Berikut adalah tabel dari hasil uji *Augmented Dickey-Fuller Test*.

Tabel 3. Uji Augmented Dickey-Fuller Test

Jenis uji	<i>P-value</i>
<i>Augmented Dickey-Fuller Test</i>	0,3707

Sumber : Open Source Software R

Dari **Tabel 3.** didapatkan nilai ADF sebesar 0,3707 yang artinya data tidak stasioner. Karena data tidak stasioner, maka diperlukan proses *differencing* agar data menjadi stasioner dan dapat dilanjutkan ke tahap uji selanjutnya.

3.2 Analisis Korelasi

Analisis korelasi digunakan untuk mengetahui seberapa kuat hubungan antara variabel X dengan variabel Y. Analisis korelasi juga digunakan untuk mengetahui apakah ada hubungan positif atau negatif antara variabel-variabel tersebut. Berikut ini disajikan hasil analisis korelasi menggunakan korelasi *pearson*.

Tabel 4. Analisis korelasi

Jenis uji	<i>P-value</i>
<i>Pearson's product-moment correlation</i>	-0,85

Sumber : Open Source Software R

Berdasarkan **Tabel 4.** diperoleh nilai *Pearson's product-moment correlation* sebesar -0,85 menunjukkan hubungan yang kuat dan negatif antara variabel rata-rata kelembapan dan rata-rata temperature. Artinya terdapat kecenderungan jika nilai variabel rata-rata kelembapan meningkat, maka nilai variabel rata-rata temperature cenderung menurun secara konstan, begitupun sebaliknya.

3.3 Uji kestasioneran data

Uji kesasioneran data adalah suatu uji yang dilakukan untuk melihat apakah data *time series* memiliki sifat stasioner atau tidak. Berdasarkan **Gambar 2.** setelah dilakukan pengecekan kestasioneran data menggunakan uji *Augmented Dickey Fuller* (ADF) didapatkan bahwa data rata-rata kelembapan tidak stasioner, sehingga diperlukan penanganan menggunakan proses *differencing*. Setelah dilakukan *differencing* satu kali pada data rata-rata kelembapan, didapatkan hasil dari pengujian ADF sebagai berikut.

Tabel 5. Uji kestasioneran data

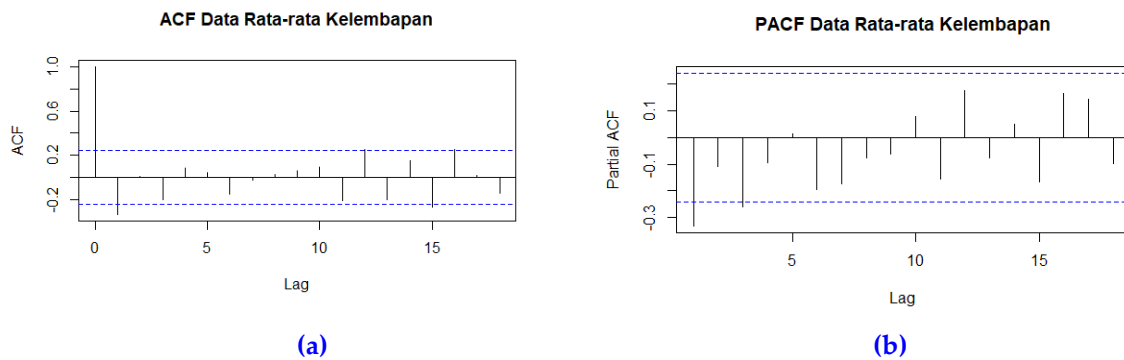
Jenis Uji	P-value	Keterangan
Augmented Dickey Fuller (ADF)	0,01	stasioner

Sumber : Open Source Software R

Dari **Tabel 5.** didapatkan hasil dari pengujian *Augmented Dickey Fuller* (ADF) setelah dilakukan *differencing* satu kali sebesar 0,01 dimana lebih kecil dari tingkat signifikansi yang digunakan yaitu sebesar 0,05 yang artinya data sudah stasioner.

3.4 Identifikasi model SARIMA

Identifikasi model dasar (non-musiman) dapat dilakukan dengan menganalisis plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) dari data yang sudah di *differencing* non musiman.

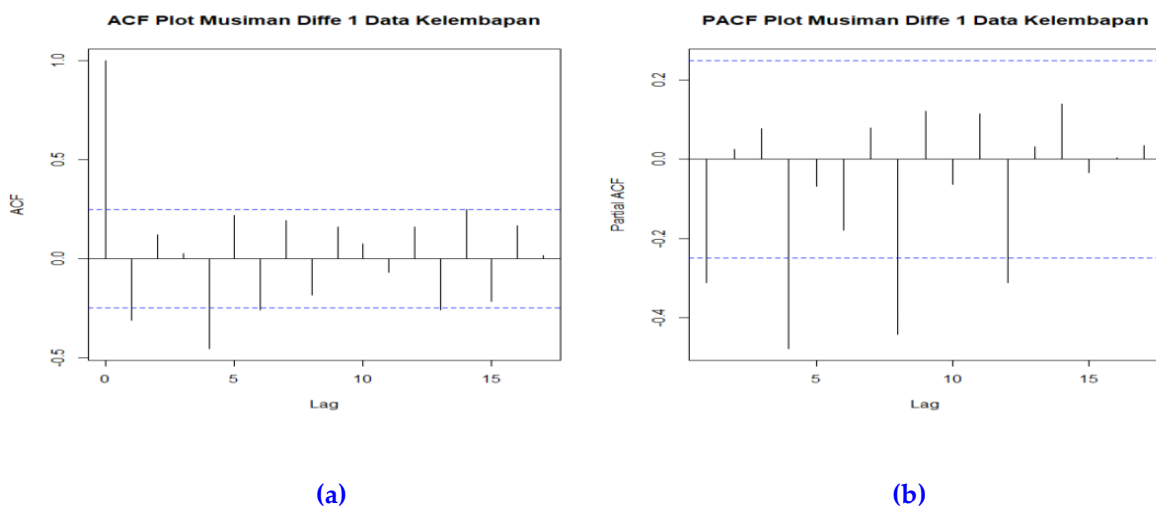


Gambar 3. (a) Plot Acf non musiman (b). Plot Pacf non musiman

Sumber : Open Source Software R

Berdasarkan **Gambar 3. (a)** dan **(b)** didapatkan *cut off* setelah lag ke 3, sedangkan pada plot Acf dapat dilihat *cut off* setelah lag ke 1, sehingga di dapatkan kemungkinan adanya proses *Autoregressive* (AR) dengan orde 3, *Integrated* (I) dengan orde 1, dan *Moving Average* (MA) dengan orde 1. Didapatkan kombinasi model ARIMA (p, d, q) = (3, 1, 1).

Setelah dilakukan identifikasi model dasar non musiman, selanjutnya dilakukan identifikasi komponen musiman dengan melihat plot ACF dan PACF dari data yang telah di *differencing* musiman. Berikut ini disajikan plot ACF dan PACF komponen musiman.



Gambar 4. (a) Plot ACF musiman (b) plot PACF Musiman

Sumber : Open Source Software R

Berdasarkan **Gambar 4 (a)** dan **(b)** Dapat dilihat bahwa plot PACF *cut off* setelah lag ke 1, sedangkan pada plot ACF *cut off* setelah lag ke 1, sehingga didapatkan kemungkinan

adanya proses *Autoregressive* (AR) dengan orde 1, *Integrated* (I) dengan orde 1, dan *Moving Average* (MA) dengan orde 1. Didapatkan kombinasi model Arima (P, D, Q) S = (1, 1, 1) 4.

Dari identifikasi model non musiman dan musiman didapatkan model SARIMA(p, d, q) (P, D, Q) s = (3, 1, 1) (1, 1, 1) 4. Dari kombinasi model tersebut didapatkan 21 model sementara SARIMA sebagai berikut.

- Model 1 = SARIMA (3, 1, 1)(1, 1, 1) 4
- Model 2 = SARIMA (3, 1, 1)(1, 1, 0) 4
-
- Model 21 = SARIMA (1, 1, 0)(1, 1, 0) 4

Setelah mendapatkan model sementara, kemudian dilakukan uji signifikansi untuk menguji apakah parameter model signifikan dan layak menjadi model terbaik atau tidak. Berikut ini disajikan hasil uji signifikan dari model SARIMA sementara yang telah di dapat.

Tabel 6. Uji Signifikansi

Model	Keterangan
(3, 1, 1)(1, 1, 1)4	Tidak signifikan
.....
(1, 1, 1)(0, 1, 1)4	Signifikan
(1, 1, 1)(1, 1, 0)4	Signifikan
.....
(0, 1, 1)(0, 1, 1)4	Signifikan
(0, 1, 1)(1, 1, 0)4	Signifikan
.....
(1, 1, 0)(0, 1, 1)4	Signifikan
(1, 1, 0)(1, 1, 0)4	Signifikan

Sumber : *Open Source Software R*

Berdasarkan **Tabel 6.** yang merupakan tabel hasil uji signifikansi dari 21 model, didapatkan 6 model yang signifikan, sedangkan 15 model tidak signifikan. Dari ke 6 model yang signifikan kemudian dilakukan pengujian untuk mencari model terbaik untuk model SARIMA. Salah satu cara untuk pemilihan model terbaik yaitu dengan melihat nilai AIC yang terkecil. Berikut ini merupakan hasil nilai AIC pada model yang signifikan.

Tabel 7. AIC Model Signifikan

Model	AIC
(1, 1, 1)(0, 1, 1)4	372,27
(1, 1, 1)(1, 1, 0)4	384,2
(0, 1, 1)(0, 1, 1)4	373,89
(0, 1, 1)(1, 1, 0)4	387,63
(1, 1, 0)(0, 1, 1)4	375,41
(1, 1, 0)(1, 1, 0)4	388,89

Sumber : *Open Source Software R*

Dari hasil **Tabel 7.** diperoleh model dengan nilai AIC terkecil yaitu pada model SARIMA (1, 1, 1)(0, 1, 1)4 dengan nilai AIC sebesar 372,27. Sehingga model SARIMA(1, 1, 1)(0, 1, 1)4 dapat dikatakan sebagai model terbaik dan dapat digunakan untuk peramalan.

3.5 Uji Asumsi Residual

Uji asumsi residual merupakan serangkaian tes yang dilakukan untuk memastikan bahwa residu memenuhi asumsi. Asumsi yang harus dipenuhi diantaranya pengujian *White Noise* dan pengujian Normalitas. Pengujian ini dilakukan pada model SARIMA terbaik $(1, 1, 1)(0, 1, 1)_4$ untuk menguji kelayakan model dan akurasi prediksi model. Berikut disajikan hasil uji *White Noise*.

Tabel 8. Uji White Noise

Test	P-value	Keterangan
<i>Box-Ljung Test</i>	0,80	Terpenuhi

Sumber : *Open Source Software R*

Berdasarkan **Tabel 8** yang memuat hasil dari *Box-Ljung Test* didapatkan nilai p-value senilai $0,80 > 0,05$, hal ini menunjukkan bahwa tidak terjadi autokorelasi, artinya asumsi *White Noise* pada model SARIMA $(1, 1, 1)(0, 1, 1)_4$ terpenuhi. Selanjutnya dilakukan pengujian normalitas, berikut ini disajikan hasil dari uji normalitas.

Tabel 9. Uji Normalitas

Test	P-value	Keterangan
<i>Kolmogorov-Smirnov</i>	0,97	Terpenuhi

Sumber : *Open Source Software R*

Berdasarkan hasil dari uji *Kolmogorov-Smirnov* pada **Tabel 9**, didapatkan nilai *p-value* senilai $0,97 > 0,05$ hal ini menunjukkan bahwa residu mengikuti distribusi normal, artinya asumsi normalitas pada model $(1, 1, 1)(0, 1, 1)_4$ terpenuhi. Karena model SARIMA $(1, 1, 1)(0, 1, 1)_4$ memenuhi kedua asumsi, maka model SARIMA $(1, 1, 1)(0, 1, 1)_4$ layak digunakan untuk meramalkan rata-rata kelembapan di Stasiun Klimatologi Jawa Tengah.

3.6 Pemodelan SARIMAX

Setelah mendapatkan ordo SARIMA, langkah selanjutnya adalah mengidentifikasi model SARIMAX. Hal ini dilakukan dengan menambahkan variabel *exogenous* yaitu variabel rata-rata temperature pada model SARIMA. Sehingga diperoleh beberapa model sebagai berikut.

Tabel 10. Uji Signifikansi Model SARIMAX

Model	Keterangan
$(3, 1, 1)(1, 1, 1)_4$	Tidak signifikan
.....
$(3, 1, 0)(1, 1, 0)_4$	Signifikan
$(3, 1, 0)(0, 1, 1)_4$	Signifikan
.....
$(2, 1, 1)(1, 1, 0)_4$	Tidak Signifikan
$(2, 1, 0)(1, 1, 1)_4$	Signifikan
.....
$(1, 1, 1)(1, 1, 1)_4$	Signifikan
$(0, 1, 1)(1, 1, 1)_4$	Signifikan

Sumber : *Open Source Software R*

Berdasarkan **Tabel 10**, diperoleh 21 model berdasarkan pemodelan SARIMAX dengan ordo SARIMA sebelumnya. Namun, terdapat beberapa model dengan nilai yang tidak signifikan. Untuk mencari model terbaik dari SARIMAX, model yang dipilih harus signifikan dan mempunyai nilai AIC dengan nilai yang terkecil. Sehingga disajikan nilai AIC dari model yang signifikan pada tabel berikut.

Tabel 11. Nilai AIC Model Signifikan

Model	AIC
(3, 1, 0)(1, 1, 0) ⁴	338,02
(3, 1, 0)(0, 1, 1) ⁴	327,84
(2, 1, 0)(1, 1, 1) ⁴	347,45
(1, 1, 1)(1, 1, 1) ⁴	323,89
(0, 1, 1)(1, 1, 1) ⁴	325,6

Sumber : *Open Source Software R*

Pada **Tabel 11**, disajikan beberapa model yang sudah signifikan, dengan nilai AIC yang tertera pada tabel diperoleh AIC terkecil dari kelima model di atas terdapat pada model (1, 1, 1)(1, 1, 1)⁴. Sehingga model SARIMAX (1, 1, 1)(1, 1, 1)⁴ adalah model terbaik dan dapat digunakan untuk peramalan.

3.7 Uji Asumsi Residual model SARIMAX

Pada model terbaik SARIMAX, perlu dilakukan uji asumsi residual untuk memastikan apakah model yang terpilih yaitu (1, 1, 1)(1, 1, 1)⁴ akan diterapkan pada uji selanjutnya. Maka dari itu, perlu adanya pengujian asumsi residual *White Noise* yang hasilnya dapat dilihat pada **Tabel 12**, dibawah ini.

Tabel 12. Uji White Noise Model SARIMAX

Test	P-value	Keterangan
Box-Ljung Test	0,5843	Terpenuhi

Sumber : *Open Source Software R*

Berdasarkan **Tabel 12**, Uji *White Noise* telah terpenuhi dikarenakan nilai *p-value* 0,5843 > 0,05 yang berarti bahwa tidak terdapat gejala autokorelasi pada model (1, 1, 1)(1, 1, 1)⁴ dan akan dilanjutkan pada uji *Kolmogorov-Smirnov* yang hasil pengujiannya disajikan pada tabel berikut.

Tabel 13. Uji Normalitas Model SARIMAX

Test	P-value	Keterangan
Kolmogorov-Smirnoov	0,8174	Terpenuhi

Sumber : *Open source Software R*

Pada **Tabel 13**, disajikan informasi bahwa uji normalitas pada model (1, 1, 1)(1, 1, 1)⁴ dengan nilai *p-value* 0,8174 yang berarti lebih dari 0,05. Sehingga nilai residual pada model terbaik SARIMAX berdistribusi normal.

3.8 Pemilihan Model Terbaik

Model SARIMA dan SARIMAX terbaik yang telah ditentukan pada tahap sebelumnya, akan dibandingkan guna mendapatkan model terbaik untuk melakukan *forecasting*. Karena kedua model terbaik telah memenuhi uji asumsi residual, maka pada tahap ini akan dibandingkan model terbaik dari SARIMA dan SARIMAX dengan nilai AIC dan MAPE yang terkecil. Berikut ini disajikan perbandingan nilai AIC dan MAPE dari model terbaik SARIMA dan SARIMAX.

Tabel 14. Perbandingan Model

Metode	Model	AIC	MAPE
SARIMA	(1, 1, 1)(0, 1, 1) ⁴	372,27	4,90201
SARIMAX	(1, 1, 1)(1, 1, 1) ⁴	323,89	2,863913

Sumber : *Open Source Software R*

Berdasarkan **Tabel 14**, kita dapat mengetahui bahwa model terbaik antara dua metode sebelumnya adalah terdapat ada pada metode SARIMAX. Model SARIMAX (1, 1, 1)(1, 1, 1)⁴ memiliki nilai AIC dan MAPE terkecil karena semakin kecil MAPE akan semakin baik untuk memberikan prediksi yang lebih akurat. Sehingga model ini adalah model terbaik yang layak digunakan untuk *forecasting* pada tahap selanjutnya.

Sehingga diperoleh persamaan SARIMAX dari model terbaik yang sudah didapatkan adalah sebagai berikut.

$$(1 - B)(1 - B^4)(1 + 0,239611B)(1 - 0,799515B^4)Y_t = 1 + 0,999997B)(1 + 0,959066B^4)e_t - 4,660900X_{tmp,r,t} \quad (7)$$

3.9 Peramalan Data

Peramalan data atau *forecasting* adalah suatu proses untuk memperkirakan nilai-nilai pada masa yang akan datang berdasarkan data historis. Dalam konteks data rata-rata kelembapan, *forecasting* memiliki arti memperkirakan tingkat kelembapan untuk waktu yang akan datang dengan menggunakan data kelembapan yang telah dikumpulkan dari periode waktu ke waktu.

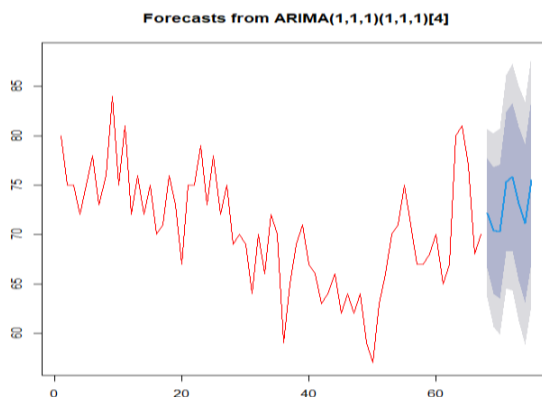
Hasil Prediksi data rata-rata kelembapan untuk periode April-Juni 2024 dengan mengambil sampel data harian menggunakan model terbaik yang digunakan pada proses peramalan ini. Peramalan dilakukan untuk periode 8 hari kedepan dengan nilai Prediksi yang diperoleh dari proses peramalan dan disajikan pada tabel dibawah ini.

Tabel 15. Forecast SARIMAX

Periode	Forecast	95% Confidence Interval	
		Lower	Upper
7/6/2024	72,15	63,64	80,67
8/6/2024	70,43	60,62	80,23
9/6/2024	70,27	59,86	80,67
10/6/2024	75,32	64,51	86,16
11/6/2024	75,81	64,35	87,26
12/6/2024	73,11	61,2	85,02
13/6/2024	71,1	58,81	83,4
14/6/2024	75,51	62,86	88,16

Sumber : Open Source Software R

Diperoleh nilai prediksi pada **Tabel 15.** yang memuat informasi tentang data prediksi dengan selang kepercayaan sebesar 95% dengan plot *forecast* dari model SARIMAX sebagai berikut.



Gambar 5. Plot Forecast SARIMAX

Sumber : Open Source Software R

Berdasarkan **Gambar 5** dapat dilihat bahwa hasil prediksi rata-rata kelembapan mulai dari tanggal 7 Juni – 14 Juni 2024 menunjukkan pola perubahan yang signifikan dalam nilai prediksi dari hari ke hari, hal ini mencerminkan pola musiman dalam kondisi cuaca atau faktor-faktor lingkungan lainnya yang mempengaruhi rata-rata kelembapan. Hal ini

memberikan indikasi tentang pola peramalan untuk masa depan. Metode SARIMAX ini hanya menginterpretasikan pola stasioner dari bagian yang dipilih saja, perlu prediksi dengan metode lain untuk mendapatkan hasil yang fluktuatif.

4. KESIMPULAN

Penelitian ini menyoroti penggunaan metode SARIMAX sebagai alat yang efektif untuk memprediksi rata-rata kelembapan dengan mempertimbangkan faktor musiman dengan pengaruh variabel *exogenous* yaitu rata-rata temperature. Data yang digunakan pada metode ini mencakup pengukuran rata-rata kelembapan dan rata-rata temperature dari Stasiun Klimatologi Jawa Tengah. Rata-rata temperature dijadikan variabel *exogenous* untuk melihat seberapa besar pengaruhnya terhadap prediksi kelembapan. Kemudian setelah dilakukan pengujian korelasi untuk mengukur seberapa kuat variabel rata-rata kelembapan dipengaruhi variabel rata-rata temperature, didapatkan hasil sebesar -0,85 yang artinya hubungan keduanya sangat kuat.

Proses pemodelan mencakup identifikasi, estimasi, dan diagnostik model SARIMAX. Validasi yang dilakukan dalam penelitian ini dilakukan dengan membandingkan hasil prediksi dengan data aktual untuk mengukur akurasi model. Hasil menunjukkan bahwa model SARIMAX dengan variabel *exogenous* rata-rata temperature memberikan prediksi yang lebih akurat dibandingkan model tanpa variabel *exogenous*. Dari pemaparan sebelumnya, didapatkan model terbaik yang digunakan untuk peramalan adalah SARIMAX (1, 1, 1)(1, 1, 1)₄ dengan nilai AIC 323,89 dan nilai MAPE sebesar 2,863913. Sehingga didapatkan persamaan model SARIMAX (1, 1, 1)(1, 1, 1)₄ sebagai berikut.

$$(1 - B)(1 - B^4)(1 + 0,239611B)(1 - 0,799515B^4)Y_t = 1 + 0,999997B)(1 + 0,959066B^4)e_t - 4,660900X_{tmpr,t}$$

Hal ini mengindikasikan adanya hubungan signifikan antara variabel rata-rata temperature dan rata-rata kelembapan pada wilayah tersebut. Prediksi yang akurat mengenai kelembapan dapat bermanfaat bagi berbagai sektor seperti pertanian, kesehatan, dan perencanaan wilayah. Dengan mengetahui pola kelembapan, dapat dilakukan langkah-langkah antisipasi terhadap dampak perubahan iklim.

DAFTAR PUSTAKA

- [1] M. H. Mukron, "Peramalan Curah Hujan di Stasiun Meteorologi Klimatologi Geofisika di Jawa Tengah Menggunakan GSTAR Kalman Filter," 2021, [Online]. Available: <http://repository.unimus.ac.id/4817/%0Ahttp://repository.unimus.ac.id/4817/6/12> BAB II.pdf
- [2] Aszani, "Penerapan Probabilistic Neural Network(Pnn) Untuk Prakiraan Cuaca Kota Pekanbaru," no. 2504, pp. 1–9, 2018.
- [3] M. Azizah, A. Subiyanto, S. Triutomo, and D. Wahyuni, "Pengaruh Perubahan Iklim Terhadap Bencana Hidrometeorologi di Kecamatan Cisarua - Kabupaten Bogor," *PENDIPA J. Sci. Educ.*, vol. 6, no. 2, pp. 541–546, 2022, doi: 10.33369/pendipa.6.2.541-546.
- [4] S. Prasetyo, Ulil Hidayat, Yosafat Donni Haryanto, and Nelly Florida Riama, "Karakteristik Suhu Udara di Pulau Jawa Kaitannya Dengan Kelembapan Udara, Curah Hujan, SOI, dan DMI," *J. Geogr. Edukasi dan Lingkungan.*, vol. 5, no. 1, pp. 15–26, 2021, doi: 10.22236/jgel.v5i1.5971.

- [5] Y. N. I. Fathulrohman and A. Saepuloh, "Alat Monitoring Suhu Dan Kelembaban Menggunakan Arduino Uno," *Jumantaka*, vol. 02, no. 1, p. 1, 2018, [Online]. Available: <https://jurnal.stmik-dci.ac.id/index.php/jumantaka/article/view/361>
- [6] R. M. Abdurrohman, "Prototipe Monitoring Suhu Dan Kelembaban Secara Realtime," *J. ICTEE*, vol. 4, no. 2, p. 29, 2023, doi: 10.33365/jictee.v4i2.3158.
- [7] Y. Novianto and Y. Nataliani, "Peramalan Curah Hujan dengan Pengelompokan Bulan Menggunakan Metode Double Exponential Smoothing dari Brown," *J. Sist. dan Teknol. Inf.*, vol. 10, no. 4, p. 347, 2022, doi: 10.26418/justin.v10i4.47167.
- [8] G. L. Marcus, H. J. Wattimanela, and Y. A. Lesnussa, "Analisis Regresi Komponen Utama Untuk Mengatasi Masalah Multikolinieritas Dalam Analisis Regresi Linier Berganda," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 6, no. 1, pp. 31–40, 2012, doi: 10.30598/barekengvol6iss1pp31-40.
- [9] I. Satria, A. Anas Manurung, and M. A. Hanafiah, "Analisis Perkembangan Produksi Tanaman Biofarmaka (Obat) di Indonesia Menggunakan Algoritma Resilient," vol. 5, no. 1, pp. 122–131, 2023.
- [10] H. K. Juliarto, I. Purnamasari, and S. Prangga, "Peramalan Peredaran Uang Kartal Di Indonesia Menggunakan Model Hybrid Sarimax-Neural Network," *J. Gaussian*, vol. 12, pp. 465–476, 2024, doi: 10.14710/j.gauss.12.4.465-476.
- [11] N. H. Latief, N. Nur'eni, and I. Setiawan, "Peramalan Curah Hujan di Kota Makassar dengan Menggunakan Metode SARIMAX," *Stat. J. Theor. Stat. Its Appl.*, vol. 22, no. 1, pp. 55–63, 2022, doi: 10.29313/statistika.v22i1.990.
- [12] L. Fitriyah, "Pemodelan Penderita Demam Berdarah Dengue (DBD) di Kabupaten Gresik Menggunakan Arima dan Regresi Nonparametrik Kernel," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [13] N. Mutammimah, P. S. Matematika, F. Sains, D. A. N. Teknologi, U. Islam, and N. Sunan, "Prediksi jumlah produksi padi di jawa timur menggunakan metode sarima with exogeneous input (sarimax)," 2022.
- [14] N. Permatasari, "Penggunaan Indeks Google Trend Dalam Peramalan Jumlah Pengunjung Taman Rekreasi Selecta Tahun 2020," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 1019–1024, 2021, doi: 10.34123/semnasoffstat.v2021i1.993.
- [15] A. A. Indrawati, *Pemodelan SARIMAX Variasi Liburan dengan Deret Input Data Iklan (Studi Kasus pada Data Mingguan Banyaknya Pengunjung Taman Rekreasi Yayasan Paggora Kediri)*. 2017.
- [16] Z. Zalila, "Peramalan Jumlah Kunjungan Wisatawan Mancanegara Ke Provinsi Jawa Tengah dengan Metode Seasonal Autoregressive Integrated Moving Average With Exogeneous Input (SARIMAX)," pp. 8–28, 2019, [Online]. Available: <http://repository.unimus.ac.id>