

Peramalan Curah Hujan di Kota Bandung dengan Metode SARIMA (Seasonal Autoregressive Integrated Moving Average)

Muhammad Ilham Hakiqi^{1*}, Arif Firmansyah², dan Restu Arisanti³

^{1,3}Departemen Statistika, Universitas Padjadjaran, Indonesia

²Dinas Komunikasi dan Informatika Kota Bandung

*Corresponding author: muhammad19424@mail.unpad.ac.id

Received: 24 January 2023

Revised: 4 August 2023

Accepted: 4 October 2023

ABSTRAK – Kebutuhan akan informasi curah hujan untuk masa mendatang, pemodelan dan peramalan curah hujan menjadi penting untuk dilakukan. Metode peramalan merupakan metode yang digunakan untuk memprediksi kondisi di masa yang akan datang berdasarkan data masa lalu. Data curah hujan adalah data runtun waktu yang berbentuk musiman suatu pola yang berulang-ulang dalam selang waktu yang tetap oleh karena itu penulis menggunakan metode SARIMA (*Seasonal Autoregressive Integrated Moving Average*) yang cocok untuk data yang berkarakteristik musiman. Penulis mengambil data curah hujan bulanan di Kota Bandung periode Januari 2016 hingga Desember 2021 untuk meramalkan curah hujan di Kota Bandung selama setahun kedepan. Lalu setelah melakukan perhitungan dengan metode SARIMA diperoleh model yang terbaik untuk peramalan curah hujan Kota Bandung yaitu model SARIMA $(0,0,0)(0,1,1)_{12}$.

Kata kunci– SARIMA, Peramalan, Curah Hujan, Kota Bandung

ABSTRACT – The need for future rainfall information, modeling and forecasting is important. The forecasting method is a method used to predict future conditions based on past data. Rainfall data is time-series data in the form of seasonality, a pattern that repeats at fixed time intervals, so the authors use the Seasonal Autoregressive Integrated Moving Average (SARIMA) method, which is appropriate for data with seasonal characteristics. The author takes monthly rainfall data in Bandung city for the period January 2016 to December 2021 to forecast rainfall in Bandung city for next year. After calculations using the SARIMA method, the best model for forecasting rainfall in the city of Bandung is then obtained, namely the SARIMA model $(0,0,0)(0,1,1)_{12}$.

Keyword– SARIMA, Forecasting, Rainfall, Bandung City

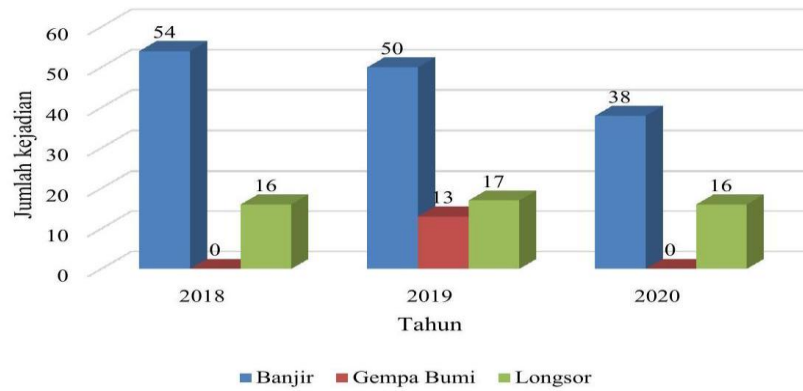
I. PENDAHULUAN

Curah hujan merupakan jumlah air hujan yang jatuh pada suatu daerah dan pada periode waktu tertentu. Banyaknya curah hujan dinyatakan dengan satuan milimeter (mm). Data curah hujan adalah data runtun waktu yang berbentuk musiman suatu pola yang berulang-ulang dalam selang waktu yang tetap [1]. Berdasarkan data curah hujan bulanan dapat dilakukan pengelompokan, menurut Badan Meteorologi Klimatologi dan Geofisika (BMKG) ada tiga kategori yaitu kategori rendah untuk curah hujan 0 – 100 mm, lalu kategori menengah untuk curah hujan 100 – 300 mm, lalu kategori tinggi untuk curah hujan 300 – 500 mm dan kategori sangat tinggi untuk curah hujan >500 mm.

Ketua umum IABI (Ikatan Ahli bencana Indonesia) Sudibyakto mengatakan bahwa, curah hujan yang berlangsung dengan intensitas sangat tinggi, dan merata kejadiannya, akan menyebabkan debit sungai dan saluran drainase Kota Bandung terlampaui. Sehingga terjadi banjir besar yang mampu menerjang apa saja yang dilewatinya. Sistem drainase Kota Bandung juga yang bertopografi miring mendukung sistem pengutusan banjir, sehingga banjir berlangsung cepat [2].

Kota Bandung merupakan ibukota Provinsi Jawa Barat yang secara astronomis terletak di antara $06^{\circ} 55'$ Lintang Selatan dan $107^{\circ} 37'$ Bujur Timur dengan luas wilayah $167,31 \text{ km}^2$ [3]. Kota Bandung merupakan salah satu kota yang tercatat sebagai daerah yang sering diterjang bencana banjir dibandingkan dengan bencana lain. Hal ini dapat dilihat pada Gambar 1. pada gambar tersebut dapat dilihat bahwa jumlah kejadian bencana banjir pada setiap kecamatan di Kota Bandung lebih banyak dibandingkan bencana lain seperti gempa bumi dan longsor.

Berdasarkan kebutuhan akan informasi curah hujan Kota Bandung untuk masa mendatang, pemodelan dan peramalan curah hujan menjadi penting untuk dilakukan. Metode peramalan merupakan metode yang digunakan untuk memprediksi kondisi di masa yang akan datang berdasarkan data masa lalu. Metode peramalan yang banyak digunakan antara lain metode ARIMA (*Autoregressive Integrated Moving Average*), SARIMA (*Seasonal Autoregressive Integrated Moving Average*), Bayesian, Fungsi Transfer, Metode *Smoothing*, dan Regresi [4]. Dari banyaknya metode untuk peramalan yang telah disebutkan, penulis memutuskan menggunakan metode SARIMA untuk peramalan curah hujan di Kota Bandung dengan tujuan untuk memperoleh nilai ramalan curah hujan di Kota Bandung. Hasil penelitian ini diharapkan dapat bermanfaat baik bagi masyarakat, pemerintah maupun pihak BMKG sebagai bahan pertimbangan dalam mempersiapkan hal-hal yang diperlukan untuk mengantisipasi kemungkinan-kemungkinan yang akan terjadi berdasarkan hasil peramalan.



Gambar 1 Jumlah bencana alam Kota Bandung 2018-2020

Tujuan yang ingin dicapai dari penelitian ini adalah mendapatkan model SARIMA yang terbaik yang akan digunakan untuk peramalan curah hujan di Kota Bandung. Hasil penelitian ini diharapkan dapat memberikan manfaat bagi pihak-pihak terkait seperti, menambah wawasan dan pengalaman dalam mengaplikasikan ilmu statistika yang sudah dipelajari selama perkuliahan, sebagai bahan pertimbangan dalam membuat kebijakan dan peraturan terkait permasalahan banjir yang dikarenakan curah hujan yang tinggi.

II. TINJAUAN PUSTAKA

A. Data Deret Berkala

Analisis deret berkala adalah bagian dari suatu proses stokastik yang merupakan barisan variabel acak yang berdasarkan pada urutan waktu. Tujuan dari analisis deret berkala meliputi dua hal yaitu memahami dan memodelkan proses stokastik yang timbul pada rangkaian observasi dan memperkirakan nilai masa depan berdasarkan data deret berkala yang telah terjadi [5].

B. Stasioneritas Data

Menurut Santoso [6], data stasioner adalah data dimana rata-rata nilai pada suatu data tidak berubah seiring dengan perubahan waktu, dengan kata lain fluktuasi data berada di sekitar nilai rata-rata dan varians yang konstan. Penentuan Stasioneritas ini sangat penting karena berkaitan dengan metode estimasi yang digunakan. Untuk melakukan melihat data stasioner dapat dilakukan dengan melihat grafik dan dengan uji ADF (*Augmented Dickey-Fuller*) dengan persamaan sebagai berikut:

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \sum \beta_i \Delta Y_{t-1} + e_t$$

Jika melihat dari grafik data dikatakan stasioner bila dari grafik terlihat rata-rata varians konstan lalu dengan uji ADF jika *p-value* lebih kecil dari taraf signifikan.

C. Differencing

Differencing merupakan cara untuk membuat data yang tidak stasioner terhadap rata-rata menjadi stasioner. Sebaliknya jika data belum stasioner terhadap varians maka harus ditransformasi dengan transformasi *Box-Cox* [7]. Data belum stasioner dalam rata-rata dihitung dengan uji ADF hasilnya data belum stasioner maka akan lakukan proses differencing dengan persamaannya sebagai berikut:

$$\Delta Y_t = Y_t - Y_{t-1}$$

D. Model ARIMA (Autoregressive Integrated Moving Average)

Model yang digunakan adalah Teknik Box-Jenkins [8]. Teknik ini memiliki tiga model yaitu AR (*Autoregressive*), MA (*Moving Average*), ARMA (*Autoregressive Moving Average*), dan model campuran ARIMA (*Autoregressive Integrated Moving Average*). Metode ini salah satu metode yang biasa digunakan untuk melakukan prediksi data deret waktu. Data yang digunakan untuk pemodelan menggunakan ARIMA harus bersifat stasioner terhadap rata-rata maupun varians, jika data bersifat tidak stasioner maka dilakukan proses *differencing* [9]. Proses *differencing* ini hanya bisa dilakukan pada model ARIMA karena model AR, MA, dan ARMA tidak mampu menjelaskan arti dari *differencing* [10], dengan model sebagai berikut:

- Model AR (p):

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t$$
- Model MA (q)

$$Y_t = \delta + e_t - \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}$$

- ARMA (p,q):

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$
- Model ARIMA:

$$Y_t = \delta + (1 + \phi_1)Y_{t-1} + \dots + (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}$$

E. Model SARIMA (Seasonal Autoregressive Integrated Moving Average)

Model ARIMA memiliki variasi yaitu model ARIMA musiman yang dapat dinyatakan sebagai SARIMA, model ini digunakan untuk data yang memiliki efek musiman yang melekat [11]. model ARIMA musiman yang dapat dinyatakan sebagai SARIMA $(p, d, q)(P, D, Q)^m$ yang terbentuk:

$$\phi_p(B)\phi_p(B^s)(1 - B)^d(1 - B^s)^D Z_t = \theta_q(B)\theta_q(B^s)\alpha_t$$

F. Estimasi Parameter

Estimasi Parameter mampu menggunakan dengan beberapa metode salah satunya adalah metode *Ordinary Least Squares* [12]. Metode ini berguna untuk meminimumkan jumlah kuadrat error. Setelah parameter diestimasi, selanjutnya dilakukan Uji signifikansi parameter dengan membandingkan p-value dengan taraf signifikan dengan Hipotesis nol (H_0) parameter tidak signifikan dan kriteria uji tolak H_0 jika p-value lebih kecil dari taraf signifikan.

G. Uji Autokorelasi

Uji autokorelasi bertujuan apakah residual bersifat *white noise* (tidak berautokorelasi) dengan uji Q-Ljung Box [13].

Hipotesis:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0 \text{ (residual tidak berautokorelasi)}$$

$$H_1: \text{Minimal ada satu } \rho_1 \neq 0 \text{ (residual ber autokorelasi)}$$

Statistik uji:

Uji Ljung-Box

H. Uji Normalitas

Untuk uji normalitas digunakan adalah dengan menggunakan uji Kolmogorov-Smirnov.

Hipotesis:

$$H_0: \text{residual berdistribusi normal}$$

$$H_1: \text{residual tidak berdistribusi normal}$$

Pengembalian keputusan yaitu apabila $p\text{-value} \geq \gamma$ maka H_0 diterima, namun jika $p\text{-value} < \gamma$ maka H_0 ditolak dengan $\gamma = 0,05$ [14].

I. Uji Homoskedastisitas

Salah satu model time series yang memiliki asumsi adanya gejala homoskedastisitas (varians yang berubah-ubah setiap waktu t) adalah model *Autoregressive Conditional Heteroscedasticity* (ARCH). Pada penelitian ini dilakukan pengujian homoskedastisitas dengan statistik uji Q-Ljung Box dengan hipotesis sebagai berikut.

Hipotesis:

$$H_0 : \text{residual bersifat homoskedastisitas}$$

$$H_1 : \text{residual bersifat heteroskedastisitas}$$

Statistik uji:

Uji Q-Ljung Box

J. Forecasting (Peramalan)

Peramalan merupakan tentang memprediksi masa depan dengan seakurat mungkin dengan semua informasi yang diperoleh seperti data masa lalu dan pengetahuan dari kejadian yang dapat mempengaruhi pada peramalan [15]. Pada penelitian ini dalam menentukan model SARIMA terbaik untuk peramalan adalah dipilih berdasarkan dengan nilai terkecil dari *Akaike's Information Criterion* (AIC) [16]. AIC dapat dihitung dengan rumus sebagai berikut:

$$AIC = -2\log \log (L) + 2m$$

dimana,

L = maksimum *likelihood*

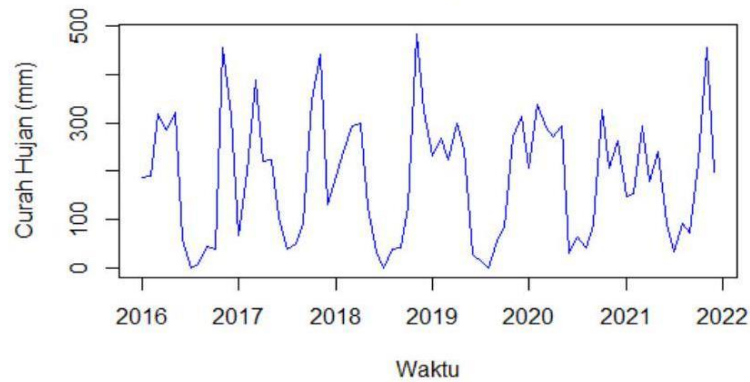
m = banyaknya parameter yang akan diestimasi dalam model AR,MA, dan koefisien regresi

III. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan data bulanan Jumlah curah hujan di kota Bandung pada Tahun 2016 sampai 2021. Data yang digunakan dikutip dari BPS Kota Bandung.

A. Identifikasi Data

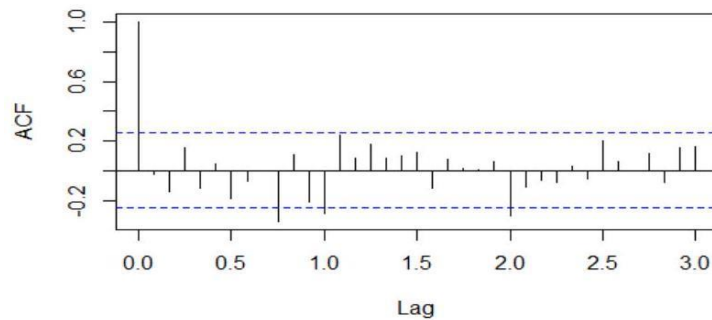
Untuk melihat data yang digunakan apakah data seasonal (musiman) atau tidak dapat dilihat dari grafik dibawah ini sebagai berikut:



Gambar 2 Grafik Jumlah Curah Hujan Bulanan di Kota Bandung

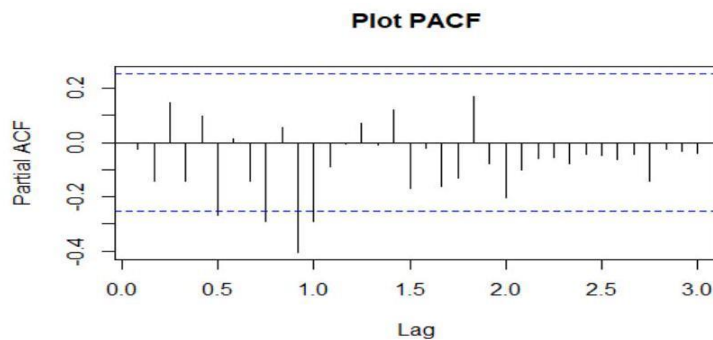
Berdasarkan Gambar 2 diperoleh bahwa data bulan jumlah curah hujan di kota Bandung merupakan data seasonal dikarenakan terdapat data yang pola yang berulang pada setiap tahunnya. Oleh karena itu dilakukan *differencing* terhadap seasonalnya. Lalu diuji dengan uji adf diperoleh hasil *p-value* $0.01439 < 0.05$ (alpha) yang artinya data sudah stasioner sehingga tidak diperlukan lagi *differencing* terhadap non-seasonal yang berarti untuk notasi D sebesar 1 dan d sebesar 0.

B. Plot ACF & PACF



Gambar 3 Plot ACF

Untuk menentukan notasi *q* dan *Q* dapat dilihat pada Gambar 3, Diperoleh hasil dari Gambar 5 bahwa pada plot ACF terjadi *cut off* di lag 1 yang berarti notasi *Q* adalah 1. Lalu tidak terjadi *cut off* di lag di sekitar lag 0 yang berarti untuk notasi *q* adalah 0.



Gambar 4 Plot PACF

Untuk menentukan notasi *p* dan *P* dapat dilihat plot PACF pada Gambar 4, diperoleh bahwa pada plot PACF terjadi *cut off* di lag 1 yang berarti notasi *P* adalah 1. Kemudian, tidak terjadi *cut off* di lag di sekitar lag 0 yang berarti untuk

notasi p adalah 0. Maka dapat diperoleh bahwa model dari SARIMA adalah SARIMA (0,0,0)(1,1,1)¹². Dari model tersebut dapat dibuat 3 variasi model.

C. Penentuan Model

Uji Signifikansi Parameter

Tabel 1 Uji Signifikansi Parameter

Model	Parameter	p-value	Keterangan
SARIMA (0,0,0)(1,1,1) ¹²	SAR(1)	0,9921	tidak signifikan
	SMA (1)	0,0002	signifikan
SARIMA (0,0,0)(0,1,1) ¹²	SMA (1)	2,00E – 04	signifikan
SARIMA (0,0,0)(1,1,0) ¹²	SAR (1)	0,0031	signifikan

Dapat dilihat pada tabel 1 bahwa dari 3 variasi model diperoleh 2 model yang hasilnya signifikan pada semua parameternya yang dimana 2 model tersebut memiliki p -value pada setiap parameter lebih kecil dari alpha (0.05).

D. Uji Asumsi

Uji Non Autokorelasi

Model yang baik adalah model residualnya bersifat *white noise* atau tidak berautokorelasi, dari hasil pengujian dengan *Q-Ljung Box* diperoleh p -value dari 2 model sebagai berikut:

Tabel 2 Hasil Uji Asumsi non Autokorelasi

Model	p-value	Keterangan
SARIMA (0,0,0)(0,1,1) ¹²	0,7632	H ₀ diterima
SARIMA (0,0,0)(1,1,0) ¹²	0,8921	H ₀ diterima

Dapat dilihat pada Tabel 2 diperoleh bahwa dari 2 variasi model semua memenuhi asumsi non autokorelasi (*white noise*) atau p -value lebih besar dari alpha (0,05).

Uji Normalitas

Uji normalitas residual pada penelitian ini menggunakan uji Kolmogorov-Smirnov diperoleh p -value dari 2 model sebagai berikut:

Tabel 3 Uji Normalitas

Model	p-value	Keterangan
SARIMA(0,0,0)(0,1,1) ¹²	0,4938	H ₀ diterima
SARIMA(0,0,0)(1,1,0) ¹²	0,6307	H ₀ diterima

Dapat dilihat pada tabel 3 bahwa diperoleh dari 2 variasi model semua memenuhi asumsi normalitas residual atau p -value lebih besar dari α (0,05).

Uji Homoskedastisitas

Mengidentifikasi efek Heteroskedastisitas dengan menggunakan uji *Q-Ljung box* diperoleh p value dari 5 model sebagai berikut:

Tabel 4 Hasil Uji asumsi Homoskedastisitas

Model	p-value	Keterangan
SARIMA (0,0,0)(0,1,1) ¹²	0,5784	H ₀ diterima
SARIMA (0,0,0)(1,1,0) ¹²	0,9763	H ₀ diterima

Dapat dilihat pada tabel 4 diperoleh bahwa dari 2 variasi model semua memenuhi asumsi homoskedastisitas atau p -value lebih besar dari α (0,05).

E. Pemilihan Model Terbaik

Setelah melewati ketiga uji asumsi diperoleh bahwa 2 variasi model yang memenuhi ketiga uji asumsi. Oleh karena itu untuk memilih model terbaik untuk peramalan maka akan dibandingkan nilai AIC pada setiap model. Untuk nilai AIC pada kedua model sebagai berikut:

Tabel 5 Nilai AIC

Model	AIC
SARIMA (0,0,0)(0,1,1) ¹²	703,84
SARIMA (0,0,0)(1,1,0) ¹²	718,82

Didapatkan bahwa dari Tabel 5 bahwa nilai AIC terkecil adalah model SARIMA (0,0,0)(0,1,1)¹² dengan nilai AIC 703,84. Maka model yang dipakai untuk peramalan adalah model SARIMA (0,0,0)(0,1,1)¹².

Peramalan

Setelah diperoleh model terbaik yaitu SARIMA (0,0,0)(0,1,1)¹² untuk peramalan jumlah curah hujan di Kota Bandung, selanjutnya dilakukan peramalan selama 12 bulan kedepan diperoleh hasilnya sebagai berikut:

Tabel 6 Hasil Peramalan

Bulan	Hasil Peramalan(mm)
Januari 2022	171,65
Februari 2022	231,28
Maret 2022	301,01
April 2022	258,36
Mei 2022	240,88
Juni 2022	57,91
Juli 2022	25,05
Agustus 2022	38,03
September 2022	65,13
Oktober 2022	189,59
November 2022	385,39
Desember 2022	356,48

IV. KESIMPULAN

Berdasarkan hasil analisa pada peramalan jumlah curah hujan di Kota Bandung pada bulan Januari 2016 hingga bulan desember 2021 dengan menggunakan model SARIMA, dapat disimpulkan bahwa : Model SARIMA yang digunakan untuk melakukan peramalan jumlah curah hujan di Kota Bandung adalah model SARIMA (0,0,0)(0,1,1)¹² karena model ini yang memiliki parameter yang signifikan, ketiga uji asumsi terpenuhi dan nilai AIC terkecil. Setelah diperoleh model yang terbaik lalu dihitung hasil peramalannya. Didapatkan bahwa:

- bulan Januari-Mei direntang 170 mm sampai 301 mm (Sedang)
- bulan Juni-September pada rentang 25 mm sampai 65 mm (Rendah)
- bulan Oktober-Desember pada rentang 189 mm sampai 385 mm. (Sedang dan Tinggi)

Dari hasil peramalan dapat dilihat bahwa pada bulan Januari-Mei dan bulan oktober-Desember adalah bulan yang harus diperhatikan oleh pemerintah kota bandung karena pada bulan tersebut berpotensi hujan dengan intensitas sedang dan tinggi sehingga pada bulan-bulan tersebut berpotensi banjir di Kota Bandung.

DAFTAR PUSTAKA

- [1] D. Ruswanti, "PENGUKURAN PERFORMA SUPPORT VECTOR MACHINE DAN NEURAL NETWORK DALAM MERAMALKAN TINGKAT CURAH HUJAN," *J. Gaung Inform.*, vol. 13, no. 1, 2020.
- [2] Gusti, "IABI: Banjir Bandung Akibat Konversi Lahan DAS Citarum | Universitas Gadjah Mada," *UGM*, 2016. <https://www.ugm.ac.id/berita/12719-iabi-banjir-bandung-akibat-konversi-lahan-das-citarum> (accessed Dec. 20, 2022).
- [3] Badan Pusat Statistik, "Statistik Indonesia 2021," Kota Bandung, 2021.
- [4] A. Lusiani and E. Habinuddin, "Pemodelan Autoregressive Integrated Moving Average (ARIMA) Curah Hujan di Kota Bandung," *Sigma-Mu*, vol. 3, pp. 9–25, 2011.
- [5] J. D. Cryer and K.-S. Chan, *Time Series Analysis*. New York, NY: Springer New York, 2008. doi: 10.1007/978-0-387-75959-3.
- [6] S. Santoso, "Business forecasting: metode peramalan bisnis masa kini dengan minitab dan SPSS," *Elek Media Komputindo, Jakarta*, 2009.
- [7] H. H. M. Hatta, F. M. Daud, and N. Mohamad, "An application of time series ARIMA forecasting model for predicting the ringgit Malaysia-dollar exchange rate," *J. Data Anal.*, vol. 1, no. 1, pp. 42–48, 2018.
- [8] A. Widarjono, "Aplikasi model ARCH kasus tingkat inflasi di Indonesia," *Econ. J. Emerg. Mark.*, vol. 7, no. 1, 2002.
- [9] S. Makridakis, S. Wheelwright, and R. J. Hyndman, *Forecasting: methods and applications*. John Wiley & Sons, 1998.
- [10] N. Salwa, N. Tatsara, R. Amalia, and A. F. Zohra, "Peramalan Harga Bitcoin Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average)," *J. Data Anal.*, vol. 1, no. 1, pp. 21–31, 2018.
- [11] S. I. Vagropoulos, G. I. Chouliaras, E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting," in *2016 IEEE International Energy Conference (ENERGYCON)*, Apr. 2016, pp. 1–6. doi: 10.1109/ENERGYCON.2016.7514029.
- [12] Tasna Yunita, "Peramalan Jumlah Penggunaan Kuota Internet Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA)," *J. Math. Theory Appl.*, pp. 16–22, Aug. 2020, doi: 10.31605/jomta.v2i1.777.
- [13] G. E. P. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *J. Am. Stat. Assoc.*, vol. 65, no. 332, pp. 1509–1526, 1970.
- [14] A. M. Al'afi, W. Widiarti, D. Kurniasari, and M. Usman, "Peramalan Data Time Series Seasonal Menggunakan Metode Analisis Spektral," *J. Siger Mat.*, vol. 1, no. 1, pp. 10–15, 2020.
- [15] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.
- [16] M. I. Rizki and T. A. Taqiyuddin, "Penerapan Model SARIMA untuk Memprediksi Tingkat Inflasi di Indonesia," *J. Sains Mat. dan Stat.*, vol. 7, no. 2, Aug. 2021, doi: 10.24014/jsms.v7i2.13168.



© 2023 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (<http://creativecommons.org/licenses/by-sa/4.0/>).