ISSN 2087-3336 (Print) | 2721-4729 (Online)

TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika
Vol. 14, No. 1, 2027, page. 37-43
http://jurnal.sttmcileungsi.ac.id/index.php/tekno
DOI: 10.37373

SARIMA-Based water consumption forecasting for distribution planning and performance evaluation at padang municipal waterworks

Nelvidawati[1*], Firdaus[2]

[1*] Teknik Lingkungan, Fakultas Teknik, Institut Teknologi Padang, Jl. Gajah Mada Kandis, Nanggalo, Kota Padang, Sumatera Barat, 25146, Indonesia
[2] Teknik Telekomunikasi, Jurusan Teknik Elektro, Politeknik Negeri Padang, Jl. Kampus Limau Manis, Kec. Pauh, Kota Padang, Sumatera Barat, Indonesia 25176

[*]Correspondence email: nelvidaus11@gmail.com, mrdauz@yahoo.com

Submitted: 21/12/2025 Revised: 13/01/2026 Accepted: 29/01/2026

Abstract

This study develops a forecasting model for water consumption at the Padang Municipal Waterworks using the Seasonal Autoregressive Integrated Moving Average (SARIMA) approach to enhance water distribution planning. Monthly consumption data from 2007 to 2024 were analyzed using time-series techniques. The results identify SARIMA $(1,1,1)(1,1,1)_{12}$ as the optimal model, effectively capturing both long-term trends and seasonal fluctuations. The model demonstrates high predictive accuracy with a Mean Absolute Percentage Error (MAPE) of 5.26%. While the Root Mean Square Error (RMSE) of 238,967.70 m^3 suggests limited robustness against extreme data anomalies, the model remains highly effective for forecasting under normal operational conditions. These findings provide a data-driven foundation for Padang City's water resource management, offering a more precise tool for anticipating demand cycles. Future research should consider incorporating external variables to better account for non-seasonal volatility.

Keywords: Water Consumption; SARIMA; Forecasting; Padang Municipal Waterworks; Resource Management.

1. Introduction

Urbanization and rapid population growth have significantly escalated the demand for clean water in developing cities [1]. Padang City, as a regional economic hub, faces substantial challenges in maintaining a sustainable water supply through its local water utility, Perumda Air Minum (PDAM) Padang. Water consumption in this area is not static; it exhibits significant fluctuations driven by seasonal cycles, consumer behavior [2], and shifting climatic conditions [3]. Such volatility often leads to an imbalance between supply and demand, which directly compromises distribution efficiency [4] and service quality [5];[6].

Effective water resource management necessitates precise forecasting to anticipate future consumption patterns [7][8]. Historical data-driven approaches provide a robust foundation for operational decision-making [9]. Recent developments in computational statistics have been widely applied to support urban policy-making in Indonesia, ranging from the prediction of crime patterns using regression models [10] to the forecasting of vital public utilities**.** In time-series analysis, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model has emerged as a powerful tool due to its specialized capability in capturing both long-term trends and recurring seasonal patterns [11].

While SARIMA has been applied in various hydrological contexts globally [12][13], its implementation for localized urban water utilities in Indonesia, particularly in Padang, remains under-explored. Most existing studies focus on macro-level data or lack a comprehensive evaluation using integrated accuracy metrics. This study addresses this gap by utilizing an extensive 17-year dataset

(2007–2024) to provide a more stable and reliable seasonal analysis. Furthermore, this research integrates mean absolute percentage error (MAPE) and root mean square error (RMSE) to rigorously validate the model's sensitivity to both average deviations and extreme anomalies.

The primary objective of this study is to develop an optimal SARIMA-based forecasting framework specifically tailored to the consumption characteristics of Padang City. By identifying deep-seated seasonal cycles and trends, this research provides a data-driven reference for PDAM Padang to optimize its distribution planning. The findings are expected to offer practical contributions to local policy-making, ensuring that water infrastructure development aligns with actual consumption trajectories to improve operational efficiency. This approach is consistent with industrial engineering principles where systematic planning is essential to optimize system layouts and production flows [14].

2. Method

a. Data description and source

This study utilizes secondary data consisting of monthly water consumption (m^3/month) from the Perumda Air Minum (PDAM) Kota Padang, covering the period from January 2007 to December 2024 (n = 216 observations). The data were sourced from the Central Bureau of Statistics (BPS) of Padang City. Prior to analysis, data preprocessing was conducted to ensure consistency and completeness, with no missing values identified in the dataset.

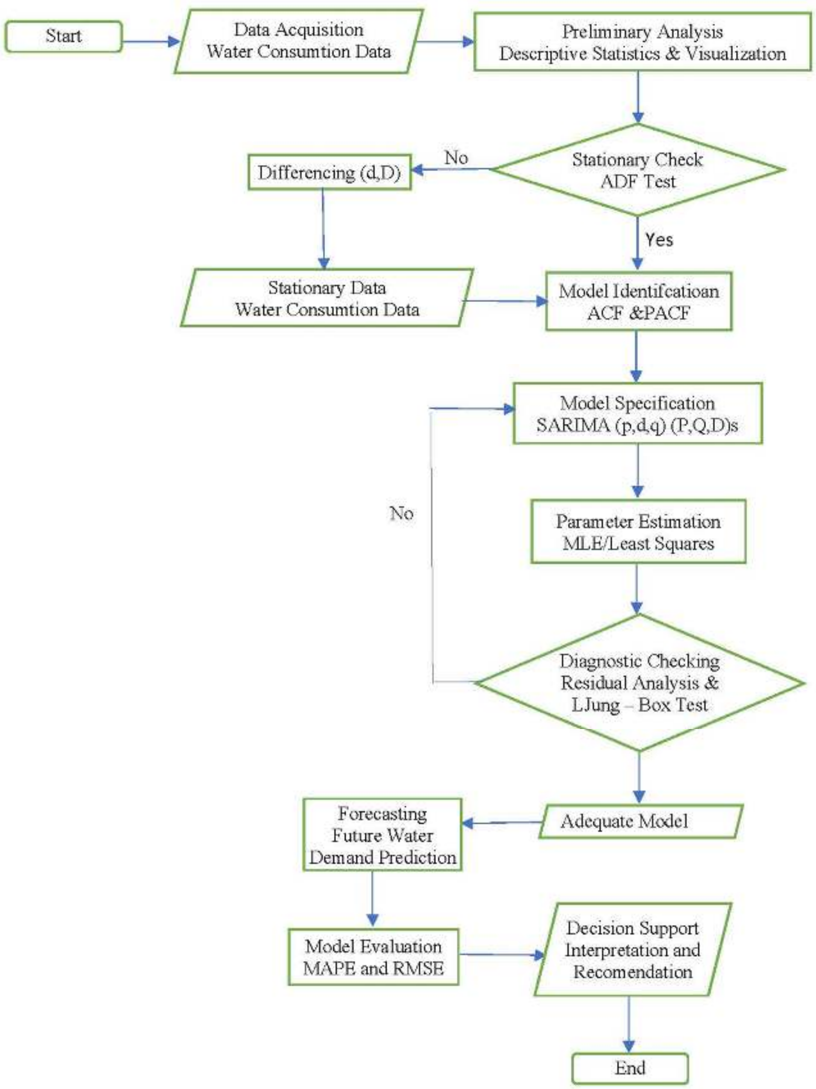

Figure 1. Flowchart SARIMA

b. Seasonal ARIMA (SARIMA) Model

The forecasting is performed using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, an extension of ARIMA designed to accommodate seasonal fluctuations. The model is denoted as SARIMA(p, d, q)(P, D, Q)$_s$ and is mathematically expressed as follows [15];[16]:

$$\Phi_P(B^s)\phi_p(B)(1-B)^2(1-B^s)^D y_t = \Theta_Q(B^s)\theta_q(B)\epsilon_t \quad (1)$$

where B is the backshift operator ($B_{yt} = y_{t-1}$), s is the seasonal period (s = 12 for monthly data), and t represents the white noise process. The parameters (p, d, q) handle the non-seasonal components, while (P, D, Q) represent the seasonal autoregressive, differencing, and moving average orders, respectively. The research follows the systematic Box–Jenkins framework, as illustrated in the flowchart in Figure 1.

c. Research procedure

The stages are summarized as follows: 1) Identification: Visual inspection of time-series plots and statistical verification using the Augmented Dickey-Fuller (ADF) test. Non-seasonal (d) and seasonal (D) differencing are applied to achieve stationarity. 2) Estimation: Potential model parameters are identified through Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. 3) Selection: The optimal model is selected based on the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values using Maximum Likelihood Estimation (MLE) in Python. 4) Diagnostic Checking: Residual analysis is performed using the Ljung–Box test to ensure the residuals satisfy white noise and normality assumptions.

d. Performance evaluation

The predictive accuracy of the model is validated using Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) [17], defined as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_t-\hat{Y}_t}{Y_t}\right| x\ 100\% \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t-\hat{Y}_t)^2} \quad (3)$$

where Y_t is the actual value and $\hat{Y}_t$ is the forecasted value. A MAPE value below 10% is categorized as highly accurate forecasting [18].

3. Results and Discussion

Time series identification and stationarity

Visual analysis of the water consumption data from 2007 to 2024 (Figure 2) reveals a persistent upward trend, with consumption doubling from 1.4 million m^3 to over 3.1 million m^3. Statistically, the original data is non-stationary, confirmed by the ADF test (p-value = 0.9812). After applying first-order non-seasonal and seasonal differencing (d = 1, D = 1), the data achieved stationarity (Figure 3), with the ADF statistic dropping to -6.0272 (p-value < 0.001).

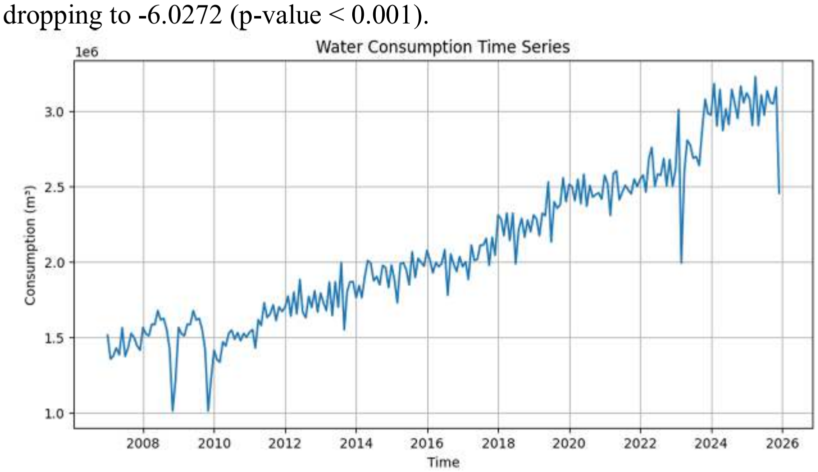

Figure 2. Water consumption time series

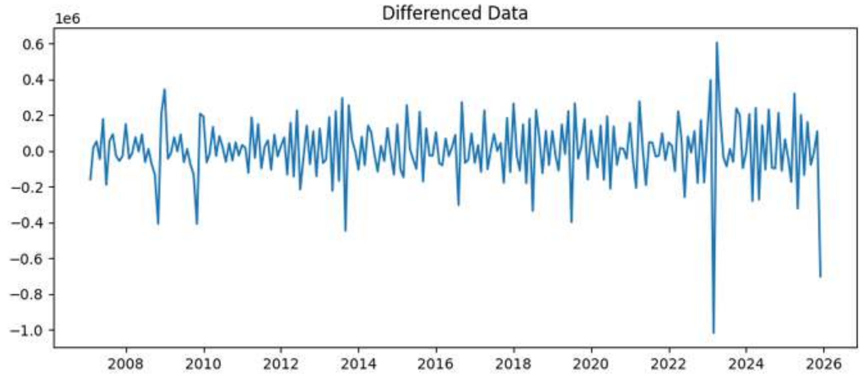

Figure 3. Differenced data

Model selection and parameter estimation

The initial identification through ACF and PACF plots (Figure 4) suggested a seasonal cycle of s = 12. Following an iterative selection process, the SARIMA(1,1,1)(1,1,1)$_{12}$ model was identified as the optimal configuration. The parameter estimation results using Maximum Likelihood Estimation (MLE) are presented in Figure 5.

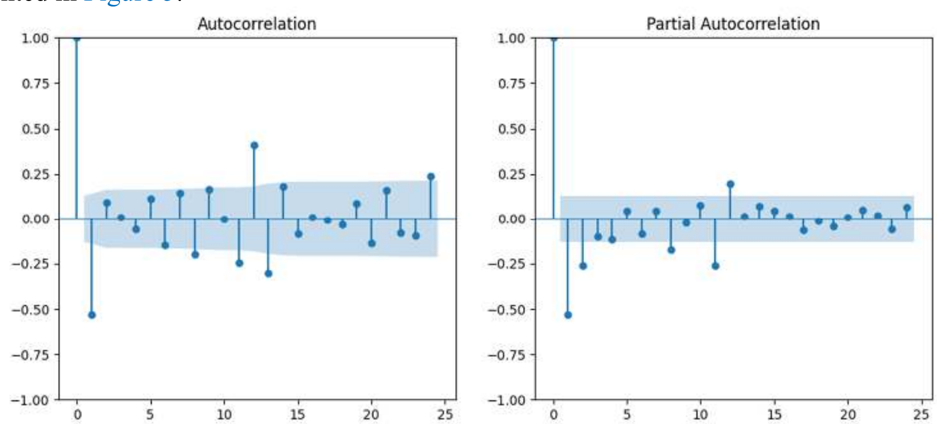

Figure 4. Autocorrelation function and partial autocorrelation function

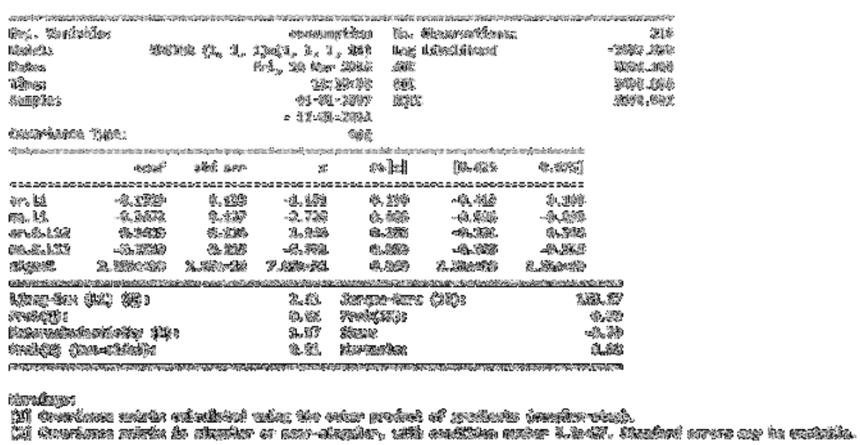
Figure 5. SARIMA Result

The model achieved an AIC of 5384.10 and a BIC of 5400.66, indicating a superior goodness-of-fit. The statistical significance of the coefficients in the output confirms that the model effectively accounts for both the trend and the complex seasonal dynamics of water demand in Padang.

Diagnostic checking

The model passed the Ljung-Box test (p-value = 0.12), indicating that the residuals are white noise (uncorrelated). Although the Jarque-Bera test (p-value < 0.05) suggests a non-normal distribution due to leptokurtic tails (Figure 6), this is common in high-volatility infrastructure data and does not invalidate the model's forecasting capability.

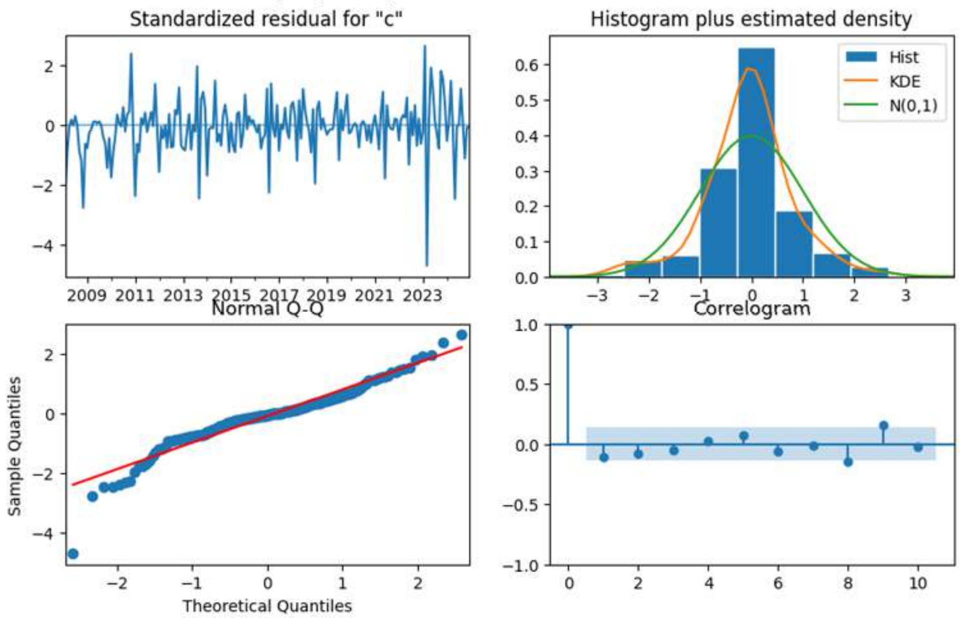

Figure 6. Diagnostic checking result

Forecasting performance and discussion

The model demonstrates exceptional accuracy under stable conditions, with a MAPE of 5.26%, classified as "Excellent". However, the RMSE of 238,967.70 m^3 highlights a significant gap during the testing phase (Figure 7).

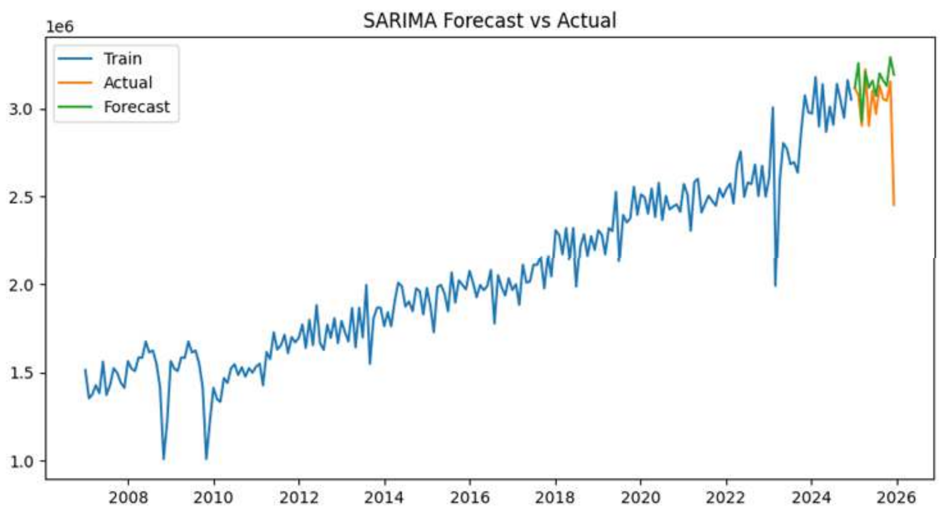

Figure 7. SARIMA forecast results and the actual data

Discussion of anomalies

The primary overestimation observed at the end of the period coincides with a flash flood disaster in Padang that severely damaged PDAM infrastructure. This event highlights a fundamental characteristic of the SARIMA model: while it excels at capturing demand-side trends, it is inherently limited in anticipating supply-side shocks. This finding suggests that for long-term utility management, statistical models must be integrated with contingency planning or hybrid approaches (such as SARIMAX) to account for extreme exogenous disruptions.

4. Conclusion

The study demonstrates that the SARIMA(1,1,1)$(1,1,1)_{12}$ model is a highly effective tool for forecasting water consumption in Padang City, achieving an excellent accuracy level with a MAPE of 5.26%. The model successfully captures the long-term upward trend and recurring monthly seasonal patterns, providing a robust data-driven foundation for PDAM Padang in strategic distribution planning and annual budgeting. However, the sensitivity of the model to the RMSE of 238,967.70 m^3 underscores a critical limitation: its inability to anticipate supply-side shocks caused by extreme exogenous events, such as the flash flood infrastructure damage observed in the data. While the SARIMA approach remains superior for forecasting under stable operational conditions, integrating exogenous variables or hybrid modeling is recommended to enhance resilience against non-periodic anomalies. Ultimately, this research offers a practical framework for more adaptive and precise water resource management in urban settings.

Reference

[1] Dinda Fitria Pida, Khadijah Nur Aini, and Cindy Amelia Putri, “Dampak Urbanisasi terhadap Perkembangan Kota di Indonesia: Tinjauan dari Aspek Ekonomi Pembangunan,” *WISSEN J. Ilmu Sos. dan Hum.*, vol. 3, no. 1, pp. 226–238, 2025, doi: 10.62383/wissen.v3i1.562.

[2] L. Marcell, R. Hadisoebroto, and T. Tazkiaturrizki, “Studi Metode Perhitungan Fluktuasi Pemakaian Air di Apartemen Tanpa Jam Pemompaan,” *Ajie*, vol. 06, no. May, pp. 65–70, 2022, doi: 10.20885/ajie.vol6.iss2.art3.

[3] D. Fiorillo, Z. Kapelan, M. Xenochristou, F. De Paola, and M. Giugni, “Assessing the Impact of Climate Change on Future Water Demand using Weather Data,” *Water Resour. Manag.*, vol. 35, no. 5, pp. 1449–1462, 2021, doi: 10.1007/s11269-021-02789-4.

[4] Z. J. Lyle, J. M. VanBriesen, and C. Samaras, “Drinking Water Utility-Level Understanding of Climate Change Effects to System Reliability,” *ACS ES T Water*, vol. 3, no. 8, pp. 2395–2406, 2023, doi: 10.1021/acsestwater.3c00091.

[5] N. Yadi, M. Belladona, and E. Dwiantoro, “Analisa Tingkat Kehilangan Air (Water Losses) pada Jaringan Distribusi PDAM Tirta Hidayah Zona 3 Kota Bengkulu Tahun 2024,” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 2280–2290, 2025, doi: 10.31004/riggs.v4i3.2302.

[6] G. Dandy, W. Wu, A. Simpson, and M. Leonard, “A Review of Sources of Uncertainty in Optimization Objectives of Water Distribution Systems,” *Water (Switzerland)*, vol. 15, no. 1, pp. 1–17, 2023, doi: 10.3390/w15010136.

[7] A. Ramazani, W. Fuadi, and R. Meiyanti, “Prediksi Jumlah Kebutuhan Pemakaian Air Menggunakan Metode Fuzzy Time Series Pada Perumda Aceh Utara,” *J. Ilm. Ilk. - Ilmu Komput. Inform.*, vol. 8, no. 2, pp. 200–209, 2025, doi: 10.47324/ilkominfo.v8i2.360.

[8] S. Alvisi *et al.*, “Battle of Water Demand Forecasting,” *J. Water Resour. Plan. Manag.*, vol. 151, no. 10, 2025, doi: 10.1061/jwrmd5.wreng-6887.

[9] B. G. Ramaddan, “Peran Data Analytics Dalam Pengambilan Keputusan Strategis Di Era Digital,” *J. Ilm. Penelit. Mhs.*, vol. 3, no. 3, pp. 749–755, 2025, [Online]. Available: https://doi.org/10.61722/jipm.v3i3.1194

[10] Fadillah Bergas, Sucipto, and Asrul Abdullah, “Prediction of the level of crime cases using multiple linear regression in the city of Pontianak,” *TEKNOSAINS J. Sains, Teknol. dan Inform.*, vol. 11, no. 2, pp. 245–256, 2024, doi: 10.37373/tekno.v11i2.1025.

[11] L. Junaedi, N. Damastuti, and A. Widodo, "Penerapan Metode Seasonal ARIMA (SARIMA) untuk Peramalan Penjualan Barang dengan Pola Musiman Tahunan," *JISEM J. Progr. Stud. Inform. Univ. Katolik Widya Mandala Surabaya*, vol. 01, pp. 38–48, 2025, [Online]. Available: https://doi.org/10.33508/jisem.v1i01.740338

[12] T. Argyanti, E. Ully, and T. Tampubolon, "Analisis peramalan curah hujan di provinsi jawa barat menggunakan pendekatan sarima," vol. 5, pp. 241–250, 2026.

[13] A. A. Maulana and H. Rosalina, "Implementasi Metode Sarimax Untuk Prediksi Curah Hujan Jangka Pendek Di Pagerageung, Tasikmalaya," *J. Sumber Daya Air*, vol. 20, no. 1, pp. 39–50, 2024, doi: 10.32679/jsda.v20i1.874.

[14] Puput Rahmawati, Indra Setiawan, and Setiawan, "Optimalisasi tata letak fasilitas produksi trolley untuk meningkatkan efisiensi produksi menggunakan metode systematic layout planning," *JENIUS J. Terap. Tek. Ind.*, vol. 6, no. 2, pp. 355–367, 2025, doi: 10.37373/jenius.v6i2.1957.

[15] H. A. Khoiri, *Analisis Deret Waktu Univariat*, no. 85. 2023. [Online]. Available: https://kwu.unipma.ac.id/buku/1823/analisis-deret-waktu-univariat-teori-dan-pengolahan-data

[16] F. Aulia, H. Yozza, and D. Devianto, "Peramalan Curah Hujan Bulanan Kabupaten Tanah Datar Dengan Model Seasonal Autoregressive Integrated Moving Average (Sarima)," *J. Mat. UNAND*, vol. 8, no. 2, pp. 37–44, 2019, doi: 10.25077/jmu.8.2.37-44.2019.

[17] H. Sulastri, G. S. Anwar, and E. N. F. Dewi, "Peramalan Stok Barang Percetakan dan ATK Menggunakan Single Moving Average," *J. Rekayasa Teknol. Inf.*, vol. 7, no. 1, p. 59, 2023, doi: 10.30872/jurti.v7i1.11876.

[18] R. Firdaus, "Prediksi Indeks Harga Produsen Pertanian Karet Di Indonesia Menggunakan Metode LSTM," *J. Fasilkom*, vol. 13, no. 01, pp. 1–6, 2023, doi: 10.37859/jf.v13i01.4851.