VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

A HYBRID FORECASTING-OPTIMIZATION FRAMEWORK FOR SEASONAL LOGISTICS USING SARIMA AND TRANSPORTATION MODEL

Wahyu Sri Utami[1*], Rr. Hajar Puji Sejati[1], Anita Fira Waluyo[1], Widya Setiafindari[1]

Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia[1]
https://uty.ac.id/[1]
wahyu.utami@uty.ac.id*, hajarsejati@uty.ac.id, anitafira@uty.ac.id, widyasetia@uty.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License

Submitted: 04 June 2026, **Revised:** 06 September 2026, **Accepted:** 09 November 2026, **Published:** 20 November 2026.

Abstract— *Demand fluctuations driven by seasonal patterns can adversely affect the performance of logistics systems if not managed effectively. Therefore, precise demand forecasting and efficient distribution planning are crucial for enhancing operational efficiency and minimizing transportation costs. This research introduces an integrated approach that incorporates demand forecasting and distribution optimization within a single approach. Historical demand data were modeled through the Seasonal Autoregressive Integrated Moving Average (SARIMA) approach to generate demand forecasts that capture seasonal patterns. Subsequently, the forecasting results were incorporated into a Transportation Model solved using Vogel's Approximation Method (VAM) to obtain a minimum-cost distribution plan. The findings demonstrate that the SARIMA model provided highly accurate forecasts, with Mean Absolute Error (MAE) values falling within the range of 8.28% to 9.69%, Mean Absolute Percentage Error (MAPE) values falling within the range of 4.53% to 5.55%, and Root Mean Square Error (RMSE) values falling within the range of 9.50% to 12.00%. The forecasting results indicate that the proposed SARIMA–Transportation hybrid framework provides an effective approach to supporting integrated demand forecasting and distribution optimization. Consequently, the framework has the potential to enhance the efficiency of logistics system management in environments characterized by seasonal demand patterns.*

Keywords: *Forecasting, logistics, SARIMA, transportation model, Vogel's Approximation Method (VAM).*

INTRODUCTION

The context of globalization and increasingly fierce competition, operational effectiveness and efficiency have become key determinants of organizational success, especially in supply chain management. Accurate demand fulfillment and efficient logistics distribution are essential for facilitating the efficient flow of goods from producers to final consumers. Demand forecasting serves as a fundamental basis for various managerial decisions, particularly in demand planning, inventory management, and order fulfillment. As a key source of information within supply chain systems, demand information is essential for coordinating operational activities, supporting effective planning, and ensuring process alignment throughout the supply chain network [1]. Nevertheless, attaining a high degree of forecasting accuracy continues to be challenging because demand patterns are inherently volatile and influenced by various sources of uncertainty. The dynamic nature of market conditions, changes in consumer preferences, and external environmental influences further complicate the demand forecasting process [2]. Therefore, the availability of accurate and reliable demand forecasts is essential, as it provides crucial analytical support for supply chain managers. Such forecasts enable more effective planning, improve resource allocation, and facilitate informed strategic as well as operational decision-making [3].

Distribution network optimization covering warehouse location, route planning, and resource allocation is closely tied to demand forecasts at each destination. Robust network design incorporates

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

demand uncertainty into optimization models to reduce total costs and prevent supply–demand imbalances [4]. Distribution network optimization cannot be effectively implemented without incorporating demand uncertainty as a key parameter within the model. Likewise, managing uncertain demand necessitates a distribution system capable of adapting efficiently to such fluctuations. This interdependence underscores the importance of research that treats demand uncertainty and distribution optimization as an integrated framework in the development of contemporary logistics strategies and models aimed at enhancing overall supply chain efficiency, responsiveness, and resilience [5].

Previous studies have investigated the existing conditions of supply chain management and logistics distribution by evaluating and comparing how various supply chain strategies influence distribution models. The findings indicate that aligning supply chain strategies with product characteristics can enhance distribution efficiency and overall performance [6]. Previous studies still show limitations in integrating the accuracy of demand forecasting with logistics distribution optimization simultaneously. Many studies tend to discuss these two topics separately. Previous studies that have discussed this issue have only focused on the challenges and approaches to demand forecasting, without directly incorporating distribution decisions or logistics optimization [7], [8], [9], [10], [11]. Meanwhile, other studies tend to focus only on discussing logistics distribution optimization [12], [13], [14].

An approach that treats demand forecasting and distribution independently is inadequate for addressing the complex interplay between demand variability and distribution decisions. This creates an opportunity to develop integrated models that align precise demand forecasts with effective distribution optimization, ultimately enhancing overall supply chain performance. The topic of distribution optimization with seasonal demand forecasting was chosen in this study because seasonal demand often causes imbalances between supply, distribution capacity, and inventory. Seasonal issues in logistics pose significant challenges because fluctuations in demand can cause imbalances in supply, inventory, and distribution capacity. Surge in demand during peak seasons or decline during off-peak seasons risks causing stockouts or overstocking, thereby increasing storage costs and lowering service levels if not managed properly [15] . In addition, in fresh produce logistics, seasonal demand can cause imbalances between supply and demand, affecting distribution networks, resource allocation, logistics costs, and market responsiveness [16]. Seasonal demand directly affects logistics operations, so accurate forecasting and distribution optimization are necessary to maintain system efficiency and responsiveness.

Within the domain of demand forecasting, numerous studies employ SARIMA and SARIMAX methodologies to generate accurate predictions using time series data [17], [18], [19]. Meanwhile, in the context of distribution optimization, various methods have been applied to achieve optimal performance. These methods are selected based on their ability to minimize costs, maximize route efficiency, and improve resource utilization, thereby providing distribution solutions with high optimal value for supply chain systems. Various transportation techniques can be applied to optimize distribution costs, including the Northwest Corner Method (NWCM), Least Cost Method (LCM), and Vogel's Approximation Method (VAM). Although these approaches can generate feasible initial solutions for transportation problems, ongoing research continues to seek alternative methods that can achieve better solutions with fewer computational iterations and improved efficiency [20].

This research develops an integrated framework in which SARIMA is employed to forecast demand, while the Transportation Model is used to optimize the subsequent distribution process. Unlike conventional approaches that perform forecasting and optimization separately, the proposed framework integrates both processes into a single decision-support framework. This integration contributes to improving the efficiency of seasonal logistics distribution planning through more accurate demand prediction and optimized allocation of transportation resources.

MATERIALS AND METHODS

This study employs a quantitative research approach based on time-series analysis, with exploratory and predictive objectives. The stages include data collection and cleaning, exploratory analysis, modeling using SARIMA, and the development of hybrid models (SARIMA and Transportation Model) that are then evaluated to generate accurate predictions. This study uses secondary data obtained from a rice logistics company through its operational database records. The study is based on 1,213 operational observations recorded from January 2023 through April 2026, including daily demand and distribution data.

CC BY NC
Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563
JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

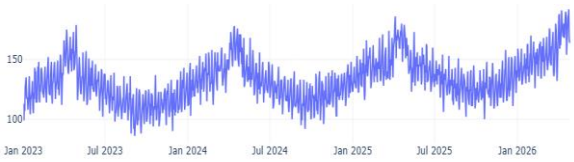
The study focuses on three warehouses and three factories operated by the company. For the forecasting stage, the daily demand data were aggregated into monthly demand series, resulting in 40 monthly observations for each warehouse (120 monthly observations in total). Distinct seasonal characteristics were observed in the aggregated demand series, providing a basis for applying the SARIMA model in the forecasting stage. The forecasted demand was subsequently used as input to the Transportation Model to determine the optimal distribution plan. Based on the data in Table 1, the team will conduct a time series analysis to forecast demand for goods at the three warehouses for the next three days.

Table 1. Product Demand Data for Each Warehouse

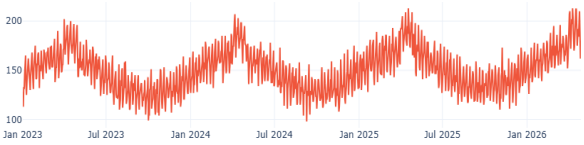
No	Data	W_A	W_B	W_C
1	2023-01-01	113	128	109
2	2023-01-02	99	113	103
3	2023-01-03	112	133	105
4	...	...	...	...
5	2026-04-28	165	180	171

Source: (Research Results, 2026)

The warehouse demand data from January 2023 to April 2026 is shown in Figure 1, Figure 2, and Figure 3.



Source: (Research Results, 2026)
Figure 1. Warehouse A Demand Data Trends



Source: (Research Results, 2026)
Figure 2. Warehouse B Demand Data Trends



Source: (Research Results, 2026)
Figure 3. Warehouse C Demand Data Trends

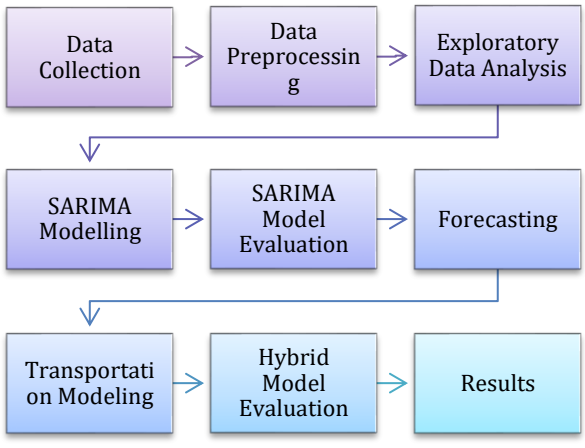
B Rice is typically harvested twice a year, during the rainy season in March–April and the dry season in August–September. The data indicate that demand peaks in April, as it aligns with the first major harvest and often overlaps with Ramadan and Idul Fitri, when household consumption increases for daily needs and festive preparations. In contrast, demand tends to decline in September due to the absence of major consumption events, and because it is close to harvest time, supply is abundant and many households rely on their own production. The company has three main distribution centers: Warehouse A, Warehouse B, and Warehouse C. The company supplies these three warehouses from three main factories: Plant A, Plant B, and Plant C. Daily Plant supply capacity: Plant A: 180 quintals, Plant B: 195 quintals, Plant C: 185 quintals. The transportation cost data is shown in Table 2.

Table 2. Shipping Cost from Plant to Warehouse

	Warehouse A	Warehouse B	Warehouse C
Plant A	Rp.4.000.000,-	Rp.7.000.000,-	Rp.8.000.000,-
Plant B	Rp.6.000.000,-	Rp.3.000.000,-	Rp.5.000.000,-
Plant C	Rp.5.000.000,-	Rp.4.000.000,-	Rp.4.000.000,-

Source: (Research results, 2026)

This study employs a quantitative approach using time series analysis, which is both exploratory and predictive in nature. The research methodology is illustrated in Figure 4.



Source: (Research Results, 2026)
Figure 4. Research Methodology

The study commences with data collection, followed by data preprocessing and feature engineering to enhance data quality. Subsequently, exploratory data analysis is performed to investigate the underlying patterns and characteristics of the dataset. The primary model used was SARIMA for forecasting, and its

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

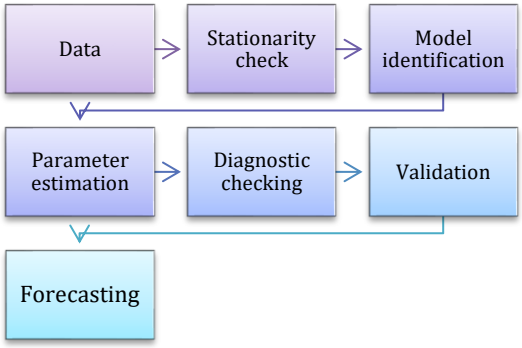
VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

performance was subsequently evaluated. Additionally, the study integrated transportation modeling and developed a hybrid model (SARIMA and transportation) to improve accuracy. The final stage involved a comprehensive evaluation and presentation of the research results.

The data preprocessing stage was undertaken to enhance data accuracy and consistency and to ensure that the dataset was appropriate for subsequent time-series analysis. At the initial stage, the Tanggal attribute was converted into datetime format to enable chronological ordering and time-series analysis. The observations were subsequently arranged in chronological order to maintain the temporal structure required for SARIMA modeling. Subsequently, missing observations and duplicate records were examined to verify the completeness and consistency of the dataset. Identified missing values were addressed using appropriate imputation methods, while duplicate records were removed to prevent redundant observations. Finally, outlier detection was performed using descriptive statistical analysis and visual inspection of the demand time series. Statistical and graphical techniques are employed to examine the dataset and identify underlying trends, seasonal patterns, correlations, and potential anomalies. Visualizations are created to better understand seasonal patterns SARIMA Modeling A time series model (SARIMA) is developed to capture temporal patterns, including seasonality and external variables.

SARIMA is a time-series statistical approach employed to model and forecast data exhibiting recurring seasonal patterns. In recent literature, SARIMA is understood as a model capable of capturing linear dependencies, trends, and seasonal cycles simultaneously through an integration of autoregressive (AR), differencing (I), and moving average (MA) terms across both seasonal and non-seasonal components [21] . The model follows the general specification (p, d, q)(P, D, Q)$_s$, enabling the representation of recurring seasonal patterns across defined time intervals, such as monthly or annual cycles. SARIMA is widely used because of its ability to generate accurate forecasts for non-stationary data after differencing [22] .

The SARIMA approach is designed to extend the conventional ARIMA model by incorporating seasonal effects, making it suitable for forecasting univariate time-series data with recurring seasonal behavior. Both seasonal and non-seasonal dynamics are represented through autoregressive and moving average terms at the corresponding seasonal lag s. The SARIMA model can be represented by the specification (p, d, q)(P, D, Q)$_s$, with its mathematical expression simplified through the use of the backward shift operator [23]. Recent research shows that SARIMA remains a robust and interpretable method for time series forecasting, even when compared to modern methods such as LSTM or other machine learning models, particularly for data with clear seasonal patterns [24]. Furthermore, several empirical studies (2023–2026) have also demonstrated that SARIMA can deliver low prediction errors and often outperforms other methods in specific cases [25] .



Source: (Research Results, 2026)
Figure 4. SARIMA Modeling Flow

SARIMA Model Form

The SARIMA model can be represented by:

$$SARIMA(p,d,q)(P,D,Q)_s \quad (1)$$

The SARIMA specification consists of two sets of parameters. The non-seasonal component is characterized by p, d, and q, which denote the AR order, differencing degree, and MA order, respectively. In contrast, the seasonal component is defined by P, D, and Q, corresponding to the seasonal AR order, differencing degree, and MA order, respectively. The parameter s denotes the length of the seasonal cycle, such as 12 for monthly observations and 4 for quarterly observations.

SARIMA Mathematical Formula

The SARIMA model can be written using the backshift operator (B) as follows:

$$\Phi_P(\mathrm{B}^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \quad (2)$$

In this model, B is a backshift operator satisfying $BY_t = Y_{t-1}$. Y_t denotes the observed value at time t, while ε_t is a random error (white noise). The function $\phi_p(B)$ represents a non-seasonal

CC BY NC
Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

autoregressive (AR) polynomial, and $\theta_q(B)$ denotes a non-seasonal moving average (MA) polynomial. Meanwhile, $\Phi_P(B^s)$ is a seasonal AR polynomial, and $\Theta_Q(B^s)$ is a seasonal MA polynomial [26] .

Polynomial Form

The components of the model can be described as follows:

Non-seasonal AR:

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (3)$$

Non-seasonal MA:

$$\phi_q(B) = 1 + \theta_1 B + \theta_2 + \cdots + \theta_q \quad (4)$$

Seasonal AR:

$$\phi_P(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - .. - \phi_P B^{P_s} \quad (5)$$

Seasonal MA:

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \cdots + \Theta_Q B^{Q_s} \quad (6)$$

Stationarity Test

Before modeling, a stationarity test was conducted:

1. Augmented Dickey-Fuller test
   p-value < 0.05 → The time series is considered stationary.
   p-value > 0.05 → The time series is considered non-stationary.
2. Kwiatkowski-Phillips-Schmidt-Shin test
   p-value < 0.05 → The time series is considered non-stationary.
   p-value > 0.05 → The time series is considered stationary.

The simultaneous use of the ADF and KPPS tests is highly recommended because their hypotheses are complementary, thereby providing stronger validation of the data's stationarity. The predictive accuracy of forecasting approaches, including ARIMA, can be assessed by quantifying the discrepancy between actual observations and their corresponding forecasts using error metrics [27] . Forecasting accuracy was determined using MAE, RMSE, and MAPE as the primary evaluation metrics [28], [29]. MSE represents the average of the squared differences between actual and forecasted observations. RMSE, in turn, is obtained by taking the square root of the resulting MSE value. Because larger prediction errors are squared, RMSE is particularly responsive to substantial deviations and potential outliers in the data [23].

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i|^2} \quad (8)$$

$$MAPE = 100\%\frac{1}{m}\sum_{i=1}^{m}\frac{|\hat{y}_i - y_i|}{|y_i|} \quad (9)$$

In this formulation, the actual observations are represented by y_i, the corresponding predictions by $\hat{y}_i$, and m denotes the total number of observations. A linear programming formulation is employed to determine how goods should be distributed among multiple destinations while satisfying supply and demand requirements at the lowest possible transportation cost [30] . The goal of this model is to optimally allocate resources so that the total distribution cost is minimized [31] . In this system, there are three main components: sources (supply), destinations (demand), and transportation costs between them [32] . The transportation model is formulated as follows:

Objective Function (Cost Minimization):

$$\min Z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \quad (10)$$

Supply constraint:

$$\sum_{i=1}^{n} x_{ij} = s_i \quad \forall_i \quad (11)$$

Demand constraint:

$$\sum_{i=1}^{m} x_{ij} = d_j \quad \forall_j \quad (12)$$

Non-negative:

$$x_{ij} \geq 0 \quad (13)$$

In the transportation model, x_{ij} denotes the amount shipped from source i to destination j. The parameter c_{ij} specifies the unit transportation cost along this route, whereas s_i and d_j represent the supply available at source i and the demand required at destination j, respectively [27].

VAM Penalty Formula:

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

$$P_i \geq c_{i(2)} - c_{i(1)}, for\ each\ row\ i \quad (14)$$

$$P_j \geq c_{(2)j} - c_{(1)j}, for\ each\ row\ j \quad (15)$$

In this notation, the first- and second-lowest costs in row i are denoted by $c_{i(1)}$ and $c_{i(2)}$, respectively, while , $c_{(1)j}$ and $c_{(2)j}$ represent the corresponding lowest and second-lowest costs in column j [28]. Cell Selection Rules. Given the following formula:

$$P^* = \max\{P_i, P_j\} \quad (16)$$

If the maximum penalty comes from row i^*, then choose:

$$j^* = \arg\min_j(c_{i^*j}) \quad (17)$$

If it is from column j^*, then:

$$i^* = \arg\min_i(c_{ij^*}) \quad (18)$$

Allocation Formula: The amount allocated to the selected cell:

$$x_{i^*j^*} = \min(s_{i^*}, d_{j^*}) \quad (19)$$

Subsequently, the values are updated [27]:

$$s_{i^*} = s_{i^*} - x_{i^*j^*} \quad (20)$$

$$d_{j^*} = d_{j^*} - x_{i^*j^*} \quad (21)$$

Termination Condition: The iteration process stops if,

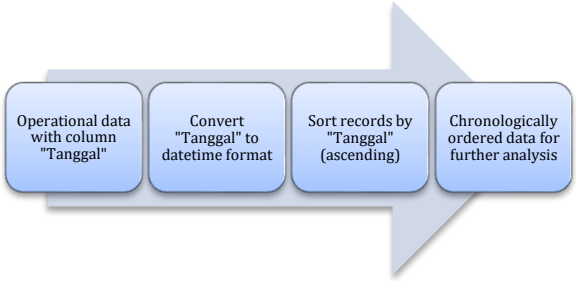
$$\sum_{i=1}^{m} s_i = 0 \ \ and \ \ \sum_{j=1}^{n} d_j = 0 \quad (22)$$

or once all requirements have been fully satisfied [29]. Total Cost of the Solution:

$$Z_{VAM} = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \quad (23)$$

RESULTS AND DISCUSSION

This section reports the findings obtained from data preparation, exploratory data analysis, data stationary, SARIMA modelling, SARIMA model evaluation, forecasting, transportation modelling, and hybrid model evaluation. Figure 5 is used to ensure that the data is neatly organized by time so that it is ready for further analysis, such as visualization or time series calculations.



Source: (Research Results, 2026)
Figure 5. Date Formatting & Data Sorting

Figure 6 presents the boxplots of demand for the three warehouses.



Source: (Research Results, 2026)
Figure 6. Boxplots of demand by Warehouse

Warehouse A shows a moderate number of outliers (more than Warehouse B but fewer than Warehouse C), with a medium-sized box and whiskers that are longer than those of Warehouse C. Warehouse B, on the other hand, has only a few outliers, along with a relatively taller box and moderately long whiskers. Even though its spread is somewhat wide, the rarity of extreme values suggests that its data is relatively consistent. In contrast, Warehouse C displays many outliers despite having a short box and short whiskers, indicating that while most data points are concentrated, frequent extreme spikes make it less stable. Overall, Warehouse B is the most stable, followed by Warehouse A, and then Warehouse C as the least stable.

The Augmented Dickey-Fuller (ADF) test results indicate that the p-values obtained for Warehouse A were 0.5068, for Warehouse B 0.4875, and for Warehouse C 0.5112. As the p-values exceed the 0.05 significance threshold, the null hypothesis (H_0) is not rejected. Thus, the time series data used for Warehouses A, B, and C are still non-stationary and still contain unit roots. Therefore, data transformation, such as differencing, is necessary before further modeling.

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

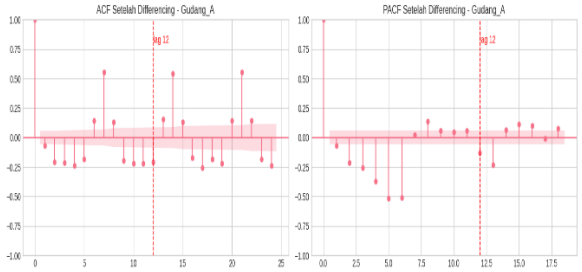
VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

Since the initial stationarity test results indicated the data were non-stationary, a first-order differencing process was performed, followed by a repeat ADF test. After the first differencing process, the ADF test results showed p-values of 0.00 for Warehouse A, 0.00 for Warehouse B, and 0.00 for Warehouse C. Since the p-values are smaller than the significance level of 0.05, the null hypothesis is rejected. Thus, the time series data used is stationary. This is further supported by the results of the Kwiatkowski-Phillips-Schmidt-Shin (KPPS) test, which yielded p-values of 0.10 for Warehouse A, 0.10 for Warehouse B, and 0.10 for Warehouse C. As the p-values exceed the 0.05 significance level, the null hypothesis is not rejected, indicating that the data are stationary. The simultaneous use of both the ADF and KPPS tests is recommended because their hypotheses are complementary, thereby providing stronger validation of the data's stationarity [33] . Thus, the data satisfy the stationarity assumption and can be used for SARIMA modeling [34] .

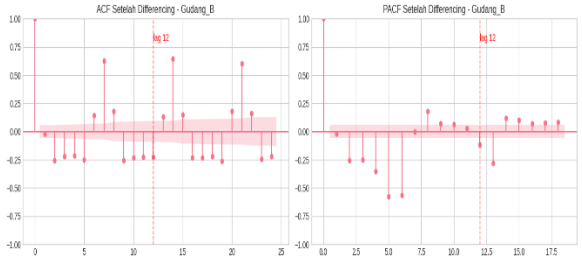
After confirming stationarity, ACF and PACF analyses guided model structure identification. Significant ACF and PACF lags guided the identification of seasonal and non-seasonal AR and MA terms. These identified configurations formed a set of candidate SARIMA models for further estimation and validation [35], [36]. Figures 7–9 present the post-differencing ACF and PACF plots for Warehouses A, B, and C

.



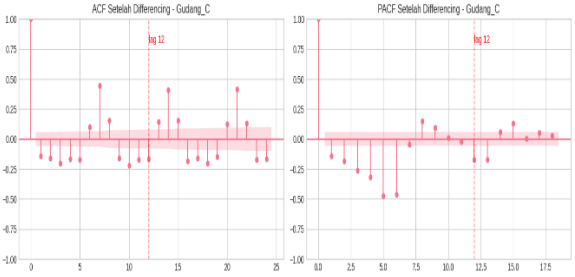
Source: (Research Results, 2026)

Figure 7. ACF and PACF plots of Warehouse A following first-order differencing



Source: (Research Results, 2026)

Figure 8. ACF and PACF plots for Warehouse B following first-order differencing



Source: (Research Results, 2026)

Figure 9. ACF and PACF plots for Warehouse C following first-order differencing.

After confirming stationarity, the SARIMA model structure was determined based on the ACF and PACF plots (Figures 7–9). Significant spikes at both non-seasonal and seasonal lags were considered in determining the AR, MA, seasonal AR, and seasonal MA orders. The observed patterns were used to formulate a set of SARIMA candidate models, which were subsequently evaluated through parameter estimation and diagnostic procedures following the Box–Jenkins framework [37] . The ACF plot reveals several statistically significant spikes at lags 7, 13–14, and 20–21. At the seasonal lag (around 12), there is a fairly prominent positive spike.

This indicates the presence of a seasonal MA (Q) component. Meanwhile, the PACF plot shows significant spikes from the initial lags up to around lag 6. At lag 12, this spike is not as strong as those at earlier lags. This suggests that the seasonal AR(P) component is likely weak or zero. Based on these patterns, several candidate SARIMA models were formulated based on the identified autocorrelation patterns. These candidate models were constructed by combining the order values of AR(p), MA(q), seasonal AR(P), and seasonal MA(Q) obtained from the interpretation of the ACF and PACF plots. Using Python, a total of 36 candidate models were obtained for Warehouse A, 16 candidates for Warehouse B, and 36 candidates for Warehouse C.

The parameters corresponding to each candidate SARIMA specification were obtained using the Maximum Likelihood Estimation (MLE) technique, while model suitability was further examined using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Among the candidate models, the specification exhibiting the smallest AIC and BIC values was selected, provided that the resulting residuals fulfilled the white-noise requirement. Table 3 summarizes the results obtained from evaluating the candidate SARIMA models across the respective warehouses.

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

Table 3. The Best SARIMA Models Based on AIC and BIC in Warehouse A, B & C

Ware house	Order	Seasonal Order	AIC	BIC
A	(2,1,2)	$(1,0,1)_{12}$	8874,099484	8909,718345
B	(2,1,2)	$(0,0,1)_{12}$	9147,743783	9178,274236
C	(2,1,2)	$(1,0,1)_{12}$	8684,340768	8719,959629

Source: (Research results, 2026)

The best model in Table 3 was selected because it had the lowest AIC and BIC values compared to all other candidates in each warehouse. The SARIMA model's forecasting performance was evaluated using MAE, RMSE, and MAPE, where lower metric values reflect greater prediction accuracy and superior model performance. Residual diagnostics were subsequently performed to verify that the residuals behaved as white noise, characterized by a mean close to zero, constant variance, and the absence of significant autocorrelation. Model adequacy was examined using residual diagnostics, autocorrelation function (ACF) plots, and the Ljung–Box test to verify that the selected model provided a reliable representation of the data [38] . The model performance and residual diagnostic results are presented in Table 4.

Table 4. Performance & Residual Diagnostic Results

Ware house	MAE	RMSE	MAPE	Ljung_Box pvalue
A	9,694903	10,821878	5,545978	$9{,}07 \times 10^{-110}$
B	9,624264	12,022988	5,109831	$5{,}71 \times 10^{-141}$
C	8,284591	9,521627	4,533927	$4{,}12 \times 10^{-80}$

Source: (Research results, 2026)

The predictive accuracy of the SARIMA models is reported in Table 4 based on MAE, RMSE, and MAPE values. Overall, the models demonstrated satisfactory forecasting accuracy, with MAPE values below 10%, reflecting a very high level of prediction accuracy. Of the three warehouses, Warehouse C achieved the best forecasting performance, with an MAE of 8.2846, an RMSE of 9.5216, and a MAPE of 4.5339. These values are the lowest compared to the other warehouses, so the SARIMA model for Warehouse C has the smallest prediction error and is capable of producing the most accurate forecasts. Meanwhile, Warehouse B has the highest RMSE value of 12.0230, indicating that the model for that warehouse produces greater variation in prediction errors compared to Warehouse A and Warehouse C. Warehouse A, on the other hand, yields an MAE of 9.6949, an RMSE of 10.8219, and a MAPE of 5.5460%, which still indicates good forecasting performance.

Nevertheless, the results of the Ljung–Box test show that all models have very small p-values: 9.07×10^{-10} for Warehouse A, 5.71×10^{-141} for Warehouse B, and 4.12×10^{-80} for Warehouse C. All of these values are far smaller than the significance level of 0.05, so the null hypothesis stating that the residuals have no autocorrelation is rejected. Thus, the residuals in all three models still contain significant autocorrelation and do not meet the white noise assumption.

Overall, the SARIMA model achieved strong forecasting performance across the three warehouses, with Warehouse C showing the smallest prediction error. Despite the satisfactory forecasting results, the Ljung–Box test detected statistically significant autocorrelation in the residuals, indicating that certain systematic patterns remained unaccounted for by the model. The optimal SARIMA model was then utilized to generate demand forecasts for each warehouse. The forecasted demand values for Warehouse A, B and C for the subsequent 3-day period are provided as an example.

Table 5. Demand Forecast Results for Warehouse A

No	Ware house	Date	Fore cast	Lower Confd. Interval	Upper Confd. Interval
1	A	2026-04-29	175	156	194
2	A	2026-04-30	178	156	200
3	A	2026-05-01	178	156	200
4	B	2026-04-29	192	170	213
5	B	2026-04-30	196	171	221
6	B	2026-05-01	196	172	221
7	C	2026-04-29	175	157	192
8	C	2026-04-30	179	159	198
9	C	2026-05-01	181	161	201

Source: (Research results, 2026)

According to the forecast generated by the selected SARIMA model, as shown in Figure 5, forecasts indicate relatively consistent demand across Warehouses A, B, and C during the following three-day period. The forecast values for Warehouse A increased slightly from 175 units on April 29, 2026, to 178 units on April 30, 2026, and 178 units on May 1, 2026, indicating a mild upward trend.

The forecast value for Warehouse B increased slightly from 192 units on April 29, 2026, to 196 units on April 30, 2026, and 196 units on May 1, 2026, indicating a slight upward trend. Similarly, in Warehouse C, the figure increased from 175 units on April 29, 2026, to 179 units on April 30, 2026, and 181 units on May 1, 2026, indicating a slight upward trend. The confidence intervals for the three related warehouses indicate that actual demand is likely to remain within the projected

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

range, while their relatively uniform width reflects a consistent level of forecast uncertainty. Overall, these results indicate stable demand conditions with no significant fluctuations expected during the forecast period.

Demand estimates obtained from the finalized SARIMA model serve as the basis for the subsequent distribution planning process. This process is formulated as a linear programming problem, with the objective of determining a cost-efficient allocation of goods while meeting all supply and demand requirements. The resulting solution provides the most cost-effective distribution plan for allocating products from the plants to the warehouses [30], [31].

Existing research has indicated that Vogel's Approximation Method (VAM) provides an effective approach for addressing transportation problems characterized by multiple sources of supply and demand destinations, particularly in generating efficient initial allocation solutions and reducing distribution costs [39] . In this study, however, the transportation model was solved using the Vogel Approximation Method (VAM) due to its ability to generate feasible and high-quality initial solutions. The use of VAM can result in an initial solution that more closely approximates the optimal transportation cost than those generated by the Northwest Corner and Least Cost methods, potentially requiring fewer computational steps for further optimization [40] .

As an illustration, the minimum transportation cost for April 17, 2026 was determined using the shipping cost matrix presented in Table 2. Based on the forecasting results, the demand quantities were estimated at 180 units for Warehouse A, 202 units for Warehouse B, and 178 units for Warehouse C. On the supply side, the available inventories were 180 units at Plant A, 195 units at Plant B, and 185 units at Plant C. Equality between the available supply and required demand indicates that the transportation problem is balanced, allowing the optimization procedure to be applied directly without the need for dummy sources or destinations. The optimal transportation results at minimum daily cost are shown in Table 6 below.

Table 6. The Optimal Transportation Results

	Warehouse A	Warehouse B	Warehouse C
Plant A	180	0	0
Plant B	0	195	0
Plant C	0	7	178

Source: (Research results, 2026)

Based on the results obtained using VAM in Table 6, to optimize shipments and minimize costs on April 17, 2026, 180 units must be shipped from Plant A to Warehouse A, 195 units from Plant B to Warehouse B, 7 units from Plant C to Warehouse B, and 178 units from Plant C to Warehouse C. The evaluation of the SARIMA–Transportation hybrid model was conducted across three main aspects. First, forecasting performance was evaluated using error metrics such as MAE, RMSE, and MAPE to measure the accuracy of demand forecasts generated by the SARIMA model.

Furthermore, the transportation model was evaluated based on its ability to fulfill warehouse demand while efficiently allocating inventory and avoiding shortages or surpluses. Third, distribution efficiency is assessed based on the total transportation costs resulting from the optimization process. Thus, the hybrid model is deemed successful if it can produce accurate demand forecasts, meet all demand and inventory constraints, and minimize transportation costs.

In this study, the proposed hybrid SARIMA–Transportation model demonstrated good performance in both demand forecasting and distribution optimization. The SARIMA model achieved high forecasting accuracy, with MAPE values ranging from 4.53% to 5.55%, indicating excellent predictive capability. The forecasted demand values are then used as input for the transportation model, which is solved using the Vogel Approximation Method (VAM).

The optimization results yield a minimum transportation cost of Rp. 2,045,000,000.00 while simultaneously satisfying all demand and warehouse inventory constraints. The optimal distribution plan assigns 180 units from Plant A to Warehouse A, 195 units from Plant B to Warehouse B, 7 units from Plant C to Warehouse B, and 178 units from Plant C to Warehouse C. Since all warehouse demand is fully met without exceeding the available inventory at each plant, the proposed hybrid model successfully integrates demand forecasting and transportation optimization. These findings demonstrate that the SARIMA–Transportation hybrid approach can support efficient logistics planning by providing accurate demand forecasts and cost-effective distribution decisions.

CONCLUSION

This research developed an integrated forecasting and logistics optimization approach, combining SARIMA for demand prediction with the Transportation Model and Vogel's Approximation Method (VAM) for determining efficient distribution plans. The proposed approach

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

successfully generated accurate demand forecasts for all warehouses, with MAPE values ranging from 4.53% to 5.55%, indicating excellent forecasting performance. These forecast results were subsequently incorporated into the transportation optimization model, which produced a cost-effective distribution plan while satisfying all warehouse demand and plant inventory constraints. The optimization process resulted in a minimum daily transportation cost of Rp. 2,045,000,000, demonstrating the practical effectiveness of integrating forecasting and optimization into a unified logistics decision-support framework.

Although the SARIMA model achieved satisfactory forecasting accuracy, the Ljung–Box test detected dependence among the residuals, implying that the selected model was unable to account for all systematic patterns present in the observed data. This finding implies that the residuals still contain valuable information that has not been adequately captured. Therefore, future studies could consider alternative SARIMA specifications, test different parameter combinations, and incorporate advanced forecasting methods, especially integrated methods that leverage both statistical forecasting and machine learning algorithms, to better represent complex temporal dynamics and improve predictive performance.

DECLARATION OF AI USE

ChatGPT (OpenAI) was used to support English language refinement and grammatical corrections. All AI-generated suggestions were carefully reviewed and revised by the authors, who assume full responsibility for the accuracy, originality, and final content of the manuscript.

AUTHOR CONTRIBUTIONS

Each author's contribution to the research and preparation of the manuscript is shown in the table below.

Author	Contribution Roles
Wahyu Sri Utami	Supervision, Conceptualization, Methodology, Formal Analysis, Data Processing, Validation, Writing—Review & Editing.
Rr. Hajar Puji Sejati	Data Analysis, Validation.
Anita Fira Waluyo	Model Programming, Validation
Widya Setiafindari	Theoretical Optimization, Validation.

FUNDING

This research was funded by the 2026 BIMA Research Grant Program under the Fundamental Regular Research scheme (Penelitian Fundamental Reguler), Indonesia.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest related to the research, writing, or publication of this article.

REFERENCE

[1] A. Walter, K. Ahsan, and S. Rahman, "Application of artificial intelligence in demand planning for supply chains: a systematic literature review," Apr. 29, 2025, Emerald Publishing, doi: 10.1108/IJLM-02-2024-0120.

[2] M. Muth, M. Lingenfelder, and G. Nufer, "The application of machine learning for demand prediction under macroeconomic volatility: a systematic literature review," Management Review Quarterly, vol. 75, no. 3, pp. 2759–2802, Sep. 2025, doi: 10.1007/s11301-024-00447-8.

[3] T. Ngoc Anh Nguyen, T. Xuan Hoa Nguyen, N. Thang Tran, T. Ha Nguyen, P. Anh Nguyen, and H. Anh Vu, "Optimizing supply chain operations using advanced Time-Series Mixer models for demand forecasting and inventory under uncertain demand," Expert Syst. Appl., vol. 296, Jan. 2026, doi: 10.1016/j.eswa.2025.128955.

[4] C. Cheng, Q. Yu, Y. Adulyasak, and L. M. Rousseau, "Distributionally robust facility location with uncertain facility capacity and customer demand," Omega (United Kingdom), vol. 122, Jan. 2024, doi: 10.1016/j.omega.2023.102959.

[5] L. Chen, T. Dong, J. Peng, and D. Ralescu, "Uncertainty Analysis and Optimization Modeling with Application to Supply Chain Management: A Systematic Review," Jun. 01, 2023, Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/math11112530.

[6] Z. Ge, "Logistics Distribution Mode Based on Supply Chain Strategy", doi: 10.23977/socsam.2022.030102.

[7] E. Aldahmani, A. Alzubi, and K. Iyiola, "Demand Forecasting in Supply Chain Using Uni-Regression Deep Approximate Forecasting Model," Applied Sciences

CC BY NC

(Switzerland), vol. 14, no. 18, Sep. 2024, doi: 10.3390/app14188110.

[8] "8. Integration of machine learning in the supply chain for decision making_ A systematic literature review _ Polo-Triana _ Journal of Industrial Engineering and Management".

[9] S. Sifuentes-Domínguez, J. M. Mejia-Muñoz, O. Cruz-Mejia, R. Pizarro-Gurrola, A. S. Domínguez-Flores, and L. Ortega-Máynez, "Predicting Demand in Supply Chain Management: A Decision Support System Using Graph Convolutional Networks," Future Internet, vol. 18, no. 1, Jan. 2026, doi: 10.3390/fi18010026.

[10] D. P. Kartikasari, T. A. T. Tambak, and A. R. Ridwanto, "Hybrid Demand Forecasting and Monte Carlo Simulation for Retail Supply Chain Inventory Optimization," JITCoS : Journal of Information Technology and Computer System, vol. 1, no. 2, pp. 74–85, Dec. 2025, doi: 10.65230/jitcos.v1i2.40.

[11] M. A. Jahin, A. Shahriar, and M. Al Amin, "MCDFN: supply chain demand forecasting via an explainable multi-channel data fusion network model," Evol. Intell., vol. 18, no. 3, Jun. 2025, doi: 10.1007/s12065-025-01053-7.

[12] N. Tahirov, N. Akhundov, S. Emde, and C. H. Glock, "Configuration of last-mile distribution networks for an encroaching manufacturer," Ann. Oper. Res., vol. 344, no. 2–3, pp. 679–720, Jan. 2025, doi: 10.1007/s10479-024-06031-3.

[13] W. Zhao, X. Dong, and M. Li, "Hybrid optimization of logistics distribution paths using neural networks and fuzzy logic," Discover Internet of Things, vol. 5, no. 1, Dec. 2025, doi: 10.1007/s43926-025-00216-3.

[14] J. Shi, "Optimization of frozen goods distribution logistics network based on k-means algorithm and priority classification," Sci. Rep., vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-72723-2.

[15] A. Borucka, "Seasonal Methods of Demand Forecasting in the Supply Chain as Support for the Company's Sustainable Growth," Sustainability (Switzerland), vol. 15, no. 9, May 2023, doi: 10.3390/su15097399.

[16] D. Besik and A. Nagurney, "Multicommodity fresh produce trade networks with quality deterioration under congestion and transportation capacities," Journal of Global Optimization, 2025, doi: 10.1007/s10898-025-01507-3.

[17] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, "Predictive Analytics for Demand Forecasting - A Comparison of SARIMA and LSTM in Retail SCM," in Procedia Computer Science, Elsevier B.V., 2022, pp. 993–1003. doi: 10.1016/j.procs.2022.01.298.

[18] Y. Zhao, "Research on E-Commerce Retail Demand Forecasting Based on SARIMA Model and K-means Clustering Algorithm," 2024.

[19] P. Negre, R. S. Alonso, J. Prieto, Ó. García, and L. de-la-Fuente-Valentín, "Prediction of footwear demand using Prophet and SARIMA," Expert Syst. Appl., vol. 255, Dec. 2024, doi: 10.1016/j.eswa.2024.124512.

[20] P. Polina, S. Ganesan, L. Karunarathne, and N. Somasiri, "Time Series Analysis for Tractor Sales using SARIMAX and Deep Learning Models," International Journal of Computer Communication and Informatics, vol. 6, no. 1, pp. 27–57, Jul. 2024, doi: 10.34256/ijcci2413.

[21] M. M. Aksoy, M. N. Mowla, M. Bilgili, E. Pinar, T. Durhasan, and D. Asadi, "Forecasting near-surface air temperature via SARIMA and LSTM: A regional time-series study," J. Atmos. Sol. Terr. Phys., vol. 275, Oct. 2025, doi: 10.1016/j.jastp.2025.106604.

[22] I. F. Benavides, J. P. Romero-Leiton, M. Santacruz, C. Barreto, V. Puentes, and J. J. Selvaraj, "Applying seasonal time series modeling to forecast marine fishery landings for six species in the Colombian Pacific Ocean," Reg. Stud. Mar. Sci., vol. 56, Nov. 2022, doi: 10.1016/j.rsma.2022.102716.

[23] V. Kramar and V. Alchakov, "Time-Series Forecasting of Seasonal Data Using Machine Learning Methods," Algorithms, vol. 16, no. 5, May 2023, doi: 10.3390/a16050248.

[24] G. Tuğba Önder, "Comparative time series analysis of SARIMA, LSTM, and GRU models for global SF6 emission management system," J. Atmos. Sol. Terr. Phys., vol. 265, Dec. 2024, doi: 10.1016/j.jastp.2024.106393.

[25] L. H. Hasibuan, S. Musthofa, D. M. Putri, and M. Jannah, "Comparison Of Seasonal Time Series Forecasting Using SARIMA And Holt Winter's Exponential Smoothing (Case Study: West Sumatra Export Data)," Barekeng, vol. 17, no. 3, pp. 1773–1784, Sep. 2023, doi: 10.30598/barekengvol17iss3pp1773-1784.

[26] S. Bosson-Amedenu, E. M. Baah, F. Ayiah-Mensah, and J. A. Addor, "Advanced time series modelling to address seasonal variability and extreme rainfall in Ghana,"

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.8563

BMC Environmental Science, vol. 2, no. 1, Jun. 2025, doi: 10.1186/s44329-025-00024-8.

[27] Y. Balagula, D. Baimel, and I. Aharon, "Comparing time series and neural network models of long memory for electricity price forecasting," Results in Engineering, vol. 29, Mar. 2026, doi: 10.1016/j.rineng.2025.108465.

[28] W. Huang, Y. Li, Y. Zhao, and L. Zheng, "Time Series Analysis and Prediction on Bitcoin," 2022. [Online]. Available: www.coingecko.com.

[29] V. Plevris, G. Solorzano, N. P. Bakas, M. El Amine, and B. Seghier, "Investigation Of Performance Metrics In Regression Analysis And Machine Learning-Based Prediction Models," 2022.

[30] H. A. Taha, Operations Research: An Introduction, 11 editions. Hoboken, New Jersey: Pearson, 2022.

[31] X. Xu, Z. Chen, Y. Song, and W. Yi, "A distributionally robust optimization model for inventory-transportation integrated optimization problem of petroleum products under uncertainty," Comput. Oper. Res., vol. 170, Oct. 2024, doi: 10.1016/j.cor.2024.106767.

[32] S. Chopra and P. Meindl, Supply Chain Management: Strategy, Planning, and Operation, 7th ed. Harlow, U.K.: Pearson, 2022.

[33] R. L. Manogna, V. Dharmaji, and S. Sarang, "Enhancing agricultural commodity price forecasting with deep learning," Sci. Rep., vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-05103-z.

[34] M. G. S. Kenyi and K. Yamamoto, "A hybrid SARIMA-Prophet model for predicting historical streamflow time-series of the Sobat River in South Sudan," Discover Applied Sciences, vol. 6, no. 9, Sep. 2024, doi: 10.1007/s42452-024-06083-x.

[35] N. A. Komara Rifai and M. A. Zhahirulhaq, "Forecasting Inflation In Indonesia Using The Autoregressive Integrated Moving Average Method," Parameter: Journal of Statistics, vol. 4, no. 1, pp. 37–45, Jun. 2024, doi: 10.22487/27765660.2024.v4.i1.17130.

[36] P. Do, C. W. K. Chow, R. Rameezdeen, and N. Gorjian, "Wastewater inflow time series forecasting at low temporal resolution using SARIMA model: a case study in South Australia," Environmental Science and Pollution Research, vol. 29, no. 47, pp. 70984–70999, Oct. 2022, doi: 10.1007/s11356-022-20777-y.

[37] U. H. Perez-Guerra et al., "Seasonal autoregressive integrated moving average (SARIMA) time-series model for milk production forecasting in pasture-based dairy cows in the Andean highlands," PLoS One, vol. 18, no. 11 November, Nov. 2023, doi: 10.1371/journal.pone.0288849.

[38] W. Yang, A. Su, and L. Ding, "Application of exponential smoothing method and SARIMA model in predicting the number of admissions in a third-class hospital in Zhejiang Province," BMC Public Health, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12889-023-17218-x.

[39] O. Access Mes ; M, E. S. M. Nababan, P. K. Nasution, and A. Syahmarani, "Kajian Metode Vogel's Approximation dalam Optimisasi Masalah Transportasi dengan Variasi Posisi Supply dan Demand," 2024.

[40] Satakshi and V. V Henry, "A novel approach to find simple, pragmatic solutions to transportation problems," International Journal of Mathematics in Operational Research, vol. 27, no. 4, pp. 496–508, Jan. 2024, doi: 10.1504/IJMOR.2024.138467.

CC BY NC