



ISSN ONLINE: 2447-0228



RESEARCH ARTICLE

OPEN ACCESS

SHORT-TERM RICE YIELD FORECASTING IN PANGASINAN: A SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) MODELING APPROACH

Eherson G. Valdez ^{*1}, and Johannis B. Rabago²

^{1,2}College of Information Technology and Computer Science, University of the Cordilleras, Baguio City, Philippines.

¹<http://orcid.org/0009-0003-0082-2661>, ²<https://orcid.org/0009-0005-3420-7143>

E-mail: egv9500@students.uc-bcf.edu.ph, jbr5495@students.uc-bcf.edu.ph

ARTICLE INFO

Article History

Received: February 26, 2026

Reviewed: March 26, 2026

Accepted: May 15, 2026

Published: September 30, 2026

Keywords:

Rice Yield Forecasting,

Sarima,

Food Security,

Time-Series Analysis.

ABSTRACT

Rice farming in Pangasinan is critical to Philippine food security but increasingly vulnerable to climate volatility. Traditional intuition-based farming calendars are insufficient against erratic weather patterns, leaving stakeholders with reactive rather than anticipatory strategies. This study aims to develop a short-term forecasting model to empower data-driven decision-making. Utilizing quarterly rice yield data from 2008 to 2025, the study employed the Box-Jenkins methodology to formulate a Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The analysis revealed a robust "sawtooth" seasonal pattern, where the fourth quarter consistently yields peak production, while the third quarter suffers a significant lean season. Following rigorous stationarity testing and differencing, the SARIMA (0,1,1) (0,1,1)₄ model was identified as the optimal predictor (AICc: 1575.88), outperforming automated selection algorithms. Diagnostic checking confirmed the model successfully captured deterministic patterns. Projections for late 2025 forecast a robust yield of 751,139.8 MT. However, the model explicitly quantifies a critical vulnerability in the third quarter of 2026, projecting a precipitous drop to 112,620.2 MT or 85% contraction from peak capacity. These findings provide agricultural authorities with the foresight needed to optimize resource allocation and mitigate the economic risks of seasonal scarcity.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

Rice farming is the cornerstone of agriculture in the Philippines, profoundly influencing national food security and local economic stability. The Province of Pangasinan is a significant contributor to the nation's rice production [1], consistently ranking among the top producers in the country. In fact, according to recent data from the Department of Agriculture [2], Pangasinan produced approximately 1.23 million metric tons of rice in 2023, securing its position as the third-largest rice-producing province nationwide. However, maintaining this level of productivity is becoming increasingly difficult due to the volatility introduced by climate variability and shifting weather patterns. Traditional farming calendars, which rely heavily on historical intuition, are proving insufficient against these erratic environmental changes. Consequently, the lack of precise yield prediction tools leaves agricultural stakeholders "blind" to future outcomes, forcing them into a reactive stance where decisions on irrigation, fertilizer application, and logistics are made only after visible stress occurs. Specifically, current practices do not leverage advanced time-series analytics, leading to forecasts that overlook critical patterns, seasonality, and the measurable impact of climate data [3]. To optimize resource allocation and mitigate the risks of environmental volatility, a shift toward data-driven decision-making is essential. This necessitates the rigorous processing of historical data to ensure it meets the technical assumptions required for advanced time-series modeling. Without the empirical foundation provided by an accurate and validated forecasting model, agricultural stakeholders lack the necessary foresight to implement anticipatory, rather than reactive, interventions. This study addresses this limitation by proposing a formal time-series modeling approach using the Seasonal Autoregressive Integrated Moving Average (SARIMA) methodology. Standard ARIMA applications have typically focused on non-seasonal trends, such as previous frameworks utilized for localized dengue outbreak forecasting [4].

Unlike standard ARIMA models which primarily model linear trends, or ARIMAX models which require external weather data that may be unavailable or delayed, SARIMA is designed specifically to model data with repeating cycles. This methodology has been extensively validated in agricultural literature as a superior tool for crops influenced by phenological cycles. For instance, recent studies have successfully applied SARIMA to forecast maize (corn) and paddy yields in regions with distinct monsoon climates, demonstrating its capacity to outperform static regression models by accurately capturing the stochastic nature of seasonal harvests [5] [6]. However, while the SARIMA model has been applied successfully in predicting seasonal phenomena globally, ranging from infectious disease outbreaks to agricultural outputs in other nations, there is a distinct gap in its localized application for specific regions like Pangasinan. To date, no formal study has calibrated these parameters specifically for the province's unique agro-climatic conditions [7]. Given that rice farming in Pangasinan follows a distinct wet and dry season calendar, the historical yield data exhibits strong autocorrelation; meaning past seasonal performance is a strong predictor of future yield, provided the model is rigorously adapted to the local context. The general objective of this study is to develop and validate a short-term rice yield forecasting model using the Seasonal Autoregressive Integrated Moving Average (SARIMA) methodology for smart agricultural planning in Pangasinan. Specifically, this study aims to identify the statistical properties and seasonal patterns of historical rice yield in Pangasinan and determine the optimal SARIMA model parameters, and generate short-term rice yield projections for the remaining quarters of 2025 and 2026. To achieve these objectives, this study will answer the following principal questions:

1. What are the statistical properties and seasonal patterns of historical rice yield in Pangasinan, and what are the appropriate SARIMA model parameters required to capture these trends?
2. How does the developed SARIMA model perform in terms of forecast accuracy and what are the projected rice yields for Pangasinan for the upcoming quarters?

The significance of this study lies in its potential to empower agricultural authorities and local policymakers to shift from reactive monitoring to proactive management of the province's rice supply. By accurately forecasting quarterly yield surges and lean periods, stakeholders can optimize the allocation of critical resources more efficiently, thereby mitigating the economic risks associated with seasonal volatility. Furthermore, this study not only contribute to the growing body of research on applied time-series modeling in agriculture but also offer a practical, quantitative solution tailored to the unique seasonal dynamics of Pangasinan.

II. METHODOLOGY

II.1 RESEARCH DESIGN

This study employs a quantitative, non-experimental research design utilizing secondary time-series data. The primary analytical framework is the Box-Jenkins methodology, which involves an iterative process of model identification, parameter estimation, diagnostic checking, and forecasting.

II.2 DATA SOURCE AND PREPARATION

The historical data for quarterly rice yield (in Metric Tons) in Pangasinan was obtained from the Philippine Statistics Authority (PSA) / OpenSTAT database [8]. The dataset covers the period from 2008 to the second quarter of 2025. Prior to modeling, the data was pre-processed to ensure data integrity. This involved imputation, wherein checking for missing values (NA) and handling them using interpolation, occurs. Followed by partitioning, wherein the dataset was split into a Training Set (80% of the data) for model formulation and a Testing Set (20% of the data) for validating the forecast accuracy.

II.2 STATISTICAL TREATMENT

II.2.1 Identification and Stationary Testing

The first step involves analyzing the statistical properties of the time series. The presence of seasonality and trend was visualized using time-series plots and decomposition. To satisfy the assumption of stationarity, the Augmented Dickey-Fuller (ADF) test was employed. If the data was found to be non-stationary, differencing (d) and seasonal differencing (D) were applied until the series became stationary. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were examined to identify potential values for the Autoregressive (p, P) and Moving Average (q, Q) terms.

II.2.2 Model Estimation and Selection

Various candidate models were generated based on the parameters identified in Phase 1. The specific model structure is denoted as:

$$\text{SARIMA (p,d,q) } \llbracket \text{(P,D,Q)} \rrbracket _s$$

Where: (p, d, q) represent the non-seasonal order, and (P, D, Q) represent the seasonal order. The s represents the seasonality period (s=4 for quarterly data). The optimal model was selected based on the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), as these metrics balance model fit with model simplicity.

II.2.3 Diagnostic Checking

The selected model underwent rigorous diagnostic testing to ensure the residuals were white noise. The Ljung-Box Q-test was utilized to test for remaining autocorrelation in the residuals. A p-value greater than 0.05 in the Ljung-Box test indicates that the model has successfully captured the data's structure.

II.2.4 Forecasting and Accuracy Evaluation

The trained model was used to generate forecasts for the test set and future quarters (2025-2026). Forecast accuracy was evaluated using standard error metrics: Root Mean Squared Error (RMSE), to measure the magnitude of the prediction errors. Mean Absolute Percentage Error (MAPE), to express the error as a percentage, offering an intuitive measure of how far off the predictions are on average.

III. RESULTS AND DISCUSSION

This section details the empirical findings regarding historical rice yield data in Pangasinan. The discussion begins with an analysis of the data's statistical properties and the determination of optimal model parameters, followed by an evaluation of predictive accuracy and the presentation of future yield projections.

III.1 STATISTICAL TREATMENT

As part of the analysis, an exploratory data analysis (EDA) was conducted on the quarterly rice yield records from 2008 to 2025 in order to identify the statistical properties of the dataset. Visual inspection of the time series, as presented in Figure 1, reveals a stochastic trend coupled with a robust seasonal component. The historical data exhibits a distinct "sawtooth" pattern, characterized by sharp oscillations between production peaks and troughs. The long-term trend line, as presented by a red dashed line, indicates a gradual linear increase in base production levels from 2008 through 2018, followed by a stabilization period in recent years. This behavior suggests that the province remains heavily dependent on the cyclical nature of the cropping calendar.

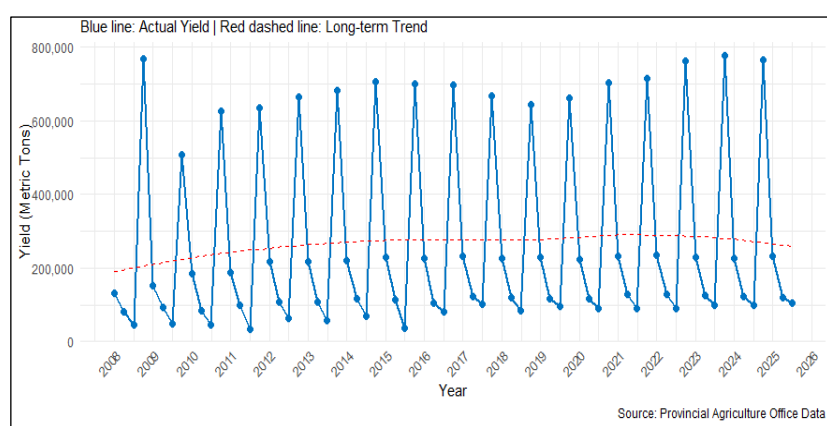


Figure 1: Historical Quarterly Rice Yield in Pangasinan (2008–2025).

Source: Authors, (2026).

Moreover, to scientifically quantify the seasonality observed in the time series, the distribution of yields was aggregated by quarter. Figure 2 presents the boxplot analysis, which provides statistically significant evidence of heteroscedasticity across quarters. The Fourth Quarter composed of October to December, is identified as the extreme outlier, consistently recording the highest production volumes. The tight interquartile range suggests high reliability in this harvest period. Conversely, the Third Quarter (July–September) exhibits the lowest median yield. This statistical trough aligns with the vegetative and reproductive phases of the rice plant in the Philippines, during which no significant harvest occurs [9].

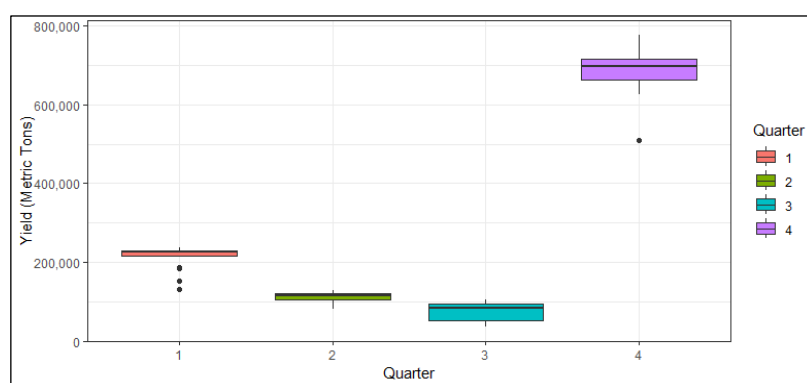


Figure 2: Historical Quarterly Rice Yield in Pangasinan (2008–2025).

Source: Authors, (2026).

To scientifically quantify the seasonality observed in the time series, the distribution of yield was aggregated by quarter as visually demonstrated in Figure 2. As seen in the figure, the Fourth Quarter (Q4) represents the primary harvest season, peaking at a median of approximately 700,000 MT and serving as a significant outlier compared to the rest of the year. In sharp contrast, the Third Quarter (Q3) reflects the critical "lean season" with a median yield plummeting below 100,000 MT. Meanwhile, the First Quarter (Q1) indicates a moderate recovery acting as a secondary harvest with an average yield hovering around 225,000 MT, followed by a contraction in the Second Quarter (Q2) to approximately 120,000 MT.

The magnitude of the difference between the Q4 peaks and Q3 troughs confirms that seasonality is the dominant statistical property of the dataset, directly supporting the methodological decision to employ a Seasonal ARIMA (SARIMA) model rather than a standard linear regression [10], which would fail to account for such rigorous periodic variance.

III.2 STATIONARITY ANALYSIS AND DATA TRANSFORMATION

To satisfy the fundamental assumptions of the ARIMA framework, the analysis began by verifying that the data exhibited stationarity. To assess this, the Augmented Dickey-Fuller (ADF) test was applied to the data at two stages: first on the raw historical yield and subsequently on the seasonally differenced data. The results of this analysis are summarized in Table 1.

Table 1: Augmented Dickey-Fuller (ADF) Test Results.

Data Transformation Stage	ADF Test	p-value	Decision	Remarks
Raw Data (Level)	-3.4362	0.05702	Fail to Reject H_0	Non-Stationary
Seasonally Differenced (Lag 4)	-6.0333	0.01000	Reject H_0	Stationary

Source: Authors, (2026).

As shown in the table 1, the test on the Raw Data yielded a p-value of 0.05702, which exceeds the standard significance level of 0.05. This failure to reject the null hypothesis statistically confirms that the original time series is non-stationary, containing significant time-dependent structures such as trends and seasonal cycles that prevent direct modeling. In order to stabilize the variance and mean, Seasonal Differencing ($D=1$) at Lag 4 was applied. The subsequent test on this Differenced Data yielded a highly significant test statistic of -6.0333 and a p-value of 0.01. Being well below the 0.05 threshold, this confirms that the transformation successfully removed the unit root. The data is now stationary, satisfying the mathematical assumptions necessary for the subsequent training phase of the SARIMA model [11]. The visual effect of this transformation is illustrated in Figure 3, where the series now oscillates around a constant mean of zero, indicating the removal of the trend and seasonal components.

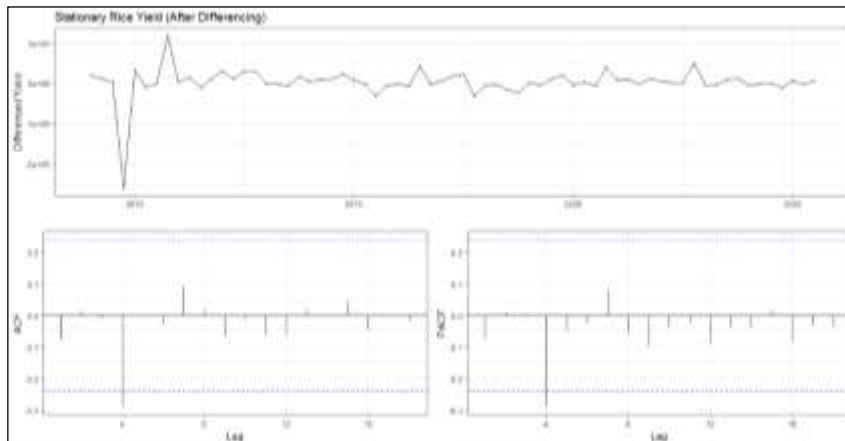


Figure 3: Stationary Rice Yield (After Differencing).

Source: Authors, (2026).

III.3 MODEL IDENTIFICATION AND SELECTION

Prior to training, it was necessary to identify the optimal model structure (p, d, q) (P, D, Q) s. Two selection strategies were rigorously evaluated using the Akaike Information Criterion (AIC) and Corrected AIC (AICc) as the primary metrics for model parsimony. First, the automated algorithm, or the `auto.arima` function, was employed to generate a baseline candidate. Concurrently, a manual model structure was specified based on the visual examination of the autocorrelation structure. The quantitative comparison of the two candidate models is summarized in Table 2. The automated function proposed a model with drift, resulting in an AICc of 1591.77. However, the manually constructed candidate yielded a significantly lower AICc of 1575.88. According to statistical conventions, a lower AICc indicates a superior model fit with fewer parameters [12] [13]. Consequently, the manual model SARIMA (0,1,1) (0,1,1)₄ was selected as the optimal predictor for Pangasinan rice yields.

Table 2: Comparison of Candidate Models.

Selection Strategy	Model Structure	AICc Value	Selection Decision
Automated Algorithm	ARIMA (0,0,0) (0,1,1) ₄	1591.77	Rejected
Manual Selection	SARIMA (0,1,1) (0,1,1) ₄	1575.88	Selected (Optimal)

Source: Authors, (2026).

To corroborate this selection, the diagnostic plots of the chosen model's residuals were generated, as displayed in Figure 4. The figure presents a comprehensive validation across three dimensions. First, the time plot demonstrates that the residuals oscillate randomly around zero with constant variance, indicating no remaining trend or volatility clustering. Second, the ACF plot confirms that no significant autocorrelation remains, as all lag correlations fall strictly within the blue dashed lines or a 95% confidence limits. Third, the histogram reveals that the forecast errors approximate a normal distribution curve centered at zero.

These visual diagnostics collectively confirm that the selected SARIMA (0,1,1) (0,1,1)₄ model has successfully captured the deterministic patterns in the data, leaving behind only random white noise.

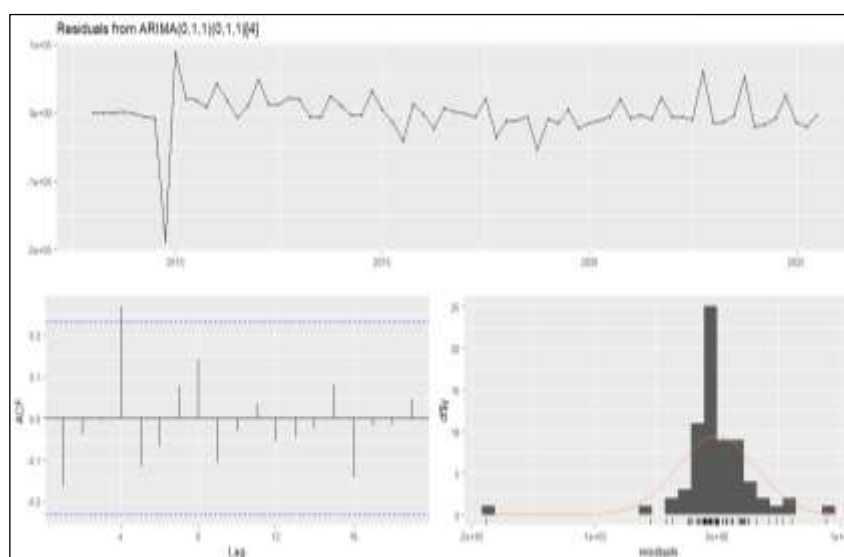


Figure 4: Residual Diagnostics of the Best Model SARIMA (0,1,1) (0,1,1)₄.
Source: Authors, (2026).

III.4 SHORT-TERM FORECAST PROJECTIONS

The validated SARIMA model was deployed to generate short-term projections for the remainder of 2025 and the full year of 2026. The forecast results, including 80% and 95% confidence intervals, are visualized in Figure 5 and detailed in Table 3.

III.4.1 Immediate Forecast and Year-on-Year Growth

For the immediate horizon, the model projects a robust yield for the Fourth Quarter of 2025, estimated at 751,139.8 MT. The 95% confidence interval for this period suggests a high degree of certainty, with the lower bound ensuring a minimum production of at least 685,260.8 MT. Looking ahead to the same period in the following year, the model forecasts a yield of 755,589.8 MT for Q4 2026. This represents a marginal year-on-year growth of approximately 4,450 MT (+0.6%), indicating that while the province's production capacity is stable and slightly increasing, it remains heavily dictated by the seasonal calendar rather than rapid exponential growth.

III.4.2 Analysis of the Seasonal Trough (Lean Season)

The 2026 projections provide a granular view of the province's agricultural vulnerability during the mid-year cycle. Following the secondary harvest in Q1 2026 (246,966.4 MT), production is forecasted to contract significantly. The model identifies Q3 2026 as the critical "lean season," with a projected yield of only 112,620.2 MT. This represents a staggering 85% decline compared to the peak Q4 production. This quantitative evidence mathematically defines the "hunger season" often experienced in rice-producing regions, confirming that during the third quarter, the province produces less than one-sixth of its peak capacity.

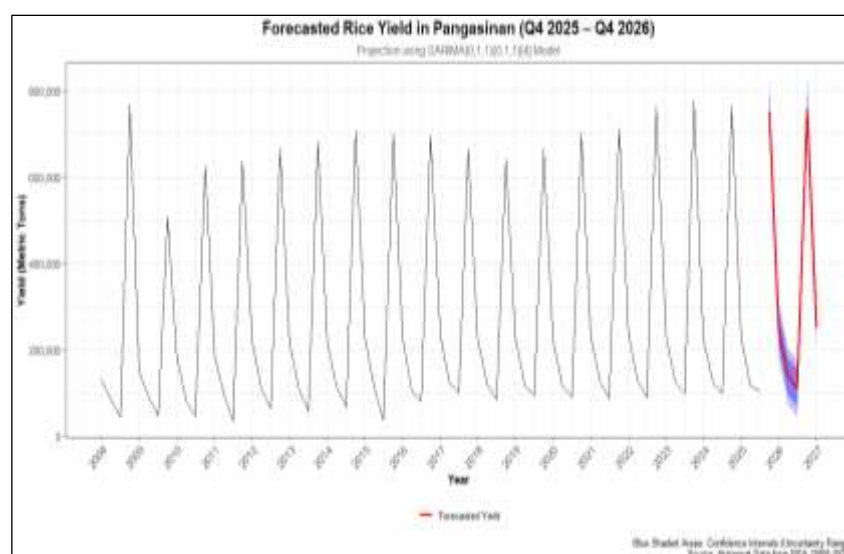


Figure 4: Forecasted Rice Yield in Pangasinan (2025-2026).
Source: Authors, (2026).

Table 3: Short-Term Rice Yield Forecast.

Period	Point Forecast (MT)	Lower 95% CI (MT)	Upper 95% CI (MT)
2025 Q4	751,139.8	685,260.8	817,018.7
2026 Q1	246,966.4	181,082.7	312,850.1
2026 Q2	138,803.6	72,919.9	204,687.3
2026 Q3	112,620.2	46,736.5	178,503.9
2026 Q4	755,589.8	687,382.6	823,797.1

Source: Authors, (2026).

III.5 FORECAST ACCURACY EVALUATION

To strictly evaluate the predictive performance of the SARIMA (0,1,1) (0,1,1)₄ model, a comprehensive set of statistical metrics was generated using the model's residuals. These metrics include the Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and the Autocorrelation of errors at lag 1 (ACF1). Table 4 summarizes this accuracy diagnostics derived from the model training set.

Table 4: SARIMA Model Accuracy Results and Evaluation Metrics.

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-892.0526	31,556.84	18,131.01	-2.1901%	10.3908%	1.0602	-0.1645

Source: Authors, (2026).

The diagnostic results indicate a robust model fit. The ME of -892.05 and MPE of -2.19% suggest a negligible bias [13], indicating that the model has a very slight tendency to over-forecast, but the magnitude is statistically insignificant relative to the total production volume, which often exceeds 700,000 MT. The RMSE of 31,556.84 and MAE of 18,131.01 reflect the stochastic nature of agricultural data. However, due to the scale-dependent of these metrics [14], it must be interpreted relative to the magnitude of the target variable. Given the high volume of rice production in Pangasinan, which often exceeding 700,000 MT, these deviations represent less than 11% of the total output, placing them well within acceptable operational limits for agricultural planning. Furthermore, the ACF1 value of -0.1645 is close to zero, confirming that the residuals have low autocorrelation [15] and resemble white noise, implying that the model has successfully extracted the available signal from the time series data [16]. To quantify the overall reliability of the model, the MAPE was utilized as the primary determinant. The model yielded a MAPE of 10.39%, which implies that, on average, the forecasted values deviate from the actual historical data by only roughly 10%. Following standard forecasting evaluation protocols, the overall Predictive Accuracy percentage was derived using the inverse of the MAPE, calculated as follows:

Accuracy (%) = 100 - MAPE.

Accuracy (%) = 100 - 10.39079.

Accuracy (%) = 89.60921 ≈ 89.61%.

With a computed predictive accuracy of 89.61%, the model demonstrates a high degree of reliability [17]. In the context of agricultural forecasting, where yield is heavily dependent on variable factors such as weather patterns and pest incidence, achieving an accuracy rate approaching 90% is considered highly significant [18] [19]. An accuracy of 89.61% validates that the SARIMA model effectively captures the complex seasonal dynamics of rice production in Pangasinan, rendering it a dependable tool for anticipating future supply trends and lean periods.

IV. CONCLUSIONS

This study successfully established a data-driven forecasting framework for Pangasinan, confirming that rice yield follows a deterministic seasonal structure rather than random fluctuations. The analysis revealed a distinct "sawtooth" pattern in historical production, where the SARIMA (0,1,1) (0,1,1)₄ model proved to be an optimal predictor with a predictive accuracy of 89.61%. This high level of accuracy validates that statistical time-series modeling can effectively substitute for intuition-based calendars, which are increasingly unreliable due to climate variability. Crucially, the model explicitly quantifies a severe "lean season" vulnerability in the Third Quarter of 2026, projecting a yield contraction to 112,620.2 MT, which is an 85% drop from peak capacity. Consequently, this identified gap between the fourth-quarter peak and the third-quarter trough serves as a critical baseline for scheduling warehouse stocking and irrigation support well in advance. Ultimately, these findings empower agricultural authorities to secure the province's food supply by transitioning from reactive observation to anticipatory, data-driven governance.

V. AUTHOR'S CONTRIBUTION

Conceptualization: Eherson G. Valdez, and Johanns B. Rabago.

Methodology: Eherson G. Valdez and Johanns B. Rabago.

Investigation: Eherson G. Valdez and Johanns B. Rabago.

Discussion of results: Eherson G. Valdez, and Johanns B. Rabago.

Writing – Original Draft: Eherson G. Valdez.

Writing – Review and Editing: Eherson G. Valdez and Johanns B. Rabago.

Resources: Eherson G. Valdez and Johanns B. Rabago.

Supervision: Johanns B. Rabago.

Approval of the final text: Eherson G. Valdez, and Johanns B. Rabago.

VI. REFERENCES

- [1] Ilocos Region. (May 15, 2023). Ilocos Region 2022 rice production up at more than 3% - PSA. DA Regional Field Office 1. Retrieved from <https://ilocos.da.gov.ph/ilocos-region-2022-rice-production-up-at-more-than-3-psa>
- [2] Department of Agriculture – Regional Field Office 1. (2024, June 18). Ilocos Region 181% rice sufficient. Philippine News Agency.
- [3] Tiwari, V., Thorp, K., Tulbure, M. G., Gray, J., Kamruzzaman, M., Krupnik, T. J., ... & Ardon, M. (2024). Advancing food security: Rice yield estimation framework using time-series satellite data & machine learning. *Plos one*, 19(12), e0309982.
- [4] Mole, P., & Palaoag, T. (2026). Exploratory data analysis and forecasting of dengue outbreaks in Pangasinan using the ARIMA model. *International Journal of Informatics and Communication Technology*, 15(1), 46–56. <http://doi.org/10.11591/ijict.v15i1.pp46-56>
- [5] Filder, T. N., Muraya, M. M., & Mutwiri, R. M. (2019). Application of seasonal autoregressive moving average models to analysis and forecasting of time series monthly rainfall patterns in Embu County, Kenya.
- [6] Urrutia, J. D., Diaz, J. L. B., & Mingo, F. L. T. (2017, March). Forecasting the quarterly production of rice and corn in the Philippines: A time series analysis. In *Journal of Physics: Conference Series* (Vol. 820, No. 1, p. 012007). IOP Publishing.
- [7] Anuada, A. M. (2022). Effect of agro-climatic factors on the yield of corn (IPB Var 6) under rainfed conditions in the Philippine. *Philippine e-Journal for Applied Research and Development*, 12, 23-31.
- [8] Philippine Statistics Authority. (n.d.). Palay and Corn: Volume of Production in Metric Tons by Ecosystem/Croptype, Quarter, Semester, Region and Province, 1987-2025. Agriculture, Forestry, Fisheries – OpenSTAT. Retrieved from: https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2E__CS/0012E4EVCPO.px/
- [9] Alberto, M. C. R., Wassmann, R., Hirano, T., Miyata, A., Hatano, R., Kumar, A., ... & Amante, M. (2011). Comparisons of energy balance and evapotranspiration between flooded and aerobic rice fields in the Philippines. *Agricultural water management*, 98(9), 1417-1430.
- [10] Majka, M. (2024). Seasonal time series analysis: Why sarima outshines arima. ResearchGate. Published September, 19, 2024.
- [11] Adams, S. O., & Bamanga, M. A. (2020). Modelling and forecasting seasonal behavior of rainfall in Abuja, Nigeria; A SARIMA approach. *Am. J. Math. Stat*, 10(1), 10-19.
- [12] Ampountolas, A. (2024). Enhancing forecasting accuracy in commodity and financial markets: Insights from garch and svr models. *International Journal of Financial Studies*, 12(3), 59.
- [13] Kobayashi-Kato, M., Fujii, E., Asami, Y., Ahiko, Y., Hiranuma, K., Terao, Y., ... & Yoshida, H. (2023). Utility of the revised FIGO2023 staging with molecular classification in endometrial cancer. *Gynecologic Oncology*, 178, 36-43.
- [14] Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts: Melbourne, Australia. Section 5.8.
- [15] Weiß, C., Scherer, L., Aleksandrov, B., & Feld, M. (2020). Checking model adequacy for count time series by using Pearson residuals. *Journal of Time Series Econometrics*, 12(1), 20180018.
- [16] Sanwa, R. L., & Chi, Y. N. (2025). GLOBAL MONTHLY COTTON PRICE TRENDS: INSIGHTS FROM ARIMA TIME-SERIES FORECASTING. *International Journal of Business & Economics (IJBE)*, 10(1), 104-124.
- [17] Srinivasan, S., Ramadass, P., Mathivanan, S. K., Panneer Selvam, K., Shivahare, B. D., & Shah, M. A. (2024). Detection of Parkinson disease using multiclass machine learning approach. *Scientific Reports*, 14(1), 13813.
- [18] Yan, L., Zhang, H. T., Goncalves, J., Xiao, Y., Wang, M., Guo, Y., ... & Yuan, Y. (2020). An interpretable mortality prediction model for COVID-19 patients. *Nature machine intelligence*, 2(5), 283-288.
- [19] C. Kästner. (2021). Model Quality: Measuring Prediction Accuracy. Medium. Retrieved from: <https://ckaestne.medium.com/model-quality-measuring-prediction-accuracy-38826216ebcb>