



Modelling and Forecasting Pulse Crop Production: An Econometric Approach to Agricultural Decision-Making

Dr. Sameerabanu P

Assistant Professor in Biostatistics, School of Allied Health Sciences, Dhanalakshmi Srinivasan University, Trichy- 621112.

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000151>

Received: 06 August 2026; Accepted: 11 August 2026; Published: 11 September 2026

ABSTRACT

This study investigates the application of econometric time series models for forecasting pulse production and its implications for agricultural policy planning. Pulses, including chickpeas, lentils, beans, and other leguminous crops, play a vital role in food security, nutritional sustainability, agricultural development, and rural livelihoods. Accurate forecasting of pulse production is essential for anticipating future supply, supporting market stability, and formulating effective agricultural policies. In this study, historical pulse production data are analyzed using the Autoregressive Integrated Moving Average (ARIMA) model to identify production trends and generate future forecasts. Model performance is evaluated using appropriate forecasting accuracy measures to assess the reliability of the predicted values. The findings demonstrate the potential of time series modelling as an effective approach for forecasting pulse production and providing quantitative evidence for agricultural planning, resource allocation, food security management, and policy decision-making.

Keywords: Pulses production; Time series forecasting; ARIMA; Econometric modelling; Agricultural policy; Forecasting.

INTRODUCTION

Pulse crops, including lentils, chickpeas, peas, and beans, are integral to global agriculture, providing essential nutrients and promoting soil health through nitrogen fixation. With the increasing demand for plant-based protein and the need to maintain sustainable agricultural practices, pulse production has gained significant attention from farmers, policymakers, and researchers. However, one of the major challenges in pulse farming lies in accurately forecasting production levels to manage supply and demand effectively.

Traditional forecasting methods rely on subjective judgement, but with the advancement of econometric models and statistical techniques, more reliable methods for production forecasting have been developed. Econometric time series models, which utilize historical data and patterns in production, provide a structured approach to forecasting. These models account for various factors such as seasonality, trend, and autocorrelation, making them suitable for agricultural applications like pulse production forecasting.

The goal of this study is to evaluate and compare the performance of econometric time series models using historical data on pulse production. The study aims to offer insights into the effectiveness of these models for predicting pulse yields, contributing to better agricultural planning, resource allocation, and market stability.

LITERATURE REVIEW

Time Series Models in Agriculture

Time series forecasting has become increasingly popular in agricultural studies for predicting crop yields, especially in the context of variable weather patterns and changing market conditions. ARIMA models have been widely applied to forecast a range of crops (e.g., wheat, rice, and maize) due to their simplicity and effectiveness in handling univariate data.

Econometric Models for Pulse Production

Studies focusing specifically on pulse production are relatively limited, but several studies have applied ARIMA and SARIMA models to forecast crop yields in general. For example, Kumar et al. (2018) applied ARIMA models to forecast wheat and rice yields in India, highlighting the robustness of time series models for predicting crop production. Similarly, Baryeh et al. (2016) used SARIMA to model seasonal patterns in food crops, although their study did not focus exclusively on pulses.

Challenges in Pulse Forecasting

Pulse crops face unique challenges that differentiate them from other cereal crops, such as climate vulnerability, soil conditions, and market price volatility. Accurate forecasting requires accounting for both long-term trends (e.g., climate change) and short-term shocks (e.g., droughts, pest infestations). These factors complicate the modeling process, making the selection of the right model critical.

Recent Developments

Newer techniques, such as machine learning algorithms and hybrid models combining econometric and artificial intelligence methods, have started to gain attention in agricultural forecasting. However, traditional econometric time series models like ARIMA and SARIMA remain powerful and widely used due to their interpretability and simplicity.

METHODOLOGY

In this study, a ARIMA model was used, consisting of identification, estimation, diagnostic checking, and forecasting. Model parameters were considered to fit the ARIMA models.

AR process of order (p) is, $Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$;

MA process of order (q) is, $Y_t = \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$; and

ARIMA process of order (p, d, q) is,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where Y_t - pulses production, ε_t 's - independently and normally distributed with zero mean and constant variance σ^2 for $t = 1, 2, \dots, n$; d - the fraction differenced while interpreting AR and MA, and ϕ 's and θ 's - coefficients to be valued.

Trend Fitting: The Box-Ljung Q statistics was used to convert the non-stationary data into stationarity data and also to validate the adequacy for the residuals. For evaluating the adequacy of AR, MA and ARIMA processes, a range of reliability statistics like R squared, Stationary R squared, RMSE, MAPE and BIC were applied. The reliability statistics viz. RMSE, MAPE, BIC and Q statistics were computed as below:

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right]^{1/2};$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \text{ and}$$

$$BIC(p, q) = \ln v^* (p, q) + (p + q) \left[\frac{\ln(n)}{n} \right]$$

where p and q - order of AR and MA processes; n - number of observations; and v^* - approximate of white noise variance σ^2 .



$$Q = \frac{n(n+2) \sum_{k=1}^n rk^2}{(n-k)}$$

where n - number of residuals and rk - residuals autocorrelation at lag k.

RESULTS AND DISCUSSION

In this analysis, the data of pulses production in Tamil Nadu were collected from the Annual book of agriculture at a glance in Tamil Nadu 2024, Ministry of Economics and Statistics, Government of Tamil Nadu for the period from 1965 to 2023 (Table 1) and were applied to fit the ARIMA model to predict the future production.

Table 1: Actual data for Pulses in Tamil Nadu (in Tonnes)

Year	Pulses Production	Year	Pulses Production	Year	Pulses Production
1965	101960	1985	275220	2005	177003
1966	112290	1986	311840	2006	290503
1967	114230	1987	283130	2007	184924
1968	106120	1988	247670	2008	167370
1969	108770	1989	333760	2009	204369
1970	133370	1990	359930	2010	245190
1971	196880	1991	351140	2011	354460
1972	197170	1992	342630	2012	212574
1973	204780	1993	276360	2013	613799
1974	114630	1994	340050	2014	766964
1975	127370	1995	233070	2015	584969
1976	176520	1996	232810	2016	337707
1977	196310	1997	244170	2017	556450
1978	218930	1998	304280	2018	551036
1979	194870	1999	290790	2019	605444
1980	176170	2000	312643	2020	472575
1981	188050	2001	270719	2021	498886
1982	189230	2002	200479	2022	503161
1983	221580	2003	201056	2023	386152
1984	248690	2004	216431		

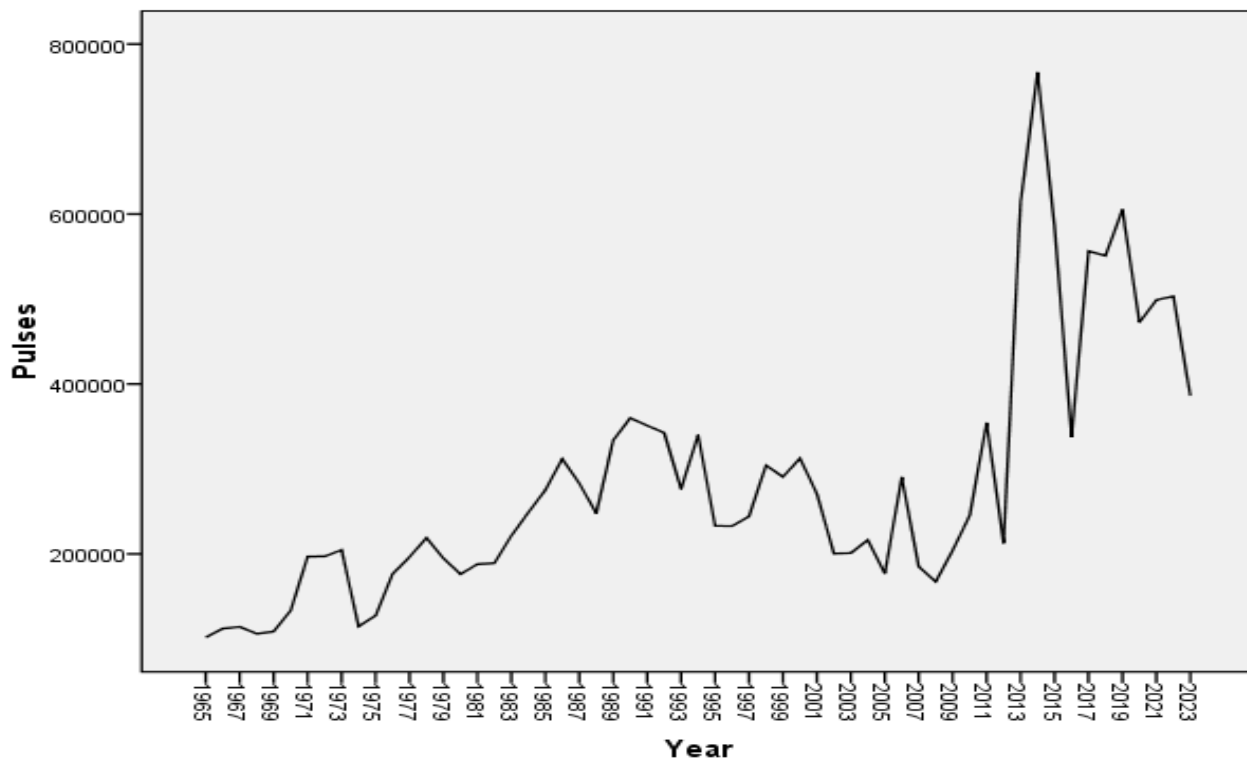


Figure 1: Time plot of Pulses Production



Figure 2: Time plot of Pulses Production with $d = 1$

Figure 1 depicts that the data used were non-stationary. The time series plot of Pulses Production from 1965 to 2023 with $d=1$ was created as shown in Figure 2 and it is stationary.

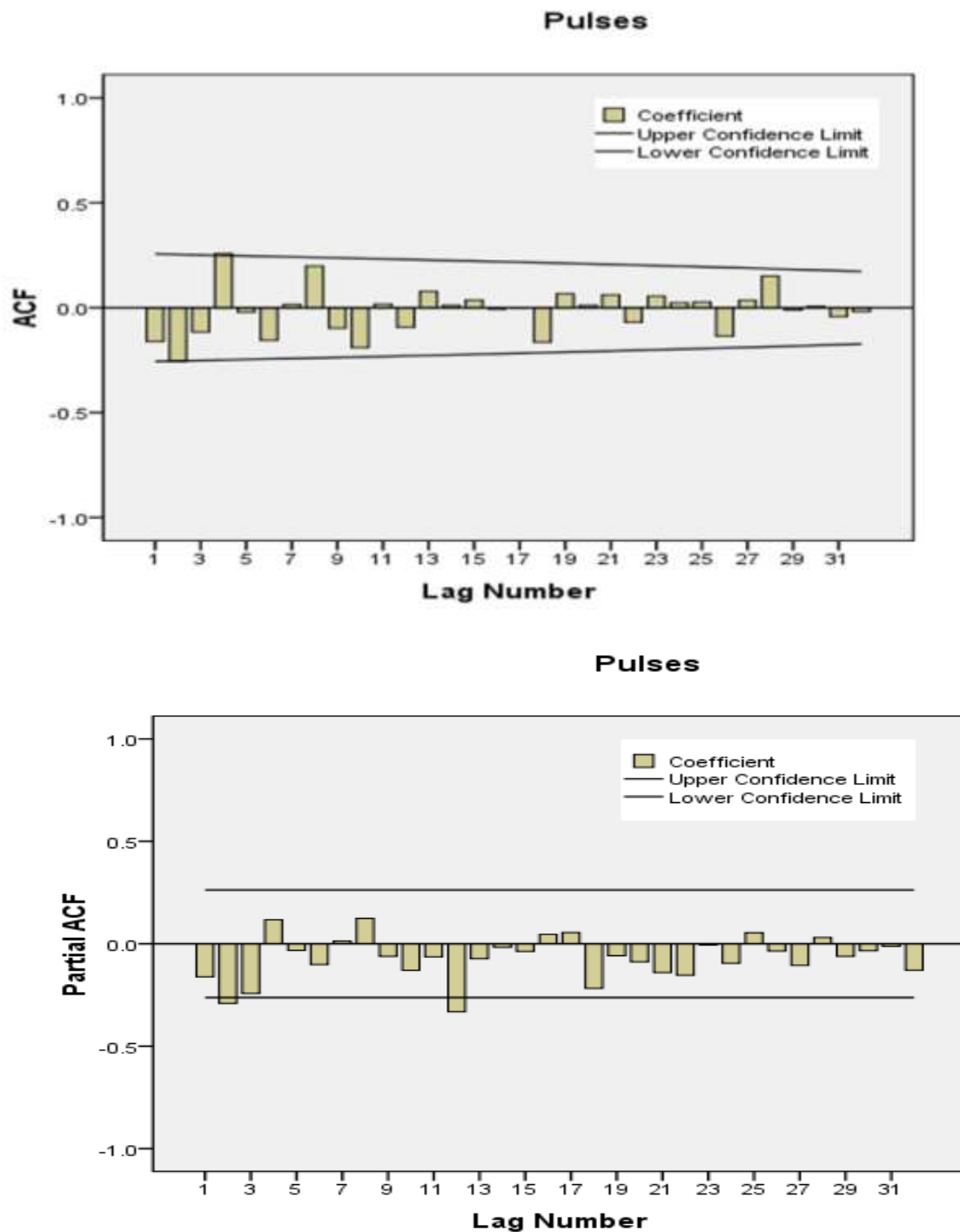


Figure 3: ACF and PACF for Pulses Production

The newly constructed variable Y_t could now be examined for Stationarity. Since, Y_t was stationary in mean, the next step was to identify the values of p and q .

Table 2: Estimated ARIMA Model Fit Statistics

Fit Statistic	(1,1,1)	(1,1,2)	(1,1,3)	(2,1,1)	(2,1,2)	(2,1,3)	(3,1,1)	(3,1,2)	(3,1,3)
Stationary R-squared	0.180	0.183	0.160	0.181	0.208	0.254	0.177	0.189	0.254
R-squared	0.672	0.673	0.664	0.672	0.683	0.702	0.671	0.675	0.701
RMSE	84963.577	85588.913	87600.237	85665.471	85061.585	83334.348	86715.184	86898.802	84156.518
MAPE	19.247	19.372	17.820	19.255	18.328	18.558	18.090	18.569	18.456
MaxAPE	72.015	78.809	70.266	72.566	75.021	70.639	74.453	79.400	70.332
MAE	55072.850	54982.124	51669.309	55050.707	53391.929	53359.913	52957.349	54691.150	53208.093
MaxAE	341669.565	347052.318	396570.672	343095.793	342790.474	345113.374	397097.557	378329.405	350763.333

Normalized BIC	22.910	22.995	23.111	22.996	23.052	23.081	23.091	23.165	23.171
----------------	--------	--------	--------	--------	--------	--------	--------	--------	--------

The ARIMA models are discussed with values differenced once ($d=1$) and the model which had the minimum normalized BIC was chosen. The various ARIMA models and the corresponding normalized BIC values are given in Table 2. The value of normalized BIC of the chosen ARIMA (1, 1, 1) was 22.910.

Table 3: Estimated ARIMA Model of Pulses Production

		Estimate	SE	t	Sig.
	Constant	5860.039	1761.804	3.326	0.002
AR	Lag 1	0.630	0.151	4.186	0.000
	Difference	1			
MA	Lag 1	0.996	0.886	1.125	0.266

The ACF and PACF of the residuals are given in Figure 5, which also indicated the ‘good fit’ of the model. Hence, the fitted ARIMA model for the Pulses Production data was

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$Y_t = 0.630 Y_{t-1} + 5860.039 - 0.966 \varepsilon_{t-1} + \varepsilon_t$$

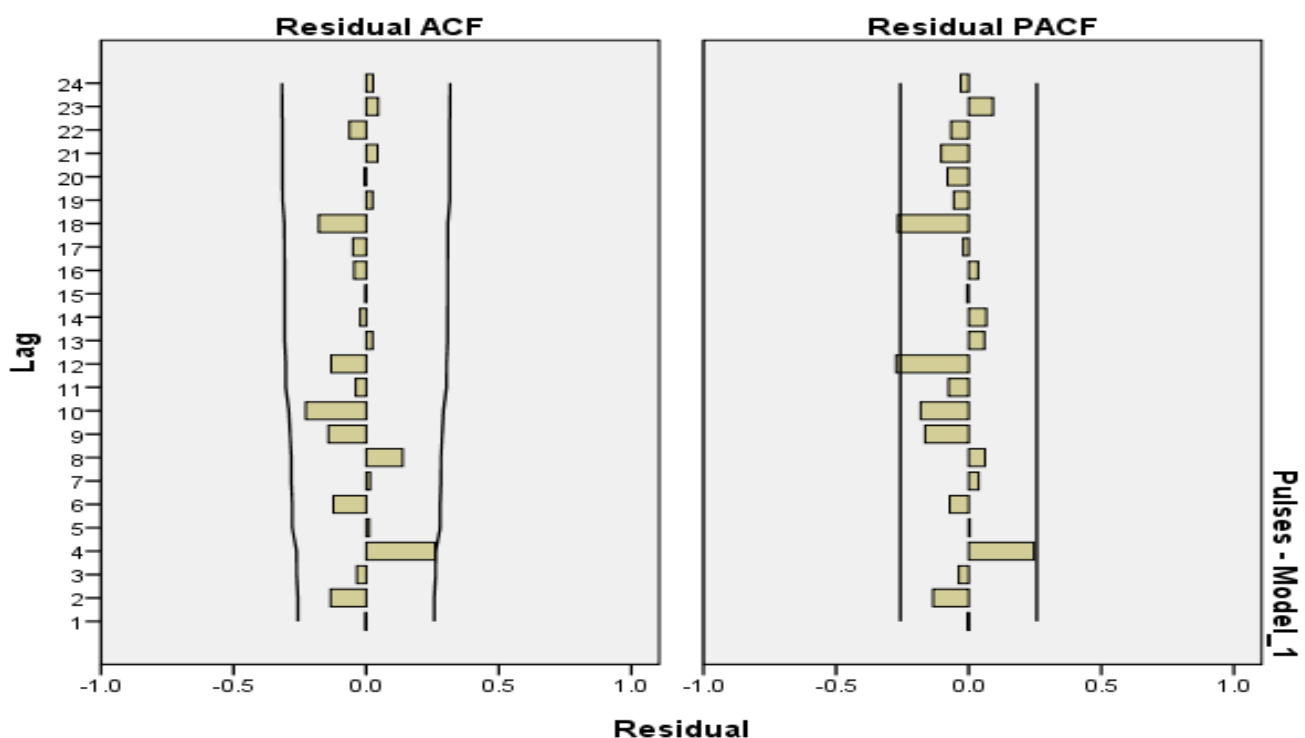


Figure 4: Residuals of ACF and PACF

Forecasting value of Pulses Production from the year 2024 to 2033 given in Table 5. To assess the forecasting ability of the fitted ARIMA (1,1,1) model, important measures of the sample period forecasts’ accuracy were computed. Figure 6 shows that the actual and forecasted value of Pulses Production data with 95% confidence limits.

Table 4: Forecast of Pulses Production

Year	Forecast
2024	415769
2025	436603
2026	451900
2027	463708
2028	473317
2029	481540
2030	488889
2031	495688
2032	502139
2033	508372

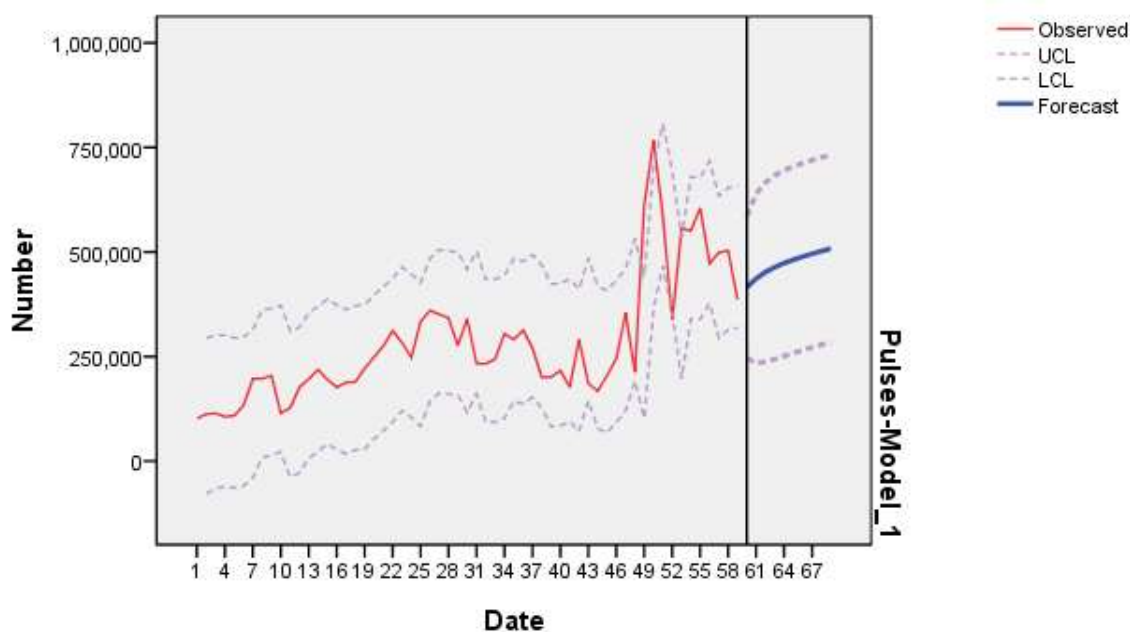


Figure 5: Actual and Estimate of Pulses Production

The results showed that, the Pulses Production would not remain stable throughout the year. The most appropriate ARIMA model for Pulses Production forecasting of data was found to be ARIMA (1,1,1). From the temporal data, it can be found that forecasted production would increase from 9698395 in 2024 to 10016689 in 2033 in India using time series data from 1965 to 2023 on Pulses Production. In future, these results will be helpful for taking necessary steps from the government authorities to increase Pulses Production in Tamil Nadu.

FINDING

The findings will compare the performance of the various time series models in terms of forecasting accuracy. The paper will also discuss the practical implications of using these models for pulse production forecasting, including their potential for use in policy-making, agricultural planning, and market prediction.

CONCLUSION

The study concludes that econometric time series models, particularly ARIMA is effective tools for forecasting pulse production. These models offer valuable insights for farmers, policymakers, and agricultural economists, contributing to more accurate planning and decision-making in the agricultural sector.



REFERENCES

1. Baryeh, E. A., et al. (2016). "Application of SARIMA Models for Seasonal Forecasting in Agriculture." *Journal of Agricultural Economics*, 47(2), 321-335.
2. Kumar, R., et al. (2018). "Crop Yield Forecasting Using ARIMA Models." *Agricultural Systems*, 100(1), 1-12.
3. FAO (2020). "World Pulse Production Statistics." Food and Agriculture Organization.
4. Shrestha, S., et al. (2019). "Predicting Crop Yields Using Time Series Models: A Case Study of Rice in Nepal." *Agricultural Systems*, 110, 25-35.
5. Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
6. Enders, W. (2014). *Applied Econometric Time Series* (4th ed.). John Wiley & Sons.