AGRO EKONOMI
ISSN 0215-8787 (print) ISSN 2541-1616 (online)
Agro Ekonomi, Vol.xx/Issue x, xxx, Page xx-xx
Available at https://jurnal.ugm.ac.id/jae/article/view/105750
doi : https://doi.org/10.22146/ae.105750

Forecasting the Annual Rice Production in Nepal Using the Box-Jenkins ARIMA Modelling Process

Suman Shrestha

Department of Agricultural Economics and Agribusiness Management, Faculty of Agriculture, Agriculture and Forestry University Chitwan, Nepal
Corresponding author : suman.pradhanag@gmail.com

ARTICLE INFO

Article History :
Submitted 28 March 2025
Revised 24 June 2025
Accepted 5 March 2026

Keywords :
ARIMA
Box-Jenkins
Forecasting
Rice production
Time series analysis

How to cite:
Shrestha, S. 2026. Forecasting the Annual Rice Production in Nepal Using the Box-Jenkins ARIMA Modelling Process. *Agro Ekonomi xx(x), xx-xx.* https://doi.org/10.22146/ae.105750

ABSTRACT

Rice is one of the main staple food crops in Nepal, yet the production level is insufficient to meet the domestic demand of the country. The study tries to determine the best-fitting ARIMA model for forecasting purpose by employing the Box Jenkins methodology, using the FAOSTAT dataset from 1961-2023. Multiple ARIMA models with the order of Autoregressive (AR) and Moving average (MA) ranging between 0 to 2 were considered. The dataset of annual rice production was found stationary at the first differencing. The ARIMA (0,1,1) model was selected as the best model based on criteria such as adjusted r square, standard error of regression, Akaike Information Criteria, and Schwarz Information Criteria. Parameter estimation revealed the significance of the first lagged value of the moving average at 1%, which indicates the model's effectiveness in explaining the data for forecasting. The diagnostic checking of the ARIMA (0,1,1) model indicated that the residuals were random and the model provided a good fit for data. The Chow break point test indicated no structural break in the model. Hence, the ARIMA (0,1,1) model can best capture the annual rice production in Nepal. Further, the model was used for in-sample and out-sample forecasting, which revealed that the model had consistent and low mean absolute percentage error. The forecast was extended for the periods from 2024-2030, which predicted an average annual growth rate of 0.983%, indicating a positive increment in the annual rice production. Policy makers should leverage accurate forecasts to enhance rice production and ensure food security through broader structural reforms and supportive policy interventions.

INTRODUCTION

Rice is considered the primary staple crop in Nepal. Looking at the import scenario, there is an increasing trend in the import of rice, despite the increase in rice production, mainly fine and aromatic rice (Gairhe et al., 2021). The rising demand of rice could be attributed to factors such as change in dietary intake of Nepalese, particularly in hilly and mountainous regions facilitated by improved road infrastructure, urbanization, and increased income, which drive the rice demand up for various rice categories (CDD, 2015). Similarly, the study conducted by Tripathi et al. (2019) revealed a shortfall of one million tons in rice production, contributing to rise in the imports. Rice is an important crop, covering 20% share of the agricultural GDP and has a significant contribution to overall GDP, nearly 5% (K. C. et al., 2021). In Nepal, rice is cultivated across 14,73,474 hectares, with production quantity of 56,21,710 metric tons, and productivity rate of 3.82 metric tons per hectare (MoALD, 2022).

Time series analysis involves investigation of equally spaced data of variable of interest in chronological order. The main objective is to identify the historical pattern within the datasets, which helps to make a forecast for the future time periods. Forecasting is utilized across several fields as an important tool for economic decision making at global level. It is an essential input for effective management of funds, resources, and time, which is critical for planning and policy formulation (Habibullah, 2003). In addition, it plays a pivotal role in vital sectors such as in business, stock markets, and weather prediction, which provides a valuable insight for planning and making sound decision (Hendikawati et al., 2020). It is used for advising on efficient surplus and deficit management, thereby stabilizing price and guaranteeing lucrative outcomes for farmers (Kumar & Baishya, 2020). Furthermore, it plays a critical role in improving policy decisions, ensuring food security, regulating import and export, and implementing pricing policies (Kumar et al., 2018). Several methods of time series forecasting have been developed over these periods. Among these, one of the most extensively used approaches is Box-Jenkins method for time series analysis. This method integrates both the auto-regressive and moving average method. Although both of these methods were used independently, the development of Box Jenkins method provides a comprehensive approach for identification and estimation of parameters, using both approaches (Dobre & Alexandru, 2008). ARIMA modelling has been extensively utilized for forecasting in several studies in agriculture. A study by Thapa et al. (2022) used the ARIMA model to forecast the area, production, and productivity of vegetables in Nepal, by employing the datasets spanning from 1977/78 to 2019/20. Ahmad et al. (2017) used this model to forecast the area, production and productivity of major cultivated crops. Similarly, Ali et al. (2015) used the datasets from 1948 to 2021 to make a forecast for production and yields of cotton and sugarcane for the 2013 to 2030 timeframe. Celik et al. (2017) applied the ARIMA model to forecast the production of groundnut in Turkey from 2016 to 2030. Nath et al. (2019) used the ARIMA modelling to predict the output of wheat in India from 2018 to 2027. All these studies highlighted the extensive applicability of the ARIMA model in forecasting. This model captures the influence of historical data, accounting for both current and past information of the error term and thus considering broad array of information confined within univariates time series throughout forecasting process.

This paper aims to use the Box-Jenkins methodology to evaluate different ARIMA family models and choose the best-fitted model. The study aims to forecast the yearly rice production in Nepal from 2024 to 2030, utilizing a dataset of 62 observations from 1961 to 2023. ARIMA modeling had been used extensively in prior research for forecasting the annual output of rice (Chaudhuri et al., 2020; Mahajan et al., 2020). Additionally, some studies have used this approach to forecast the price of rice (Anggraeni et al., 2019; Ohyver & Pudjihastuti, 2018). However, in the context of Nepal, studies that deal with forecasting using the ARIMA approach remain limited, particularly for rice production. It is crucial to identify the best model that predicts the yearly production and also contribute to the existing literature through validation of the ARIMA model put forward by prior studies to forecast the annual rice production. In Nepal, as the rice sector being the predominant part of agriculture, it is crucial to understand the trends of yearly production and predictive modeling, which could provide significant insights about the coming trends. This contributes to informed decision-making which helps to take proactive steps to achieve sustainable rice sufficiency. Forecasting plays a crucial role in planning and policy formulation, since accurate forecasting could assist in identifying the potential shortfall in supply, which helps to guide policymakers in import regulation, storage planning, and pricing strategies. In addition, it might enable the policymakers to bridge the gap between the demand and supply, which minimizes the risk of food crisis and reduces inflation as well as import dependency. Hence, by employing the simple ARIMA modeling within the Box-Jenkins framework, this study investigate the best-fitted model for making an in-sample prediction and employ the selected model to extrapolate predictions for the forthcoming years, thereby offering a valuable insight for decision-making in the agricultural sector.

METHODS

For our study purposes, we used the secondary datasets obtained from FAOSTAT, spanning from the periods of 1961 to 2023. For performing the analysis, we used the EViews student version free software. We followed the steps involved in the Box-Jenkins methodology for the empirical development and analysis of the appropriate ARIMA model. An appropriate ARIMA model was used to forecast the annual rice production for the period 2024-2030. The Box-Jenkins methodology, which was developed by George Box and Gwilym Jenkins in 1968 and known as Autoregressive Integrated Moving Average (ARIMA) model, is a powerful econometric tool used for forecasting. The term ARIMA is used interchangeably with the general ARIMA process, which is employed specifically for time series analysis, forecasting and control (Hemavathi & Prabakaran, 2018). The ARIMA process places a significant weight on the recent observation rather than distant observation, which makes it very suitable in short term forecasting. The distinctive characteristics of the ARIMA model are entrenched in the principle of “Let data speak for themselves”. In contrast to several regression models which rely on the exogenous variables, the ARIMA model is exclusively dependent on historical data. The primary assumption of ARIMA model is that there exists a past pattern which will continue to exist in the future and these hidden patterns are captured by these models in time series analysis and used for forecasting (Mgaya, 2019).

The construction of the ARIMA models comprises of three main components. The first component is called the AR component, which stands for the Autoregressive term, and it indicates a linear model where the dependent variable is regressed on its own lagged values. It can be mathematically represented in equations as follows:

$$Yt = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots \dots + \phi_p y_{t-p} + \epsilon_t \text{..............................}(1)$$

In the aforementioned equation, δ denotes the intercept, $y_{t-1}, y_{t-2}, \dots, y_{t-p}$, refer to the lagged values of y_{-t} at specified time period, $\phi_1, \phi_2, \dots \phi_p$, represent the autoregressive coefficients for each lag, and ϵ_t is the error term. If the model includes a lag up to the period of p, the order of the AR process will be p. The next component of the ARIMA model is the Moving average (MA). It is also a linear model in which the dependent variable is predicted by the value of current and previous observations error terms. The MA model is mathematically represented as follows:

$$Y_t = \mu + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots\dots + \theta_q\epsilon_{q-1} + \epsilon_t \text{.........} (2)$$

In the above equations, μ represents the mean, $\theta_1, \theta_2, \dots \theta_q$, signify the moving average coefficients, which determine how past error term influence the current value and ϵ_t represents the white noise error term at time t. The order of the MA process is denoted as q, that is, the model includes a lag up to q periods.

The ARIMA technique works effectively when utilized on time series data. Initially, the study involves testing for the presence of a unit root or stationarity. ACF and PACF plots were employed to investigate the presence of stationarity or non-stationarity in the time series data. Additionally, Augmented Dickey Fuller test can be used to verify the stationarity. If the results indicate that the unit root is present in the data at level, then the first differencing operator, which is denoted as d, is used to make our data series stationary. One of the key concepts in the analysis of time series is stationarity referring to stability properties of a time series (Ryan et al., 2023). A stationary process must satisfy the following criteria:

$$E(Y_t) = \mu_i E[(Y_t-\mu)^2] + \sigma^2; Cov\ (Y_t, Y_{t+k} = \gamma\ (k) \text{................}(3)$$

The above equation inferred that mean, variance and autocovariance should remain constants over time for the data series to be stationary. The ARIMA model is expressed as (p, d, q) as follows:

$$Yt = \delta + \{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p}\} + \{\theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots + \theta_q\epsilon_{t-q}\} + \epsilon_t \text{..........}(4)$$

Box Jenkins is a sequential process. The initial step in the process is the identification of the model, where we determine the value of p, d, and q for the ARIMA modelling. Firstly, we determine the value of d, which represents the order of differencing required to make the stationary time series. After achieving the stationarity, we then determine the autoregressive order (p) using the Partial Autocorrelation Function (PACF) and the moving average order (q) using the Autocorrelation Function (ACF). Based on these values, we obtain one or more candidates for ARIMA modelling.

Then, we estimate the parameters of candidates for ARIMA model. From these several models, the best model is selected based on the several criteria's such as volatility, standard error of regression, adjusted R-square, Akaike Information Criterion (AIC), and Schwarz Information Criterion (SIC). The model which yields least volatility, which is measured by sigma squared, is considered the best. Lower sigma squared value is associated with the enhanced predictive efficiency of the estimated model. Additionally, the best model is characterized by the highest R -square value, and lowest value of standard error of regression, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC). The model which stood as best among majority of these mentioned criteria is selected and the estimation of the parameters are carried out using the maximum likelihood (ML) estimation. This method is one of the best and widely recognized for parameter estimation (Box et al., 2013). A good model is expected to minimize the sum of squared residuals, have uncorrelated residuals, and should exhibit parsimonious. It is important to keep in mind that the parsimonious model yield better forecast compared to the overparametrized model. The third step in the process is the diagnostic checking, which investigates the validity of the chosen model. This step is similar to the validation of model in the non-linear least square regression. The residual ϵ_t should represent the white noise, which refer that the residuals should be independently and identically distributed with a constant mean and variance. The best fitted model must adhere to these assumptions of residual distribution. If the stated assumptions are violated then the modification is necessary and a new appropriate model is chosen. We can perform the residual analysis by employing the Ljung-Box test (Q). If the value of Q≤ chi-square, the Autocorrelation Functions (ACF) represent while noise, denoting that chosen model is appropriate. However, if Q > chi-square, the Autocorrelation function (ACF) are not white noise and the model is deemed inappropriate, which suggest the need for refinement of model by repeating the model building cycle.

Finally, after the chosen model had cleared all the diagnostic check, we performed a two-stage validation process to confirm its predictive accuracy and robustness. In first stage, we carried out in-sample good of fit test, which involved using the model estimated on full data period (1961-2023) to observe how well it fits the historical data on which it was trained. In second stage, we performed more rigorous approach, i.e., out of sample validation to prove the robustness on the unseen data. For this purpose, we partitioned data sets into a training set (1961-2017) and a testing set (2018-2023). We re-estimated ARIMA (0,1,1) model using only training data and it was used to forecast the testing period. It assists to simulate a real-world forecasting scenario and provide strong evidence of the model predictive accuracy. For these both stages, we quantified the performance of model using the Mean absolute percentage error (MAPE), which is estimated as follows.

$$MAPE = \frac{1}{n}\sum\frac{(|actual - forecast|)}{actual}x100 \quad \text{.....................(5)}$$

In the above equation 5, n represent the sample size, actual refer to actual data value while forecast refer to the predicted value. After estimation of the predictive accuracy, the best fit model is employed to estimate the rice production for the next seven years i.e., from 2024 to 2030. This forecast was obtained by using the model trained on complete data set (1961-2023) to ensure all available historical information was utilized. The flow chart of the sequential processes which are involved in the Box-Jenkins Methodology is illustrated in Figure 1 .

RESULT AND DISCUSSION

Identification of the Model

The initial step involved in the Box-Jenkins methodology is the identification of the model. During this step, it is essential to check for the stationary of the data. A unit root test is employed to ascertain whether the given time series data is non-stationary under the auto-regressive model. For this purpose, this study utilizes the Augmented Dickey Fuller test. The null hypothesis of this test posits the presence of a unit root. The correlogram illustrates the Autocorrelation Function (ACF) and

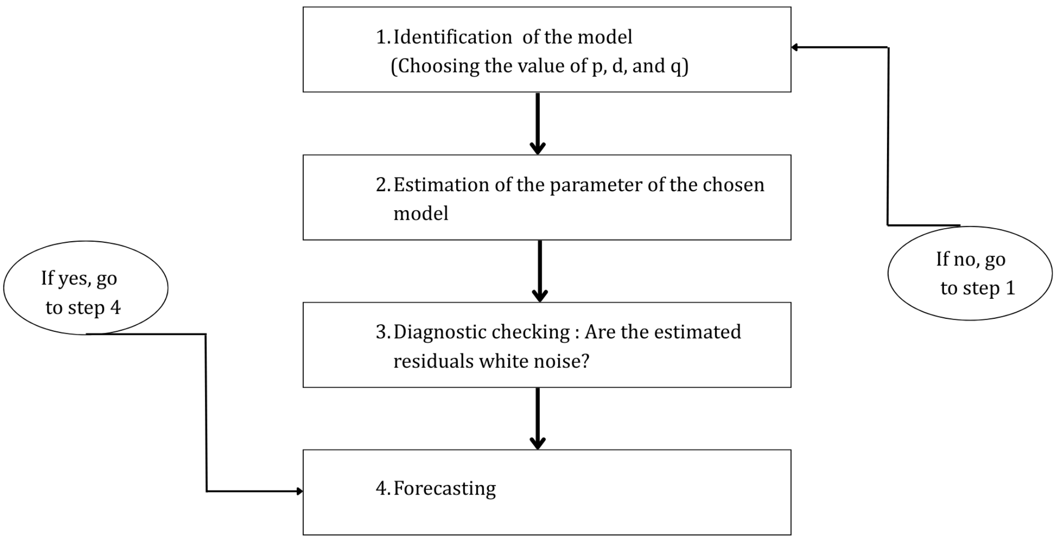

Figure 1. Step of Box-Jenkins's methodology

Partial Autocorrelation Function (PACF) of the level time series data representing the annual rice production for the sampled period has also been plotted to indicate the presence of unit root in data series. Before the unit root testing, it is important to visualize the raw data of annual rice production over the sampled period to observe the hidden pattern or trend in the data series. Figure 2 illustrates the plot showing the annual production of rice from 1961-2023. Figure 2 shows the clear upward trend in the production of rice over these sampled periods, indicating the presence of unit root in the data series.

The correlogram comprising both the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) is illustrated in Figure 3, which represents the level data series of annual rice production using 20 lag periods. From Figure 3, we can observe that the ACF plot shows a gradual decline up to 20 lags and extends beyond the 95 percent confidence band. Similarly, we can see that the PACF plot also exhibits a rapid drop after the initial lag. These observations indicate the presence of non-stationarity of level data series of annual rice production in the sampled periods. The Augmented Dickey Fuller test, which is reported in Table 1, further confirmed the presence of unit root in the data series, as the computed p-value was found higher than the significance level of 5%. Hence, the result suggested no sufficient evidence of rejecting the null hypothesis, which posits the presence of unit root in data series of annual rice production over these sampled periods.

From the result obtained above, it showed that the data series was found non-stationary at level. The next step in the process is to make the data series stationary by using the differencing operator over the sampled period, that is, 1961-2023. After the first differencing of data series at level, we performed another test for the presence of unit root in the differenced data series. The

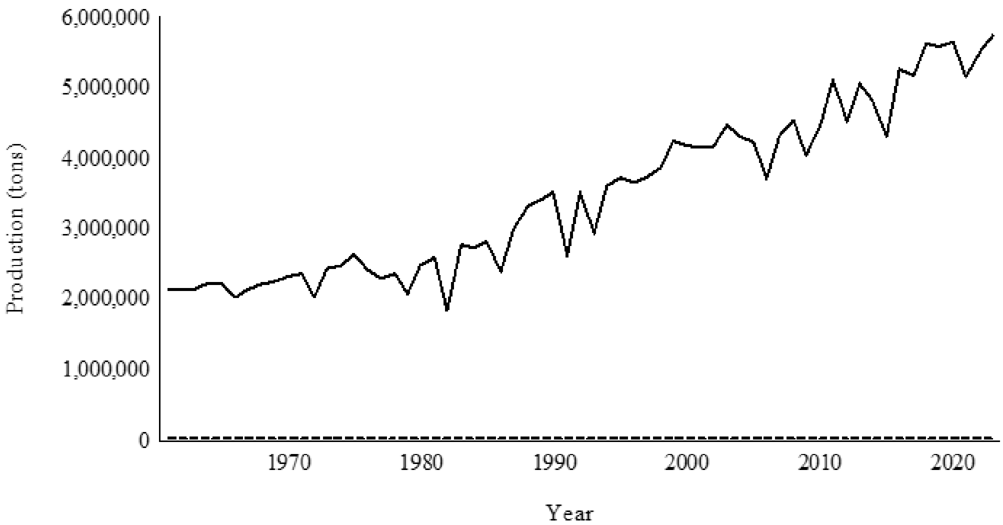

Figure 2. Annual production of rice in tons from 1961-2023
Source: FAOSTAT (2025)

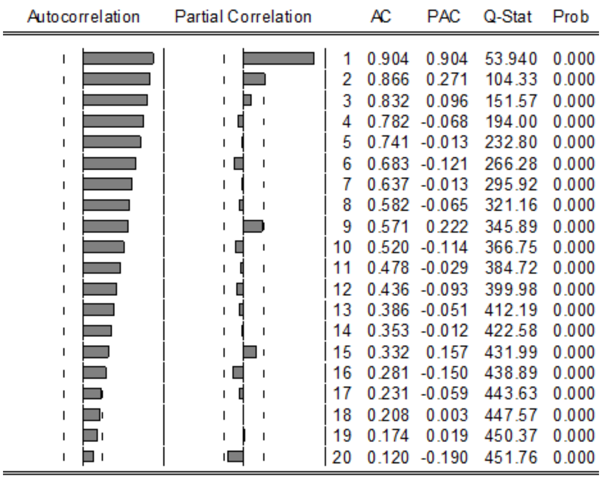

Figure 3. ACF and PACF plots for the annual rice production up to 20 lags

Table 1. Unit root test for the series at level using Augmented Dickey Fuller test.

		T statistics	Prob
Augmented Dickey Fuller test statistics		0.303526	0.9767
Test critical values	1% level	-3.544063	
	5% level	-2.910860	
	10% level	-2.59090	

Source: Data processed (2025)

result of the ACF and PACF plots are illustrated in Figure 4. The plots show that the majority of the ACF and PACF are within the 95 percent confidence band up to the specified 20 lag periods, which indicated that the data series is stationary. Additionally, for testing the presence of unit root, we employed Augmented Dickey Fuller test. The result of this test showed that the associated p-value of the test was below the significance level of 1%, which provides the substantial evidence to reject the null hypothesis of presence of unit root. This confirmation validated the conclusion that the first differencing of level data series of annual rice production renders the data series stationary. Hence, the order of differencing which is represented by notation d was found to be 1.

The next step involves determining the value of p and q, which represent the autoregressive and moving average order, respectively. The ACF and PACF plots were utilized to determine the order for the AR and MA terms. The ACF plot for the first differenced data series for annual rice production revealed significant autocorrelation at lag order 1. The ACF plot exhibits a prominent noticeable decay and a dampen sine wave pattern after lag 1. At lag 1, there is noticeable and significant negative spike in the ACF plot. This suggests that all the autocorrelations fall within the 95 percent confidence band afterward. Therefore, the tentative order of the moving average process (q) was determined to be 1. Again, the PACF plot was used to determine the AR order. Significant spike we observed at lag 1 and lag 2 coming out from the 95 percent confidence band, whereas other

Table 2. Unit root test at first difference annual using Augmented Dickey Fuller test

		T statistic	Prob
Augmented Dickey Fuller test statistics		-8.851445	0.000
Test critical values	1% level	-3.544063	
	5% level	-2.910860	
	10% level	-2.593090	

Source: Data processed (2025)

partial correlations were found within the band. The partial autocorrelations were -0.505 and -0.290 at lag orders 1 and 2, respectively and it was found to be significantly different from zero. Hence, the tentative order for the autoregressive process (p) was found to be1 and 2. However, it is crucial to note that the order of autoregressive parameter (p) and moving average (q) are just the tentative values, and the determination of the parameter orders in the Box-Jenkins methodology is inherently subjective. To incorporate a wide range of ARIMA models under testing for choosing the best suited model, we proposed the multiple ARIMA (p, d, q) models for the range value of p and q. This infers that the value of both p and q ranges between the lowest order of 0 to highest order of 2. Similar approach was utilized by (Rana, 2019) to select the best model among the different ARIMA models to forecast the GDP movement of Nepal using the similar modelling approach i.e., ARIMA.

Estimation of the parameters of the ARIMA model

In identifying the most suitable model specification for forecasting, various models with autoregressive and moving average order ranging between 0 to 2 were first estimated. The correlogram of first differenced data series of annual rice production suggests the inclusion of both the lag value of annual rice production and lags of the error term in the estimation. Table 3 illustrates the results of the estimated models within the ARIMA family, which allows for the comprehensive evaluation of performance of the model based on these stated specified criteria. Among all the fitted values of p and q, the ARIMA (0,1,1) model emerges as the most appropriate model based on the criteria such as adjusted R-square, standard error of regression, Akaike Information Criterion (AIC), and Schwarz Information Criterion (SIC), as presented in Table 3. ARIMA (0,1,1) had the highest adjusted R-squared value (0.340975), and the lowest value for standard error of regression (317199.5), Akaike Information Criterion (AIC) (28.23082), and Schwarz Information Criterion (SIC) (28.33374). Multiple models had identical sigma squared values. However, the ARIMA (0,1,1) was selected for further analysis as this model was considered apt among all other counterpart models based on the majority of the specified criteria, indicating relatively better fit and simplicity which is vital for robust forecasting with minimal overfitting. However, it is necessary to perform parameter significance test as well as the residual diagnostic test to validate it as the best model for

Table 3. Identification of the best ARIMA model based on several criteria.

Model	Sigma squared	Adj. R2	SER	AIC	SIC
ARIMA (0,1,1)	9.57E+10	0.340975	317199.5	28.23082	28.33374
ARIMA (1,1,0)	1.12E+11	0.230760	342698.6	28.37889	28.48181
ARIMA (1,1,1)	9.54E+10	0.331810	319397.4	28.25991	28.39715
ARIMA (2,1,0)	1.50E+11	-0.032302	396994.7	28.66846	28.77138
ARIMA (0,1,2)	1.50E+11	-0.031922	396921.8	28.66812	28.77104
ARIMA (2,1,1)	9.57E+10	0.330120	319801.1	28.26229	28.39952
ARIMA (1,1,2)	9.57E+10	0.329712	319898.5	28.26299	28.40022
ARIMA (2,1,2)	1.49E+11	-0.042222	398897.7	28.69360	28.83084

Source: Data processed (2025)

forecasting. The comprehensive review of the estimation of parameter of ARIMA (0,1,1) model along with corresponding standard error, t-statistics, and the p-value is demonstrated in Table 4. Other tests include several key fit statistics such as adjusted R-square, standard error of regression, F statistics, Akaike information Criterion (AIC), and Schwarz Information Criterion (SIC). From Table 4, it can be inferred that the coefficient of MA (1) is negative and highly significant at 1% significance level. The negative coefficient implies a strong short term correction mechanism in yearly production of rice. That is, when an external shock occurs in one year, production in the coming year tends to adjust in the opposite direction, which negates the initial fluctuation. Such fluctuation is consistent with the nature of rice production in Nepal, which is subjected to climatic variability, monsoon irregularities, and changes in policies. However, the MA (1) process in the model indicates that the external shock is not persistent, and the production exhibits a short run resilience or mean reverting behavior. This finding is pivotal to policy makers which suggests that even though rice production may be vulnerable to temporary shocks, it possesses an aptitude to self-correct with no long-term structural issues. Furthermore, the results indicate the model goodness of fit to the sampled data period from 1961 to 2023, which suggests that the model effectively captures the data and is well suited for the forecasting of annual rice production. The Durbin Watson test statistic was at 2.093, which suggests that there is no presence of autocorrelation.

During the selection of the best model, it is vital to consider the principle of parsimony. A parsimonious model creates better results compared to an overparametrized model. Increasing the number of independent variables in the model enhances the

Table 4. Estimation of parameters of the Arima (0,1,1) model

Variable	Coefficient	St. Error	t-statistics	Prob
C	58071.05	12683.45	4.578492	0.000
MA (1)	-0.707316	0.092966	-7.608317	0.000
SIGMASQ	9.57E +10	1.91E+10	5.016287	0.000
R-squared	0.362582	Mean dependent variable		58325.81
Adjusted R-squared	0.340975	S.D. dependent variable		390734.2
S.E. of regression	317199.5	Akaike information criterion		28.23082
Sum squared residual	5.94E +12	Schwarz criterion		28.33374
Log likelihood	-872.1553	Hannan-Quinn criterion		28.27123
F-statistics	16.78047	Durbin-Watson statistic		2.093994
Prob (F-statistic)	0.000002			
Inverted MA Roots	0.71			

Source: Data processed (2025)

model fit along with increase in r-square, however, such model is counteracted by the decrease in the adjusted R-square, which could be zero or negative when a model includes numerous irrelevant variables. Additionally, incorporation of too many variables in the model consumes the degree of freedom, which results in the minimum contribution of variables to the significance of the dependent variable.

Diagnostic checking

It is necessary to conduct the diagnostic checking for the validation of goodness of fit, before proceeding with forecasting with the final model. In selecting the best model, it is assumed that the residuals exhibit the white noise and remain uncorrelated. The adequacy of the Box-Jenkins model is evaluated by examining the residuals. The autocorrelation and partial autocorrelation are statistically zero for the randomly distributed residuals. If not, then it indicates that the chosen model is inadequate. Autocorrelation and partial autocorrelation were calculated up to 20 lags for testing. Ljung-Box Q statistics were employed for the evaluation of serial correlation by considering the model residual autocorrelation for 20 lag periods. The correlogram of the residuals and squared residuals, which also includes the autocorrelation and partial autocorrelation for the ARIMA (0,1,1) model along with corresponding p-values are illustrated in Figure 5 and 6 respectively. The correlogram of both residuals and squared residuals for the ARIMA (0,1,1) model appears flat, which suggests that the model had effectively captured all the relevant information of the data. All p-values lie above the 5% significance level. The Ljung Box test statistics indicated that the autocorrelation and partial autocorrelation of model residuals are not significantly different from zero. It implies that all the spikes are within the 95 percent confidence band. This indicates that the residuals are normal and the model is good fit for data. Finally, the best fitted ARIMA (0,1,1) model incorporates a zero-lag order of autoregression, a first differing operator to

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.050	-0.050	0.1650	
		2	0.020	0.017	0.1910	0.662
		3	0.016	0.018	0.2079	0.901
		4	-0.041	-0.040	0.3243	0.955
		5	0.105	0.101	1.0934	0.895
		6	0.038	0.050	1.1981	0.945
		7	0.016	0.017	1.2157	0.976
		8	-0.113	-0.120	2.1562	0.951
		9	0.167	0.167	4.2504	0.834
		10	-0.082	-0.076	4.7635	0.854
		11	0.013	-0.002	4.7758	0.906
		12	-0.013	-0.033	4.7894	0.941
		13	-0.001	0.044	4.7894	0.965
		14	0.058	0.025	5.0700	0.974
		15	0.146	0.167	6.8713	0.940
		16	0.004	-0.007	6.8725	0.961
		17	-0.174	-0.145	9.5352	0.890
		18	-0.108	-0.188	10.586	0.877
		19	0.103	0.156	11.558	0.869
		20	-0.100	-0.154	12.500	0.863

Figure 5. Correlogram of residuals for ARIMA (0,1,1) model

make the level data series of annual rice production stationery, and a one lag order of moving average component. The subsequent forecasting of annual rice production will be based on the ARIMA (0,1,1) model.

The inverse root of the autoregressive (AR) and moving average (MA) characteristics used to assess the stability of the selected ARIMA (0,1,1) model is shown in Figure 7. If all the roots are less than one and fall inside the unit circle, the model is considered to be stable, otherwise the impulse response standard error will not be considered valid. Figure 7 illustrates that the inverse root of the selected model is stable as the inverse root of the characteristics polynomial MA root lies inside the unit circle, indicating the value of 0.67, which is less than one.

The graph of the actual, fitted, and the residual values of the selected ARIMA (0,1,1) model is presented in Figure 8. The orange solid line represents the first differenced value of the annual rice production. The solid green line represents the fitted value, while the solid black dotted line illustrates the model residuals. The actual, fitted, and the residual graph

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.081	-0.081	0.4292	0.512
		2	0.009	0.003	0.4349	0.805
		3	0.166	0.168	2.2871	0.515
		4	-0.003	0.025	2.2876	0.683
		5	0.117	0.119	3.2459	0.662
		6	-0.040	-0.051	3.3587	0.763
		7	0.032	0.019	3.4309	0.842
		8	-0.135	-0.180	4.7728	0.782
		9	0.194	0.198	7.5909	0.576
		10	0.003	0.003	7.5915	0.669
		11	-0.147	-0.085	9.2773	0.596
		12	0.098	0.009	10.038	0.613
		13	-0.114	-0.083	11.098	0.603
		14	-0.168	-0.228	13.427	0.493
		15	0.061	0.062	13.742	0.545
		16	-0.097	-0.067	14.554	0.558
		17	-0.091	-0.004	15.279	0.575
		18	-0.048	-0.082	15.483	0.629
		19	-0.112	-0.109	16.640	0.614
		20	0.051	0.092	16.886	0.660

Figure 6. Correlogram of squared residual for ARIMA (0,1,1) model

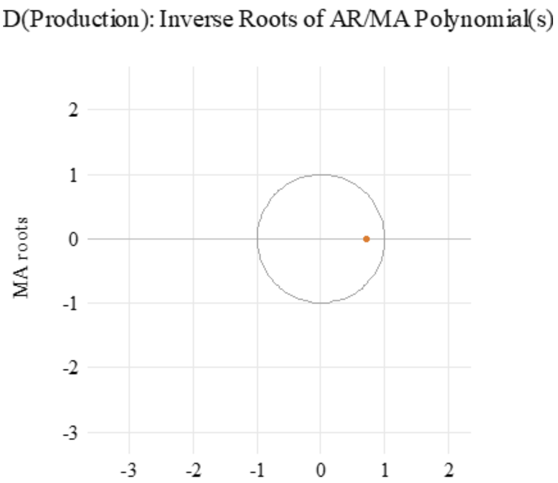

Figure 7. Inverse root of AR/MA Polynomial

visually represent the actual and the fitted values of the dependent variable, along with the residuals of the regression analysis. The actual values equal the sum of predicted values and the residuals. Residuals are limited to within plus or minus one standard deviation. From the visual inspection, it can be inferred that there is a good fit between the actual and the fitted values of the dependent variable, since the majority of the residuals are within the bound of the plus or minus one standard deviation.

Forecasting

It is vital to check for the structural break in the model before proceeding with the forecast for the annual rice production. Due to factors such as changes in policies and unforeseen shocks to an economy, it is possible that the data series could exhibit the structural breaks. For this assessment, the chow break point test was used. The model in effect uses the F-test which involves fitting equations separately for each subsample to determine if there is a significant difference in the estimated

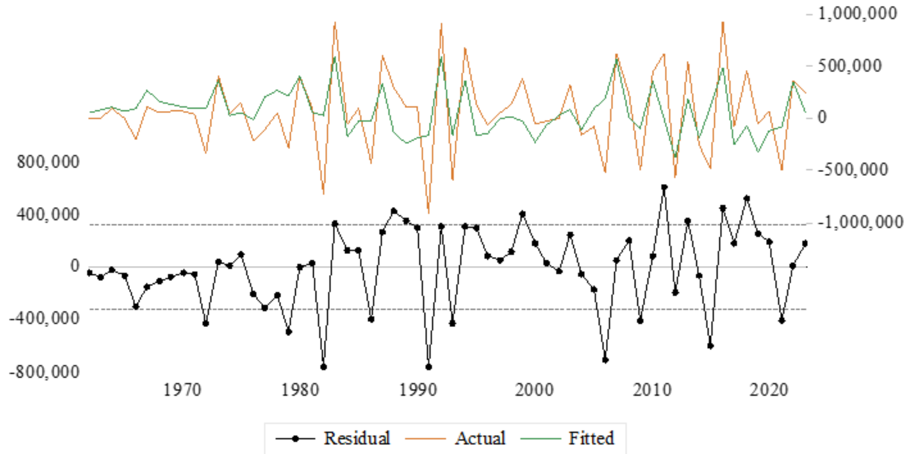

Figure 8. Actual, fitted, and residual graph of first differenced value of annual rice production

Table 5. Result of the Chow breakpoint test for the structural break in year 2000

F-statistics	0.301	Prob. F (3,56)	0.824
Log likelihood ratio	1.687	Prob. Chi-Square (3)	0.640

Source: Data processed (2025)

equations. The existence of the significant difference indicates the presence of structural break in the relationship. To perform the chow break point test, it is important to partition the data series into at least two subsamples by selecting the break point. For selecting the break point for this study, the year 2000 was purposively chosen from the sample period of 1961 to 2023. The result of the Chow test is illustrated in Table 5. Table 5 shows that all the test statistics such as F-statistics and log-likelihood ratio are not statistically significant at 5 percent level. It supports the null hypothesis that there is no evidence of structural break. Hence, it can be inferred that there is no structural break in the selected ARIMA (0,1,1) model.

The main objective of fitting the ARIMA model is to have a prediction for the time series data. Following the validation of the model through the process of diagnostic checking, the next step is to generate the in-sample forecasts for the annual rice production spanning from 1961-2023. The in-sample forecasts for the annual rice production for these periods are illustrated in Figure 9. Additionally, we found the theil's coefficient to be 0.073, which is close to zero. It implies that the forecasted value of the time series data is closer to the observed values in the data series, making the forecast better. This further supports the effectiveness of the selected ARIMA (0,1,1) model in capturing and predicting the annual rice production. The findings are consistent with GC and Yeo (2020), who also reported that ARIMA (0,1,1) model was the best fitted model for forecasting the area and yield of rice in Nepal. In forecasting the area and production of vegetables in Nepal using Box Jenkins framework, Thapa et al. (2022) found that ARIMA (0,2,1) served as the best model. Likewise, Gautam and Adhikari

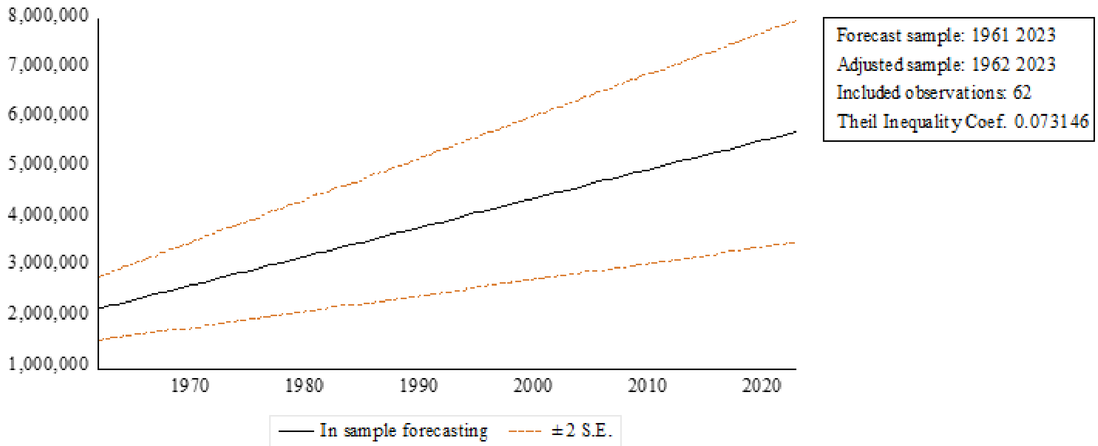

Figure 9. *In sample forecast of rice production*

(2024) reported ARIMA (2,1,3) as the appropriate model for forecasting the production of wheat in Nepal.

Before generating the final forecast for the period 2024-2030, a two-stage validation was performed to confirm the reliability of ARIMA (0,1,1) model. The predicted values of in-sample forecasting from the year 2018-2023 are reported in Table 6, along with the lower and upper bound value of annual rice production at 95 percent confidence interval. The assessment of in-sample goodness of fit of the chosen model, by comparing the actual production value within sample prediction from 2018-2023 was carried out to estimate the Mean absolute percentage error (MAPE), which was found to be 2.96%. This indicates that the chosen model deviates less than 3% from the actual value indicating strong predictive accuracy and model reliability. In addition, a more rigorous approach was followed to confirm the model robustness and predictive ability. We trained the chosen ARIMA model for time period between 1961-2017 and used it to make the forecast for the test period between 2018-2023, which are presented in Table 7. The out of sample MAPE was found 5.06%, which indicates the consistency between both the in-sample MAPE and out-of-sample MAPE. This confirms that the model is not overfitted to the data. This provides strong evidence for model robustness and suitability for practical application in forecasting.

Table 6. In-sample Goodness of Fit test for the ARIMA (0,1,1) model

Year	Actual rice production	Forecasted rice production	95% Confidence Interval Lower bound	Upper bound	MAPE
2018	5610011	5418050	3314766	7521334	2.96%
2019	5550878	5476121	3347206	7605036	
2020	5621710	5534192	3379652	7688732	
2021	5130625	5592263	3412104	7772422	
2022	5486500	5650334	3444562	7856106	
2023	5724200	5708405	3477026	7939785	

Source: Data processed (2025)

Table 7. Out-of-sample forecast validation for the ARIMA (0,1,1) model

Year	Actual rice production	Forecasted rice production	95% Confidence Interval Lower bound	Upper bound	MAPE
2018	5610011	5173315	3156867	7189764	5.06%
2019	5550878	5227093	3185923	7268263	
2020	5621710	5280870	3214979	7346762	
2021	5130625	5334648	3244034	7425261	
2022	5486500	5388425	3273089	7503761	
2023	5724200	5442203	3302144	7582262	

Source: Data processed (2025)

Table 8. Forecasted values of rice production from periods 2024-2030

Year	Forecasted value	Lower bound (-2 se)	Upper bound (+ 2 se)	Growth rate
2024	5766476	3509495	8023458	-
2025	5824547	3541969	8107126	1.007%
2026	5882618	3574448	8190788	0.997%
2027	5940689	3606933	8274446	0.987%
2028	5998760	3639421	8358099	0.978%
2029	6056831	3671915	8441748	0.968%
2030	6114903	3704413	8525392	0.959%
Average Annual Growth rate				0.983%

Source: Data processed (2025)

After the validation of the selected model through diagnostic checking and estimating predictive power, the ARIMA (0,1,1) model was selected to make a point forecast for the annual rice production from 2024-2030. The predicted values of annual rice production are presented in Table 8, along with the lower and upper bound of forecasted value. The average annual growth rate of rice production was found to be 0.983%, which indicated that the yearly increment of annual rice production is likely to be increased by 0.983%, during these periods. This rate is marginally higher compared to the average annual population growth rate of Nepal. i.e., 0.92%, which raises concerns regarding the nations objective in becoming self-sufficient in rice production. A study by Gairhe et al. (2021) reported an annual increment in rice import quantity by 24.48%, illustrating the gap between the domestic production and consumptions requirement. This widening gap calls upon the intervention to enhance the domestic production by reducing inefficiencies. The model only provides the short-term

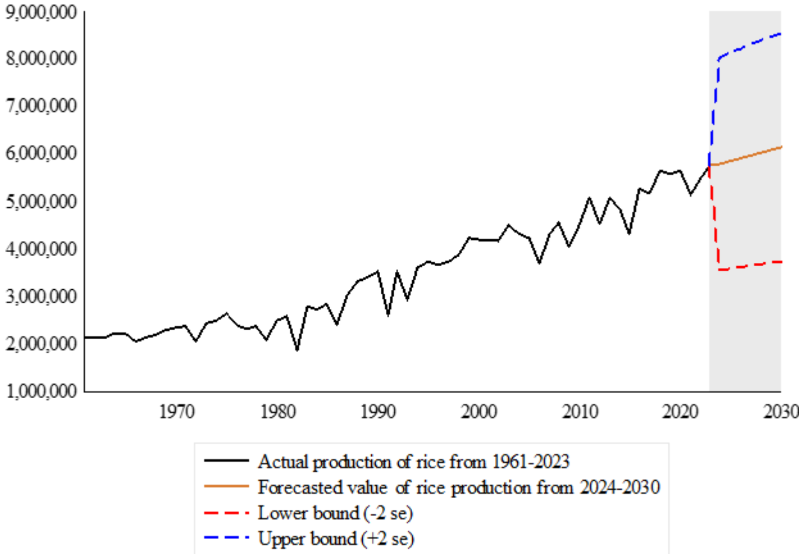

Figure 10. Actual and forecast value of rice production

forecast, but the findings emphasize the need for broader structural reforms and policy interventions to achieve the goal of self-sufficiency. Likewise, GC and Yeo (2020) on forecasting the area and yield of rice from 2018 to 2030 in Nepal found that the average annual growth rate of area and yield is 0.47% and 0.73% percent respectively during these periods. While the projected growth in production offers some optimism, it is insufficient to bride the widening gap between the nations demand and supply without more comprehensive reforms and investment in agriculture. Hence, it is imperative not only to strengthen the production system but also adopt the sustainable strategies to ensure long term food security.

In the Figure 10, the actual data series i.e., annual rice production is plotted over the sampled periods from 1961-2023, which is depicted by the solid black color line. The solid orange line in the shaded region represents the forecasted value of annual rice production from 2024-2030. The red dotted line represents the lower bound, while the blue dotted line represents the upper bound of annual rice production over these forecasted periods.

CONCLUSION AND SUGGESTION

The main objective of the study was to establish the best fitted model through identification of the most appropriate model to explain the non-seasonal time series data, i.e., annual rice production, by using the Box-Jenkins methodology. For this purpose, the data sets of annual rice production from 1961-2023 were collected from FAOSTAT and analyzed. We estimated the multiple ARIMA model with the order of AR and MA term ranging between 0 to 2. By using several metrics such as sigma square, adjacent R^2, standard error of regression, Akaike Information Criteria (AIC) and Schwarz Information Criteria (SIC), we selected the ARIMA (0,1,1) as the best model. In the selected ARIMA (0,1,1) model, the estimated parameters were found to be statistically significant at 1% level. This implied that the model specification, with the differencing of order 1 and autoregressive and moving average order of 0 and 1 respectively, effectively captures the underlying pattern in the data. Prior to using the selected model for forecasting, the model underwent diagnostic checking, and the selected model overcome the goodness of fit test as well as the stability test. The predictive accuracy of the chosen model was found high, indicating the strong predictability power and model reliability. The proposed model performed best for both in-sample and out-sample forecasting, and the best fitted model was used to extrapolate the forecasting of annual rice production from 2024-2030. The result indicated that the average yearly growth rate of annual rice production during these forecasted periods is likely to be 0.983%, which indicate the positive increment in value of annual rice production. In conclusion, the application of Box-Jenkins's methodology to make time series forecasting provide the robust forecast for the annual rice production over these specified periods. The confidence in forecasting is supported by the fact that all the observed values of annual rice production fall within the 95 percent confidence interval estimates. This suggested that the selected model effectively captured the variability in the data and produces the prediction that closely align with the real observation values, which makes the proposed model more reliable for forecasting purpose. It is crucial for policy makers to view the accurate forecasting as the guiding tools to draft out policies, so effective policy interventions could potentially enhance the production level as well as achieve food security. It is crucial to note that the ARIMA modeling has some limitations. It does not use other variables rather uses its own lagged value, which result in the inadequacy of the model to capture the technological interventions made to increase the annual rice production during these forecasted periods. On the other hand, ARIMA model can only be effectively applied, when the length of the data series exceeds 50 years. This study was mainly curtailed to select the best fitted ARIMA model to make forecasting of annual rice production from 2024-2030. However, there are several other approaches to make the forecasting of annual time series data such as neural network autoregressive model, vector autoregressive model (VAR), and holt double exponential smoothing. Future studies could explore the predicting efficacy of these stated several forecasting tools.

REFERENCES

Ahmad, D., Chani, M. I., & Humayon, A. A. (2017). Major Crops Forecasting Area, Production and Yield Evidence from Agriculture Sector of Pakistan. *Sarhad Journal of Agriculture*, *33* (3), 385–396. http://dx.doi.org/10.17582/journal.sja/2017/33.3.385.396

Ali, S., Badar, N., & Fatima, H. (2015). Forecasting Production and Yield of Sugarcane and Cotton Crops of Pakistan for 2013-2030. *Sarhad Journal of Agriculture*, *31*(1), 1–10.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2013). Time series analysis: Forecasting and control: Fourth edition. In *Time Series Analysis: Forecasting and Control: Fourth Edition*. wiley. https://doi.org/10.1002/9781118619193

CDD. (2015). *Rice Varietal Mapping in Nepal: Implications for Development and Adoption*.

Celik, S., Karadas, K., Eyduran, E., & Iqbal, F. (2017). Forecasting the Production of Groundnut in Turkey Using ARIMA Model. *The Journal of Animal & Plant Sciences*, *27*(3), 920–928. http://www.thejaps.org.pk/Volume/2017/27-03/index.php

Dobre, I., & Alexandru, A. A. (2008). Modelling unemployment rate using Box-Jenkins's procedure. *Journal of Applied Quantitative Methods*, *3*(2), 156–166.

Gairhe, S., Gauchan, D., & Timsina, K. P. (2021). Temporal Dynamics of Rice Production and Import in Nepal. *Journal of Nepal Agricultural Research Council*, *7*, 97–108. https://doi.org/10.3126/jnarc.v7i1.36932

Gautam, B., & Adhikari, S. (2024). Forecasting Wheat Area and Production in Nepal using Autoregressive Integrated Moving Average Model (ARIMA). *Innovare Journal of Agricultural Science*, *12* (2), 16–20. https://doi.org/10.22159/ijags.2024v12i1.50055

GC, A., & Yeo, J.-H. (2020). Rice Production of Nepal in 2030: A Forecast using Autoregressive

Integrated Moving Average Model. *Journal of South Asian Studies*, *25*, 31–58.https://doi.org/10.21587/jsas.2020.25.4.002

Habibullah, M. S. (2003). The Rationality Of Economic Forecasts: The Cases Of Rubber, Oil Palm, Forestry And Mining Sector. *AGRO EKONOMI*, *10*(1), 67–79. https://doi.org/10.22146/agroekonomi.16788

Hemavathi, M., & Prabakaran, K. (2018). ARIMA Model for Forecasting of Area, Production and Productivity of Rice and Its Growth Status in Thanjavur District of Tamil Nadu, India. *International Journal of Current Microbiology and Applied Sciences*, *7*(2), 149–156. https://doi.org/10.20546/ijcmas.2018.702.019

Hendikawati, P., Subanar, Abdurakhman, & Tarno. (2020). A survey of time series forecasting from stochastic method to soft computing. *Journal of Physics: Conference Series*, 012019. https://doi.org/10.1088/1742-6596/1613/1/012019

K. C., B., Pandit, R., Kandel, B. P., Kanchan, K. K., K. C., A., & Poudel, M. R. (2021). Scenario of Plant Breeding in Nepal and Its Application in Rice. In *International Journal of Agronomy* (Vol. 2021). Hindawi Limited. https://doi.org/10.1155/2021/5520741

Kumar, P. S., Dwivedi, S., Ali, L., & Arora, R. K. (2018). Forecasting Maize Production in India using ARIMA Model. *Agro Economist*, *5*(1), 1–6. https://doi.org/10.30954/2394-8159.01.2018.1

Kumar, R. R., & Baishya, M. (2020). Forecasting of Potato Prices in India: An Application of ARIMA Model. *Economic Affairs (New Delhi)*, *65*, 473–479. https://doi.org/10.46852/0424-2513.4.2020.1

Mgaya, J. F. (2019). Application of ARIMA models in forecasting livestock products consumption in Tanzania. *Cogent Food and Agriculture*, *5* (1). https://doi.org/10.1080/23311932.2019.1607430

MoALD. (2022). *Statistical Information on Nepalese Agriculture. Ministry of Agriculture and Livestock Development, Government of Nepal.*

Nath, B., Dhakre, D., & Bhattacharya, D. (2019). Forecasting wheat production in India: An ARIMA modelling approach. *Journal of Pharmacognosy and Phytochemistry*, *8*(1), 2158–2165.

Rana, S. B. (2019). Forecasting GDP Movements in Nepal Using Autoregressive Integrated Moving Average (ARIMA) Modelling Process. *Journal of Business and Social Sciences Research*, *4*, 1–20. https://doi.org/10.3126/jbssr.v4i2.29480

Ryan, O., Haslbeck, J. M. B., & Waldorp, L. (2023). *Non-Stationarity in Time -Series Analysis: Modeling Stochastic and Deterministic Trends.* https://doi.org/10.1080/00273171.2024.2436413

Thapa, R., Devkota, S., Subedi, S., & Jamshidi, B. (2022). Forecasting Area, Production and Productivity of Vegetable Crops in Nepal using the Box-Jenkins ARIMA Model. *Turkish Journal of Agriculture - Food Science and Technology*, *10*, 174–181. https://doi.org/10.24925/turjaf.v10i2.174-181.4618

Tripathi, B. P., Bhandari, H. N., & Ladha, J. (2019). Rice Strategy for Nepal. *Acta Scientific Agriculture*, *3*(2), 171–180