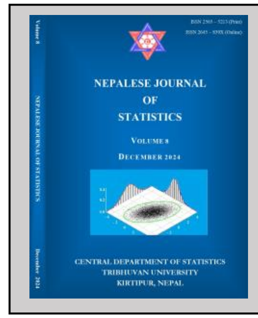
Predicting Cultivation Area, Production, and Yield of Maize in Nepal: An ARIMA Model Approach

Bibas Poudel

Submitted: 22 October 2024; Accepted: 9 December 2024
Published online: 27 December 2024

DOI: https://doi.org/10.3126/njs.v8i1.73170



ABSTRACT

Background: Maize is the primary cereal crop of Nepal after rice. It is the major component of feed for the livestock and poultry sectors. The current maize yield is unable to meet its increasing demand in Nepal. Hence, substantial quantities of maize are being imported to fill the gap. Accurate forecasting of maize cultivation area, production, and yield is critical for successful market stabilization and sustainable agricultural practice promotion.

Objective: The study aims to predict the cultivation area, production, and yield of maize in Nepal from 2023/24 to 2029/30 using appropriate Autoregressive Integrated Moving Average (ARIMA) models.

Materials and Methods: The study uses time series data from 1963/64 to 2022/23 covering maize area (ha), production (Mt), and yield (Mt/ha), obtained from the Ministry of Agriculture and Livestock Development and Agriculture Information and Training Center. The Box-Jenkins methodology-based ARIMA model was used for modeling and forecasting the future time series data. The estimated models were further diagnosed to validate no significant autocorrelation among residuals.

Results: The Box-Jenkins methodology demonstrated ARIMA (4, 1, 0), ARIMA (1, 1, 1) and ARIMA (1, 1, 1) models for forecasting maize cultivation area, production, and yield, respectively. The study predicts a 4.84% increase in maize cultivation area, a 6.83% rise in production, and a 3.17% improvement in yield from 2023/24 to 2029/30. However, these increases are not projected to meet Nepal's rising maize demand.

Conclusion: The study findings are relevant for ensuring import/export management and implementing the price policy in Nepal. The research highlights the need for technological advancements and improved management practices in maize production to ensure long-term sustainability.

Keywords: Augmented Dickey-Fuller test, Box-Jenkins methodology, Box-Ljung test, correlogram, inverse root plot, mean absolute percentage error.

Address correspondence to the author: Department of Agricultural Economics and Agribusiness Management, Agriculture and Forestry University, Chitwan, Nepal
Email: poudelbibas23@gmail.com

INTRODUCTION

Maize is the primary cereal crop of Nepal after rice in terms of cultivated area, production, and consumption (K.C. et al., 2015; Marahatta, 2021). Around 7.6% of the agricultural gross domestic product is contributed by maize in Nepal (AGDP) (MoALD, 2023). Currently, maize is cultivated on 0.98 million hectares (ha) of land with the production of around 3.1 million metric tons (Mt/ha) (MoALD, 2023). The maize yield in Nepal is 3.15 Mt/ha; far lower than the global yield of 5.71 Mt/ha (FAO, 2023; MoALD, 2023). Maize plays a crucial role in Nepal's food security and economy (Doody & Pradhan, 2022). The current requirement of maize in Nepal is around 45 kg per capita per year (Pokhrel, 2020). However, the annual consumption of maize is around 43 kg per capita (Upadhyay et al., 2018). 13% of the total dietary calorie requirement per day per capita is fulfilled by the maize in Nepal (DoFTQC, 2012). Maize accounts for the major food source for the people residing in the hilly region of Nepal (Timsina et al., 2016). Similarly, it serves as the primary component of feed for the livestock and poultry sector (Ghimire et al., 2018; Khanal et al., 2022). In Nepal, the yearly demand for maize is rising at a pace of 4-6% (Panday, 2019). The existing maize yield of 3.15 Mt/ha and the annual maize yield growth rate of 0.5% haven't been able to meet the growing demand for maize in Nepal. According to the Feed Association of Nepal, nearly 25% of the 391,538 Mt of maize required for poultry feed has been produced domestically (Gairhe et al., 2021; Choudhary et al., 2022; Koirala et al., 2020). According to the Department of Customs, Nepal imports nearly 435,217 Mt of maize which costs over $130 million (MoF, 2023).

Forecasting of agricultural products has been crucial for controlling imports/exports, enforcing pricing regulations, and supporting policy decisions (Badmus & Ariyo, 2011; Sharma et al., 2018). Appropriate prediction also addressed the allocation of land, selection of improved varieties, sufficient agricultural inputs supply, modern technology adoption, and environmental issues (Mahapatra & Dash, 2020; SenthamaraiKannan & Karuppasamy, 2020). Additionally, a sudden fall or rise in agricultural production affects the farmer's income, which further hampers the marketable surplus (Tripathi et al., 2014). Accurate forecasting indicates that excess and deficit should be managed appropriately to stabilize prices and guarantee farmers' profitability (Kumar & Baishya, 2020; Thapa et al., 2022). There are several econometric models that are appropriate for forecasting various issues, such as agricultural area and production (Ali et al., 2015). The cultivation area and productivity are fundamentally forecasted using a number of methodologies, such as remote sensing and simulation modeling (Tripathi et al., 2014). However, forecasting is occasionally seen to be essential prior to crop harvest or, in certain situations, even before crop planting (Thapa et al., 2022). The Autoregressive Integrated Moving Average (ARIMA) is one of the most popular forecasting models for time-series data (Badmus & Ariyo, 2011). Its statistical performance for univariate time series data and the well-known Box-Jenkins methodology have contributed to its appeal (Mahapatra & Dash, 2020; Ray & Bhattacharyya, 2020; Yasmin & Moniruzzaman, 2024). Additionally, it considers the non-zero autocorrelation between the time series data's subsequent values (Kumar & Anand, 2014).

Debnath et al. (2013) forecasted India's cotton output, production, and area between 2011 and 2020. Kumar and Anand (2014) projected India's sugarcane production from 2013 to 2017. Hossain and Abdulla (2017) projected Bangladesh's potato production between 2014 and 2023. Jadhav et al. (2017) used the ARIMA model to validate agricultural price forecasts and show how useful they are for main crops. From 2017 to 2022, Sharma et al. (2018) projected India's maize production. India's wheat output was predicted by Nath et al. (2019) to increase from 2018 to 2027. Mahapatra and Dash (2020) projected India's black gram productivity from 2016/17 to 2018/19. Madlul et al. (2020) used an ARIMA model to forecast Iraq's wheat crop's productivity, area, and output from 2019 to 2029. Lwaho and Ilembo (2023) used the ARIMA model to predict Tanzania's maize output from 2022 to 2031. These studies demonstrate the ARIMA model's suitability for predicting future values of agricultural products. The study aims to predict the cultivation area, production, and yield of maize in Nepal from 2023/24 to 2029/30 using appropriate ARIMA models. The model incorporates all types of information limited to the univariate time series during forecasting and capturing the impact of its historical data. This study plays a significant role in effective planning and policy-making to meet the future demand for maize. The study's findings will help Nepal's maize production become more resilient and sustainable.

MATERIALS AND METHODS

Data collection and analysis

The "Statistical Yearbook" produced by the Ministry of Agriculture and Livestock Development (MoALD) and “Krishi Diary” released by Agriculture Information and Training Center (AITC) provided the time-series data for the study (MoALD, 2023; AITC, 2023). The time series data from 1963/64 to 2022/23, making a total of 60-year data points, includes the target variables as cultivated area (ha), production (Mt), and yield (Mt/ha) of maize. Prior to analysis, this time series data were separated into two categories. For model estimation, 85% of the time series were utilized as training data, while the remaining 15% were used as test data to cross-validate the suggested model. Thus the ARIMA model was predicted using the data from 1963/64 to 2013/14, and the model was cross-validated using the data from 2014/15 to 2022/23. Using the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model and Stata software, the best-fit model for the time-series data was examined. The area, production, and yield of maize in Nepal were predicted using the best-fitted model from 2023/24 to 2029/30.

Autoregressive integrated moving average (ARIMA) model

Future values of univariate time-series data are analyzed and predicted using an ARIMA model (Kannan & Karuppasamy, 2020; Yasmin & Moniruzzaman, 2024). It consists of three components: autoregressive (AR) of order 'p', differencing of degree 'd', and moving average (MA) of order 'q', represented as the ARIMA (p,d,q) model (Box & Jenkins, 1970; Anderson, 1977). The econometric expression for the ARIMA model of order (p,d,q) is as follows:

$$Y_t = \mu + \sum_{i=1}^{p} \theta_i Y_{t-1} + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} + \varepsilon_t \quad (1)$$

where Y_t = variable at time 't', μ = mean of the series, $\theta_1, \theta_2, \ldots, \theta_p$ = AR model parameters, $\alpha_1, \alpha_2, \ldots, \alpha_q$ = MA model parameters, $\varepsilon_t, \varepsilon_{t-1}, \ldots \varepsilon_{t-q}$ = white noise residuals.

Box-Jenkins methodology

To forecast future time series values, the ARIMA 'Box-Jenkins' approach makes use of stationary time-series data (Box & Jenkins, 1970). It involves the methodological approach to identify, estimate, diagnose, and predict future values (Box & Jenkins, 1970). The methodology included the following steps:

Model identification

a) Unit root test/stationary test

Before the estimation, it is essential to analyze the asymptotic characteristics of the time series data to avoid spurious results (Gujarati et al., 2012; Poudel, 2024). In this study, the stationary nature of time series data was examined using the Augmented Dickey-Fuller (ADF) test. According to the Augmented Dickey-Fuller (ADF) test, a variable (Y_t) can be expressed in the following regression form (Dickey & Fuller, 1979):

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \quad (2)$$

where ε_t = error term, $\Delta Y_{t-i} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$……..

The null hypothesis (H_0) and alternate hypothesis (H_a) are stated as:

H_0: Y_t is a non-stationary variable; has a unit root i.e. $\varepsilon_t = 0$

H_1: Y_t is a stationary variable; lacks a unit root i.e. $\varepsilon_t < 0$

The null hypothesis (H_0) is rejected when the calculated absolute value of the t-statistics exceeds the absolute Augmented Dickey-Fuller (ADF) or McKinnon's critical values (5%) and is represented by I(0) (MacKinnon, 1991). Then, the model became the ARMA (p,q) model. Alternatively, the series is categorized as integrated of order '1', represented as I(1), if it became stationary after taking the first difference. In the ARIMA model, the 'd' degree of differentiating transforms the provided non-stationary time series into stationary.

b) The moving average (MA)

It is a smoothing method for analysis and forecasting time series data, specifically for those that do not show a trend. An MA (moving average) model incorporates 'q' lags in its regression, denoted as MA (q) (Khan et al., 2020). Mathematically,

$$Y_t = \mu + \beta_0 \mu_t + \beta_1 \mu_{t-1} + \beta_2 \mu_{t-2} + \cdots + \beta_q \mu_{t-q} \quad (3)$$

where μ = constant. $\beta_0, \beta_1, \ldots .. \beta_q$ = model parameters.

c) The autoregressive model (AR)

An AR (autoregressive) model incorporates 'p' lags in its regression, denoted as AR (p) (Hamjah, 2014). Mathematically,

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \cdots + \alpha_p(Y_{t-p} - \delta) + \mu_t \qquad (4)$$

where, δ = Mean of Y, $\alpha_1, \alpha_2, \ldots, \alpha_p$= model parameters, μ_t = white noise term.

The autocorrelation function (ACF) and the partial autocorrelation function (PACF) are plotted to observe where the spikes became significant (Fattah et al., 2018). Box and Jenkins (1970) method to identify the model with their orders is shown in Table 1.

Table 1. Procedure for identifying the ARIMA model (Gujarati et al., 2012).

Model	ACF	PACF
AR (p)	Exponential decay and/or sine function	Significant spikes through p lags
MA (q)	Significant spikes through q lags	Exponential decay
ARMA (p, q)	Exponential decay	Exponential decay

Model estimation

It involves identifying the model parameters as ARIMA (p, d, q). The inclusion of many parameters can lead to over fitting of the model.

Table 2. Different optimal model selection criteria (Montgomery et al., 2015).

Selection criteria	Description
Significance of ARMA component	If the p-value < 0.05; coefficients are significant.
Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)	Smaller the value of AIC and BIC; better the model
Log likelihood	Higher the value of likelihood; better the model
$Sigma^2$	Lower the value of $Sigma^2$; better the model

Diagnostic test

It is essential to conduct diagnostic tests to validate the goodness of fit of the model before forecasting (Yasmin & Moniruzzaman, 2024). The residual of the appropriate ARIMA (p,d,q) model exhibits white noise and is uncorrelated (Hyndman & Khandakar, 2008). If the diagnostic tests are not verified, revise the model identification step before forecasting the data (Gujarati et al., 2012; Yasmin & Moniruzzaman, 2024). Diagnosis can be done by performing the following tests:

a) White noise graph

If the residuals are confined around the mean with constant variance, it resembles a white noise structure (Yasmin & Moniruzzaman, 2024).

b) Correlogram plots of the residual

Correlogram (Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)) plots of residuals are used to evaluate the model fit (Rahman et al., 2016; Yasmin & Moniruzzaman, 2024). When the correlogram plots of the residuals show no significant spikes and all correlations fall within the 95% confidence limits, then the model resembles white noise (Debnath et al., 2013).

c) Box-Ljung test (Portmanteau test for white noise)
It is used to detect the presence of serial correlation among the residuals (Ljung & Box, 1978). If the p-value is below 0.05, the null hypothesis is rejected indicating the residuals are not white noise.
d) Inverse Roots of AR/MA Polynomials
The model's stationary/invertible status is shown by the inverse roots of AR/MA polynomials. If all eigenvalues are less than one and fall inside the unit circle, the model is said to be stable/invertible (Fauzi & Abu Bakar, 2022).

Model adequacy

The ultimate test of every model is based on its ability to make accurate predictions about the future. Following the selection of the best-fitting ARIMA model, the model's adequacy should be further examined. The model's goodness of fit was estimated using the test dataset, which covered the years 2014/15 to 2022/23. For the same time period, the expected value was predicted from the chosen ARIMA model. Lastly, model adequacy was estimated using mean absolute percentage error (MAPE). It utilizes all data and has the least variation between samples (Jadhav et al., 2017). The mathematical expression of MAPE is given below:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{Y}_t - Y_t}{Y_t}\right| \times 100$$

where Y_t = Actual values, $\hat{Y}_t$ = Predicted value, N = Number of observation, Forecasting.
After performing the diagnostic test, the ARIMA model was used to predict the future values of the time series (Kannan & Karuppasamy, 2020; Yasmin & Moniruzzaman, 2024).

RESULTS AND DISCUSSION

Graphical description of cultivated area, production, and yield of maize in Nepal

The trend of cultivated area (ha), production (Mt), and yield (Mt/ha) of maize from 1964 to 2023 is illustrated in Figure 1. The area shows a relatively stable trend with a slight increase over the years. Similarly, the production (Mt.) has a gradual increase up to 1988 followed by a marked rise up to 2023. Lastly, the yield (Mt/ha) remains steady initially, increases significantly from the mid-1990s, and has a sharp upward trajectory from 2010 to 2023. The graph highlights that while the cultivated area has not changed drastically, the production and yield have improved in the last decade, indicating advancements in cultivation practices and technology adoption (Dhakal et al., 2022).

Model identification

ADF test for Stationary

At first difference, all the variables i.e. area, production, and yield are stationary at the 1% significance level as shown in Table 3. Thus, the degree of differentiation 'd' for all variables was found to be 1.

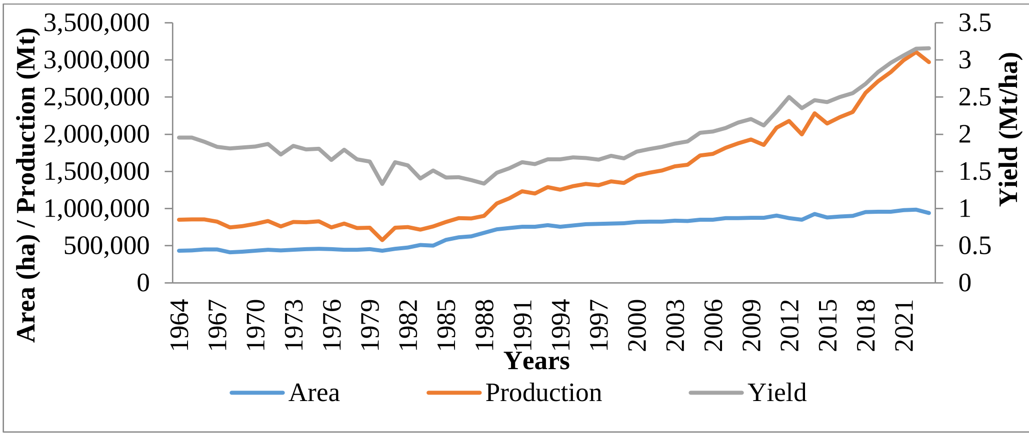


Fig. 1. Trend of cultivated area (ha), production (Mt), and yield (Mt/ha) of maize in Nepal from 1964 to 2023.

Table 3. Augmented Dickey-Fuller (ADF) test.

Variables	Level		Difference		Remarks
	Intercept	Intercept + trend	Intercept	Intercept + trend	
Area	0.028	-1.853	-4.316***	-4.289***	I(1)
	(0.9608)	(0.6786)	(0.004)	(0.003)	
Production	2.375	-1.467	-5.440***	-7.359***	I(1)
	(0.999)	(0.8403)	(0.000)	(0.000)	
Yield	0.593	-0.715	-5.730***	-7.539***	I(1)
	(0.9874)	(0.9721)	(0.000)	(0.000)	

Note: *** shows significance at 1%; parenthesis and non-parenthesis are t-statistics and MacKinnon p-value respectively.

Correlogram plot

The correlogram plots of the cultivated area, production, and yield in the first order are shown in Figure 2 and Figure 3. The ACF bar of the area parameter has no significant spikes, indicating the time series data has an MA order of 0, i.e. q = 0. Similarly, the ACF plots of both production and yield variables show a significant spike at lag 1, indicating the MA orders of 1, i.e. q = 1. Similarly, the PACF plot of area variables shows significant spikes at lags 3 and 4, indicating the time series data has two possible AR order i.e. p = 3 and p = 4. The PACF plots of both production and yield variables shows significant spikes at lag 1 indicating AR order as p = 1. The higher-order lags might be due to random fluctuations or noise (Box & Jenkins, 1970). Hence they are neglected. Hence the possible models for the area variable will be ARIMA (3, 1, 0) and ARIMA (4, 1, 0). Out of the two models, one will be selected as best-fitted model based on model selection criteria, which will be done in further steps. Similarly, the best-fitted models for both production and yield variables are found to be ARIMA (1, 1, 1).

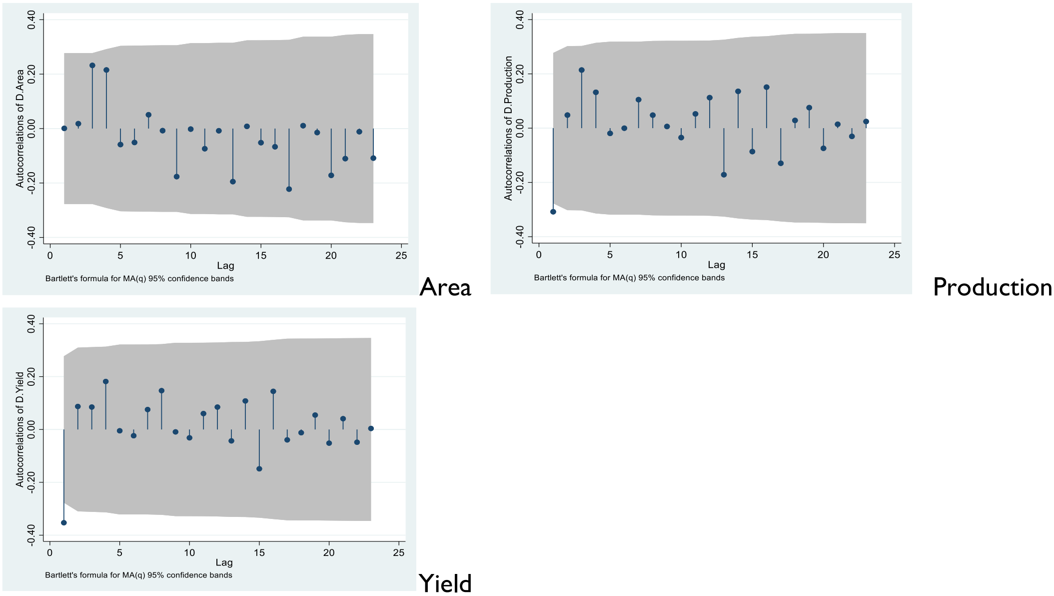


Fig. 2. Autocorrelation function (ACF) plot at first difference.

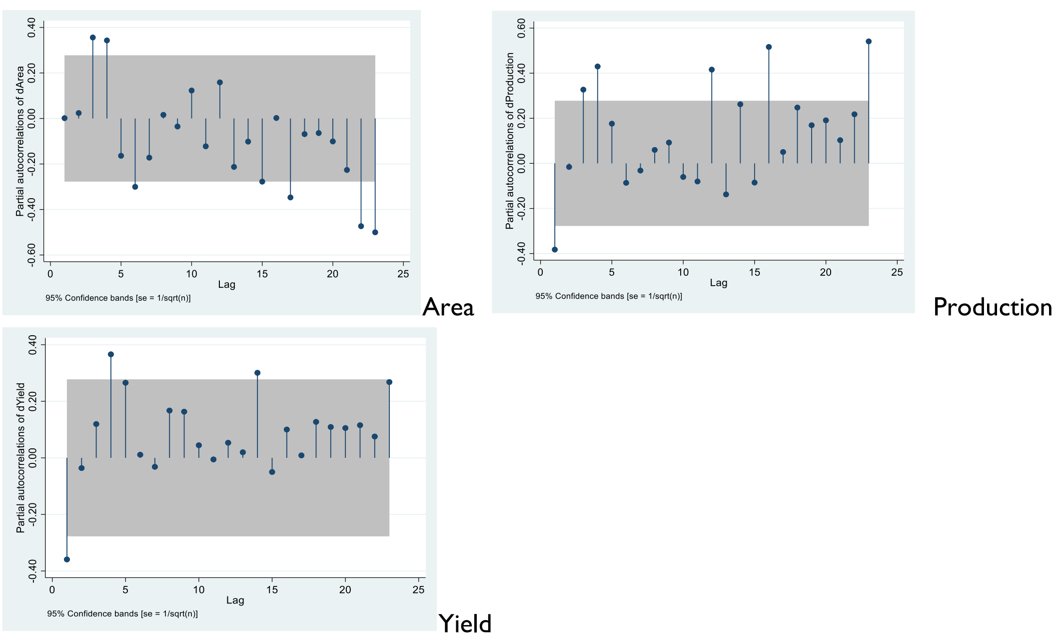


Fig. 3. Partial autocorrelation function (PACF) plot at first difference.

Model estimation

ARIMA (4, 1, 0) has the most significant coefficients, highest log likelihood, lowest AIC, lowest BIC, and lowest Sigma^2 than ARIMA (3, 1, 0) of the area variable. Hence, the ARIMA (4, 1,

0) was the suitable model for forecasting the area variable. Since both the production and yield time series data have only one ARIMA model, i.e., ARIMA (1, 1, 1), hence it was their respective appropriate model for forecasting.

Table 4. Model selection for area, production, and yield of maize.

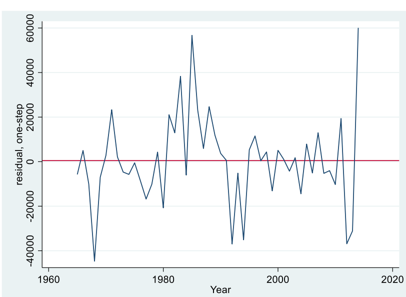
Characteristics		Area		Production	Yield
		ARIMA (3, 1, 0)	ARIMA (4, 1, 0)	ARIMA (1, 1, 1)	ARIMA (1, 1, 1)
Coefficient	Constant	9632.783**	8736.813	27410.65**	.0096
	AR(1)	-.0589	-.1308	-.3577	-.3031
	AR(2)	.0157	-.0538		
	AR(3)	.3288*	.3211**		
	AR(4)		.3604*		
	MA(1)			-.0174	-.0561
Sigma^2		21013.61***	20019.71***	77573.84***	.0955
AIC		1147.53	1145.241	1275.94	-84.7942
BIC		1157.09	1156.713	1283.588	-77.1461
Log likelihood		-568.7648	-566.6206	-633.9698	46.3971

Note: ***, **, and * show the significance at 1%, 5%, and 10% respectively.

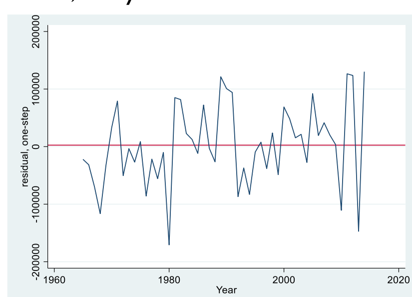
Diagnostic test

White noise graph

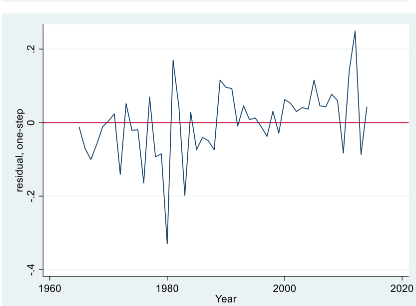
Figure 4 shows the residuals of each variable used in the study. Since all the residuals are confined around the mean with constant variance, they resemble a white noise structure.



Area



Production

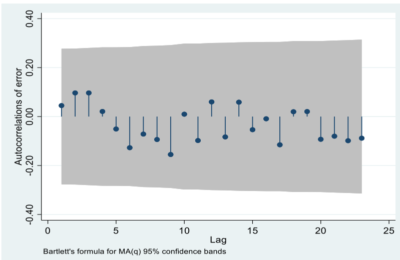


Yield

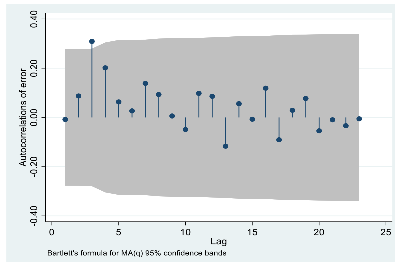
Fig. 4. Residual plots.

Correlogram plots of the residual

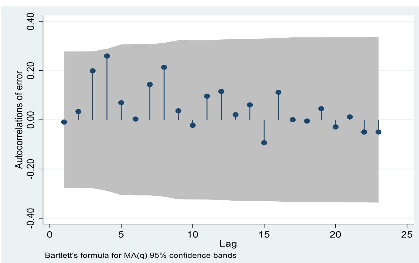
The ACF and PACF plots of the residual for the respective ARIMA model are shown in Figure 5 and Figure 6 respectively. The correlogram of the ACF for the residuals is flat which indicates that all information has been captured. Similarly, the correlogram of the PACF for the residuals is not so flat. It was showing some significance. But because parsimony is the watchword, that significance will not be considered.



Area

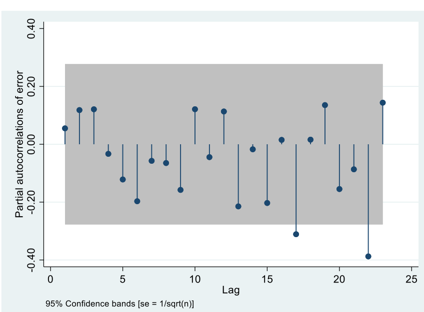


Production

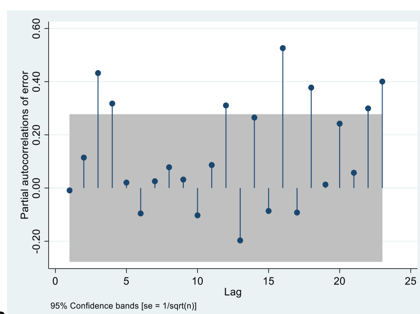


Yield

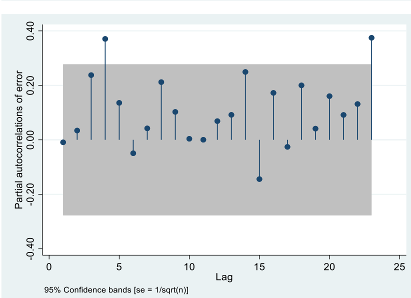
Fig. 5. Autocorrelation function (ACF) plot for the residuals.



Area



Production



Yield

Fig. 6. Partial autocorrelation function (PACF) for the residuals.

Box-Ljung test (Portmanteau test for white noise)

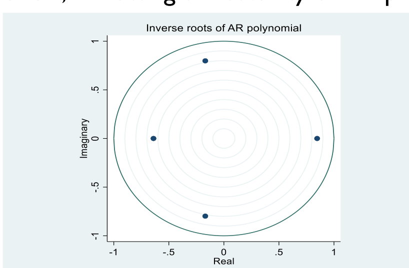
The Box-Ljung Q statistics and corresponding p-values are shown in Table 5. There was no autocorrelation among all the residuals of the fitted ARIMA models.

Table 5. Box-Ljung statistics of different ARIMA models.

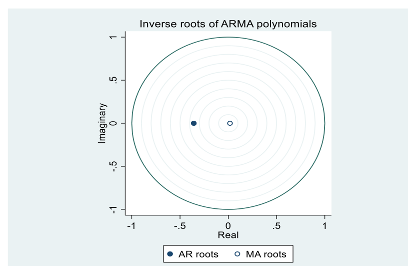
Variables	Model	Box-Ljung statistics	
		Q statistic	p-value
Area	ARIMA (4, 1, 0)	10.5801	0.9871
Production	ARIMA (1, 1,1)	15.1183	0.8903
Yield	ARIMA (1, 1, 1)	14.6488	0.9068

Inverse roots of AR polynomial (s)

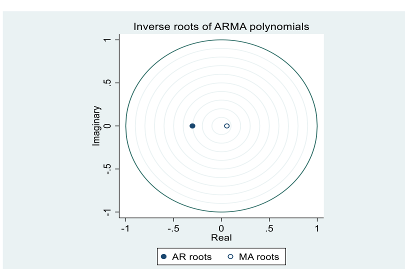
The inverse root of AR plots for each ARIMA model shows that AR roots lie inside the unit circle, indicating the stability of AR plots.



Area



Production



Yield

Fig. 7. Inverse root plots of area, production, and yield.

Hence, after performing the diagnostic tests, the appropriate ARIMA models for area, production, and yield are ARIMA (4, 1, 0), ARIMA (1, 1, 1), and ARIMA (1, 1, 1), respectively.

Model adequacy

Using the ARIMA (4, 1, 0) for area and ARIMA (1, 1, 1) for both production and yield, the actual data from test series and predicted values from the model were compared from the year 2015/16 to 2019/20 (Table 6). The lower value of Mean Absolute Percentage Error (MAPE) i.e. 2.21%, 4.36%, and 2.63% for prediction of the area, production and yield, respectively, implies very good model accuracy for further prediction. Thus, this cross-validation showed that the ARIMA (4, 1, 0), ARIMA (1, 1, 1), and ARIMA (1, 1, 1) are suitable for forecasting the values for area, production, and yield, respectively.

Forecasting

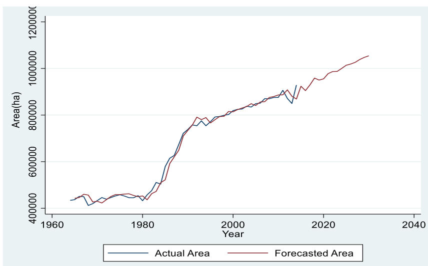
After determining the appropriate model and measuring the adequacy of the models, the selected models are rendered fit for further forecasting. The predicted values of the cultivation area, production, and yield of maize from 2023/24 to 2029/30 are presented in Table 7. The cultivation area shows a percentage growth of approximately 4.84% over the forecasted year. Similarly, the total production follows a similar positive trend representing a notable increase of about 6.83% over the forecasted period. The yield also corresponds to a percentage increase of approximately 3.79%. Overall, during this period a consistent upward trajectory in cultivation area, production, and yield emphasizes steady progress in agricultural development during the forecasted period.

Table 6. Model adequacy of respective ARIMA models from 2014/15 to 2022/23.

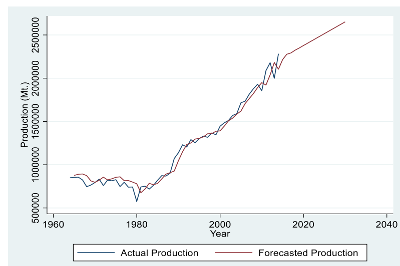
Variables	Year	Actual value	Predicted Value	MAPE
Area	2014/15	882,395	915,413.5	2.21%
	2015/16	891,583	879,814.7	
	2016/17	900,288	907,528.5	
	2017/18	954,158	918,076.1	
	2018/19	956,447	931,969.1	
	2019/20	957,650	967,039.6	
	2020/21	979,776	980,241.9	
	2021/22	985,565	999,155.3	
	2022/23	940,256	990,026.9	
Production	2014/15	2,145,291	2,255,338	4.36%
	2015/16	2,231,517	2,237,442	
	2016/17	2,300,121	2,240,334	
	2017/18	2,555,847	2,338,517	
	2018/19	2,713,635	2,546,685	
	2019/20	2,835,674	2,739,361	
	2020/21	2,997,733	2,857,853	
	2021/22	3,106,397	3,012,817	
	2022/23	2,969,222	3,134,632	
Yield	2014/15	2.4312	2.4485	2.63%
	2015/16	2.5028	2.4699	
	2016/17	2.5548	2.5072	
	2017/18	2.6786	2.5724	
	2018/19	2.8372	2.6768	
	2019/20	2.9610	2.8326	
	2020/21	3.0596	2.9649	
	2021/22	3.1518	3.0669	
	2022/23	3.1578	3.1597	

Table 7. Forecasted value of area, production, and yield from 2023/24 to 2029/30.

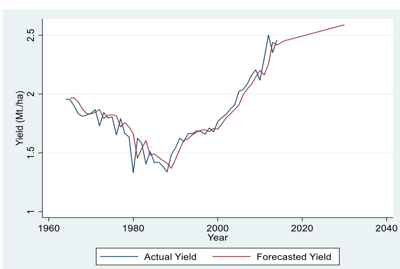
Year	Area (ha)	Production (Mt)	Yield (Mt/ha)
2023/24	1,000,080	3,175,248	3.1862
2024/25	1,011,739	3,209,469	3.2038
2025/26	1,016,822	3,246,994	3.2256
2026/27	1,024,565	3,282,812	3.2454
2027/28	1,034,061	3,319,511	3.2662
2028/29	1,041,903	3,355,755	3.2865
2029/30	1,048,533	3,392,235	3.3070



Area



Production



Yield

Fig. 8. Actual and forecasted value from 1960/61 to 2029/30.

CONCLUSION

ARIMA (4, 1, 0), ARIMA (1, 1, 1), and ARIMA (1, 1, 1) were found to be appropriate models for modeling and forecasting the cultivation area, production, and yield of maize from 2023/24 to 2029/30 in Nepal. Various diagnostic tests have shown no significant autocorrelation in the residuals of the respective models that show the reliability of these models. The respective models forecast an increase of 4.84%, 6.83%, and 3.79% in area, production, and yield over the next seven years from 2024 to 2030. To control imports/exports and carry out a pricing strategy, this forecasting will be crucial to research and policymaking in Nepal. Although the production and yield show significant increases in the upcoming years, their increments are not at the desired level that will lead to the sustainability of maize in Nepal. The production projection in the next seven years will not lead to a cut-off of maize imports in Nepal. Hence the concerned authorities are recommended to consider technological advancements in maize production. Moreover, the application of better management practices, price support programs, and enhanced cooperation between farmers and research experts will result in the sustainability of maize in Nepal. Developing and executing action plans based on anticipated trends in maize cultivation area, production, and yield leads to the sector growth and expansion effectively.

ACKNOWLEDGEMENTS

The author would like to thank the Agriculture Information and Training Center (AITC) and Ministry of Agriculture and Livestock Development (MoALD) for providing access to data from their annual report through their official website for the study. The Central Department Research Committee, Central Department of Statistics, TU, is also acknowledged by the authors for their feedback and recommendations for publishing the article.

CONFLICT OF INTEREST

The author declared that there is no conflict of interests.

AUTHOR CONTRIBUTION

The author performed all the duties regarding the manuscript preparation. He conceptualized the study, performed data collection and analysis, and prepared the manuscript.

FUNDING

The author received no funding for the study.

DATA AVAILABILITY

The data used in the study are available in the supplementary materials section of the journal.

ETHICAL STATEMENT

The research is based upon secondary data obtained from Ministry of Agriculture and Livestock Development and Agriculture Information and Training Center through their annual reports. No ethical approval was obtained for the study.

REFERENCES

AITC. (2023). *Agriculture and Livestock Diary 2081*. Ministry of Agriculture and Livestock Development (MoALD). Lalitpur: Agriculture Information and Training Center (AITC).

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.

Anderson, O. D. (1977). The interpretation of Box-Jenkins time series models. *Journal of the Royal Statistical Society, 26*(2), 127-145. https://doi.org/10.2307/2987959

Badmus, M. A., & Ariyo, O. S. (2011). Forecasting cultivated areas and production of maize in Nigerian using ARIMA model. *Asian Journal of Agricultural Sciences, 3*(3), 171-176.

Box, G., & Jenkins, G. (1970). *Time series analysis: Forecasting and control.* Holden-Day.

Choudhary, D., Khanal, N. P., & Pandit, N. R. (2022, October 24). *From seed to feed*. International Maize and Wheat Improvement Center (CIMMYT). Retreived from https://www.cimmyt.org/news/from-seed-to-feed/

Debnath, M. K., Bera, K., & Mishra, P. (2013). Forecasting area, production and yield of cotton in India using ARIMA model. *Journal of Space Science & Technology, 2*(1), 16-20.

Dhakal, S., Sah, S. K., Amgain, L. P., & Dhakal, K. H. (2022). Maize cultivation: Present status, major constraints and farmers's perception at Madichaur, Rolpa. *Journal of Agriculture and Forestry University, 5*, 125-131.

Dickey, D. A., & Fuller, W. A. (1979, June). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(36), 427-431. https://doi.org/10.2307/2286348

DoFTQC. (2012). *Food consumption tabel for Nepal.* Ministry of Agriculture Development, Department of Food Technology and Quality Control (DoFTQC). Kathamandu: Goverment of Nepal.

Doody, A., & Pradhan, B. (2022, October 12). *Elite maize seeds handed over to seed sector stakeholders in Nepal*. International Maize and Wheat Improvement Center (CIMMYT). Retreived from https://www.cimmyt.org/news/elite-maize-seeds-handed-over-to-seed-sector-stakeholders-in-nepal/

FAO. (2023). *Food and Agriculture Organization Database (FAOSTAT)*. Food and Agriculture Organization (FAO). Retreived from https://www.fao.org/faostat/en/#data/QCL/visualize

Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management, 10*. https://doi.org/10.1177/1847979018808673

Fauzi, F. D., & Abu Bakar, A. S. (2022). Rice production forecasting in Malaysia: A Box-Jenkins and ARIMA model approach. *8th Annual ECOFI Symposium*.

Gairhe, S., Timsina, K. P., Ghimire, Y. N., Lamichhane, J., Subedi, S., & Shrestha, J. (2021). Production and distribution system of maize seed in Nepal. *Heliyon,* 7(4), 1-12. https://doi.org/10.1016/j.heliyon.2021.e06775

Ghimire, Y. N., Timsina, K. P., Devkota, D., Gautam, S., Choudhary, D., Poudel, H., et al. (2018). Dynamics of maize consumption and its implication in maize technology demand in Nepal. In B. Prasanna, A. Das, & K. Kaimenyi (Ed.), *13th Asian maize conference and expert consultation on maize for food, feed, nutrition and environmental security*, Ludhiana, India.

Gujarati, D. N., Porter, D. C., & Gunasekar, S. (2012). *Basic econometrics* (5th ed.). McGraw-Hill/Irwin.

Hamjah, M. A. (2014). Forecasting major fruit crops productions in Bangladesh using Box-Jenkins ARIMA model. *Journal of Economics and Sustainable Development, 5*(7), 96-107.

Hossain, M., & Abdulla, F. (2017). Forecasting potato production in Bangladesh by ARIMA model. *Journal of Advanced Statistics, 2*(4), 125-136.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 27*(3), 1-22. https://doi.org/10.18637/jss.v027.i03

Jadhav, V., Chinnappa Reddy, B. V., & Gaddi, G. M. (2017). Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science and Technology, 19*(5), 981-992.

KC, G., Karki, T. B., Shrestha, J., & Achhami, B. B. (2015). Status and prospects of maize research in Nepal. *Journal of Maize Research and Development, 1*(1), 1-9. https://doi.org/10.3126/jmrd.v1i1.14239

Kannan, S., & Karuppasamy, K. M. (2020). Forecasting for agricultural production usisng ARIMA model. *PalArch's Journal of Archaeology of Egypt/Egyptology, 17*(9), 5939-5949.

Khan, D., Ullah, A., Bibi, Z., Ullah, I., Zulfiqar, M., & Khan, Z. U. (2020). Forecasting area and production of guava in Pakistan: An econometric analysis. *Sarhad Journal of Agriculture, 36*(1), 272-281. http://dx.doi.org/10.17582/journal.sja/2017/33

Khanal, P., Dhakal, R., Khanal, T., Pandey, D., Devkota, N. R., & Nielsen, M. O. (2022). Sustainable livestock production in Nepal: A focus on animal nutrition strategies. *Agriculture, 12*, 679. https://doi.org/10.3390/agriculture12050679

Koirala, K., Rijal, T., KC., G., Khan, S., Mahato, D., Manandhar, S., et al. (2020). Evaluation of maize hybrids in Terai and inner Terai ecological belt of Nepal. *Journal of Agriculture and Forestry University, 4*, 109-116. https://doi.org/10.3126/jafu.v4i1.47055

Kumar, M., & Anand, M. (2014). An application of time series arima forecasting model for predicting sugarcane production in India. *Studies in Business and Economics, 9*(1), 81-94.

Kumar, R. R., Baishya, M. (2020). Forecasting of potato prices in India: An application of ARIMA model. *Economic Affairs, 65*(4), 473-479. https://doi.org/10.46852/0424-2513.4.2020.1

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297-303. https://doi.org/10.1093/biomet/65.2.297

Lwaho, J., & Ilembo, B. (2023). Unfolding the potential of the ARIMA model in forecasting maize production in Tanzania. *Business Analyst Journal, 44*(2), 128-139.

MacKinnon, J. (1991). Critical values for cointegration tests. In R. a. Engle, *Long run economic relationships*. Oxford University Press.

Madlul, N. S., Al-Najjar, E. Y., Baker, Y. T., & Irhaim, F. I. (2020). Using the ARIMA models to predict wheat crop production in Iraq. *Int. J. Agricult. Stat. Sci, 16*(1), 121-127.

Mahapatra, S. K., & Dash, A. (2020). ARIMA model for forecasting of black gram productivity in Odisha. *Asiatic Society for Social Science Research, 2*(1), 131-136. https://doi.org/10.46700/asssr/2020/v2/i1/196124

Marahatta, S. (2021). Effect of plant densities and fertilizer rates on grain yield of spring maize in inner Terai condition. *Agronomy Journal of Nepal, 5*, 63-78. https://doi.org/10.3126/ajn.v5i01.44784

MoALD. (2023). *Statistical Information on Nepalese Agricutlure 2021/22.* Ministry of Agriculture and Livestock Development (MoALD). Singhadharbar, Kathamandu: Goverment of Nepal.

MoF. (2023). *Nepal foreign trade statistics.* Ministry of Finance, Department of Customs (DoC). Tripureshwor, Kathmandu: Goverment of Nepal.

Montgomery, D. C., Jennings, C., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. In D. J. Balding, G. M. Fitzmaurice, G. H. Givens, H. Goldstein, G.

Molenberghs, D. W. Scott, et al., *Wiley series in probability and statistics* (2nd ed.). John Wiley.

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

Panday, D. (2019, December 16). Future of maize: Introduce hybrid varieties. Kathmandu: The Himalayan Post.

Pokhrel, S. (2020). Food availability and consumption in relation to developinh strategies for sustained production and supply in Nepal. *Journal of Agriculture and Forestry University, 4*, 13-28. https://doi.org/10.3126/jafu.v4i1.47024

Poudel, B. (2024). Estimation of aggregate consumption function for Nepal using Ardl model. *Food & Agribusiness Management, 5*(2), 79-84.

Rahman, N. M., Hasan, M. M., Hossain, M. I., Baten, M. A., Hosen, S., Ali, M. A., et al. (2016). Forecasting rice area and production in Bangladesh using Box-Jenkins approach. *Bangladesh Rice Journal, 20*(1), 1-10.

Ray, S., & Bhattacharyya, B. (2020). Time series modeling and forecasting on pulses production behavior of India. *Indian Journal of Ecology, 47*(4), 1140-1149.

SenthamaraiKannan, K., & Karuppasamy, K. M. (2020). Forecasting for agricultural production using ARIMA model. *Palarch's Journal Of Archaeology Of Egypt/Egyptology, 18*(7), 5939-5949.

Sharma, P. K., Dwivedi, S., Ali, L., & Arora, R. K. (2018). Forecasting maize production in India using ARIMA model. *Agro-Economist, 5*(1), 1-6.

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*(2), 174-181. https://doi.org/10.24925/turjaf.v10i2.174-181.4618

Timsina, K. P., Ghimire, Y. N., & Lamichane, J. (2016). Maize production in mid hills of Nepal: From food to feed security. *Journal of Maize Research and Development, 2*(1), 20-29. https://doi.org/10.3126/jmrd.v2i1.16212

Tripathi, R., Nayak, A. K., Raja, R., Shahid, M., Kumar, A., Mohanty, S., et al. (2014). Forecasting rice productivity and production of Odisha, India, using Autoregressive Integrated Moving Average Models. *Advances in Agriculture, 2014*(1), 1-9. http://dx.doi.org/10.1155/2014/621313

Upadhyay, N., Ghimire, Y. N., Sharma, B., Acharya, Y., Gairhe, S., & Sapkota, S. (2018). Factors affecting adoption of maize varieties in Nepal. *Journal of Institution of Agriculture and Animal Science, 35*(1), 39-45. https://doi.org/10.3126/jiaas.v35i1.22512

Yasmin, S., & Moniruzzaman, M. (2024). Forecasting of area, production, and yield of jute in Bangladesh using Box-Jenkins ARIMA model. *Journal of Agriculture and Food Research, 16*(101203). https://doi.org/10.1016/j.jafr.2024.101203

Reference to this paper should be made as follows:
Poudel, B. (2024). Predicting cultivation area, production, and yield of maize in Nepal: An ARIMA model approach. *Nep. J. Stat, 8,* 61-78.