ZIBELINE INTERNATIONAL™ PUBLISHING
ISSN: 2682-7786 (Online)
CODEN: BDAIDR

Big Data in Agriculture (BDA)

DOI: http://doi.org/10.26480/bda.02.2025.85.92

CrossMark

RESEARCH ARTICLE

TREND ANALYSIS AND FORECASTING OF LARGE CARDAMOM AREA, PRODUCTION AND PRODUCTIVITY IN NEPAL USING ARIMA MODEL

Gaurav Thapa* **and Neha Sah Teli**

Department of Agricultural Economics, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, UP, India
Corresponding Author Email: gaurav.thapa@bhu.ac.in

ARTICLE DETAILS

Article History:

Received 23 August 2024
Revised 18 September 2024
Accepted 03 October 2024
Available online 05 November 2024

ABSTRACT

This study examines the trends and future projections of large cardamom (*Amomum subulatum Roxburgh*) area, production and productivity in Nepal. Utilizing secondary data from the Nepalese Ministry of Agriculture and Livestock Development and the Federation of Nepalese Chambers of Commerce and Industry, the research spans from 1994-95 to 2021-22 year. The analysis employs the Mann-Kendall test for trend detection and the ARIMA model for forecasting. The results reveal a stable increase in the area under cultivation and overall production, though productivity per hectare remains relatively constant. Large cardamom is identified as a significant contributor to Nepal's agricultural GDP and a crucial export commodity. The findings provide valuable insights for policymakers and stakeholders to develop strategies that support sustainable growth and address challenges such as climate change and market volatility. Advanced statistical models like ARIMA prove essential in enhancing agricultural forecasting and decision-making.

KEYWORDS

Large Cardamom, ARIMA Model, Trend Analysis, Forecasting, Mann-Kendall Test, Agricultural GDP, Export, Climate Change, Market Volatility, Agricultural Forecasting, Sustainable Growth, Policy Development

1. INTRODUCTION

Large Cardamom (*Amomum subulatum Roxburgh*) is a perennial herbaceous sciophytic crop of Zingiberaceae family, under the order Scitamineae grown in north facing hill slope (Shrestha, 2018; Shrestha et al., 2018), also known as 'Black Cardamom', or 'Nepalese Cardamom', or 'Alaichi' in Nepali or 'Badi Alaichi' in Hindi; renounced as black gold (Peter and Shylaja, 2012; Kalauni and Joshi, 2019).

Large cardamom thrives in cool, humid, and shaded areas between 800 to 2,100 meters (m) above sea level. The plant is sensitive to climate, with optimal conditions for the best production being temperatures ranging from 4 °C to 20 °C, annual rainfall between 2,000 millimeter (mm) to 2,500 mm, and humidity levels exceeding 90 percent (%) (FAO, 2024). Plant is thought to have originated in Sikkim and Darjeeling states in India, as well as in the eastern hills of Nepal. By 1830, farmers in four Nepali districts — Ilam, Taplejung, Panchthar, and Bhojpur — were cultivating it (Shrestha, 2018; ITC, 2017). Introduced to the Ilam district in 1865, large cardamom saw its commercial cultivation only beginning in the 1950s (FBC, 2008; Rijal, 2013). Over 80,000 households are involved in large cardamom farming in 55 districts of hills and few mountains ecological regions of Nepal contributing to 0.3177% to total agricultural Gross Domestic Product (GDP) of Nepal (MoALD, 2023; NSO, 2022-23). Nepal produces 8,714 metric tons (mt.) from 15,975 hectare (ha) of land while around 78 percent is produced by top five districts of Nepal which are Taplejung, Panchthar, Sankhuwasabha, Illam and Khotang respectively as shown in Figure 1 (MoALD, 2023).

Large cardamom is a high value low volume crop in Nepal with highest export value in terms of foreign currency exported 53,67,443 kilogram (kg) worth of Rs. 4,81,34,65,000 and quantity imported was 175 kg worth of Rs. 3,04,000 (Sharma et al., 2016; MoALD, 2023). Only 2 to 5 percent of the overall large cardamom production is used domestically, with India being the primary importer, buying about 99 percent of Nepal's large cardamom exports. The remaining 1 percent is exported to other markets such as Singapore, Pakistan, the UAE, Canada, China, and Bangladesh (ITC, 2021). Due to its delightful fragrance, the fruit is a crucial component in blended spices for enhancing the flavor of various food dishes in both industrial and culinary settings, as well as for medicinal purposes (Gudade et al., 2013). The research indicates that large cardamom exhibits pesticidal properties as well (Satyal et al., 2012) seed oil components, such as 1, 8-cineole, display properties that act against nematodes and insects, particularly targeting fruit fly activity (Bist and Bhatt, 2021).

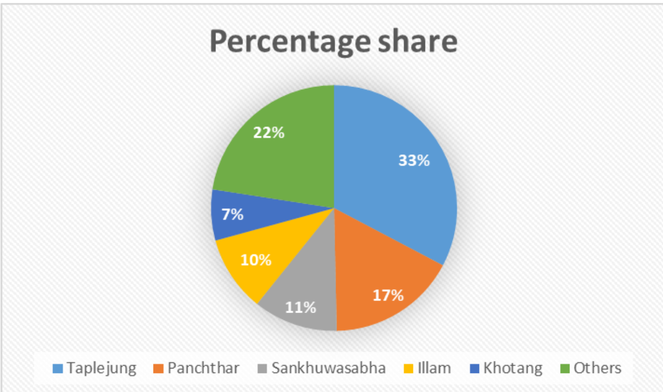


Figure 1: Top five large cardamom producing districts of Nepal

Though, primary agricultural exports of Nepal include lentil (constituting 29.6 % of total exports), cardamom (7 %), wheat (6.7 %), and tea (4 %) where large cardamom rank as the second most significant commodity in

Quick Response Code	Access this article online	
	Website: www.bigdatainagriculture.com.my	**DOI:** 10.26480/bda.02.2025.85.92

Cite The Article: Gaurav Thapa and Neha Sah Teli (2025). Trend Analysis and Forecasting of Large Cardamom Area, Production and Productivity in Nepal Using Arima Model. *Big Data in Agriculture, 7(2): 85-92.*

terms of export volume (CBS, 2013). Cardamom production is decreasing, and it was mainly planted on outward-sloped terraces with improper mulch placement. Farmers typically irrigate using sprinklers fed by spring water from upstream, but during the visit, few operational sprinklers were found. Trees planted long ago may compete for nutrients, and the lack of effective soil moisture conservation practices on sloping land could contribute to declining crop productivity (ICIMOD, 2016). Factors such as aged plantations, varied agricultural systems, a decrease in pollinator numbers, less flowering, altered climate patterns, and natural disasters contribute to the decline in cardamom productivity. The study concludes that despite continuous price increases, growing large cardamom remains essential for sustaining the livelihoods of small farmers in mountainous regions due to rising domestic demand and excessive spending on imports (Negi et al., 2018).

The cost of large cardamom is influenced by the market in Siliguri, West Bengal, India, leading to significant price fluctuations. Farmers consistently express dissatisfaction, noting that lowering the selling price of large cardamom results in significant losses for them (Timsina et al., 2012). Large cardamom production significantly contributes to people's livelihoods, serving as the primary household income source and offering seasonal employment to thousands from farm to market. Cultivating large cardamom has allowed farmers to boost household income by at least three times compared to traditional crops. A precise prediction can indicate suitable strategies for managing surpluses and deficits, stabilizing prices, and guaranteeing profitability for farmers (Kumar, 2020). Forecasting holds significant importance in strengthening policy decisions, securing food supplies, overseeing import/export activities, and enforcing pricing policies (Badmus and Ariyo, 2011; Sharma et al.,2018). Additionally, proper forecasting can address issues such as land use planning, food safety measures, selection of high-yielding crop varieties, provision of training for enhanced agricultural practices, ensuring sufficient supply of inputs, adoption of cutting-edge technologies, and addressing security and environmental concerns (Mahapatra and Dash, 2020; Tripathi et al., 2014).

ARIMA model is commonly used to predict future needs for internal consumption and export, enabling the adoption of suitable measures (Muhammad et al., 1992; Shabur and Haque, 1993; Sohail et al., 1994; Iqbal et al., 2005). Pakistan's wheat cultivation faces challenges like traditional farming methods, low yields, and shortages of inputs and irrigation water, leading to fluctuating production and unstable prices. Farmers often suffer losses due to inadequate marketing facilities and lack future production insights. Therefore, forecasting the area, yield, and production of large cardamom is crucial. The ARIMA model, by analyzing past trends, provides accurate predictions, aiding farmers and policymakers in making informed decisions to stabilize production and prices, thus supporting agricultural planning and economic growth (Iqbal et al., 2005). Large cardamom plays a crucial role in enhancing rural livelihoods and the economy (Chettri et al., 2024). The objective of this study is to analyze past trends and provide accurate predictions to farmers and policymakers, thereby stabilizing production and prices, while also boosting productivity and exports.

2. Methodology

The research draws upon secondary data sourced from the statistical records of the Nepalese Ministry of Agriculture and Livestock Development, provided by Government of Nepal. The dataset spans the years 1994-95 to 2021-22 and was utilized for the analysis of trends and the forecasting of large cardamom's area, production, and productivity in Nepal. Additionally, information pertaining to the quantity exported and revenue generated has been extracted from the Federation of Nepalese Chambers of Commerce and Industry, covering the period from 2009-10 to 2022-23.

Quantitative data was collected and inputted into MS-Excel 2013, where it underwent analysis to extract valuable insights using both R programming and MS-Excel. The results were visually presented through tables, graphs, line charts, and pie charts to effectively illustrate the findings.

2.1 Data Analysis Technique

2.1.1 Mann-Kendall Test

The Mann-Kendall (MK) test serves the purpose of statistically evaluating whether a variable exhibits a monotonic upward or downward trend over time. A monotonic upward (downward) trend indicates that the variable consistently increases (decreases) over time, although the trend may not necessarily be linear. Unlike parametric linear regression analysis, which assesses whether the slope of the estimated linear regression line differs from zero, the MK test does not assume normal distribution of residuals from the fitted regression line. As a non-parametric (distribution-free) test, the MK test provides a robust method for trend analysis without relying on specific distributional assumptions (Mann, 1945; Kendall, 1975; Gilbert, 1987). The Mann-Kendall (MK) test is primarily regarded as an exploratory analysis tool and is best utilized for identifying stations where significant or substantial changes occur over time. It enables quantification of these findings, allowing for a better understanding of the magnitude and significance of observed trends (Hirsch et al., 1982).

The Mann-Kendall (MK) test is conducted by arranging the data in the chronological order in which they were collected over time, $x_1, x_2, \ldots, x_n$, which denote the measurements obtained at times, 1, 2, …, n, respectively. Determined the sign of all n(n-1)/2 possible differences $x_j - x_k$, where j > k. These differences are $x_2 - x_1, x_3 - x_1, \ldots, x_n - x_1, x_3 - x_2, x_4 - x_2, \ldots, x_n - x_{n-2}, x_n - x_{n-1}$. Let sign $(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$, that is,

Sgn $(x_j - x_k) = 1$, if $x_j - x_k > 0$

$(x_j - x_k) = 0$, if $x_j - x_k = 0$ or, if the sign of $(x_j - x_k)$ cannot be determined due to non-detects

$(x_j - x_k) = -1$, if $x_j - x_k < 0$

Compute,

$$S = \sum_{k-1}^{n-1} \sum_{j-k+1}^{n} sgn\ (x_j - x_k)$$

value is calculated as the difference between the number of positive differences and the number of negative differences. A positive result suggests that observations collected later in time are generally larger than those made earlier, while a negative result indicates the opposite trend, with later observations typically being smaller than earlier ones.

If n> 10, by referring to a table of probabilities, if the obtained probability (S) is lower than the significance level (α), which represents the probability of mistakenly concluding a trend exists when it does not, the null hypothesis is rejected, and it is concluded that a trend exists. In cases where the sample size (n) cannot be found in the table, usually due to tied data values, the closest available value to the calculated S statistic in the table is used instead.

If n > 10,

Compute the variance of S as follows:

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)\right]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Compute the MK test statistic, Z_{MK}, as follows:

$$Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}} \text{ if } S > 0$$

$$= 0 \text{ if } S = 0$$

$$= \frac{S+1}{\sqrt{VAR(S)}} \text{ if } S < 0$$

A positive (negative) value of Z_{MK} indicates that the data tend to increase (decrease) with time.

The calculated Z-value was then compared with the corresponding Z-value from the standard Z-probability table, typically at a significance level of 5 percent. This comparison helped determine whether to accept or reject the hypothesis and draw conclusions regarding the presence of a monotonic pattern. A positive Z-value suggested an increasing trend, while a negative Z-value indicated a decreasing trend (Alhaji et al., 2018). In this investigation, the determination of whether a monotonic trend existed in the provided time series data relied on the obtained p-value from the Mann-Kendall test. If the p-value fell below the significance level (α = 5%), which is typically 0.05, the null hypothesis was rejected, indicating the presence of a monotonic trend in the data. Conversely, if the p-value exceeded 0.05, the null hypothesis was retained, suggesting the absence of a monotonic trend in the time series data.

The determination of the strength of the monotonic relationship between the two variables was based on the Kendall correlation coefficients, commonly referred to as Kendall's tau. Kendall's correlation coefficient ranges from -1 to +1, where the sign of the coefficient indicates the direction of the relationship slope, whether it is increasing or decreasing. The magnitude of the coefficient represents the intensity or strength of the relationship between the variables (Helsel and Hirsch, 1992).

2.1.2 Sen's Slope

Originally devised by Sen, this test was designed to assess statistical linear relationships and determine the magnitude of trends in long-term temporal data (Agarwal et al., 2021). Sen's slope is preferred for detecting linear relationships as it remains unaffected by outliers present in the data (Ray et al., 2021). In cases where time series data exhibit linear trends, the slope or trend magnitude is often estimated using the least squares or simple linear regression technique. However, the least squares estimation of the regression coefficient (i.e., slope) is susceptible to significant errors, and the confidence interval derived from it is influenced by the non-normality of the parent distribution (Sen, 1968). Moreover, if there are significant errors or outliers present in the data, the slope calculated using this method may diverge significantly from the actual slope (Gilbert, 1987).

The linear model assumed was

$Y = \alpha + \beta t$

Where,

β= slope

α = Intercept

Y = time i.e., independent variable

The median is less impacted by significant errors or outliers compared to the weighted average, making median-based regression coefficient calculations more robust than those obtained from the least squares method.

$$\beta_{ij} = \frac{y_j - y_i}{j - i}$$

β_{ij} = all the slopes of lines connecting each pair of points (t_i, y_i) and (t_j, y_j) and $t_i \neq t_j$

If the series comprises n values, then we have precisely obtained $N = \frac{n(n-1)}{2}$ slope estimates β_{ij}. The N values representing slopes are sorted in ascending order of magnitude, and the median, which serves as the Sen Estimator of a slope, is then calculated.

$$\beta = \begin{Bmatrix} \frac{1}{2}\left(\beta_{\frac{N}{2}} + \beta_{\frac{N+2}{2}}\right), if N\ is\ even \\ \beta_{\frac{N+1}{2}}, \quad if\ N\ is\ odd \end{Bmatrix}$$

The intercept for each time step t was computed using the Sen's method as follows:

$\alpha_t = Y_t - \beta t$

The respective intercept value (α) is calculated as the mean of all αt values (Alhaji et al., 2018). A positive slope (β) value obtained from Sen's method indicates an upward trend, while a negative value signifies a downtrend. The regression line was plotted to visualize the trend, utilizing the slope and intercept values derived from Sen's calculations. This method was employed to analyze and identify linear trends within time-series data concerning the production, area, and productivity of large cardamom in Nepal.

For this test, each subsequently measured value is compared to all earlier measured values, resulting in a total of $(n-1)2$ possible pairs of data, where n represents the total number of observations. The null hypothesis (H_0) states that there is no monotonic trend in the time-series, while the alternative hypothesis (H_1) asserts the presence of a monotonic trend (Meals et al., 2021).

2.1.3 ARIMA Test

The ARIMA approach, pioneered by Box and Jenkins in 1976, has garnered significant popularity across various fields, with research confirming its efficacy and versatility (Hoff, 1983; Pankratz, 1983; Vandaele, 1983; Biswas et al., 2014). The ARIMA model serves as a method of extrapolation used for forecasting, similar to other techniques of its kind. It relies solely on historical time series data pertaining to the variables being forecasted (Singh et al., 2015). This model serves as a form of regression analysis, assessing the relationship strength between a dependent variable and its independent variable. It encompasses three distinct processes: autoregressive (AR) of order p, differencing of degree d to achieve time-series stationarity, and moving average (MA) of order q, denoted as ARIMA (p,d,q). The Box-Jenkins methodology of the ARIMA model employs non-stationary time-series data to forecast future values. Through differentiation within the ARIMA model, known as the 'd' degree of differencing, non-stationary time-series are transformed into stationary ones. Typically, a stationary stochastic process is utilized for forecasting future point values within a time series (Verbeek, 2004; Thapa et al., 2022).

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

The equation above suggests that alterations in the time origin do not impact the mean, variances, and autocovariances. The ARIMA model with order (p, d, q) can be represented in econometric model form as depicted below.

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

Where,

Y_t is given variable at time t, is the mean of series, the $\theta 1, \ldots \theta_p$ are the parameters of the AR model, the $\alpha_1, \alpha_2, \ldots, \alpha_q$ are the parameters of the MA model and the $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$ are white noise residuals. Similarly, the mathematical form of the ARIMA model of order (p, d, q) can be expressed as:

$$\emptyset_p(L)(1-l)^d Y_t = c + \theta_q(L)\varepsilon_t$$

Lag operator; some uses as Backshift operator, d = order of difference operator, p = order of non-seasonal AR operator; and q = order of non-seasonal MA operator (Verbeek, 2004; Thapa et al., 2022).

The stationarity requirements and invertibility of the time-series were satisfied only when all roots of the characteristic equations $\emptyset_p(L) = 0$, $\emptyset_p(L) = 0$ lied outside the unit circle (Thapa et al., 2022, Box and Jenkins, 1976).

2.1.4 Model prediction and validation

2.1.4.1 Identification

Identification involves techniques used to determine the values of p, q, and d in an ARIMA (p, d, q) process. These values are determined through the examination of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). In an ARIMA (p, d, q) process, the theoretical PACF exhibits non-zero partial autocorrelations at lags 1 through p and zero partial autocorrelations at all other lags, while the theoretical ACF displays non-zero autocorrelations at lags 1 through q and zero autocorrelations at all other lags. The nonzero lags observed in the sample PACF and ACF are provisionally accepted as the parameters p and q, respectively. In the case of a non-stationary series, differencing is applied to achieve stationarity, with the number of differencing operations determining the value of d. Consequently, for a stationary dataset, d equals 0, and the ARIMA (p, d, q) model can be represented as ARMA (p, q) (Singh et al., 2015).

2.1.4.2 Estimation

Once the suitable model was chosen, the model parameters were estimated utilizing the Maximum Likelihood (ML) method. ML is widely recognized as one of the most effective techniques for parameter estimation within an ARIMA model Thapa et al., 2022, Box and Jenkins, 1976).The primary methodologies for fitting Box-Jenkins models involve non-linear least squares and maximum likelihood estimation. These estimation techniques were employed using R Studio.

2.1.4.3 Diagnostic Checking

The optimal ARIMA model fit was determined using two performance criteria: the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The model with the lowest value of either criterion was selected as the best model. Although both AIC and BIC were considered, AIC was prioritized as the primary criterion for selecting the best model (Verbeek, 2004). In addition to the AIC and BIC criteria, the goodness of fit was assessed by examining the autocorrelation function (ACF) of the residuals from the fitted model. The autocorrelation of residuals was evaluated using the Box-Ljung Q test, a statistical tool that tests for the presence of autocorrelation in residuals. The Box-Ljung test statistic is calculated as follows:

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

Where n, represents the number of samples, r_k denotes the estimated autocorrelation of the series at lag k = 1,2, ..., m and 'm' signifies the number of lags under consideration. An insignificant p-value at a specified level of significance confirms the absence of correlation in the residuals. Thus, this approach enables us to consider the model as the best fit (Gujrati and Porter, 2009).

2.1.4.4 Model Adequacy

The outlined procedure constitutes the essence of the Box-Jenkins methodology. Following the selection of the best-fitted ARIMA model, we assessed the model's adequacy using the test dataset spanning from 2094-95 to 2021-2022 to evaluate its goodness of fit. The selected ARIMA model was then utilized to forecast predicted values for the same time frame. Model adequacy was evaluated using the Mean Absolute Percentage Error (MAPE).

The formula to calculate the MAPE is as follows:

$$MAPA = \left(\frac{1}{n}\right) \times \sum(|actual - forecast|/|actual|) \times 100$$

Where, n = Sample size, actual = the actual data value and forecast = the forecasted data value.

2.1.4.5 Model Forecasting

The model was employed to forecast the production, area, and productivity of large cardamom in Nepal for the upcoming ten years, spanning from 2023 to 2032, encompassing the production period.

3. RESULTS

3.1 Trend of Area, Production and Productivity of Large Cardamom

3.1.1 Trend of large cardamom area in Nepal

The positive slope coefficient (211.74) indicates a positive trend in the area over the years as shown in Fig 2. The average increase of approximately 211.74 Ha per year suggests a general upward trend in the area. The calculated growth percentage is approximately 2.41% per year. This means that, on average, the area increases by about 2.41% each year according to the linear regression model. R^2 estimated was 0.8009 represents the proportion of the variance in the dependent variable (area) that is predictable from the independent variable (year) suggests that 80.09% of the variability in the area can be explained by the linear regression model.

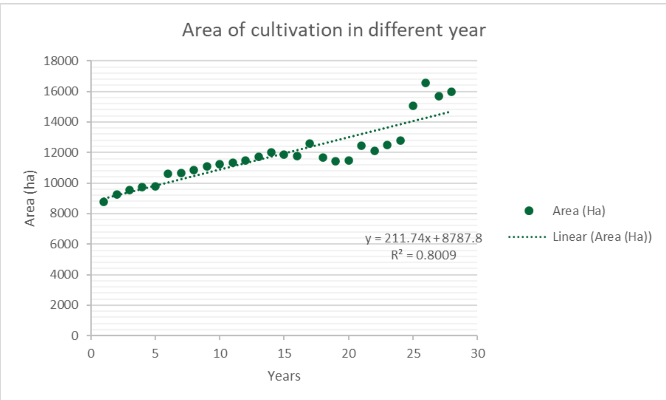


Figure 2: Trend showing large cardamom cultivation area (ha) in Nepal from year 1994-95 to 2021-22

3.1.2 Trend of Large Cardamom Production (mt) in Nepal

The positive coefficient of the year (x) indicates a positive linear relationship between the year and production as shown in Fig 3. Specifically, the coefficient 127.58 suggests that, on average, the production increases by approximately 127.58 Mt per year. The trend in the production data is upward, indicating a positive linear growth in production over the years. The R^2 value of 0.53 also suggests that approximately 53% of the variability in production can be explained by the linear relationship with the year. Model indicates a positive relationship between area and production, suggesting that as the area increases, the production tends to increase.

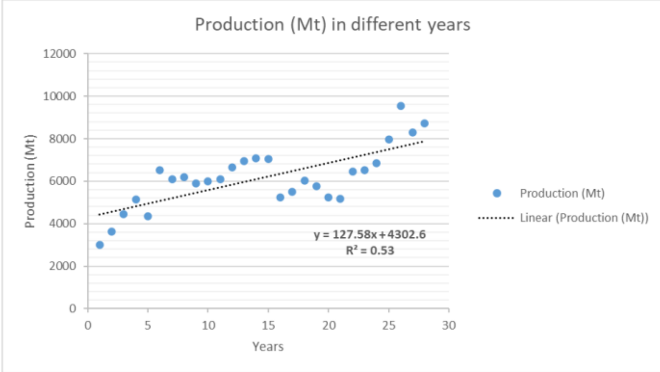


Figure 3: Trend showing large cardamom production (mt) Nepal from year 1994-95 to 2021-22

3.1.3 Trend of productivity (mt/ha) of large cardamom in Nepal

The data shows a weak positive trend in productivity over the years, but the low R^2 value is quite low (0.0493), indicating that only a small proportion (approximately 4.93%) of the variability in productivity can be explained by the linear relationship with the year as shown in Fig 4. This suggests that other factors not accounted for in the linear model may influence productivity. The positive slope coefficient suggests a positive relationship between the year and productivity. However, the small magnitude of the coefficient (0.0018) indicates a very gradual increase in productivity over the years.

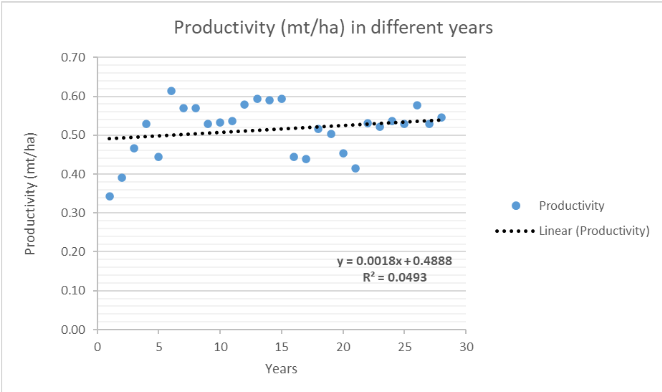


Figure 4: Trend showing large cardamom productivity in Nepal from 1994-95 to 2021-22

3.1.4 Mann Kendall and Sen's Slope

For Area the low p-value suggests strong evidence to reject the null hypothesis, implying that the parameter 'S' (possibly representing the area of cardamom cultivation) is significantly different from zero in the time series data. This is a measure of how many standard deviations an observation is from the mean. A z-value of 6.3814 indicates a significant deviation from the expected value. The positive estimate of 'tau' (0.8571429) indicates a positive trend or pattern in the time series data for cardamom cultivation area in Nepal. The positive Sen's slope (192.0143) suggests a positive trend in the area of cardamom cultivation over time in Nepal. This indicates a general increase in cardamom cultivation area over the observed period. The 95 percent confidence interval (146.6316 to 249.4444) provides a range of plausible values for the true Sen's slope. Since this interval does not include zero, it further supports the notion that there is a significant positive trend in the cardamom cultivation area. This indicates a general increase in cardamom cultivation area over the observed period.

Similarly for production also, small p- value (0.0002572) suggests strong evidence against the null hypothesis, indicating significant findings. This value measures how many standard deviations a data point is from the mean. In this case, a z-value of 3.655 indicates a significant deviation from the expected value. The positive value of 'tau' (0.4920635) and Sen's slope (130.7385) suggest a positive trend in cardamom production over time in Nepal which indicates a statistically significant positive trend in cardamom production in Nepal over the observed time series, with strong evidence against the null hypothesis. The values of 'tau' and Sen's slope provide specific measures of trend, while 'S' and 'varS' offer insights into the central tendency and variability of the data.

For productivity z-value (0.84953) suggests a relatively small deviation from the expected value, indicating limited evidence against the null hypothesis. The p-value (0.3956) being greater than 0.05 indicates that there is not enough statistical evidence to reject the null hypothesis of the true parameter 'S' being equal to 0. The positive value of 'tau' (0.1164021) and Sen's slope (0.001308747) suggest a positive, but potentially weak, trend in cardamom productivity over time in Nepal, suggests that there is not strong evidence to conclude a significant trend in cardamom productivity in Nepal over the observed period. The positive values of 'tau' and Sen's slope indicate a potential positive trend, but the p-value and z-value suggest caution in making strong conclusions.

Table 1: Model fit statistics for Mann- kendall and sen's slope

Parameter	z-value	p-value	Kendall Tau	Sen's slope	Trend
Area (ha)	6.3814	1.755e-10	0.8571	192.0143	Increasing
Production (mt)	3.655	0.0002572	0.4920635	130.7385	Increasing
Productivity (mt/ha)	0.84953	0.3956	0.11	0.0013	Weak increasing trend

At the 95% confidence interval

3.2 ARIMA Model Analysis for Large Cardamom

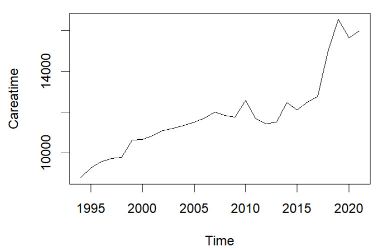
In this study, we employed the ARIMA model to forecast the area, production, and productivity of large cardamom over a ten-year period as shown in table 3. The data was successfully transformed into a time series object with yearly frequency, facilitating the analysis. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were calculated to determine the order of the ARIMA model (p, d, q) as shown in Figure 5. ACF measures the correlation between observations at different time lags, while PACF measures the correlation between observations that are a specific number of time periods apart, after adjusting for the effects of intervening observations. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to select the optimal ARIMA model. The model fit statistics, including AIC, BIC, and log likelihood, indicate a good fit for the data as shown in table 2. For Area; ARIMA (0, 1, 0) model with drift predicts a steady increase in the cultivated area of large cardamom over the next decade. The drift term suggests a consistent upward trend. As for production; ARIMA (0, 1, 0) model indicates stable production levels. This model does not include a drift component, reflecting the assumption that production changes are due solely to random shocks rather than a consistent trend. Since for productivity ARIMA (1, 0 ,0) model with a non-zero mean forecasts a slight increase in productivity. The positive AR(1) coefficient suggests that past productivity levels have a significant influence on future values. The Box-Ljung test p-values for all models are above the 0.05 threshold, indicating no significant autocorrelation in the residuals and confirming the adequacy of the models.

The ARIMA model parameters for each of the series (area, production, and productivity) are summarized in the Table 2:

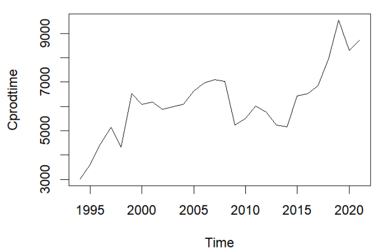
Table 2: Model fit statistics of Arima model

Parameters	**Area**	**Production**	**Productivity**	
ARIMA model	ARIMA (0,1,0) with drift	ARIMA (0,1,0)	ARIMA (1,0,0) with non-zero mean	
Coefficients	Drift	-	ar1	mean
	266.4074	-	0.5659	0.5091
s.e.	121.2376		0.179	0.0239
sigma^2	412126	702278	0.003456	
log likelihood	-212.34	-220.05	40.46	
AIC	428.69	442.1	-74.92	
AICc	429.19	442.26	-73.92	
BIC	431.28	443.39	-70.93	
Box -Ljung Test				
Coefficients	Area	Production	Productivity	
X- squared	2.6672	3.7897	5.1297	
p-value	0.7511	0.5801	0.4003	

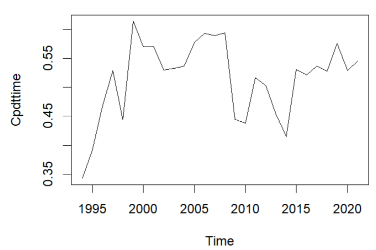
At 5% level of significance



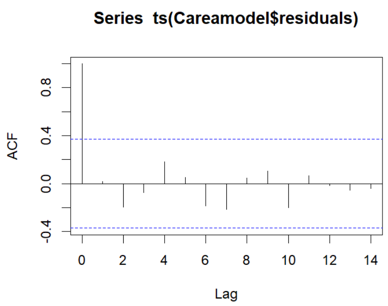
Plot of time series area data



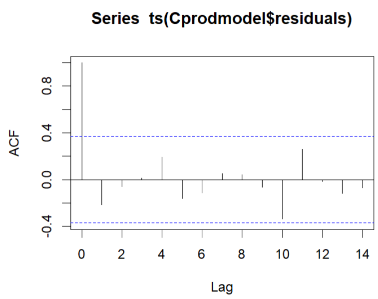
Plot of time series Production date



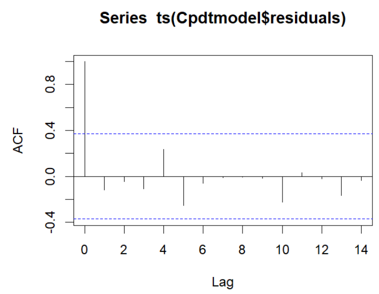
Plot of time series Productivity data



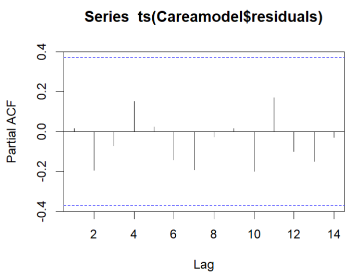
Autocorrelations (ACF) of area series after differencing



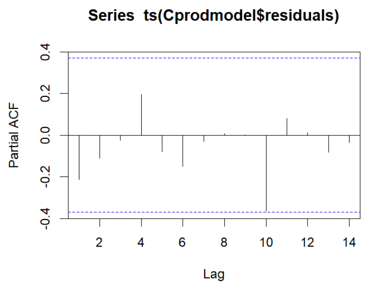
Autocorrelations (ACF) of production series after differencing



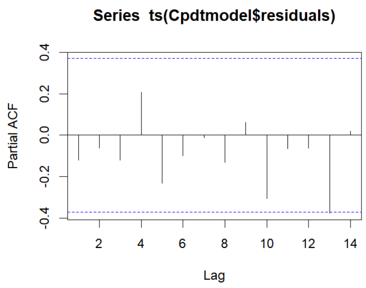
Autocorrelations (ACF) of productovity series after differencing



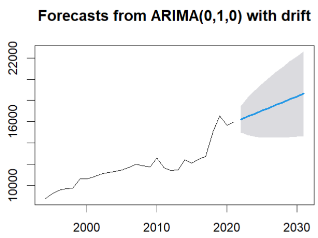
Partial Autocorrelations (PACF) of area series after differencing



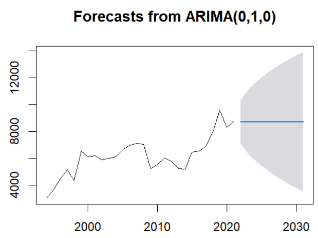
Partial Autocorrelations (PACF) of production series after differencing



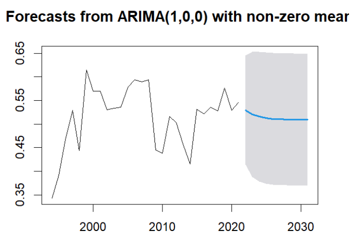
Partial Autocorrelations (PACF) of productivity series after differencing



Area forecasted for ten years



Production forecasted for ten years



Productivity forecasted for ten years

Figure 5: Plot of time series data, ACF, PACF after differencing and forecast of area, production and productivity of Large Cardamom

The ARIMA models were used to forecast the area (in hectares), production (in metric tons), and productivity (in metric tons per hectare) of large cardamom for the next 10 years. The forecasted values, along with the 95% confidence intervals, are presented below:

Table 3: Forecasting of Area (ha), Production (mt) and Productivity (mt/ha) of large cardamom for ten years by ARIMA model

Year	Area			Production			Productivity		
Point	Forecast	Lo 95	Hi 95	Forecast	Lo 95	Hi 95	Forecast	Lo 95	Hi 95
2022	16241.41	14983.17	17499.65	8714	7071.511	10356.49	0.529674	0.414452	0.644896
2023	16507.81	14728.4	18287.23	8714	6391.17	11036.83	0.520731	0.38834	0.653122
2024	16774.22	14594.89	18953.56	8714	5869.125	11558.87	0.51567	0.378234	0.653107
2025	17040.63	14524.15	19557.11	8714	5429.022	11998.98	0.512807	0.373793	0.65182
2026	17307.04	14493.53	20120.54	8714	5041.283	12386.72	0.511186	0.371672	0.650701
2027	17573.44	14491.4	20655.49	8714	4690.74	12737.26	0.510269	0.370595	0.649944
2028	17839.85	14510.87	21168.84	8714	4368.382	13059.62	0.50975	0.370024	0.649476
2029	18106.26	14547.42	21665.1	8714	4068.339	13359.66	0.509457	0.369714	0.649199
2030	18372.67	14597.95	22147.38	8714	3786.533	13641.47	0.50929	0.369543	0.649038
2031	18639.07	14660.17	22617.97	8714	3519.993	13908.01	0.509196	0.369447	0.648945

By these forecasts of production, area, and productivity of large cardamom have several practical applications. For policymakers, these forecasts provide critical data for planning and implementing agricultural policies. They can help in determining the amount of subsidy or financial support needed, planning for export quotas, and ensuring that there is no overproduction that could lead to a fall in prices. For farmers, understanding the projected trends in productivity and area can guide their planting decisions, investment in fertilizers and pesticides, and other crop management practices. By anticipating market trends, farmers can optimize their production to meet market demands, thereby maximizing their profits. Furthermore, the forecasts are invaluable for exporters and traders. They can use the data to plan their marketing strategies, negotiate better prices, and explore new markets. This proactive approach can help stabilize the market prices and ensure a steady supply chain, benefiting all stakeholders involved in the large cardamom cultivation and marketing.

3.3 Trend Analysis of export earnings and quantity exported of large cardamom

The Figure 6 shows the quantity of large cardamom exported (in kilograms) and the amount earned (in NPR) for each fiscal year from 2009/10 to 2021/22. The linear regression equation suggests a positive relationship between the quantity of large cardamom exported and the amount earned. As the quantity increases, the amount earned is expected to increase according to the equation. The coefficient of determination (R^2) is a measure of how well the linear regression model explains the variability in the dependent variable (amount earned) based on the independent variable (quantity exported). Here low value of R^2 = 0.0318 value indicates that the linear model does not provide a strong explanation for the variability in the amount earned. Other factors not included in the model may influence the earnings like as price fluctuations.

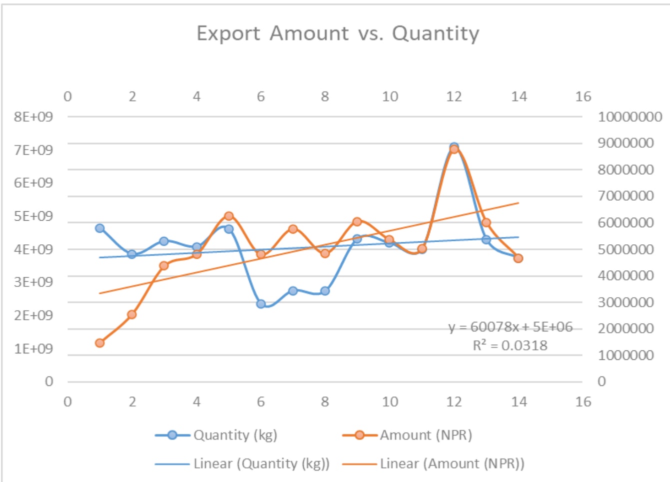


Figure 6: Trend analysis of export earnings and quantity exported of large Cardamom

4. Conclusion

The research successfully forecasts the future trends of large cardamom area, production and productivity in Nepal, highlighting its critical role in the agricultural sector. The ARIMA model, applied to historical data, offers reliable projections, indicating a consistent increase in cultivation area and production volume over the next decade. However, the productivity per hectare is projected to remain stable, suggesting that improvements in farming practices and resource management are necessary to boost yields. The study emphasizes the importance of large cardamom as a key export commodity, contributing significantly to Nepal's agricultural GDP and providing livelihoods for thousands of households. Despite the positive outlook, challenges such as climate change, aging plantations, and market fluctuations pose risks that need to be addressed through adaptive strategies and policy interventions. This research underscores the potential of advanced statistical methods in agricultural forecasting, offering valuable tools for policymakers and stakeholders to make informed decisions. Ensuring the sustainability and profitability of large cardamom farming requires continuous monitoring, investment in research and development, and support for farmers in adopting innovative agricultural practices. This paper lays the groundwork for future research to explore strategies for expanding the cultivation area and production of large cardamom, identifying factors influencing its production dynamics, and achieving higher productivity levels.

Acknowledgement

I would like to express my heartfelt gratitude to the Ministry of Agriculture and Livestock Development, Nepal, and international organizations such as the Food and Agriculture Organization (FAO) and the International Centre for Integrated Mountain Development (ICIMOD). Their comprehensive and accessible databases provided the essential data and resources necessary for this study. The invaluable access to these datasets was crucial for the successful completion of my research.

Additionally, I extend my deepest appreciation to my family and friends for their unwavering encouragement and support. Their motivation and inspiration have been indispensable throughout this journey.

References

Agarwal S., Suchithra A. and Singh S. P., 2021. Analysis and interpretation of rainfall trend using Mann-Kendall's and Sen's slope method. Indian Journal of Ecology 48 (2), Pp. 453–457.

Alhaji, U. U., Yusuf, A. S., and Edet, C. O., 2018. Trend Analysis of Temperature in Gombe State Using Mann Kendall Trend Test. JSRR 20, Pp. 1–9.

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

Bist, P., and Bhatt, P., 2021. Review On Status Of Large Cardamom (Amomum Subulatum Roxb.) Production And Marketing In Nepal. Food and Agri Economics Review (FAER), 1(2), Pp. 121-123.

Biswas, B., Dhaliwal, L. K., Singh, S. P., and Sandhu, S.K., 2014. Forecasting wheat production using ARIMA model in Punjab. International Journal of Agricultural Sciences, 10(1), Pp. 158-161.

Box, G. E. P., and Jenkins, G. M., 1976. Time Series Analysis, Forecasting and Control.

Central Bureau of Statistics (CBS), 2013. National Population Census 2011. Kathmandu: Central Bureau of Statistics.

Chettri, Santosh., Pandey, Aseesh., Gaira, Kailash., and Joshi, Rajesh., 2024. Large Cardamom-Based Agroforestry System for Livelihood Security in Sikkim Himalayas, Apple Academic Press , Pp: 425-432.

FAO, 2024. Food and Agricultural Organization. Codex Alimentarius Commision. Retrieved from https://www.fao.org/fao-who-codexalimentarius/sh-proxy/es/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FMeetings%252FCX-736-07%252FCRDs%252Fsc07_crd06_Rev.pdf

Full Bright Consultancy (FBC), 2008). Product Chain Study: Cardamom. Kathmandu: Ministry of Agriculture and Cooperatives, Department of Agriculture.

Gilbert, R.O., 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, New York.

Gudade, B. A., Chhetri, P., Gupta, U., Deka, T. N., and Vijayan, A. K., 2013. 6. Traditional Practices Of Large Cardamom Cultivation In Sikkim And Darjeeling. Life Sciences Leaflets, 43, Pp. 62-to.

Gujrati, D., and Porter D.C., 2009. Basic Econometrics (Fifth ed.). Irvin: McGraw-Hill.

Helsel, D.R., and Hirsch, R.M., 1992. Statistical Methods in Water Resources. Elsevier.

Hirsch, R.M., J.R. Slack, and R.A. Smith, 1982. Techniques of trend analysis for monthly water quality data, Water Resources Research 18(1), Pp. 107-121.

Hoff, J.C., 1983. A practical guide to Box-Jenkins forecasting.

ICIMOD, 2016. Nepal Water Conservation Foundation and International Center for Mountain Development.

Iqbal, N., Bakhsh, K. H. U. D. A., Maqbool, A. S. I. F., and Ahmad, A. S., 2005. Use of the ARIMA model for forecasting wheat area and production in Pakistan. Journal of Agriculture and Social Sciences, 1(2), Pp. 120-122.

ITC, 2017. Nepal National Sector Export Strategy: Large Cardamom. International Trade Center (ITC), Geneva.

ITC., 2021. Nepal National Sector Export Strategy: Large Cardamom. International Trade Center (ITC), Geneva.

Kalauni, D., and Joshi, A., 2019. Production economics, marketing and critical success factors of Large Cardamom in Bhojpur, Nepal. Cogent Food and Agriculture, 5(1), Pp. 1701358.

Kendall, M.G. 1975. Rank Correlation Methods, 4th edition, Charles Griffin, London.

Kumar, R.R., 2020. Forecasting of Potato Pvegetabless in India: An Application of Arima Model. Economic Affairs, 65(4), Pp. 473–479. https://doi.org/10.46852/0424-2513.4.2020.1

Mahapatra SK, Dash A., 2020. ARIMA Model for Forecasting of Black Gram Productivity in Odisha. Asiatic Society for Social Science Research, 2(1), Pp. 131. https://doi.org/10.46700/asssr/2020/v2/i1/196124

Mann, H.B., 1945. Non-parametric tests against trend, Econometrica 13, Pp. 163-171.

Meals, D. W., Spooner, J., Dressing, S. A., Harcum, J. B. 2021. Statistical Analysis for Monotonic Trends. National Nonpoint Source Monitoring Program, 23. https://www.epa.gov/sites/production/files/201605/documents/tech_notes_6_dec2013_trend.pdf

MoALD, 2023. Statistical Information on Nepalese Agriculture, 2021/22. Government of Nepal. Ministry of Agriculture and Livestock Development, Planning and Development Cooperation Coordination Division, Statistics and Analysis Section, Singhdurbar, Kathmandu, Nepal.

Muhammad, F., M. Siddique, M. Bashir and S. Ahmad, 1992. Forecasting rice production in Pakistan–using ARIMA Models. J. of Animal Plant Sci., 2, Pp. 27–31.

Negi, B. K., Joshi, R. k., and Pandey, A., 2018. Status of large cardamom farming systems in the changing scenario of modern economics of Sikkim, Himalaya. International Journal of Science and Nature, 7 (1), Pp. 189-199.

NSO, 2022-23. National Account Statistics of Nepal. Retrieved from https://moald.gov.np/wp-content/uploads/2023/08/Agricultural-Statistics-Pocket-book-2023-1.pdf

Pankratz, A., 1983. Forecasting with univariate Box-Jenkins.

Peter, K. V., and Shylaja, M. R., 2012. Introduction to herbs and spices: Definitions, trade and applications. In Handbook of herbs and spices, Pp. 1-24. Woodhead Publishing.

Ray S., Das S. S., Mishra P. and Al Khatib A. M. G., 2021. Time series SARIMA modelling and forecasting of monthly rainfall and temperature in the south Asian countries. Earth Systems and Environment 5, Pp. 531–546.

Rijal, S., 2013. Impact of climate change on large cardamom-based livelihoods in Panchtar District, Nepal. The Third Pole: Journal of Geography Education, 13 (33-38), Pp. 36-37).

Satyal, P., Dosoky, N. S., Kincer, B. L., and Setzer, W. N., 2012. Chemical compositions and biological activities of Amomum subulatum essential oils from Nepal. Natural product communications, 7(9), Pp. 1934578X1200700935.

Sen, P. K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American statistical association, 63(324), Pp. 1379-1389.

Shabur, S.A. and M.E. Haque, 1993. An analysis of rice price in Mymensing town market pattern and forecasting. Bangladesh J. of Agric. Econ., 16: 120–3.

Sharma, G., Partap, U., Dahal, D. R., Sharma, D. P., and Sharma, E., 2016. Declining large-cardamom production systems in the Sikkim Himalayas: climate change impacts, agroeconomic potential, and revival strategies. Mountain Research and Development, 36(3), Pp. 286-298.

Sharma, P.K., Dwivedi. S., Ali, L., Arora, R.K. , 2018. Forecasting maize production in India using ARIMA model. Agro-Economist, 5(1), Pp. 1–6. https://doi.org/10.30954/2394-8159.01.2018.1

Shrestha, J., Prasai, H. K., Timsina, K. P., Shrestha, K. P., Pokhrel, D., Poudel, K., and Yadav, M., 2018. Large cardamom in Nepal: Production practice and economics, processing and marketing. Pakhribas, Dhankuta: National Commercial Agriculture Research Program, Nepal Agriculture Research Council, Nepal.

Shrestha, K., 2018. Profitability of large cardamom enterprise in Nepal: Evidence from financial analysis. Journal of Agriculture and Natural Resources, 1(1), Pp. 76-89.

Singh, L. N., Darji, V. B., and Parmar, D. J., 2015. Forecasting of wheat production and productivity of Ahmedabad region of Gujarat state by using ARIMA models. Indian Journal of Economics and Development, 3(6), Pp. 1-7.

Sohail, A., A. Sarwar and M. Kamran, 1994. Forecasting total food grains in Pakistan. Department of Math and Stat, University of Agriculture, Faisalabad. J. Eng. Appl. Sci., 13, Pp. 140–6.

Thapa, R., Devkota, S., Subedi, S., and 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), Pp. 174-181.

Timsina, K.P., Shrestha, K.P, Pandey, S, and Paudel, I.P., 2012. Value Chain Analysis of Cardamom (Ammomum subalatum Roxb.): A case of Taplejung District, Nepal. Agriculture Development Journal, 9.

Tripathi R, Nayak AK, Raja R, Shahid M, Kumar A, Mohanty S, Panda BB, Lal B, Gautam P., 2014. Forecasting Vegetables Productivity and Production of Odisha, India, Using Autoregressive Integrated Moving Average Models. Advances in Agriculture, Pp. 1–9. https://doi.org/10.1155/2014/621313

Vandaele, W., 1983. Applied time series and Box-Jenkins models", Academic Press, New York. Pp. 417.

Verbeek, M., 2004. A Guide to Modern Econometrics (Second Edition). West Sussex, England: John Wiley andSons, Ltd.