Original Research Article

Econometric Modelling and Forecasting of Tea Production and productivity

Abstract

India is world's second leading tea producer, and biggest consumers in the worldwide. The main objective of this investigation is to identify the Box-Jenkins method an Autoregressive integrated moving average (ARIMA) models that can be used to make predictions the production and yield of tea in India. In thisstudy considered the published yearly secondary data of tea production and yield in Indiaperiod of 1970-71 to 2021-22. In accordance to the Sigma square, RMSE, MAPE, MAE, AIC and SIC the most appropriate modelsforprediction the tea production, and yield in India are ARIMA models (1, 1,1) and (0, 1, 1). Acceptability of the selected models has been verifiedwiththe Ljung-Box test, ACF, PACF criterion,and white noise followed by residual diagnostics. Comparison between the original data series and the estimated data series in the similar manner indications that the fitted model performsgood statistically and is suitable for forecasting 19 year (period of 2022-23 to 2040-41) the tea production and yield in India i.e., the models are good prediction during and beyond the forecasting period; and we have estimated that the annual amount of tea production and yield achieved in the year 2022-23 is 13500.17 lakh tonnes and 2119.73 kg/ha respectively, which will accelerate to reach 16728.34 lakh tonnes and 2444.34 kg/ha in the year 2041-41 respectively.

Keyword: Econometric Modelling, Time Series Analysis, Forecasting, TPY, ARIMA, AIC

JEL Code:C12, C22, C52. C53

1. Introduction

Tea,due to its cultivation and harvesting system, it is a unique harvest compared to any other common crops (Bekhit, 2006; Muganda et al., 2021). It's kindof type husbandry harvest which demonstrationsextensiveadaptableness and cultivates in a range of weathers and top soil in dissimilar parts of the worldwide (Hossain and Abdulla, 2015; Deka et al., 2021) and it's the world's oldest non-alcoholic caffeinated beverage. (Mondal et al., 2004). At the present time tea is the most populardrink in the all over worldand its refreshes the mind and gives energy. In recent time without a cup of tea, a common man cannot be think of starting the day.Approximately everyconsumers consumes a cup of tea one or two time tack daily (Khisa, 2001;Hossain and Abdulla, 2015). Due to its medicinal effects, tea is helpful in eliminating many diseases like reducing heart diseases, anti-obesity, reducing the risk of atherosclerosis, preventing tumor cell growth and reducing obesity etc. (Wang *et al*., 2010; Dhekale et al., 2014).Midst of the major tea producing countries are China, India, Sri Lanka, Kenya and Indonesia, these five countries are contributes of 80% tea exports of worldwide and globally 77% of tea production (Gupta and Dey, 2010; Dhekaleet al., 2014; Gunathilaka and Tularam, 2016; Hameem et al., 2016).India is the leading tea producercountryof the worldwide (Chaudhry et al., 2017);and it's contributes 1% of GDP of India through primary sector (Gupta and Dey, 2010). The major tea producer states of the country are West Bengal, Assam, Kerala and Tamil Nadu. West Bengal provides tea from Dooars, Terai and Darjeeling, which is 24 per cent contributes to India's total tea production (Hazarika and Muraleedharan, 2011;Dhekale et al., 2014).

Indian tea industry is one of the earliest agro-based organised sector of India. Because it is affords straight employment to more than one million workforces largely coming from undeveloped and economically feebler sections of the society, in which a large numbers are women (Jain, 2011; Dhekale et al., 2014; Mech, 2017).India is played ansignificantcharacter in tea basket of the world (Mishra et al., 2012);and area under tea cultivation is about 5.8 lakh/ hectares, production 13.25 million tonnes and yield 2210 kg/ha in the FY 2017-18. Which increased the cultivable area, production and yield respectively 6.4 lakh/ ha, 13.25 million tonnes and 2210 kg/ha respectively in FY 2021-21; and FY 2021-22 in 13.44 million tonnes.India is the second leading producer of tea in the worldwide with a production of 13.44 million tonnes. India is the leading producer of black tea in the worldwide and also the biggest consumer of tea. There are many producing regions of black, green tea and specialty tea in the country;as Darjeeling in the North, Assam in the North East are the most popular regions.

The Predicted trends and behaviour are investigated by applying Box and Jenkins methods the Autoregressive integrated moving average (ARIMA) model (Box and Jenkins, 1976; Gujarati et al., 2009; Reddy, 2000). The preceding analysis clearly demonstrates that the ARIMA model is a major forecasting tool for linear time series observations. In present study attempts to select the best fit ARIMA model forprediction of tea production and productivity in India. Six performance measures Sigma square, RMSE, MAPE, MAE, AIC and BIC are used to determine forecast accuracy. We formulate in this paper a unit-root test of ARIMA (p, d, q) through float invalid unit-root measurement compared to the trend constant ARMA (p, q) alternating cycles, and compare previous tests or relevant hypotheses and the request to arrange time through previous measurement and character has been accepted. In this paper discusses the future tea production and yield (TPY) of extending from year 2022-23 to 2040-41. In this paper contributes to the existing studies in threefold. First, A very few studies has been done on forecasting the tea production of India. Therefore we consider the teaproduction and yield of India. Second, the historical data series has been taken from 1970-71 to 2021-22 and the datasets are obtained from the website ofReserve Bank of India - Database (rbi.org.in).Third, we apply ARIMA model to forecast the tea (PY) from 2022-23 to 2040-41. With the help of this paper, formers and policy makers will be benefited as they get the idea about the forecast of tea production and yield; and policy decision take accordingly.

1.1. Literature Review

(Islam et al., 2020)Tea production and internal consumption were predicted with Auto-Regressive Integrated Moving Average (ARIMA) and Jarque-Bera criterion was used for the adequacy of the fitted model followed by residual analysis, and it was found that a comparison has been made between the two series and results have emerged that similar trends were visible in the original series and the projected series. (Rahman, 2017) In this study results shows that among the eleven ARIMA models, ARIMA (1, 1, 2) was the most suitable model to predict tea production and estimated value of tea production in Bangladesh.(Hettiarachchi and Banneheka, 2013)The ability to predict was tested by unit price of tea at Colombo auction with generalized least square and artificial natural network (ANN) model found that ANN performance of slightly well. (Kumarasinghe and Peirisl, 2018) Their study focused on modelling the annual tea production to be produced in Sri Lanka and forecasting the tea production volume to be produced in the future.(Hazarika, 2010) They used Autoregressive Integrated Moving Average (ARIMA) model for tea production in Assam with the help of Box-Jenkins approach and predicted the future production quantity. (Mishra et al., 2012) Studies tea production and exports of India. Results was shown that the selected ARIMA (1, 1, 4), (1, 1, 3) modelswas fairlyand suitable for the tea production and exports. (Ghosh, 2017)Studies forecasting tea exports in India have shown that the ARIMA model (1, 1, 0) is well suited for forecasting tea exports. (Sankaran, 2014) In their study, Seasonal Auto Regressive Integrated Moving Average (SARIMA) model was used to forecast the daily demand of fresh vegetables in the wholesale market of Mumbai (India). The study results was shows that the developed model can be used to make predictions for future value an mean absolute percentage error (MAPE) of 14 percent. (Verma et al., 2016) The researcher used an ARIMA model to predictionof coriander prices for mandi in Rajasthan. The researcher observed that such studies will help our farmers to choose appropriate crop options. (Induruwage et al., 2016)Researcher was used SECM and VEC models and predict for future prices of several months and tested the accuracy of the forecasts using MSE and MAPE; and both methods confirmed that the seasonally unadjusted models produced more exactness prediction than other obtained models throughout the study period. (Gijo, 2011) The Box-Jenkins Seasonal Auto Regressive Integrated Moving Average Model was used to predict monthly demand for tea; and they found that such studies are very imperative for planning more efficiently of the production activities. (Weerapura, and Abeynayake, 2013) developed a forecasting of tea production using Time Series model in Shri Lanka. With this background, this study was conducted to identify the trend and appropriate time series models to forecast black tea production according to altitude; and found that appropriate models are ARIMA (1, 0, 1), (1, 1, 1) and (1, 1, 0) respectively for forecasting tea production in Shri Lanka. (Upadhyay, 2013) In this study used ARIMA model using time series data for 1996/97 -2011/2012 with the aid of Box and Jenkins method; and Bayesian information criterion (BIC) and mean absolute percentage error (MAPE) were used to obtain suitable models and forecast the export and import of wood based panels in India. (Balogh et al., 2009) In this study, international tourist arrivals in India and Thailand for the years 2007–2010 were forecast with the help of ARIMA model,, and the results of the study revealed that tourism trends will increase in both the countries.(Celik et al., 2017) An attempt was made to forecast peanut production volume between the years 2016 and 2030 with the help of Autoregressive Integrated Moving Average (ARIMA) model using data from 1950-2015 period in Turkey. And in the study, out of the six ARIMA models studied, ARIMA (0, 1, 1) was found to be the appropriate model to forecast peanut production volume for the next 15 years.(Gopinath et al., 2019) Coffee production volumes in India between 2015 and 2025 were forecast using autoregressive integrated moving average (ARIMA) models. For which historical monthly time series data (period 2010 – 2015) were used and in the study, ARIMA (2, 1, 1) was found to be the most suitable model for forecasting coffee production volume.

We documented a detailed literature on predicting the various data series from 1966 to 2017. The academicians carried study on forecasting of coriander prices, demand of tea, vegetables, production, consumptionand exports of tea. A few studies are available of forecasting of tea production in India and no study are available for period from 2022-23 to 2040-41. Therefore, there is a gap and motivated us to undertake study on forecasting the tea production in India.This paper is structured as followed: In Section 1 introduction is given about the tea production and ARIMA model with present's detailed review of literatures. Section 2 provides research methodology (objectives of the study, research design anddata collection, Section 3econometric modelling of the study, Section 4 Result and discussion and Section 5 provides respectively conclusion and references.

2. Research Methodology

2.1.*Objective of the Research Study*

The main objectives of this research study are as following:

1. To test stability of the accumulatedtime series data for the study, i.e., TeaProduction and Yield (TPY).
2. To recognise the rank of difference at which the time series will become stability.
3. To estimate the ARIMA (p, d, q) model and investigation.
4. To predict future values of tea production and yieldwith the estimated ARIMA model.

2.2.*Research Design*

In order to fulfil the above objectives of the study, exploratory research design and stochastic time series model have been adopted. Exploratory research interprets the already available information and emphasizes on the analysis and interpretation of the available secondary data. Sigma square, AIC, BIC, RMSE, MASE and MAE model are being used for selecting the best ARIMA model and predicting the time series.

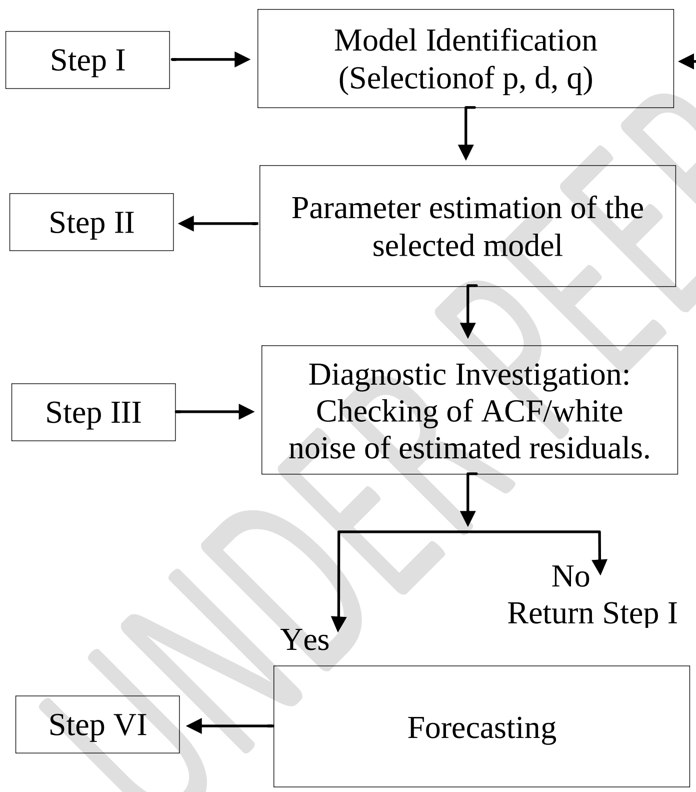
2.3.*Data Collection*

Time series data of tea production & yield (productivity) in India has been collected by the official website Reserve Bank of India - Database (rbi.org.in) for the period of 1970-71 to 2021-22 (52 observations). In this series of data, the researcher has beentried to predict the future values for period of 2022-23 to 2040-41 (19 observations).

2.4.*The Box-Jenkins Methodology (BJ)*

ARIMA is abestmodern methods for prediction of non-stationary time series. In compare to regression model, ARIMA model consents time series to be explain by its stochastic error termsand previous or lag values. The models developed by this approach are usually called ARIMA models because its use a combination of autoregressive (AR), integration (I)and moving average (MA) to produce the forecast (Box and Jenkins, 1976; Rahman A. 2017).The Box-Jenkins method consist four steps as below (Gujarati et al., 2015).

Figure 1 The Box-Jenkins methodology, source: (Gujarati et al., 2015)



3. Econometric Methods for model selection

To choosing the most suitable ARIMA model, several statistical methods are applied, such as Akaike information criterion (AIC), Mean absolute percentage error (MAPE), Schwarz (Bayesian) information criterion (SIC), Mean absolute error (MAE), Root mean square error (RMSE)and Ljung-Box test; thusmodels are formulation as follows:

RMSE is used as a standard criterion to measure model performance for forecasting of the time series. When applied RMSE, the prime assumption is that errors must be unbiased and follow a normal distribution. It provides a complete picture of the error distribution and its value should be relatively low(Draxler, 2014). RMSE id calculated using the following formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}_i\right)^2}{n}} \qquad \text{Eq. (1)}$$

MAPE is a forecasting method or measure of forecasting accuracy. With the help of which we are able to make accurate predictions. It usually expresses the forecasting accuracy of a model as a percentage; whose value is maximum (Reddy, 2000; Tofallis, 2015). Generally we calculate MAPE using the following formula:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{X_i - \overline{X}_i}{X_i}\right| \qquad \text{Eq. (2)}$$

MAE measures the level of variation in the model and expresses how much it differs from the original series and the predicted model (Reddy, 2000). We can calculate MAE using the following formula:

$$MAE = \sum_{i=1}^{n}\left|\frac{X_i - \overline{X}_i}{X_i}\right| \qquad \text{Eq. (3)}$$

AIC is an important criterion for selecting an appropriate ARIMA model. This helps in selecting the appropriate model among a set of different models. With the help of which we are able to select the appropriate ARIMA model, it is based on the probability function (Priya et al., 2015).

AIC = (-2log L + 2m)

Where: m= p + q, L = likelihood function; -2log L = {n (1+ log 2□) + n log $σ^2$}; $σ^2$ = MSE. Finally, we can write the AIC formula as:

AIC= {n (1+ log 2□) + n log $σ^2$ + 2m}Eq. (4)

SIC is also known as a Bayesian Information Criterion (BIC) (Reddy, 2000).The calculation formula of BIC can be written as follows:

$$BIC = \log\left(\frac{rss}{n}\right) + \frac{k}{n}\log n \qquad \text{Eq. (5)}$$

Where: "*rss*" = Residual sum of squares; *k* = the number of coefficients estimated, i.e., 1 + p + q +P + Q; *n* = number of observations

The Ljung-Box test is applied to the residuals after fitting the predictive model to the data. This test tells whether the errors in the forecasting model are correlated with each other. If the p-value (p = 0.05) is less than the significance value then it indicates serial correlation in the selected model (Box and Jenkins, 1978; Gujarati et al., 2015; Khandelwal and Mohanty, 2021). We can calculate the Ljung-box (Q-statistics) through the following formula:

$$Q = n(n+2)\sum_{i=1}^{k} \frac{r_i^2}{n-i} \qquad \text{Eq. (6)}$$

3.1. *AR, MA, and ARIMA Modelling of Time Series Data*

In this section, some old and some new ideas are presented of time series (Gujarati et al., 2015).In this study, we are working on data series of tea production and yield of India, and time series graphs are given in figure 2part (a) Original series and part (b) first difference (DTEA_P and DTEA_Y).Recall that TEA_P and TEA_Y data series is non-stationary on the level, but it is stationary in after the first difference. If data series is stationary, so it can be model in different ways.

3.2.*An Autoregressive (AR) Process*

If Yt represents time t, then we can write the model as:

$(Y_t - \varphi) = \alpha_1 (Y_t - \varphi) + \varepsilon_t$ Eq. (7)

Where: φ = the mean of Y; ε_t= an uncorrelated abnormal error term with zero mean and stable variance σ^2 (white noise).

Then if,$Y_t$follow a first order difference autoregressive (AR1) process. So, the value of *Y* at time t depends on the value at a one past time. Thus we can consider the following model for the study:

$(Y_t - \varphi) = \alpha_1 (Y_{t-1} - \varphi) + \alpha_2 (Y_{t-2} - \varphi) + \varepsilon_t$ Eq. (8)

Then if,Y_t follow a second order difference autoregressive (AR 2) process. In which, the value of Y at time t depends on the values at two previous times, Y values are expressed around their mean value φ.Thus we can follow this model for the study:

$(Y_t - \varphi) = \alpha_1 (Y_{t-1} - \varphi) + \alpha_2 (Y_{t-2} - \varphi) + \ldots\ldots + \alpha_p (Y_{t-p} - \varphi) + \varepsilon_t$ Eq. (9)

Where, $Y_t = p^{th}$ - order autoregressive process.

3.3.*A Moving Average (MA) Process*

An MA (q) is a moving average procedure, which is follows the expression. In other word, a moving average process is a linear conjunction of white noise error terms. Assume that we model *Y* as:

$Y_t = \theta + \beta_0 u_t + \beta_1 u_{t-1}$ Eq. (10)

Where: θ = constant, u = white noise stochastic error term.

Thereafter, if*Y* follow a first order moving average (MA 1) process. So we know that *Y* follows the expression i.e. the value of Y at time t depends on the value at the next time (or next one value). In this conditions, the model can be write as:

$Y_t = \theta + \beta_0 u_t + \beta_1 u_{t-1} + \beta_1 u_{t-2}$ Eq. (11)

Formerly, if it is (Y)follow an MA (2) process. So, the model can be write as:

$Y_t = \theta + \beta_0 u_t + \beta_1 u_{t-1} + \beta_1 u_{t-2} + \ldots\ldots.. + \beta_q u_{t-q}$ Eq. (12)

3.4.*An Autoregressive and Moving Average (ARMA) Process*

As we know that a model Y can have the characteristics of both AR and MA, and hence it is ARMA. Thus, if Y follows an ARMA(1, 1) process, we can write it as:

$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_1 u_{t-1} + \varepsilon_t$Eq. (13)

3.5.*An Autoregressive Integrated Moving Average (ARIMA) Process*

ARIMA is a linear regression model, which we use for forecasting time series. And it uses its own intervals for forecasting. If any time series that exhibits trend and does not pass the white noise test can be modelled for forecasting with an ARIMA model. The ARIMA model is mainly expressed by the following 3 terms: p, d, q. (Box and Jenkins, 1978; Gujarati et al., 2015).
Where: p = order of the AR term; d = order of differencing and q = order of the MA term

3.6.*ARIMA Model Identification of TEA*

In general, a non-stationary series is made stable by differentiating it over 'd' bars, and is then assumed to be now integrated of order 'd'; And which is represented by the term i(d). If the original series is stationary then it is d=0. And we call this type of model as ARMA model instead of ARIMA model. The time series data TEA_PY used for the present study is stabilized after first-order difference. Since there is no need to further differentiate the series, d=1 (first difference) is adopted for the ARIMA (p, d, q) model. We then examined the correlogram to obtain appropriate numbers for 'p' (in AR) and 'q' (in MA) in the model (figure 3part (a) and part (b)). Study of the correlogram shows that there is no notable spikes in the ACF and PACF of the time series TEA_P, and TEA_Y, and the residuals of the selected ARIMA models are white noise. Therefore there is no need to consider AR (p) and MA (q) further,and selected models are fulfil all the criterion, like that (at least σ^2, maximum adjusted R-square, comparatively lower least standard error of regression, lowest value of AIC, SIC (BIC) values, comparatively low RMSE, MAPE and MAE).Therefore, this model has been selected as the best predictive model. Which has been used to forecast future values of tea production and yield. Table 2 shows that selected best fitted ARIMA model with parameters and table 3shows that estimation results of several parameters of ARIMA models for TEA_P and TEA_Y. The forecasting equation for the above models are as follows:
The equation for tea production:

$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_1 u_{t-1} + \varepsilon_t$ Eq. (14)

The equation for tea yield:

$Y_t = \theta + \beta_1 u_{t-1} + \varepsilon_t$Eq. (15)

3.7. *Estimation of the ARIMA Model*

Based on the estimate results of ARIMA (1, 1, 1) models for tea production and ARIMA (0, 1, 1) models for tea yield, and time series forecasting models for tea production and yield has been created, and on substituting values in Eq. (14) and Eq. (15), the functional form of the models is as follows: (giventable 3)

Model for TeaProduction-

Y_t= 179.3563–0.9786Y_{t-1}+ 0.8895u_{t-1} +ε_t Eq. (16)

Model for TeaYield-

Y_t= 18.033– 0.1195u_{t-1}+ε_t Eq. (17)

4. Results & Discussion

4.1. *Stationary Test (ADF Test)*

The results of Argument Ducky Fuller (ADF) test at 1storder difference represent in table 1 presents that the calculated t–statistics value of tea production and yield are respectively = -8.453495 and -7.640717 and p-value are respectively = 0.0000 and 0.0000; which are slighter than estimatedpossibility values at 1%, 5% and 10% levels of significance. Consequently, we failed to agree to take the null hypothesis for the unit root. It means that the data series DTEA_P, and DTEA_Y is not comprising the unit root and consequentlydata series are stationary. The graphical presentation offigure 2part (a) represent plot of original data series and part (b)represent plot of stationary data series.Figure 3part (a), andpart (b) represent plot of the Correlogram (ACF and PACF) of the time series data DTEA_P and DTEA_Y for 1st order difference (lags 1 to 24).

Table 1 Unit Root test- ADF test (at 1st Difference)

DV	Augmented Dickey-Fuller test static (ADF)	t-statistic	Prob.*
Tea	Production	-8.453495	**0.0000**
	Yield	-7.640717	**0.0000**

* MacKinnon (1996) one-sided p-values. **Source:** Author’s Calculation using EViews12.

Table 2 Selected Best Fitted Model (Model Selection Criterion)

Variable	ARIMA	σ^2(Volatility)	Adjusted R^2	S.E. of Regression	AIC	BIC	RMSE	MAPE	MAE
Production	(1,1, 1)	98378.34	0.057	326.73	14.503	14.655	760.38	7.6186	625.77
Yield	(0,1,1)	5378.696	-0.027	75.596	11.546	11.659	133.88	6.9272	118.85

Source: Author’s calculation using EViews12

Table 3 Estimation Parameter for Tea (Production and Yield)

Parameters	AR1	MA1	Intercept	Log Likelihood
		Production (1, 1, 1)		
C	-0.9786	0.8895	179.3563	
S.E.	0.1189	0.2127	45.1947	-365.85
p-value	0.0000	0.0001	0.0002	
Sigma square (σ^2) = 98378.34; AIC = 14.503923				
		Yield (0, 1, 1)		
C	-	-0.1195	18.033	
S.E.	-	0.1401	9.0751	-291.42
p-value	-	0.3983	0.0526	
Sigma square (σ^2) = 5378.696; AIC = 11.54601				

Source: Author's calculation using EViews12

Table 4 Result of the Ljung-Box Test

Model	Lag	χ^2 value	p-value
ARIMA (1, 1, 1)	24	20.580	0.547*
ARIMA (0, 1, 1)	24	22.234	0.506*

Source: Author's calculation using EView12; Note: (*) Insignificant at 0.05% level of significant

4.2. Diagnostic Checking

4.2.1. Residual Diagnostic and Autocorrelation (lag 1 to 24)

We use automatic ARIMA forecasting for calculation of models and of the temporal models presented of the previous part, ARIMA (1, 1, 1) model for tea production withAR1 = -0.9786, MA1 = 0.8895, intercept (θ) = 179.3563, σ^2= 98378.34, Adjusted R^2= 0.057, Standard Error of Regression = 326.73, AIC = 14.503, BIC = 14.655, RMSE = 760.38, MAPE = 7.6186, MAE = 625.77, and log likelihood = -365.85.ARIMA (0, 1, 1) model for tea yield with MA1 = -0.1195, intercept (μ) = 18.033, σ^2= 5378.696, Adjusted R^2= -0.027, Standard error of regression = 75.596, AIC = 11.546, BIC = 11.659, RMSE = 133.88, MAPE = 6.9272, MAE = 118.85, and log likelihood = -291.42has recognised as the preeminent models for forecasting (giventable 2 and table 3). However, before predicting the time series with the tentative models, it is essential to perform diagnostic investigation to avoid overfitting the ARIMA model. The diagnostic analysis has divided in the following part:

- Parameters of the ARIMA models, viz., lowest value of Sigma square (σ^2 Volatility) and Standard error of regression, highest values of Adjusted R square and the lowermost value of the AIC, BIC, RMSE, MAPE and MAE criterions is chosen as the best fitted model for TPY (giventable 2), and the selected ARIMA model has been converted into equation form eq. (16) and eq. (17).

- After fitting a suitable ARIMA models, the goodness of fit can be estimated by plotting the ACF of the residuals of the fitted models. If most of the residuals are within the sample autocorrelation coefficient range (±1.96/√N), where: N = number of observations, then the residuals are white noise showing that the model has adequate and suitable.
- The followingfigure 4part (a) and part (b) gives the ACF of residuals for TPY. The null hypothesis for this test is there is no autocorrelation in residual, and p-value is insignificant which indicates that there is no autocorrelation. Hence, it can be summarised that the residuals are not correlated with each other or in other words, so we it can be said that residuals gotfrom the models are independent from each other.Figure 5 and figure 6 represent plot of residual graph of tea production and yield (residual, actual & fitted) respectively.
- The Ljung-Box test result of production and yield, ARIMA (1, 1, 1) and ARIMA (0, 1, 1) respectively are shows insignificant at the 5% level of significance (given table 4).
- Here, the goodness of fit of the ARIMA (1, 1, 1) and ARIMA (0, 1, 1) models can be checked out through ACF of residualsand white noise. Mostly, a uniform Correlogram over, and done withnegligiblethorns is outstandingand unspoiled. So we gout for forecasting the above models (figure 4part (a)andpart (b)).

4.3. Forecasting the series using selected ARIMA model

4.3.1. Forecasting Results of ARIMA models

This research study is based on annual amount of the TEA_P and TEA_Y,and covering the period of 1970-71 to 2040-41 (71 observations); of which 52 observations ranging from 1970-71 to 2021-22 were historical data and 19 observations ranging the period of 2022-23 to 2040-41 is forecasted value of tea production and yield (giventable 5), and figure 7exhibits the forecasting results of ARIMA (1, 1, 1) for tea production, andfigure 8represent graph of the predicting result of ARIMA (0, 1, 1) model for tea yield, which is observed as the best suitable model for predicting the future value of TEA_P and TEA_Y. We have projected that the yearly amount of tea production, and yield achieved in the year 2022-23 from 13500.17 lakh tonnes, and 2119.73 kg/hectare respectively to 16728.34 lakh tonnes, and 2444.34 kg/hectare respectively in the year 2040-41 will reach (table 5). The forecasting data series line of continuous rising throughout the forecasted period of 2022-23 to 2040-41. Figure 7 and figure 8shows that graphical presentations of forecasted data, and actual data of tea Production and Yield. The software is named the forecasted data series as TEA_PRODUCTION_F, and TEA_YIELD_F; and the forecasted annual amount of the series from FY 2022-23 to 2040-41 (19 observations) with ± 2* Standard Errors (giventable 5).

Table 5 Forecast Value (FV) of Tea Production and Yield(with ±2*Standard Error)

Year	**Tea Production (FV)**	Tea Production (Upper Limit)	Tea Production (Lower limit)	Standard Error	**Tea Yield (FV)**	Tea Yield (Upper limit)	Tea Yield (Lower limit)	Standard Error
2022-23	13500.17	19927.65	7072.68	3213.74	2119.73	3467.89	771.57	674.08
2023-24	13678.04	20190.99	7165.09	3256.47	2137.77	3505.19	770.34	683.71
2024-25	13858.85	20476.65	7241.05	3308.90	2155.80	3542.46	769.14	693.33
2025-26	14036.78	20740.01	7333.56	3351.61	2173.83	3579.70	767.97	702.93
2026-27	14217.53	21025.12	7409.94	3403.79	2191.87	3616.91	766.82	712.52
2027-28	14395.53	21288.53	7502.53	3446.49	2209.90	3654.09	765.70	722.09
2028-29	14576.23	21573.11	7579.32	3498.45	2227.93	3691.26	764.61	731.66
2029-30	14754.27	21836.57	7671.96	3541.15	2245.97	3758.39	763.55	741.21
2030-31	14934.90	22120.66	7749.14	3592.88	2264.00	3765.50	762.50	750.75
2031-32	15113.01	22384.18	7841.63	3635.58	2282.04	3802.59	761.48	760.28
2032-33	15293.58	22667.79	7919.37	3687.12	2300.07	3839.66	760.48	769.79
2033-34	15471.75	22931.40	8012.09	3729.82	2318.10	3876.70	759.50	779.30
2034-35	15652.27	23214.56	8089.99	3781.14	2336.14	3913.73	758.55	788.79
2035-36	15830.48	23478.25	8182.72	3823.88	2354.17	3950.74	757.61	798.28
2036-37	16010.96	23760.96	8260.96	3875.00	2372.20	3987.72	756.68	807.76
2037-38	16189.22	24024.75	8353.69	3917.76	2390.24	4024.69	755.79	817.23
2038-39	16369.65	24307.05	8432.26	3968.69	2408.271	4061.64	754.90	826.68
2039-40	16547.95	24570.93	8524.97	4011.48	2426.30	4098.57	754.03	836.13
2040-41	16728.34	24834.82	8603.85	4062.24	2444.34	4135.49	753.18	845.56

Source: Author's calculation using EViews12

Figure 2Represent Graph of Tea Production and Yield

(a) Original Series(b) Stationary Series

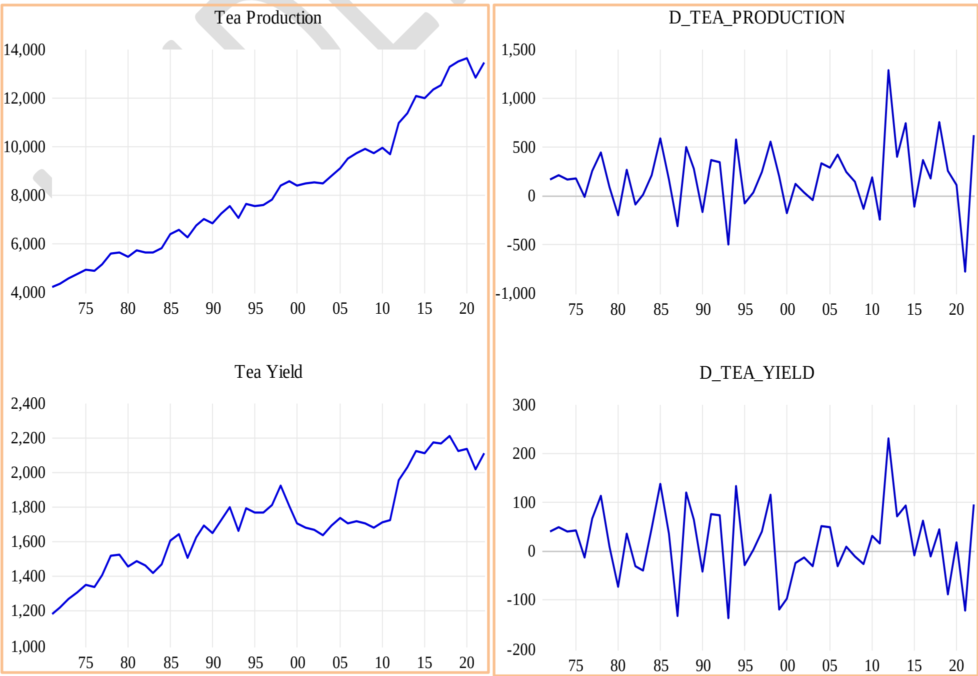


Figure 3Represent plot of Correlogram 1st order difference (leg 1 to 24)

(a) Production(b) Yield

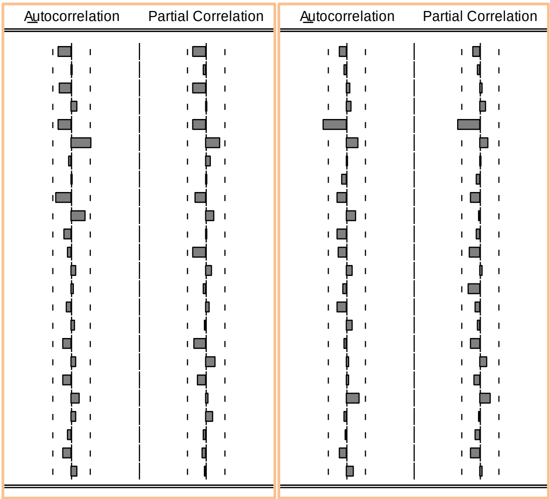


Figure 4Represent plot of Autocorrelation (leg 1 to 24)

(a) Tea Production(b) Tea Yield

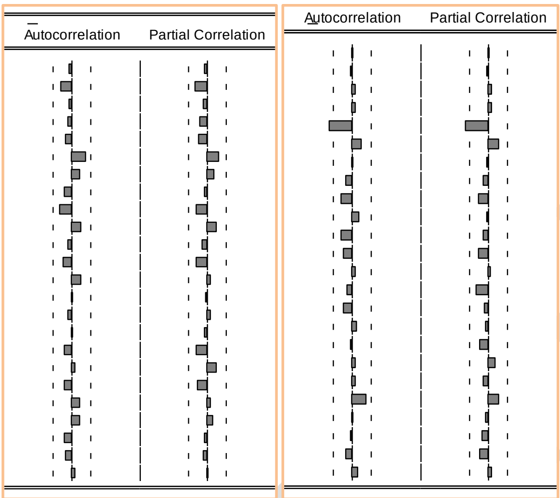


Figure 5Represent plot of Residual graph Tea Production (Residual, Actual & Fitted)

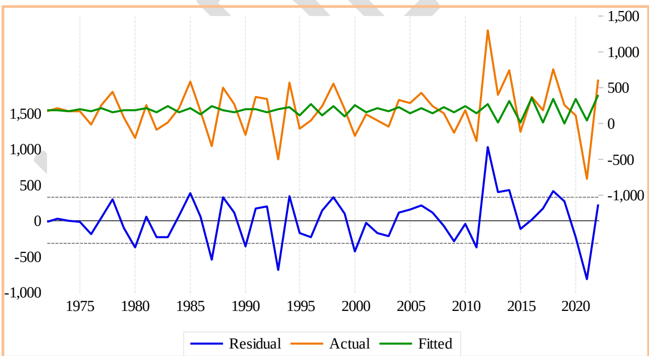


Figure 6Represent plot of Residual graph Tea Yield (Residual, Actual & Fitted)

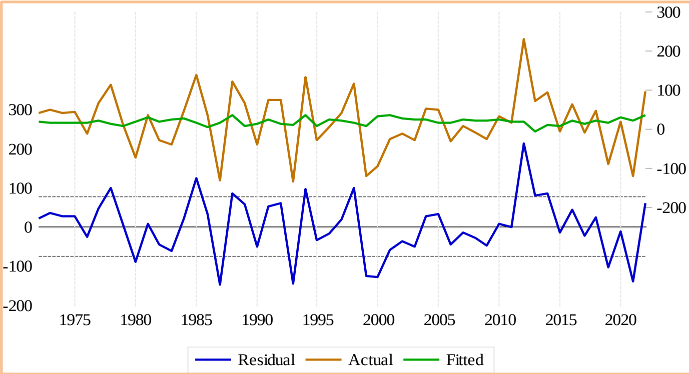


Figure 7RepresentForecast graph of Tea Production

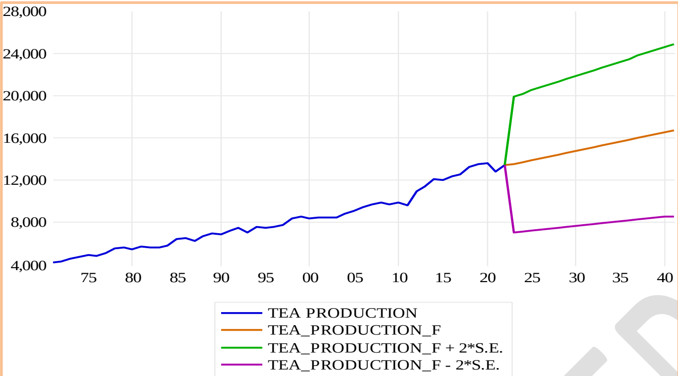
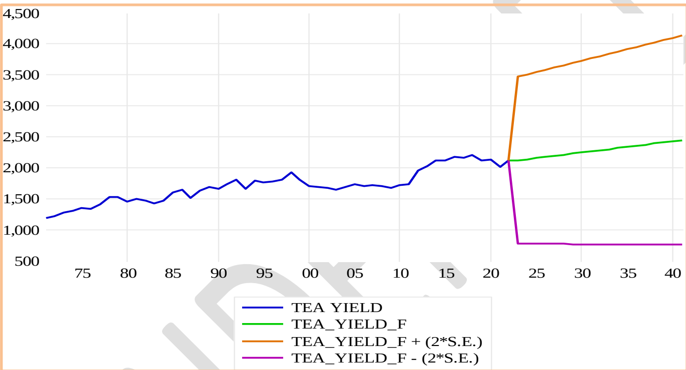


Figure 8Represent Forecast graph of Tea Yield



Conclusion

Tea consumption in Indiarising day-by-day, in cause of the rapidlyrisingof population. The main purpose of this research study has to make a models for tea production and yield of India, with data period of 1970-71 to 2021-22 in India an attempt to prediction tea production and productivity quantity ofamong the FY 2022-23 to 2040-41 with Autoregressive integrated moving average (ARIMA) models. In this study, for tea production and yield (TPY) respectively ARIMA (1, 1, 1) and ARIMA (0, 1, 1), the mostapplicable models has been found among the twenty related ARIMA models; andhence the tea production and yield quantity has been forecast for the next 19 years; we have estimated that the annual volume of tea production and yield achieved in the FY 2022-23 is

13500.17 lakh tonnes and 2119.73 kg/ha respectively, which will accelerate to reach 16728.34 lakh tonnes and 2444.34 kg/ha in the year 2041-41 respectively. Forecasting results of the ARIMA (1, 1, 1) and (0, 1, 1) the volume of tea production and yield shows an increasing trend, and they can help in determining a better policy for increasing tea production in India.

Data Availability Statement

The required secondary data used in this paper has available at Reserve bank of India website: Reserve Bank of India - Database (rbi.org.in).

References

Balogh, P., Kovacs, S., Chaiboonsri, C.C. and Chaitip, P., (2009). Forecasting with X-12-ARIMA: International tourist arrivals to India and Thailand. *Applied Studies in Agribusiness and Commerce, AGRIMBA*. 3.43-63. Google Scholar

Bekhit, M. Y., (2006). Levels of Essential and Non- Essential Metals in Leaves of the Tea Plant (Camellia, sinensis L.) and Soils of Wushwush Farms, Ethiopia, http://etd.aau.edu.et/dspace/bitstream/123456789/307/1/Michael%20Yemane.pdf

Box, G. P. E., & Jenkins, G. M., (1976). *Time Series Analysis: Forecasting and Control*. Holden Day, San Francisco.

Box, G. P. E., & Jenkins, G. M., (1978). *Time Series Analysis: Forecasting and Control*, Revised edition, Holden Day, San Francisco.

Celik, S., Karadas, K. and Eyduran E., (2017). Forecasting the production of groundnut in turkey by using ARIMA model, *The Journal of Animal & Plant Sciences*, 27(3), Page: 920-928. Google Scholar

Chaudhry, Sneha, Negi, Y. S., Shukla, Rakesh K., (2017). A Time Series Analysis of Auction Prices of Indian Tea, *International Journal of Research in Economics and Social Sciences (IJRESS)*, Vol. 7 Issue 6, June, pp. 100-111. Google Scholar

Deka, Sakuntala, Hazarika, P. J. and Patowary, A. N., (2021). Tea production in Assam: Forecasting with ARIMA model, *ASR, Dibrugarh University*, Vol. 33 (1),March. Google Scholar

Dhekale, B. S., Sahu, P. K., Vishwajith, K.P., Mishra, P. and Noman, Md., (2014). Modelling and forecasting of tea production in West Bengal, *Journal of Crop and Weed*, 10(2):94-103. Google Scholar

Draxler, T. C., (2014). Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Argument against avoiding RMSE in the Literature. *Geoscientific Model Development*, 7(3), 1247-1250. Google Scholar

Ghosh, Sudeshna, (2017). Forecasting Exports of Tea from India: Application of ARIMA Model, Journal of Commerce and Trade, October, Vol. XII, No. 2. Google Scholar

Gijo, E. V., (2011). Demand forecasting of tea by seasonal ARIMA model, *International Journal of Business Excellence*, 4(1), 111-124. Google Scholar

Gopinath, M., Sweka S. & Vishalakshi, S., (2019). Forecasting of Coffee Production in India Using ARIMA Model, *International Journal of Science, Engineering and Management (IJSEM)*, Vol 4, Issue 10, October. Google Scholar

Gujarati, D. N., Porter, D. C., &Gunsekar, S., (2009). *Basic Econometrics*, Revised edition, McGraw Hill Education, New Delhi.

Gujarati, D. N., Porter, D. C., & Gunsekar, S., (2015). *Basic Econometrics*, Revised edition, McGraw Hill Education, New Delhi.

Gunathilaka, R. P. Dayani and Tularam, Gurudeo A., (2016). The Tea Industry and a Review of Its Price Modelling in Major Tea Producing Countries, *Journal of Management and Strategy*, Vol. 7, No. 1. DOI: http://dx.doi.org/10.5430/jms.v7n1p21Google Scholar

Gupta, R. and Dey, S. K., (2010). Development of a productivity measurement model for tea industry, *ARPN J. Eng. App. Sci.*, 5:16-25. Google Scholar

Hameem, M. R., Dharmadasa, R. A. I. P. S., Jayawardana, K. M. K., (2016). Time Series Analysis of Colombo Tea Auction Price Data, *Proceedings of the 6th Research Symposium of Uva Wellassa University*, January, 28-29. Abstract: Google Scholar

Hazarika, J., (2010). Development of ARIMA Model for Tea Production of Assam, India: A Case Study, *Int. J. Agricult. Stat. Sci.*, *6*, 11-18. Google Scholar

Hazarika, M. and Muraleedharan M., (2011). Tea in India: An Overview, *Two and a Bud*, 58:3-9. Google Scholar

Hettiarachchi, H. A. C. K. and Banneheka, B.M.S.G., (2013). Time series regression and artificial natural network approaches for forecasting unit price of tea at Colombo auction, *J.Natn.Sci.Foundation Sri Lanka*, 41 (1), 35-40. Google Scholar

Hossain, Md. Moyazzem and Abdulla, Faruq, (2015). Forecasting the tea Production of Bangladesh: Application of ARIMA model, *Jordan Journal of Mathematics and Statistics (JJMS)*, 8(3), pp. 257- 270. Google Scholar

Induruwage, D., Tilakaratne, C. D. and Rajapaksha, S. R. M. S. P., (2016). Forecasting Black Tea Auction Prices by Capturing Common Seasonal Patterns, Sri Lankan Journal of Applied Statistics, Vol (16-3). Google Scholar

Islam, MA, Sumy, MSA, Uddin, MA, and Hossain, MS, (2020). Fitting ARIMA Model and Forecasting for the Tea Production, andInternal Consumption of Tea (Per year) and Export of Tea, *International Journal of Material and Mathematical Sciences*, 2(1), 8–15, DOI: https://doi.org/10.34104/ijmms.020.08015Google Scholar

Jain, S., (2011). Indian tea industry - an analysis using concept of forward engineering, *ABES-IT Group of Institutions*. Google Scholar

Khandelwal, S., Mohanty, D., (2021). Stock Price Prediction Using ARIMA Model, *International Journal of Marketing & Human Resource Research*, Vol. 2, No. 2, April.

Khisa, P. and Iqbal, M., (2001). Tea Manufacturing in Bangladesh: Problems and Prospects, *4th International Conference on Mechanical Engineering*, December 26-28, Dhaka, Bangladesh, 6, 85-91. Google Scholar

Kumarasinghe, H.P.A.S.S. and Peiris1, B.L., (2018). Forecasting Annual Tea Production in Sri Lanka, *Tropical Agricultural Research*, Vol 29 (2): 184 – 193. Google Scholar

Mech, Annesha, (2017). Status of Tea Production in Assam: Past Trends and its Future Projections, *JOLRC*, Volume III, 45-56. Google Scholar

Mishra, P., Sahu, P. K., Bajpai P. and Niranjan, H. K., (2012). Past Trends and Future Prospects in Production, and Export Scenario of Tea in India, *International Review of Business and Finance*, Volume 4, Number 1, pp. 25-33. Google Scholar

Mondal, T. K., Bhattacharya, A., Laxmikumaran, M. and Ahuja, P. S., (2004). Recent advances of tea (Camellia sinensis) biotechnology, *Plant Cell, Tissue and Organ Culture*, 76, 195–254. Google Scholar

Muganda, Consolata A., Stanley, Sewe and Onsongo, Winnie, (2021). Modeling Effects of Climatic Variables on Tea Production in Kenya Using Linear Regression Model with Serially Correlated Errors, Asian Journal of Probability and Statistics, 13(2): 56-75, 2021; DOI: https://doi.org/10.9734/AJPAS/2021/v13i230306Google Scholar

Priya, S R K., Bajpai, P. K. & Suresh, K. K., (2015). Stochastic models for sugarcane yield forecasting. *Indian Journal of Sugarcane Technology,* 30(01): 1-5, June.

Rahman, A., (2017). Modeling of Tea Production in Bangladesh Using Autoregressive Integrated Moving Average (ARIMA) Model, *Journal of Applied & Computational Mathematics*, 6 (2), DOI: https://doi.org/10.4172/2168-9679.1000349Google Scholar

Reddy, C. Viswanatha, (2000). Modelling and Forecasting: NCDEX AGRIDEX, *SCMS Journal of Indian Management*, July-September, 114-130. Google Scholar

Sankaran, S. (2014). Demand forecasting of fresh vegetable product by seasonal ARIMA model, *International Journal of Operational Research*, 20(3), 315-330. Google Scholar

Tofallis, C., (2015). A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation, *Journal of Operations Research Society*, 66(8), 1352-1362. Google Scholar

Upadhyay, V. K., (2013). Modelling and forecasting export and import of Indian wood based panel using ARIMA models. *Elixir International Journal, Statistics*, 63, 18145-18148. Google Scholar

Verma, V. K., Kumar, P., and Singh, H., (2016). Use of ARIMA modelling in forecasting coriander prices for Rajasthan, *International J. Seed Spices*, 6(2), 40-45. Google Scholar

Wang H., Wen Y., Du Y., Yan X., Guo H., Rycroft J., Boon N., Kovas E. and Mela, D., (2010). Effects of Catechin Enriched Green Tea on Body Composition, *Obesity*, 18: 773–79.Google Scholar

Weerapura, W.H.E.B.P. and Abeynayake, N.R., (2013). Forecasting of Tea Production Using Time Series Models, *Proceedings of 12th Agricultural Research Symposium*, 351-355. Google Scholar