Status of Tea Production in Assam: Past Trends and its Future Projections

Annesha Mech*

Abstract

Assam tea industry is India's largest tea industry producing about 6,14,570 tonnes of tea per year. Assam alone produced 652.95 million kg of tea in 2015 and contributed about 52 per cent of total tea production in India. The present study endeavours to analyse the growth and trend of tea production in Assam. The growth performance of tea in Assam was found to be highly uneven. The production (3.76 per cent) as well as yield (2.62 per cent) was recorded highest during the period 1970-78. The growth in area was highest during the period 1997 to 2005 which was mainly due to the increase in the number of small tea gardens. ARIMA model has been applied to assess the temporal variation in production of tea in Assam. It has been used to forecast the production scenario of tea sector for the next ten years considering the past data pattern from 1970-71 to 2014-15. It was found that there will be increase in area, production and yield of tea in next ten years which shall be a good sign for the beverage market.

Key words: Autoregressive Integrated Moving Average Model (ARIMA), compound growth rate, forecasting, tea production, time series analysis

1.0 Introduction

The tea industry of Assam plays a significant role in the state as well as in the national economy. Assam alone produced 652.95 million kg of tea in 2015 which is about 52 per cent of the total tea production in India. The tea plantation is the largest employer among the organized sector of the state employing about more than 6.86 lakh persons daily. Tea is the most widely used beverage which gains its popularity because of its attractive aroma, taste, cheapness among all beverages. It has many pharmacological effects too, like suppressing tumour cell growth, reducing cardiovascular diseases, anti obesity and decreases risk of atherosclerosis (Wang et al., 2010). But, the tea industry in India is undergoing some severe crises in the form

* Assistant Professor, Department of Economics, Dakha Devi Rasiwasia College, Chabua, Assam.

of decline in yield, increase in pests and diseases, reduction in fertilizer response as the pest are becoming more resistant. And moreover extreme and erratic weather conditions are also contributing to this crisis in the form of longer dry periods; heavy downpours over many days, frequent hail and cyclonic storm. This has adversely affected the tea sector by increase in pest population, change in monthly crop cycles unpredictably, decline in production as the bushes are getting stressed (Dutta, 2014). A similar picturesque of crisis is observed in case of tea sector in Assam too. Though Assam is witnessing a spurt in the growth of small tea gardens during the last three decades, but in terms of share in output their contribution remains marginal. The large tea growers are not only in the grip of price recession but also face problem of glut in the export market as better quality tea is supplied by countries like Sri Lanka, Cuba, Kenya and China. Assam Branch of India Tea Association (2016) reveals that tea prices over the last ten years has not been able to keep pace with inflationary trends as input cost have increased at a Compound Annual Growth Rate (CAGR) of more than 10 per cent per annum while tea prices have grown at a CAGR of 6 per cent to 7 per cent. As 70 per cent of the input costs comprising of wage, fuel, fertilizer etc. are fixed in nature and the selling prices of tea are benchmarked with auction prices the producer is unable to pass the burden of increasing cost of production on the consumer. Assam tea fetched an average price of Rs 164.11 per kg in 2016 which is less than the previous year price of Rs 169.78 per kg (Ghosal,2016). A considerable number of tea gardens of the state have gone sick over the period due to age old gardens, scanty rainfall, increasing trend in the cost of production, general fall in the price of tea, rise in the bed of Brahmaputra, frequent pest attacks, lack of infrastructure, modernization, lukewarm attitude of tea planters to the tea garden labourers and inefficient management (Husain & Hazarika, 2010).

Tea, one of the most important cash crops of Assam contributes a great share to state income. Tea industry has been undergoing major changes with time over the period with change in ownership pattern and improvement of tea cultivation practices. In order to assess the temporal variation in production of tea in Assam an attempt has been made to study the overall performance of tea production and forecast the production scenario considering the past data pattern.

2. 0 Material and Methods

In order to study the overall performance of tea production in Assam and to forecast the production scenario, time series data on area, production and productivity of tea for the period 1970-71 to 2014-15 were collected from Statistical Hand Book published by the Directorate of Economics and Statistics, Government of Assam. To explain the status of tea production of Assam, descriptive measures of central tendency and dispersion like arithmetic mean, standard deviation and coefficient of variation were used along with compound growth rate. Annual compound growth rates of area, production and productivity were calculated by using the following log linear function (Dandekar,1980, Singh et al., 2001) to analyse the past trends.

$$Y_t = A(1+r)^t \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where, Y = the value for which growth rate is to be calculated; t = time in years; r = growth rate.

Taking log on both sides of equation (1)

$$\text{Log } Y_t = \log A + t \log (1+r)$$

$$Y = a+bt$$

Putting Log Y_t = Y, log A = a and log (1+r) = b

Now, the value of r is readily obtained.

Since, $b = \log(1+r)$

Antilog $b = 1+r$

Δ

$$r = (\text{antilog } b) - 1$$

Finally, the compound growth rate is estimated by the following equation:

$$r = [(\text{antilog } b) - 1] \times 100$$

Coefficient of Variation was calculated by using the equation (Sharma,1977; Singh et al., 2001):

$$\Delta Q = A_0\, \Delta Y + Y_0\, \Delta A + \Delta A.\Delta Y \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

where, A_0 and Y_0 are area and yield per hectare in base year and ΔQ, ΔY and ΔA are changes in production, yield per hectare and area respectively between base year and 't' th year. The three terms on the right hand side of the equation when divided by ΔQ provide estimates of the contributions of average yield, area and their interactions to the increase in tea production.

The annual data on tea cultivated area, production and yield for the period from 1970-71 to 2014-15 were used for forecasting their future values using Autoregressive Integrated Moving Average (ARIMA) models. The ARIMA methodology is also known as Box-Jenkins methodology which is concerned with identifying the stochastic process of the time-series and predicting the future values accurately. Unlike regression models, the BJ-type time series models allow Y_t to be explained by past, or lagged, values of Y itself and stochastic error terms (Gujarati

& Sangeetha, 2009). A stochastic process is either stationary or non-stationary. As the most time series are non-stationary in nature so the first step of Box-Jenkins model is to reduce the non-stationary series to stationary series by taking the first order differences as the ARIMA model refer to only stationary time series. ARIMA is an extrapolation method which requires historical time-series data of underlying variable namely area, production and yield of tea (Mishra et al., 2013; Dhekale et al., 2014; Padhan, 2012). The procedure for ARIMA modeling are expressed as follows:

Autoregressive Process (AR): Let Y_t be a discrete time series variable which takes different values over a period of time. The AR (p) model of Y_t which is the generalization of autoregressive model can be expressed as :

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + \alpha_p Y_{t-p} + \varepsilon_t$$

where, Y_t is the response variable at time t; $Y_{t-1,} Y_{t-2} \ldots\ldots\ldots\ldots Y_{t-p}$ are the respective variables at different time with lags; $\alpha_{0,} \alpha_{1,} \ldots\ldots\ldots\ldots.. \alpha_p$ are the coefficients and u_t is the error term.

Moving Average (MA) model (q) which is again the generalizations of moving average model may be specified as:

$$Y_t = \mu_t + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + \beta_q \varepsilon_{t-q} + \upsilon_t$$

where μ_t is the constant mean of the series; $\beta_1 \ldots\ldots\ldots\ldots\ldots\ldots \beta_q$ are the coefficients of the estimated error term; ε_t is the error term.

Combining both the models to form the ARIMA model which is written as:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + \alpha_p Y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + \beta_q \varepsilon_{t-q} + \upsilon_t$$

If Y_t is stationary at level or I(0) or at first difference I(1) determines the order of integration which is called as ARIMA model. The details of the estimation and forecasting process are discussed below.

3.0 Results and Discussion

3.1 Descriptive statistics

Descriptive statistics are relatively simple and very convenient summaries of distributions of values. Before analyzing the descriptive statistics, a visual inspection of the time series plots of the variables i.e, area under tea, production and productivity are performed. Time series plots are often the first step for exploratory analysis and it is seen that the plots are showing volatile pattern, with ups and down over a period of time; some of their trends may be non-linear and non-normal (Figures 1 to Figures 3). This is also reflected in descriptive statistics which has been shown in Table 1.

Figure 1 Time series plot of area under tea (in hectares)

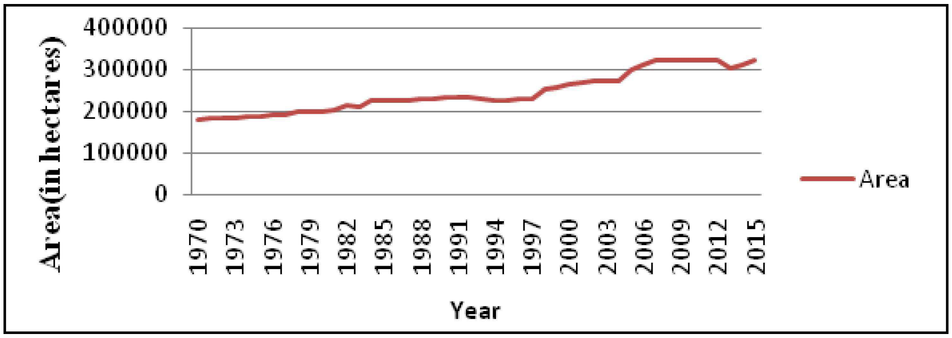


Figure 2 Time series plot of production of tea (in '000kg)

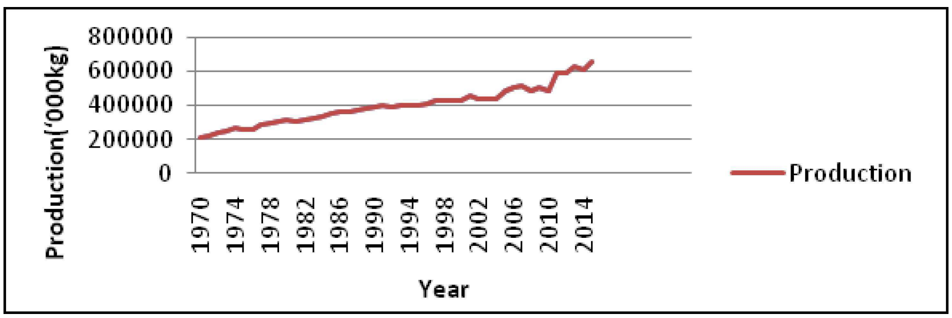


Figure 3 Time series plot of yield of tea (in kg per hectare)

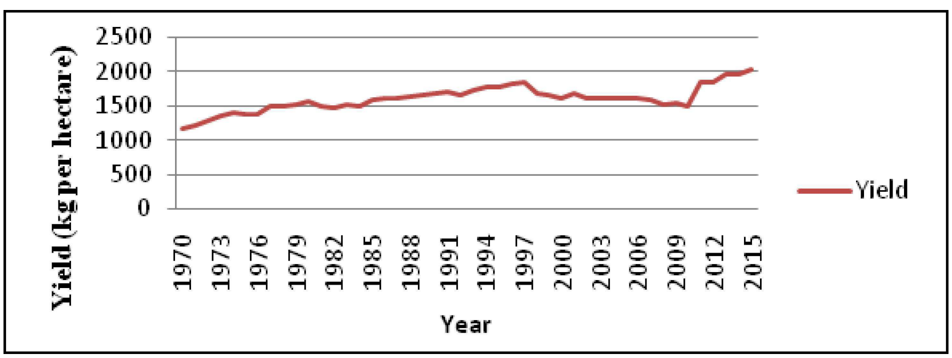


Table 1 Descriptive Statistics for assessing the overall performance of tea in Assam

Descriptive Statistics	Area (in hectares)	Production ('000kg)	Yield (kg per hectare)
Mean	243230.02	383741.22	1587.49
Maximum	322005	629050	1952
Minimum	180065	153587	1178
Standard Deviation	46569.83	105304.67	160.20
Coefficient of Variation	19.15	27.44	10.09
Kurtosis	-0.99	-0.16	0.50
Skewness	0.50	0.16	-0.31
CGR (per cent)	1.18	2.27	1.08

The descriptive statistics in Table 1 demonstrates the average, maximum and minimum values along with other statistical properties. Area under tea in Assam has varied between 180065 hectares to 322005 hectares registering a compound growth rate of 1.18 per cent. The measures of skewness and kurtosis were used to study the different aspects of a frequency distribution. Skewness is positive while kurtosis is negative indicating there has been an increase in the area under tea and it remains steady throughout the study period. The average production of tea during the study period is 383741.22 thousand kilogram with a compound growth rate of 2.27 per cent. Skewness is positive while Kurtosis is negative indicating there has been increase in the production of tea during the early half of the study period and it remains steady for a long duration. Average productivity of tea in Assam is 1587.49 kg per hectare and ranges between 1178 kg per hectare to 1952 kg per hectare The productivity of tea has recorded a 1.08 per cent compound growth rate during the whole period. Positive kurtosis and negative skewness indicates an uneven increase in productivity of tea throughout the period. There has been growth in area under tea due to increase in the number of small tea gardens, but in terms of growth performance in production and yield the picture remains gloomy (Mishra et al., 2014).

3.2 Growth of area, production and yield of tea in Assam: Past trend analysis

The growth performance of tea in Assam has been highly uneven. In terms of trend in growth of production of tea it was highest during the period 1970 to 1978. The major contribution to growth in production was rise in productivity. The highest increase in production (3.76 per cent) as well as yield (2.62 per cent) was recorded during the period 1970-78. The growth in area was highest during the period 1997 to 2005 which was probably due to the increase in the number of small tea plantations (Mishra et al., 2008). Though growth rate in area in this period has been recorded highest but production was low and the yield has shown a negative growth (-1.42 per cent). One important reason for declining productivity growth of tea is the inadequate replanting of bushes (Bhowmik, 2002).

Table 2 Growth of Area, Production and Yield of tea in Assam

Period	Area	Production	Yield
1970-1978	1.10	3.76	2.62
1979-1987	1.52	2.08	0.64
1988-1996	0.06	1.26	1.22
1997-2005	3.02	1.57	-1.42
2006-2015	0.34	2.66	2.31

3.3 Forecasting of area, production and yield of tea

Autoregressive Integrated Moving Average (ARIMA) model is used for future projection regarding area, production and yield of tea for the next ten years. The annual data on tea cultivated area, production and yield for the period from 1970-71 to 2014-15 were used for forecasting their future values. ARIMA model is generally applied for stationary time series data. Augmented Dickey Fuller (ADF) Test is applied for checking the time-series properties of stationary and non-stationary. The null hypothesis is that the variable contains a unit root and the alternative is that the variable was generated by a stationary process. The ADF Tests results are estimated with level

Table 3 Augmented Dickey Fuller unit root tests

Variables	**Level**		**Difference**	
	Intercept only	**Trend and intercept**	**Intercept only**	**Trend and intercept**
Area	-0.054(0) [0.9539]	-1.883(0) [0.6632)]	-5.557(0) [0.000]*	-5.514(0) [0.000]*
Production	-1.533(0) [0.5173]	-4.579(0) [0.001]*	-11.008(0) [0.0001]*	-10.921(0) [0.0001]*
Yield	-1.486(0) [0.5404]	-2.180(0) [0.5009]	-7.625(0) [0.0001]*	-7.527(0) [0.0001]*

** indicates 1 per cent significance level.*

Figures within (...) denotes lag length which is employed for estimating regression model of ADF tests.

Figures within [...] denotes probability values of the test statistics.

and first difference both including and excluding a trend term along with constant values in regression model (Pradhan, 2012). The estimated results are depicted in Table 3.

The results depict that the variables used in the study i.e, area under tea, production and productivity of tea were found to be non-stationary at level in most of the cases, both including and excluding a trend term along with constant or intercept values. This implies that the variables have a time-varying mean or a time-varying variance or both. As time series is nonstationary, its behaviour can be studied only for the time period under consideration and so it is not possible to generalize it to the other time periods. For the purpose of forecasting, nonstationary time series are of little practical value (Gujarati & Sangeetha, 2007). So, the non-stationary series were converted to stationary series through first difference. This is represented at various significance level.

3.4 Box-Jenkins ARIMA forecasting model

ARIMA forecasting model is applied for stationary data and involves four steps which are discussed below along with the estimated results:

Identification: The first step of applying Box-Jenkins forecasting model is to identify the appropriate order of ARIMA (p,d,q) model. Identification of ARIMA model implies selection of order of AR(p), MA(q) and I(d). The order of d is estimated through I(1) or I(0) process of unit root stationary tests. The chief tools used in model specification and selection of order p and q involve plotting of autocorrelation function(ACF) and partial autocorrelation function (PACF) against the lag length (Gujarati and Sangeetha, 2009). The autocorrelation function specify the order of moving average process, q and partial autocorrelations select autoregressive of order, p. The ACF shows autocorrelation coefficients at different lag length with 95 per cent confidence interval whether they are statistically significantly different from zero or not (Padhan, 2012). The significance level of individual coefficients is measured by Box-Pierce Q statistics and for all the coefficients jointly together by Ljung-Box statistics. The results of ACF and PACF are reported in Table 4.

Estimation of the model: After identifying the appropriate p and q values, the next stage is estimation of the parameters of the autoregressive and moving average terms included in the model. The regression model is estimated with simple ordinary least square methods. After estimation of the model, significance of each coefficient are tested by applying 't' test and jointly together by 'F' test. The adjusted R^2 provides whether the model is a good model or not.

Diagnostic checking: The next step is to diagnosis whether the chosen ARIMA model fits the data which is done by checking the residual term obtained from the ARIMA model by applying the same ACF and PACF functions. Adjusted R^2, Bayesian Information Criterion (BIC) values and Mean Absolute Percentage Error (MAPE) values are used to obtain the optimal ARIMA model.

Table 4 Best fitted ARIMA models for area, production and yield of tea in Assam

Variables	ARIMA (p, d, q) Model	Adj R^2	MAPE	BIC Values	Ljung-Box Q Test
Area (in hectares)	0,1,0	0.93	1.702	17.75	17.88(.46)
Production('000kg)	1,1,1	0.77	7.992	21.89	8.39(.93)
Yield (kg per hectare)	1,1,1	0.84	2.91	8.84	11.74(.76)

The adjusted R^2 is used to check whether the model is a good model or not. It is used to indicate how well the independent factor explains the variations in the dependent variable. As the values of adjusted R^2 in all the models are higher than 0.75 which reflects the fact that the models are good fit. Mean Absolute Percentage Error (MAPE) is the most commonly used statistics in all types of forecasting. It measures the average relative size of the absolute forecasting error as a percentage of the corresponding demand value, irrespective of whether the forecasting error is positive or negative (Lewis, 2012). As the value of MAPE in all the models are less than 10 per cent which depicts that the models are very good fit and has the potential for good forecasting. Another criterion that is used for model selection is the Bayesian Information Criterion (BIC). If the BIC value for a variable equals or falls below zero, the data provide little support for including the variable in the model. BIC difference of 0-2 is defined as weak; 2-6 as positive; 6 to 10 as strong and greater than 10 as very strong (Pampel, 2000). As the BIC values are greater than 10 in case of two models for area and production, so it depicts that the model are very strong. Ljung-Box test was run to provide statistical evidence of a good fit.As the p- value in the brackets are significantly larger than 0.05 and as such we can state that there are strong evidence that the models are good fit to the residuals. Hence, the ARIMA models are good fit as expected.

Forecasting: After completion of the three steps of forecasting, the final step is to obtain the forecasted values by estimating the appropriate model devoid of problems. The forecasted values obtained from ARIMA model are reported in table 4. for next 10 years i.e, till 2025. Forecasting figures indicates that there will be increase in area , production and yield of tea in Assam in the next ten years.

Table 5 Forecasting of Area, Production and Yield of Tea in Assam

Year	Area (in hectares)	Production (in '000kg)	Yield (kg per hectare)
2016	325620	633235	1945
2017	328784	650840	1980
2018	331949	668444	2014
2019	335113	686048	2047
2020	338277	703653	2080
2021	341441	721257	2112
2022	344606	738861	2144
2023	347770	756466	2175
2024	350934	774070	2205
2025	354098	791674	2236

Figure 4 Observed and forecasted values for area under tea in Assam (1970-71 to 2024-25)

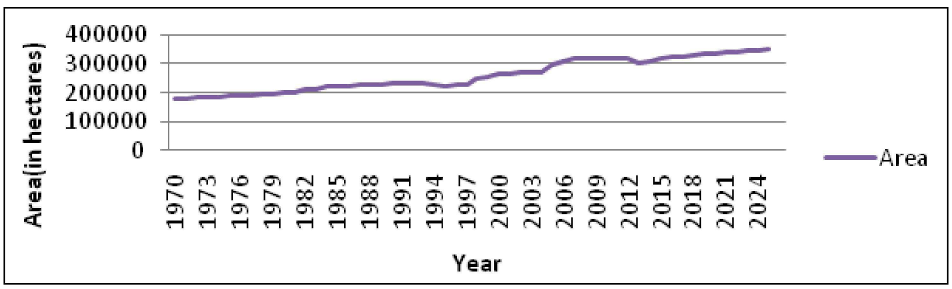


Figure 5 Observed and forecasted values for production of tea in Assam (1970-71 to 2024-25)

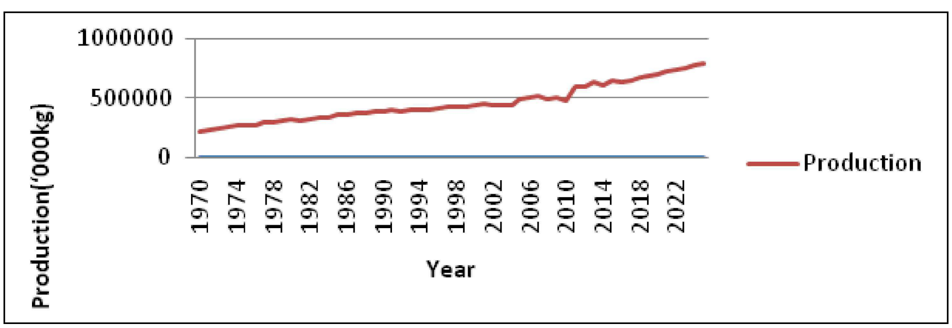
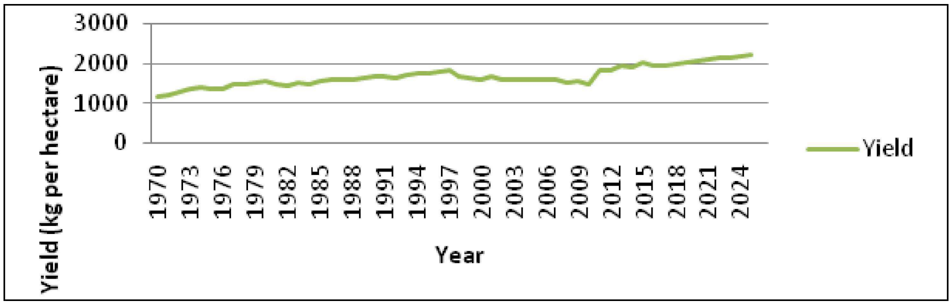


Figure 6 Observed and forecasted values for yield of tea in Assam (1970-71 to 2024-25)



4.0 Conclusion

ARIMA model has been applied to assess the temporal variation in production of tea in Assam. It has been used to forecast the production scenario of tea sector for the next ten years considering the past data pattern. It was found that there will be increase in area , production and yield of tea in Assam in the next ten years which is a good sign for the beverage market where the demand for tea is increasing day by day with the increase in population consuming tea. This will also increase revenue earnings of the state and shall also contribute to the national exchequer in the form of foreign exchange earnings through its exports. As Assam is the largest tea producing state of India, government need to take some measures for developing the tea sector of Assam by organizing programmes to educate the small tea growers regarding how to increase the productivity per hectare, how much nutrients and fertilizers are to be applied per hectare so that the fertility of the soil remains intact. Moreover the government should take special care so that the tea growers receive a stable price throughout the year for which initiatives like determination of Minimum Support Price for tea, development of a cooperative market for tea at the local, district and state level.

References

Bhowmik, S. K. (2002). Productivity and labour standards in tea plantation sector in India. *Labour and social issues in plantations in South Asia: role of social dialogue*, 133-166.

Dandekar, V.M. (1980). Introduction to seminar on data base and methodology for the study of growth rates in agriculture. *Indian Journal of Agricultural Economics, 35*(2), 1-12.

Dhekale, B.S., Sahu, P.K., Vishwajith, K.P, Mishra, P & Noman, M. (2014). Modeling and forecasting of tea production in West Bengal. *Journal of Crop and Weed, 10*(2), 94-103.

Dutta, R. (2014). Climate change and its impact on tea in Northeast India. *Journal of Water and Climate Change, 5*(4), 625-632.

Gujarati, D.N. & Sangeetha. (2007). *Basic Econometrics*. New Delhi: Tata McGraw Hill Education Private Limited.

Ghosal, S. (2016). *Assam Tea Industry under pressure due to rising costs, price stagnation*. Retrieved from https://economictimes.indiatimes.com/news/economy/agriculture/assam-tea-industry-under-pressure-due-to-rising-costs-price-stagnation/articleshow/54045561.cms

Husain, M. & Hazarika, S.D. (2010). Assam Tea Industry and its Crisis. *Dialogue*, *12*(1).

Lewis, C.D. (2012). *Demand Forecasting and Inventory Control.* (n.p.): Woodhead Publishing Limited.

Mishra, D.K., Upadhyay, V. and Sarma, A. (2008). Crisis in the Tea Sector A Study of Assam Tea Gardens. *The Indian Economic Journal, 56*(3), 39-56.

Mishra, P., Vishwajith, K.P, Dhekale, B.S. & Sahu, P.K. (2013). Instability and forecasting using ARIMA model in area, production and productivity of onion in India. *Journal of Crop and Weed, 9*(2), 96-101.

Mishra, D.K., Upadhyay, V. and Sarma, A. (2014). *Unfolding Crisis in Assam's Tea Plantations: Employment and Occupational Mobility.* New Delhi: Routledge.

Padhan, P.C. (2012). Application of ARIMA Model for Forecasting Agricultural Productivity in India, *Journal of Agriculture and Social Sciences, 8*, 50-56.

Pampel, F.C. (2000). *Logistic Regression A Primer*. New Delhi: Sage Publications.

Sharma, K.L. (1977). Measurement of the area, Yield and prices in the increase value of crop output in India. *Agricultural Situation in India*, *32*, 349-351.

Singh, S.B., Sarma, K.B., Goswami, N.S., Dutta, K.K. & Singh, B.K. (2001). Production and Productivity Analysis of Rice in North- East India. *Indian Journal of Hill Farming, 14*(1), 39-44.

Wang, H., Wen, Y., Du, Y., Yan, X., Guo, H., Rycroft, J., Boon, N., Kovas, E., & Mela, D. J. (2010). Effects of catechin enriched green tea on body composition. *Obesity*, *18*(4), 773-779.