*Journal of Social Sciences and Humanities Review (JSSHR)*
Vol. 10, No. x (x-x)
ISSN: 2279-3933

**Original Article**

# Modelling and Forecasting of National Tea Production in Sri Lanka

**P.W.S. Fernando[1], Rekha Nianthi[2], Shyamantha Subasinghe[3]**

[1]Department of Geography, University of Ruhuna, Sri Lanka
[2,3]Department of Geography, University of Peradeniya

[1]pwsurandi@geo.ruh.ac.lk

[1]https://orcid.org/0000-0003-4776-0595

## Abstract

The tea plantation is of paramount significance in the Sri Lankan economy, contributing to the Gross Domestic Product (GDP) and foreign exchange earnings. Accurate national tea production forecasting is essential for policy formulation, strategic planning, and the sustainable development of the tea industry. Therefore, this study attempts to develop a robust statistical model to accurately forecast national tea production in Sri Lanka. Historical annual national tea production (in million kilograms) data from 1972 to 2022 were collected through the Tea Board of Sri Lanka. To identify the trend and pattern in tea production, the Autoregressive Integrated Moving Average (ARIMA) time series forecasting model was adopted. The model was developed using the Box-Jenkins methodology, which involves model identification, parameter estimation, and residual checking. Results revealed that the ARIMA (0,2,2) was selected as the best-fitted model for predicting national tea production in Sri Lanka. ARIMA (0,2,2) model forecasts future tea production for the next five-year period (2023-2027). Furthermore, it indicates that national tea production will gradually decline from 262.51 Mt and 231.11 Mt in the next five years (2023-2027). The model indicates a statistically significant decline in the national tea production in Sri Lanka. Therefore, it is essential to have long-term structural planning and management strategies for the sustainable tea plantation in Sri Lanka.

**Keywords:** ARIMA, Forecasting, Sri Lanka, Tea Production, Time Series Analysis

ISSN: 2279-3933

**Original Article**

## INTRODUCTION

Tea plantations are the commodity crop in Sri Lanka that generated $1.3 billion in annual exports in 2021 (Central Bank of Sri Lanka, 2023). Sri Lankan Tea plantation has a notable history beyond the British colonial era, and large-scale tea plantations were introduced in the 19th century, replacing coffee plantations devastated by the *Hemileia vastatrix* disease (Fernando et al., 2023a). The Sri Lankan tea industry occupies a prominent role in the global market, ranking as the fourth-largest tea producer in China, India, and Kenya (Export Development Board,2022). Sri Lankan tea plantations provide direct and indirect employment in Tea Growing Regions (TGR). It contributes 2% of Gross Domestic Product (GDP) and generates approximately 1 million direct or indirect occupational opportunities for communities. Tea plantations contribute 15% of total export revenue and 58% of the total agricultural export earnings (Hilal & Ismail, 2019).

In Sri Lanka, tea plantations were divided into three main Tea Growing Regions (TGR): High-Growing Regions (HGR) (1200 m upwards), Mid Growing Regions (MGR) (600 m-1200 m), and Low Growing Regions (LGR) (below 600 m) (Dharmadasa et al., 2018). Geographically, it covers 14 administrative districts in Sri Lanka (Tea Board, 2014). According to the operational process of the tea plantation, it has been divided into the smallholding sector and the Estate plantation sector governed by the regional plantation companies (Tea Board, 2014). The smallholding tea sector is a prominent sector that contributes 70% of tea production to the national supply chain in Sri Lanka (International Labour Organization, 2018). Sri Lanka manufactures several types of tea, such as black, green, instant, and white. Black tea generates 15% of Sri Lanka's net foreign income, ranking it the country's top source of foreign exchange (Esham and Garforth, 2013). However, tea production is influenced by different factors: climate, soil, land use, land cover, labour efficiency and shortage, global market and demand, socioeconomic dynamics, policy and government regulations.

Tea production estimation is significant for any TGR to acquire production planning, policy formulation, market management, pricing strategies, and financial planning. Different types of traditional and modern methods can utilise the forecasting of tea production in any TGR worldwide. However, time series analysis is a statistical method that researchers employ to examine and model data that have been collected over time. Time series analysis assists in the two main outcomes: understanding the significant past patterns and trends, as well as forecasting future values. Auto-Regressive Integrated Moving Average (ARIMA) is a key time series statistical forecasting method used to achieve these objectives (Montgomery et al., 2015). The ARIMA model is a suitable method for identifying the pattern of time series tea production data and forecasting production for the next few years (Islam et al., 2020).

Time series analysis involves analysing and modelling sequential data points recorded over time (Box, 2013). It is widely applied in several research areas, including economics, finance, environmental science, and engineering, to understand underlying patterns and make future predictions (Hyndman & Athanasopoulos, 2018). A time series consists of observations measured at consistent time intervals, and its analysis involves identifying trends, seasonal variations, cyclic behaviors, and irregular fluctuations (Wilson, 2016).

The occurrence of stationarity, or the constancy of statistical characteristics like mean and variance over time, is an essential assumption in time series forecasting. However, many time series in real existence behave non-stationarily because of outside factors, which require changes like differencing or decomposition methods (Wei, 2006). Additionally, time series models must account for potential autocorrelation, where past values influence future observations, making proper model selection and validation essential for accurate forecasting (Shumway & Stoffer, 2017). Forecasting using time series methods is crucial for decision-making in both academic research and practical applications. More advanced methods, including machine learning techniques and deep learning algorithms, have gained popularity due to their ability to capture complex, nonlinear relationships within time series data (Hewamalage et al., 2021).

In Sri Lanka, Tea production plays a significant role in the socioeconomic development, and it has become a leading agricultural export commodity and a key source of foreign exchange earnings. The tea plantation sector provides employment and livelihoods for rural populations, including the estate workers and smallholding tea farmers in TGR in Sri Lanka. Climate variability, loss of the tea land suitability and market fluctuations caused a decrease in the tea yield and production (Fernando et al., 2023b). These conditions create the key challenges for national tea production planning, export management, and policy formulation as of the renowned tea producer in the global market. Therefore, it is required for reliable tea production forecasting for a reliable decision-making process and sustainable management. However, several studies have examined the tea production trends over the last decades in Sri Lanka, but limited studies have incorporated the advanced time-series modelling techniques for accurate forecasting in recent years. Therefore, this study adopted the Autoregressive Integrated Moving Average (ARIMA) model to examine tea production modelling and forecasts, thereby providing significant recommendations to stakeholders such as policymakers, plantation sector management, and stakeholders for strategic planning and long-term sustainability of the Sri Lankan tea sector.

## LITERATURE REVIEW

ARIMA is a time series data forecasting method. This method can forecast temporal variations in tea production and trends. Main tea-producing countries like India, Kenya, Sri Lanka, Vietnam, and other countries have used this method to forecast trends (Mahanta, 2023 ;Borah & Amrin, 2022;Niranjan et al., 2022; Deka et al., 2022; Mech, 2017) . The Indian tea production forecast has been predicted based on the ARIMA model. Bangladesh is a tea-producing country representing the South Asian region. ARIMA model was used to conduct tea production forecasting in tea-growing areas in Bangladesh (Mila et al., 2022; Islam et al., 2020; Hossain & Abdulla, 2015). However, this forecasting method is vital for the future assessment and formulation of policies related to tea production in tea-growing countries or areas.

Sri Lanka is also vital in tea production, exporting, and consumption. However, temporal variations in tea production have created several severe conditions in TGR in Sri Lanka. However, future tea forecasting can be made by incorporating the central bank reports' annual statistics on tea production. As the country's main tea-exporting and agricultural-based economy, it is essential to identify future trends or destinations for tea plantations.

Several scholars have examined and predicted the tea production trends in the Main Tea Growing Region (MTGR) in Sri Lanka. Abeynayake and Weerapura (2013) conducted a study to forecast tea production using the ARIMA between 1963 and 2011 in the MTGR in Sri Lanka. According to the study, the HGR and MGR tea production showed a declining trend (1963-2011), and the LGR (1963-2011) showed an increasing trend in a particular period. Furthermore, they pointed out the declining trend in the whole TGR in Sri Lanka. Abeynayake and Weerapura (2013) stated that there has been a declining trend in tea production from 1963 to 2011. Kumarasinghe and Peiris (2018) predicted that there would be an increasing trend in tea production from 2011 to 2015 and that the trend would increase in 2020. The results of the forecasting process showed that the average annual production of tea from 2011 to 2015 will increase by 4.08% in 2020. Moreover, production increases of 5.15% in LGR, 2.6% in MGR, and 4.5% in HGR areas were predicted for 2020. However, both scholars have used the Box and Jenkins ARIMA model with annual tea production data. After 2020, there were no studies for the projection of future tea production and its trend in Sri Lanka.

**Original Article**

## RESEARCH METHODOLOGY

From 1972 until 2022, the time series annual data on Sri Lanka's tea production have been obtained from the tea board. Box-Jenkins' approach is recognised and widely applied for forecasting stationary or non-stationary time series data. Checking the stationarity or non-stationarity of time series data sets, identifying the ARIMA model, estimating parameters, diagnostic checking, and forecasting are the main steps to build the ARIMA model using the Box-Jenkins method (Mila et al.,2022). According to the study, tea production data indicated non-seasonality. ARIMA formulation can be included in the following way. The ARIMA equation (1) can be represented as follows.

$$Y_t = \delta_0 + \delta_1 Y_{t-1} + \delta_1 Y_{t-2} + \cdots + \delta_{\mathcal{P}} Y_{t-\mathcal{P}} + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots + \varphi_{q\varepsilon_{t-q}} \text{ ....... (1)}$$

$Y_t$ indicates the total tea production in the data set. The equation has three main components. ARIMA (p,d,q): where q stands for moving average lags, d for order of difference, and P for autoregressive process order. This can be indicated as follows (2).

$$Y_t = \delta_0 + \delta_1 Y_{t-1} + \delta_1 Y_{t-2} + \cdots + \delta_{\mathcal{P}} Y_{t-\mathcal{P}} + \varepsilon_t \text{ .... (2)}$$

The above model d mean is the differences in time series data sets, and those data should be converted to a stationary data set. Then, that part can be presented as follows.

The annual tea production data show no evidence of a unit root (d=0): $y_t = Y_t$

First difference (d=1): $y_t = Y_t - Y_{t-1}$

Second difference (d=2): $y_t = (Y_t - Y_{t-1}) - (Y_t - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$

In the above $Y_t$ indicates the original series, and yt indicates the differences in time series data. The Moving Average or MA term (q) integrates the relationship between the random error term and the observation. The following is an indication of that model (3).

$$Y_t = \delta_0 + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots + \varphi_{q\varepsilon_{t-q}} \text{ ........................................ (3)}$$

$Y_t$ = dependent variable at time t
$Y_{t-1}$, $Y_{t-2}, \ldots, Y_{t-p}$ =response variable at time lags
$\delta_0$= constant intercept
$\delta_1, \delta_2, \ldots, \delta_p$ = coefficients of each parameter p
$\varphi_1, \varphi_2, \ldots, \varphi_p$ = coefficients of each parameter q
$\varepsilon_{t-1} + \varepsilon_{t-2} + \cdots, \varepsilon_{t-q}$ = The response Yt incorporates errors from earlier periods

$\varepsilon_t$ = Error term at time t

### Examination of Data Stationarity

The following hypotheses can be used to determine the stationary nature of a data series.

*H0*: The annual tea production data exhibits non-stationary characteristics, such as a stochastic trend.

*H₁*: The annual tea production data exhibit stationarity, characterised by constant mean and variance over time.

The stationary condition (not significant at a 0.05 significance level) was examined using the Augmented Dicky Fuller (ADF) test using stationary vs. non-stationary time series. Stationarity does not persist because all series' p-values are large.

## Identifying the Appropriate Tea Production Model

The Box-Jenkins technique first requires determ ining the accurate values of p, d, and q to develop the model. In order to identify the order of Auto Regressive (AR) and moving average (MA) terms, the study examined at the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) graphs. The three primary components of the non-seasonal ARIMA model can be used to summarise it (Prabhakaran, 2021).

P, d, and q are the main components of the ARIMA (p,d,q) model.

p- The number of autoregressive terms
d- The number of nonseasonal differences
q- The number of moving-average terms.

## Residuals Analysis of the ARIMA Model

Residual analysis or error estimation is required after estimating the time series of the ARIMA model. A mean close to zero and no obvious correlation structure show that the model has successfully captured the underlying pattern of the data, confirming that the model errors are random, according to diagnostic testing. The residuals are equal to the difference between the observations and the corresponding fitted values

## Forecasting of the Future Pattern

After checking, the residuals model can be used for tea production forecasting using the following equation (Hyndman & Athanasopoulos, 2018).

$$\text{MAPE} = \frac{100\%}{n} \sum_{t-1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \quad \text{.......................} (4)$$

*At* = denotes the actual observation
*Ft*= denotes the forecast generated by the mode

## RESULTS AND DISCUSSION

Tea plantations are a major agricultural sector in Sri Lanka that contributes to the socioeconomic well-being of the country and the livelihoods of Sri Lanka (Mohotti and Shaymalie 2024). In Sri Lanka, national tea production has fluctuated in the last five decades due to physical, socioeconomic, and policy factors. Therefore, analysing the temporal variations of national and region-wise tea production, trends, and forecasting are essential for identifying future upliftment in the tea industry.

Descriptive statistics is one of the major categories in statistics and can be used for quantitative data analysis in the study (Mertens, Pugliese and Recker,2017). However, descriptive statistics allow us to recognise distinct observations and their distributions. Table 1 indicates the mean, maximum, first quarters, median, third quarters, and maximum values of tea production from 1972 to 2022 at the national and regional levels.

Table 1. Summary Statistics for National Tea Production 1972- 2022 (Mn Kg)

| Variable | Mean | Maximum | First Quartile | Median | Third Quartile | Maximum |
|---|---|---|---|---|---|---|
| Total Production (Mn Kg) | 260.52 | 178.90 | 211.30 | 258.40 | 305.80 | 340.20 |
| High Growing Region (Mn Kg) | 74.64 | 53.70 | 72.60 | 75.90 | 79.30 | 87.00 |
| Mid Growing Region (Mn Kg) | 53.68 | 37.90 | 49.00 | 52.60 | 56.10 | 74.90 |
| Low Growing Region (Mn Kg) | 131.85 | 53.20 | 74.10 | 138.20 | 183.60 | 210.00 |

*Note*. Author compiled based on the data of the Central Bank Reports, 1972-2022

Table 1shows that the total tea production varied from 260.52 Mn Kg to 340.20 Mn Kg between 1972 and 2022. The mean value of total national tea production was 260.52 Mn Kgs. In addition, MTGRs contributed different proportions of tea production to the national production. Total tea production of the HGR varied by 74.64 Mn Kg to 87 Mn Kg. MGR total tea production also varied between 53.68 Mn Kg and 74.90 Mn Kg within the previous half-century. In addition, LGR contributes to a considerable proportion of national tea production. That amount also varied between 131.85 Mn Kg to 210 Mn Kg within the last 50 years in Sri Lanka. The time series analysis can delineate temporal patterns and trends in national tea production.

*Journal of Social Sciences and Humanities Review (JSSHR)*
Vol. 10, No. x (x-x)
ISSN: 2279-3933

**Original Article**

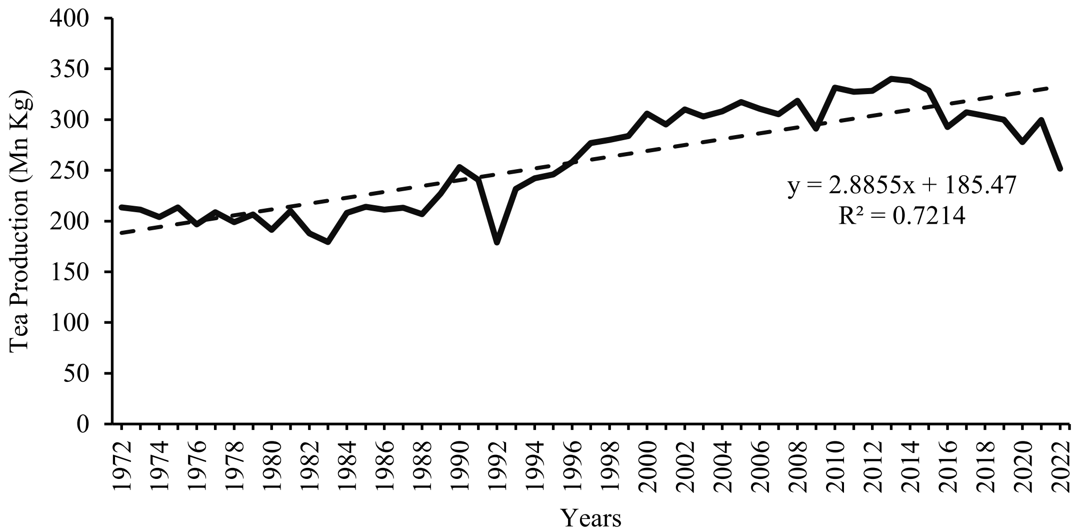


Figure 1: Trend of National Tea Production in Sri Lanka from 1972 to 2022
*Note*. Author compiled based on the data of the Central Bank Reports, 1972-2022

Figure 1 shows the total tea production of Sri Lanka from 1972 to 2022. It illustrates an increasing pattern of fluctuations between periods. According to the figure, there is no seasonal effect in the national tea production data. From 1972 to 2022, total tea production fluctuated for some years. In 1990, tea production reached a record high of 233 Mn Kg, up from an average of 205 Mn kg for the previous decade ( Central Bank of Sri Lanka, 1990). As a result, favourable weather conditions and reduced civil conflicts contributed to the growth of the tea production trend in 1990. Before 1990, tea production had declined owing to unfavourable weather conditions and the escalation of civil war in the country.

After 1993, national tea production in Sri Lanka gradually increased; after 2009, it illustrated a decreased pattern. After 1990, the tea production pattern gradually increased due to the growth of smallholding tea plantations, new planting projects assisted by the Tea Small Holdings Development Authority (TSHDA) under the Asian Development Bank, replantation, and the maximum usage of fertiliser from privately owned tea lands. These subsidy schemes were implemented by the TSHDA in 1990 ( Central Bank of Sri Lanka, 1990). Due to the unfavourable weather conditions, the amendment of fertilizer subsidies and strikes of labourers caused a decline in tea production in 2007.

In 2009, tea production declined 9.1% to 290 Mn Kg from 2008. Dry weather conditions have affected the decline in tea production in 2009 and 2016 in considerable amounts in Sri Lanka. Tea production declined by 7.1% in 2020 due to severe weather

ISSN: 2279-3933

**Original Article**

conditions and manpower supply difficulties caused by COVID-19 containment measures (Central Bank Report 2020). However, in 2022, there was a significant decrease in tea production of 16.0% over the previous year. This decline was caused mainly by acute shortages of agrochemicals and fertilisers throughout the country (Central Bank of Sri Lanka, 2023).

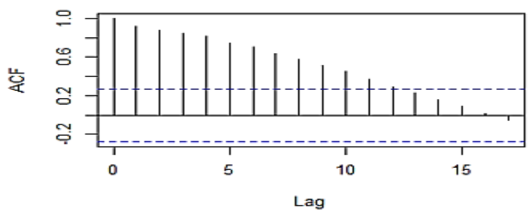


Figure 2. ACF of National Tea Production before the stationary
Note. Author compiled based on the data of the Central Bank Reports, 1972-2022

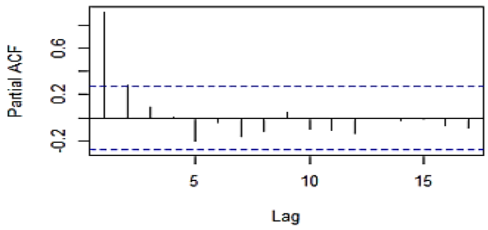


Figure 3. PACF of National Total Tea Production before the stationary
Note. Author compiled based on the data of the Central Bank Reports, 1972-2022

Figure 2 and Figure 3 illustrate the data's slow declining trend of total national tea production, showing the correlation variable with lag. Lags are explained by previous data and its correlations. According to the figure, the projected correlations in the ACF do not decay to zero these data should be transformed into the stationary mood.

The ACF and PACF correlograms were applied to find the potential models for each of the four-time series. A correlogram, also called as an autocorrelation graph, is a graphical representation of serial correlation in time data (Hyndman, 2002a,b). To determine possible ARIMA model ordering, the correlograms (ACF and PACF) for each of the four time series are shown in Figures 2 and 3. Figures 2 and 3 indicate that the data series is nonstationary. Therefore, applying the ARIMA model, those time series of tea production data should be transformed into stationary series.

According to the study, the threshold significance value or P-value exceeds 0.05. (P= >0.05). That means is these data sets are non-stationary. Therefore, time series data should be transformed into a Stationary series. To convert the stationary, "differencing" techniques can be applied to whole data sets. Differencing means that consecutive observations are excluded from one another in this method, which might help to reduce any trend or seasonality in the data. By transforming the data to its natural logarithm, certain non-stationary time series can become stationary (Mila et al., 2022).

The correlation between two observations obtained at different times within a time series is recognised as autocorrelation. Correlations are shown in the present data sets, illustrating that past values influence present values. The correlation between the two data points in a time

*Journal of Social Sciences and Humanities Review (JSSHR)*
Vol. 10, No. x (x-x)
ISSN: 2279-3933

**Original Article**

series is recognised as autocorrelation. In another way, the name is derived from the fact that the time series data are self-correlated. When referring to these associations, we use the word "lags." ACF determines a time series relationship among observations for an identified collection of interruptions (Kim et al., 2024). Significant correlations in time series data sets can be applied to the ACF to find which lags are obtained. Furthermore, ACF can be used to evaluate variability and stationary patterns of time series. PACF indicated that direct value lagged in the time series data set. According to the study, tea production data, ACF and PACF figures (Figure 4 & 5) can be included to identify the time series data, whether stationary or not.

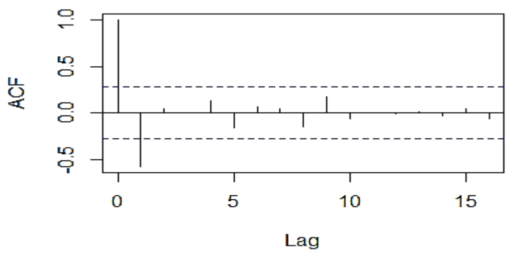


Figure 4: ACF of National Tea Production
Note. Author compiled based on the data of the Central Bank Reports, 1972-2022

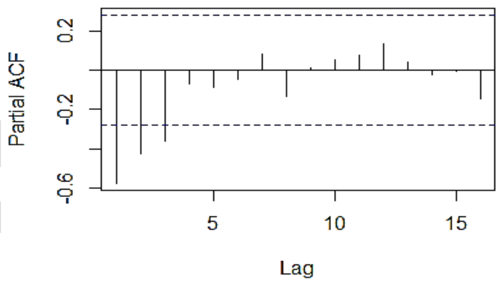


Figure 5: PACF of National Tea Production
Note. Author compiled based on the data of the Central Bank Reports, 1972-2022

| Data series | Resulted P-value for the original data | Resulted from P-value after First Order Differencing | Resulted in P-value after Second Order Differencing |
|---|---|---|---|
| National Tea Production | 0.99 | 0.1021 | 0.01* |

Table 2. Augmented Dicky Fuller (ADF) Test Statistics Results
*Note.* The P values indicated * are significant at 0.05 level

Table 2 indicates the results of the ADF test statistics for national tea production. The ADF test determines whether a lagged value (for example, a value at a previous time point Y [t-1]) and a linear trend can account for the change in Y. If there is a linear trend in the study time series data, it will be deemed non-stationary. But the lagged value is unable to express how Y varies over time. (Yaffee 1999). According to the results of the original tea production data, national wise the P-value is greater than 0.05 (P >0.05). The total national production P- value is 0.99 (P =0.99), and the value exceeds the P- P-value. Results show that P-values were greater than 0.05; the alternative ($H_1$) hypothesis will be rejected. That means data is non-stationary. The

*Journal of Social Sciences and Humanities Review (JSSHR)*
Vol. 10, No. x (x-x)
ISSN: 2279-3933

**Original Article**

nonstationary data set was not applied to the ARIMA model. Therefore, those data sets should "difference" until stationary time series are acquired.

National tea production is still nonstationary; therefore, we attempted a second differencing to those data sets. The second differencing indicated that the calculated P-value is smaller than the tabulated 0.05 P value (0.01) at a 5% level of significance. National tea production data sets become stationary after the second differencing.

According to the study, the next step is the identification of the best-fitted ARIMA model for national tea production in Sri Lanka. A non-seasonal ARIMA model can be summarised by using the three main components. P, d, and q are the main components of the ARIMA (p,d,q) model.
p- The number of autoregressive terms
d- The number of nonseasonal differences
q- The number of moving-average terms

Best fitting models for national tea production in Sri Lanka based on the ACF, PACF correlogram and Akaike Information Criterion (AIC) value, can be included in the following ways. Based on the AIC values, it can be the best ARIMA model for nationally. Diagnostics checking is also required to estimate the best-fitted model. Hirotugy Akaike (Cavanaugh & Neath, 2019) introduced the Akaike Information Criterion. The lowest AIC model was selected as the best model.

Table 3. Selected ARIMA Model for National Total Tea Production

| National Total Tea Production | Model | AIC value |
|---|---|---|
| Model 1 | ARIMA (2,2,2) | 437.15 |
| Model 2 | ARIMA (0,2,0) | 481.07 |
| Model 3 | ARIMA (1,2,0) | 461.28 |
| Model 4 | ARIMA (1,2,2) | 435.14 |
| **Model 5** | **ARIMA (0,2,2)** | **433.34** |
| Model 6 | ARIMA (0,2,3) | 435.15 |
| Model 7 | ARIMA (1,2,1) | 437.09 |

Note. Author compiled based on the data of the Central Bank Reports, 1972-2022

Table 3 shows the estimation of ARIMA (p, d, q) Models for the national tea production in Sri Lanka. There are 7 models selected to estimate the national tea production based on the last 50 years' time series of tea production data in Sri Lanka. Table 3 shows that AIC values are important factors in selecting the best model for estimating national tea production data. AIC values show the prediction error of the assessed models. However, the minimum value of the AIC model represents the goodness statistical model for prediction (Project Pro,2023). Based on the model estimation, minimum AIC values are obtained from model 5 compared with other models. Therefore, ARIMA (0,2,2) was

Vol. 10, No. x (x-x)
ISSN: 2279-3933

**Original Article**

selected as the best-fitted model for predicting national tea production in Sri Lanka.

After estimating the ARIMA model's time series, a residual analysis or error estimation is required (Yaffee, 1999). The mean of the errors or residuals is close to zero, and there is no significant correlation between them. The residuals equal the difference between the observations and the corresponding fitted values. To determine whether the fitted model's residual analysis is sufficient, the variation between actual and forecasted or expected values is referred to as residual terms (Annamalai & Jhonson,2023).

Residuals are significant in checking whether the ARIMA model has enough details in the data. If predicted ARIMA model ACF and PACF values represent a small amount, they can be appropriate for forecasting the series. Residuals are significant in that they lie between substantial boundaries. That means there is no autocorrelation in the forecast values (Shumway & Stoffer, 2017). An acceptable forecasting method will represent residuals with two main properties. The Residuals correlate, and then the information was left in the residuals. Those should be used in computing forecasts.

Residuals have a mean value other than zero, and the forecasts are biased.

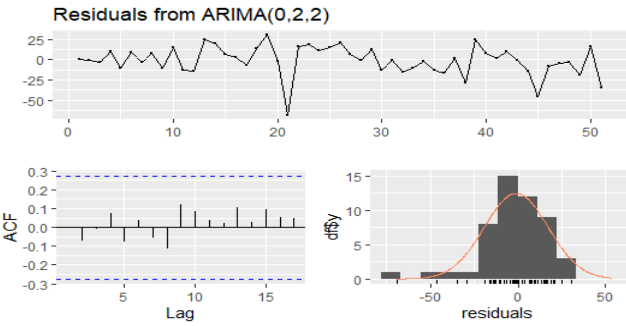


Figure 6. ACF of residuals of fitted model ARIMA (0,2,2) for National Total Tea production

Figure 6 illustrates the Residuals for the ARIMA (0,2,2) model for the national tea production in Sri Lanka. According to the model value of ACF, the white noise series, their error distribution is also normal. Therefore, the ARIMA (0,2,2) model can be accepted for forecasting the national tea production.

The model error should be calculated before forecasting the tea production for the next 05 years. Therefore, different methods can be used to estimate a model's errors. There are different time series error forecasting metrics in statistics, such as scale-depending metrics, percentage error metrics, relative error metrics and scale-free error metrics (Rink, 2021). The Mean Absolute Percentage Error (MAPE) is the percentage error metric that can estimate the residual analysis in a time series. This is a popular method for calculating the error percentage to compare the model accuracy. Mean Absolute Percentage Error (MAPE) for all the forecast series is listed in Table 4

Table 4. Mean Absolute Percentage Error (MAPE) for selected models

**Original Article**

| National Tea Production | MAPE (%) |
|---|---|
| | 5.38 |

MAPE values should be less than 10%, which indicates selected ARIMA models are highly suitable for forecasting tea production in Sri Lanka. According to the results of the model, it has generated a short-term tea production forecast up to 2027. With a value of 0% signifying a precise forecast and increasing values indicating decreased accuracy, the MAPE metric quantifies forecast precision.

## Tea Production Forecasting for 2023-2027 in Sri Lanka

Tea production forecasting is an essential component of the ARIMA model. Based on the previous annual tea production data, future tea projections in Sri Lanka can be predicted as follows. Table 5 illustrates the forecast statistics for national tea production.

Table 5. Total Tea Production Forecasting for 2023 to 2027

| Year | National Total Tea Production (Mt) |
|---|---|
| 2023 | 262.51 |
| 2024 | 254.66 |
| 2025 | 246.81 |
| 2026 | 238.96 |
| 2027 | 231.11 |

*Note.* Author compiled based on the data of the Central Bank Reports, 1972-2022

Table 5 illustrates the forecasting of national tea production (Mt) for the years 2023 to 2027 in Sri Lanka. It indicates that national tea production will gradually decline in the next five years (2023-2027). Tea production will vary between 262.51 Mt and 231.11 Mt during the next five years. That declining trend is shown in Figure 7.

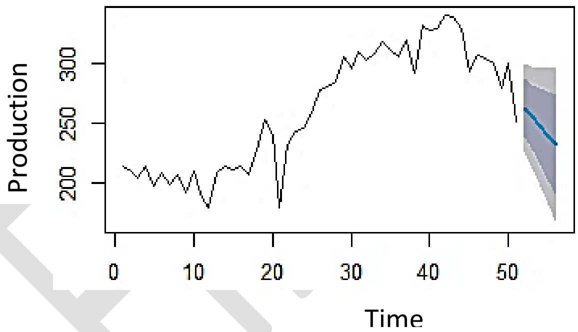


Figure 7. Tea Production Forecasting for 2023 to 2027 in Sri Lanka
*Note.* Author compiled based on the data of the Central Bank Reports, 1972-2022

Figure 7 indicates the predicted values and trend of national tea production in Sri Lanka over the next five years. According to the forecast figures, there was a gradual decline in national tea production in Sri Lanka. Climate change and its vulnerability directly influenced the decline of tea yield and quality in TGR in Sri Lanka (Wijeratne et al. 2007). Furthermore, in 2021 in Sri Lankan government's policy on the ban of chemical fertilizers and pesticides has led to a decline in the tea yield in the TGR of Sri Lanka. World Food and Agriculture Organisation (FAO) emphasised there were direct correlation between the policy and decline of the tea production in Sri

Original Article

Lanka (FAO,2022). Economic crisis of Sri Lanka and lack of skilled labourers are identified as another causative factor by world bank in 2023. Therefore, the projected decline of tea production should be addressed by the government and relevant stakeholders for the sustainability of the tea industry in Sri Lanka.

## CONCLUSION

Tea plantation is a key commodity crop that enhances Sri Lanka's economic profile and provides livelihoods to the rural communities in tea-growing regions. Sri Lankan tea plantations mainly depend on physical, socioeconomic, and policy factors. Physical factors like rainfall, temperature, humidity, soil, and soil pH are significant factors determining tea bush growth. In addition to labour, management practices, policies, and regulations are highly incorporated into the sustainable tea plantations in Sri Lanka. However, tea production forecasting is essential for the policy formulation and sustainable development of the tea industry. Time series data analysis technique is a robust statistical method for the identification of historical trends, patterns and forecasting of the tea production in Sri Lanka. Therefore, the ARIMA model has been adopted as the analytical method with the annual tea production data in Sri Lanka from 1972 to 2022. The linear trend analysis indicated a gradual increasing trend of the annual tea production between the last five decades in Sri Lanka. Moreover, the ARIMA (0,2,2) model has been identified as the best-fitted model for forecasting the tea production for the next five years (2023-2027) in Sri Lanka. The ARIMA (0,2,2) model illustrated that within the next five years, national tea production will decrease from 262.51 MT to 231.11 MT in Sri Lanka. Escalating the fertilizer and other related inputs due to the economic crisis, the organic fertilizer policy was introduced by the government in 2021, labour shortage, old tea plantation and climate change impacts were identified as causative factors affecting the tea plantation and sustainable tea industry in Sri Lanka. Therefore, it is required that the long-term strategic planning for the upliftment of the sustainable tea plantation in Sri Lanka be implemented.

## References

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

Annamalai, N., & Johnson, A. (2023). Analysis and forecasting of area under cultivation of rice in India: Univariate time series approach. ***SN Computer Science***, *4*(2), Article 193. https://doi.org/10.1007/s42979-022-01604-0

Borah, P., & Amrin, S. (2022). Future Projection of Production of Tea In Assam : By Using ARIMA Model. *IJFANS International Journal of Food and Nutritional Sciences, 11(3)*, 1474–1482.

Box, G. E. P. (2013). *Time Series Analysis, Forecasting and Control*. John Wiley & Sons

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics, 11*(3). https://doi.org/10.1002/wics.1460

Central Bank of Sri Lanka. (1991). *Annual report 1990*

Central Bank of Sri Lanka. (2020). *Economic and Social Statistics of Sri Lanka 2020* (Vol.42).

Central Bank of Sri Lanka. (2021). *Annual report 2020*. https://www.cbsl.gov.lk/en/publications/economic-and-financial-reports/annual-reports/annual-report-2020

Central Bank of Sri Lanka. (2023). *Annual report 2022: National output, expenditure, income and employment*. https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/publications/annual_report/2022/en/6_Chapter_02.pdf.

Deka, S., Hazarika, P. J., & Patowary, A. N. (2022). Tea production in Assam: Forecasting with ARIMA model. *Annual Scientific Report Dibrugarh University*, 33(1), 48–56.

Dharmadasa, M., Zubair, L., Nijamdeen, A., & Najimuddin. N. (2018). *Review of Tea Industry in Sri Lanka for Climate Analysis*. Dilmah Conservation. https://www.dilmahconservation.org/pdf/review-of-tea-industry-in-sl.pdf.

Esham, M., & Garforth, C. (2013). Climate change and agricultural adaptation in Sri Lanka: A review. ***Climate and Development,*** **5**(1), 66–76. https://doi.org/10.1080/17565529.2012.762333

Export Development Board. (2022). *Industry capability report: Tea sector (Camellia sinensis)*. Sri Lanka Export Development Board. https://www.srilankabusiness.com/ebooks/industry-capability-report-tea-2021.pdf

Fernando, P. W.S., Nianthi, R. & Subasinghe, S. (2023). Evaluating Tea Land Suitability and Potential Challenges in the High Grown Region (HGR) of Sri Lanka: A Case Study of Kothmale DS Division, Sri Lanka. *IOP Conference Series: Earth and Environmental Science* 1266(1). https://doi.org/10.1088/1755-1315/1266/1/012016

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting, 37(1),* 388–427. https://doi.org/10.1016/j.ijforecast.2020.06.008

Hilal, M., & Ismail, M. (2019). Sri Lanka's tea economy: Issues and strategies. *Journal of Politics and Law, 13*(1), 1–10. https://doi.org/10.5539/jpl.v13n1p1

Hossain, M. M., & Abdulla, F. (2015). Forecasting the tea production of Bangladesh: Application of ARIMA model. *Jordan Journal of Mathematics and Statistics*, 8(3), 257–270.

Hyndman, R. J. (2002a). ARIMA Processes. In H. Daellenbach & R. Flood (Eds.), *The informed student guide to management science (pp. 27–28).* Cengage Learning Business Press.

Hyndman, R. J. (2002b). *Computer-assisted time series analysis* (3rd ed.). Monash University, Department of Econometrics and Business Statistics. https://robjhyndman.com/papers/catsa.pdf

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp3/

International Labour Organization. (2018). *Future of Work for Tea Smallholders in Sri Lanka*. ILO Country Office for Sri Lanka and the Maldives.

Islam, M. A., Sumy, M. S. A., Uddin, M. A., & Hossain, M. S. (2020). Fitting ARIMA model and forecasting for the tea production, and internal consumption of tea (per year) and export of tea.

**Original Article**

*International Journal of Material and Mathematical Sciences*, 2(1), 8–15.

Kim, J., Kim, H., Kim, H., Lee, D., & Yoon, S. (2024). A comprehensive survey of time series forecasting: Architectural diversity and open challenges. *arXiv*. https://doi.org/10.48550/arXiv.2411.05793

Kumarasinghe, H.P.A.S.S., & Peiris. B.L. (2018). Forecasting annual tea production in Sri Lanka. *Tropical Agricultural Research*, 29(2), 184–93.

Mahanta, K.K. (2023). A time series analysis of tea production in South Bank of Assam from 1961 to 2013 Using ARIMA Model. *International Journal of Agricultural and Statistical Sciences* 19(02), 725–32.

Mech, A. (2017). Status of tea production in Assam: Past trends and its future projections. *Journal of Open Learning and Research Communication (JOLRC)*, 3, 45–56.

Mertens, W., Pugliese, A., & Recker, J. (2017). *Quantitative data analysis: A companion for accounting and information systems research*. Springer International Publishing. https://doi.org/10.1007/978-3-319-42700-3

Mila, F. A., Noorunnahar, M., Nahar, A., Acharjee, D. C., Parvin, M. T., & Culas, R. J. (2022). Modelling and Forecasting of Tea Production, Consumption and Export in Bangladesh. *Current Applied Science and Technology* 22(2), 1–20.

Mohotti, K. M. & Shaymalie, H W. (2024). Economic and Social Evolutions of the Tea Industry in Sri Lanka : Lessons for Sustainability in Responsible Business and Way Forward. *Sri Lanka Economic Journal*, 21(Special Issue), 154–170.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015*). Introduction to time series analysis and forecasting (2nd ed.).* John Wiley & Sons.

Niranjan, H. K., Kumari, B., Raghav, Y. S., Mishra, P., Al Khatib, A. M. G., Abotaleb, M., & Supriya. (2022). Modeling and forecasting of tea production in India. *Journal of Animal and Plant Sciences*, 32(6), 1598–1604. https://doi.org/10.36899/JAPS.2022.6.0569

Prabhakaran, S. (2021). ARIMA Model - Complete guide to time series forecasting in Python. *Time Series*. 1–43. https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/.

ProjectPro. (2023). *Solved big data and data science projects*. https://www.projectpro.io

Rink, K. (2021, October 21). *Time series forecast error metrics you should know.* Medium. https://medium.com/data-science/time-series-forecast-error-metrics-you-should-know-cc88b8c67f27 (researchgate.net preview link)

Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (4th ed.). UCLA Statistics & Data Science. http://www.stat.ucla.edu/~frederic/tsa4

Sri Lanka Tea Board. (2014). *Annual report 2014*. https://srilankateaboard.lk/uploads/2021/04/Annual-Report-2014.pdf

Wei, W. W. S. (1990). *Time series analysis: Univariate and multivariate methods*. Addison Wesley. https://doi.org/10.2307/2289741

Wijeratne, M. A., Anandacoomaraswamy, A., Amarathunga, M. K. S. L. D., Ratnasiri, J., Basnayake, B. R. S. B., & Kalra, N. (2007). Assessment of Impact of Climate Change on Productivity of Tea (Camellia Sinensis L.) Plantations in Sri Lanka. *Journal of the National Science Foundation of Sri Lanka*, 35(2), 119–26.

Wilson, G. T. (2016). Time series analysis: Forecasting and control (5th ed.) by G. E. P. Box, G. M. Jenkins, G. C. Reinsel & G. M. Ljung. *Journal of Time Series Analysis*, 37(5), 709–711. https://doi.org/10.1111/jtsa.12194

Yaffee, R. A., & McGee, M. (1999). *An introduction to time series analysis and forecasting: With applications of SAS and SPSS*. Academic Press.