Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

Tea crop Yield Prediction using Time Series Models in Kerala

Sreenivasulu Arigela[1] , M. Bhupathi Naidu[2] , M. Pedda Reddeppa Reddy[3] ,K. Murali[4] , G.Mokesh Rayalu[5]**

[1]Research Scholar, Dept. of Statistics, S.V University, Tirupati
[2]Professor of Statistics, Directorate of Distance Education, S.V University, Tirupati
[3]Associate professor, Dept. of Statistics, S.V Arts College, Tirupati
[4]Academic Consultant, Dept. of Statistics, S.V University, Tirupati
Corresponding Author **
[5]Assistant Professor Grade 2,Department of Mathematics,
School of Advanced Sciences,VIT,Vellore
mokesh.g@gmail.com

ABSTRACT

Kerala's agricultural sector, of which the tea business is a prominent subset, is an important contributor to the state's economy. In this research, we use time series models to forecast the harvest of Kerala's tea plants. We build a solid framework that combines time series modeling approaches like ARIMA and SARIMA by utilizing past production data, weather patterns, and pertinent socio-economic factors. The results of our study should help tea farmers, policymakers, and other stakeholders make better decisions about crop management, resource allocation, and future market strategy. In order to promote long-term growth and resilience in Kerala's tea business, we plan to implement sophisticated time series analysis to boost the accuracy of yield projections for the state's tea crops.

Keywords: Tea Crop, ARIMA, SARIMA, Forecasting.

INTRODUCTION

Kerala, also known as "God's Own Country," is a state in southern India that is famous for its beautiful scenery, rich culture, and prosperous agricultural industry. State residents and the nation as a whole owe a great deal to the agricultural sector of the state's economy. Tea, one of many crops grown in Kerala, has a unique place in the state's history and economy. Kerala's tea estates are well-known for their verdant expanses, from which come leaves of exceptional quality that are used to make some of the best tea in the world.

In Kerala, growing tea is more than a means to an end; it's a way of life that has been passed down from generation to generation. The state's distinctive topography and climate make it perfect for tea plantations, with its rolling hills and plentiful rainfall. Climate change, environmentally friendly agricultural methods, and consumer demand are just a few of the factors that can make or break the tea industry.

Predicting the yield of tea harvests is crucial for efficient resource management, strategic market planning, and long-term growth, just as it is for any other agricultural business. With reliable forecasts, tea farmers and other stakeholders may better plan planting and harvesting seasons and manage resources. It also helps authorities guarantee the long-term health and growth of the industry.

This research dives into the field of tea crop production prediction utilizing cutting-edge time series models to fill this urgent knowledge gap. In order to accurately predict future tea crop yields in Kerala, we want to use historical data, weather patterns, and socio-economic indices. In order to provide accurate forecasts, this framework makes use of a number of time series modeling approaches, including AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA).

This study is significant because it may increase the profitability of tea farming, which would have a ripple effect throughout Kerala's agricultural sector. In order to foster resilience, sustainability, and expansion in the tea business, it is essential that farmers and other stakeholders have access to relevant information. This research was conducted with the intention of strengthening the future of Kerala's tea industry while also preserving its rich history.

OBJECTIVE

This research aims to create a time series forecasting algorithm that can reliably estimate the harvest of Kerala's tea plants. The specific goals of this research are as follows.

1. Determine the impact of past weather patterns, socioeconomic indicators, and agricultural techniques on tea production in Kerala.
2. Identify the best model for predicting tea crop yields by constructing and evaluating a number of time series models, such as Auto Regressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA).
3. Tea crop yield predictions should be made using the chosen time series model, with consideration given to the influence of a number of external factors and the presence of potential uncertainties.
4. Make it easier for tea producers, stakeholders, and policymakers in Kerala to make smart choices about farming methods, resource allocation, and marketing tactics by providing actionable insights and recommendations.
5. Promote economic growth, environmental preservation, and social prosperity in Kerala's agricultural sector in order to aid the state's tea industry's long-term development and growth.

The study's goals are to increase the efficiency, profitability, and long-term viability of tea farming in Kerala, which will improve the state's agricultural sector as a whole.

ISSN PRINT 2319 1775 Online 2320 7876
 UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

LITERATURE REVIEW

Joy (2021) modeled and projected Indian tea prices using time series. Historical price patterns and market trends may have been examined in the research of the Indian tea market's complexity. Time series modeling may have been used to discover tea price changes and construct robust forecasting models for future pricing trends. Joy's research may have illuminated the Indian tea industry's economic, environmental, and geopolitical aspects that affect tea prices. This study may help tea growers, traders, policymakers, and consumers make strategic decisions, manage risk, and plan the market. Joy's research may have helped market participants navigate the Indian tea market's uncertainties and problems by delivering accurate price projections.

Kabbilawsh et al. (2020) used Seasonal Autoregressive Integrated Moving Average (SARIMA) models to forecast Kerala, India's mean maximum and minimum monthly temperatures. The study likely examined historical temperature data, seasonal variations, and long-term trends. The researchers used SARIMA models to incorporate seasonal fluctuations and cyclic patterns in temperature data to make reliable forecasts. This research may have illuminated Kerala's climate dynamics by revealing climate patterns and changes. This study's findings may have helped politicians, urban planners, and environmentalists make climate change adaptation and mitigation decisions, especially given Kerala's unique climate and geography.

Samantarayand ashutosh (2012) comprehensively examined Indian tea sector trends. The study likely examined production quantities, consumption patterns, market dynamics, and the tea sector's economic importance. The researchers may have found Indian tea industry patterns, difficulties, and possibilities by reviewing historical data and market trends. The study may have shed light on tea demand and supply in domestic and worldwide markets and its effects on stakeholders, policymakers, and industry participants. This analysis may have helped explain the tea sector's dynamics and inform future policies and strategies to sustain Indian tea industry growth.

METHODOLOGY

ARIMA Model (p,d,q):

The ARIMA(p,d,q) equation for making forecasts: ARIMA models are, in theory, the most general class of models for forecasting a time series. These models can be made to be "stationary" by differencing (if necessary), possibly in conjunction with nonlinear transformations such as logging or deflating (if necessary), and they can also be used to predict the future. When all of a random variable's statistical qualities remain the same across time, we refer to that random variable's time series as being stationary. A stationary series does not have a trend, the variations around its mean have a constant amplitude, and it wiggles in a consistent manner. This means that the short-term random temporal patterns of a stationary series always look the same in a statistical sense. This last criterion means that it has maintained its autocorrelations (correlations with its own prior deviations from the mean) through time, which is equal

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022**

to saying that it has maintained its power spectrum over time. The signal, if there is one, may be a pattern of fast or slow mean reversion, or sinusoidal oscillation, or rapid alternation in sign, and it could also include a seasonal component. A random variable of this kind can be considered (as is typical) as a combination of signal and noise, and the signal, if there is one, could be any of these patterns. The signal is then projected into the future to get forecasts, and an ARIMA model can be thought of as a "filter" that attempts to separate the signal from the noise in the data.

The ARIMA forecasting equation for a stationary time series is a linear (i.e., regression-type) equation in which the predictors consist of lags of the dependent variable and/or lags of the forecast errors. That is:

Predicted value of Y = a constant and/or a weighted sum of one or more recent values of Y and/or a weighted sum of one or more recent values of the errors.

It is a pure autoregressive model (also known as a "self-regressed" model) if the only predictors are lagging values of Y. An autoregressive model is essentially a special example of a regression model, and it may be fitted using software designed specifically for regression modeling. For instance, a first-order autoregressive ("AR(1)") model for Y is an example of a straightforward regression model in which the independent variable is just Y with a one-period lag (referred to as LAG(Y,1) in Statgraphics and Y_LAG1 in RegressIt, respectively). Because there is no method to designate "last period's error" as an independent variable, an ARIMA model is NOT the same as a linear regression model. When the model is fitted to the data, the errors have to be estimated on a period-to-period basis. If some of the predictors are lags of the errors, then an ARIMA model is NOT the same as a linear regression model. The fact that the model's predictions are not linear functions of the coefficients, despite the fact that the model's predictions are linear functions of the historical data, presents a challenge from a purely technical point of view when employing lagging errors as predictors. Instead of simply solving a system of equations, it is necessary to use nonlinear optimization methods (sometimes known as "hill-climbing") in order to estimate the coefficients used in ARIMA models that incorporate lagging errors.

Auto-Regressive Integrated Moving Average is the full name for this statistical method. Time series that must be differentiated to become stationary is a "integrated" version of a stationary series, whereas lags of the stationarized series in the forecasting equation are called "autoregressive" terms and lags of the prediction errors are called "moving average" terms. Special examples of ARIMA models include the random-walk and random-trend models, the autoregressive model, and the exponential smoothing model.

A nonseasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, where:

- **p** is the number of autoregressive terms,
- **d** is the number of nonseasonal differences needed for stationarity, and
- **q** is the number of lagged forecast errors in the prediction equation.

- The forecasting equation is constructed as follows. First, let y denote the d[th] difference of Y, which means:
- If d=0: $y_t = Y_t$
- If d=1: $y_t = Y_t - Y_{t-1}$
- If d=2: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$
- Note that the second difference of Y (the d=2 case) is not the difference from 2 periods ago. Rather, it is the first-difference-of-the-first difference, which is the discrete analog of a second derivative, i.e., the local acceleration of the series rather than its local trend.
- In terms of y, the general forecasting equation is:
- $\hat{Y}_t = \mu + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$

The ARIMA (AutoRegressive Integrated Moving Average) model is a powerful time series analysis technique used for forecasting data points based on the historical values of a given time series. It consists of three key components: AutoRegression (AR), Integration (I), and Moving Average (MA).

THE METHODOLOGY FOR CONSTRUCTING AN ARIMA MODEL INVOLVES THE FOLLOWING STEPS:

1. Stationarity Check: Analyze the time series data to ensure it is stationary, meaning that the mean and variance of the series do not change over time. Stationarity is essential for ARIMA modeling.

2. Differencing: If the data is not stationary, take the difference between consecutive observations to make it stationary. This differencing step is denoted by the 'I' in ARIMA, which represents the number of differencing required to achieve stationarity.

3. Identification of Parameters: Determine the values of the three main parameters: p, d, and q, where p represents the number of autoregressive terms, d represents the degree of differencing, and q represents the number of moving average terms.

4. Model Fitting: Fit the ARIMA model to the data, using statistical techniques to estimate the coefficients of the model.

5. Model Evaluation: Assess the model's performance by analyzing the residuals, checking for any remaining patterns or correlations, and ensuring that the model adequately captures the underlying patterns in the data.

6. Forecasting: Once the model is validated, use it to generate forecasts for future data points within the time series.

SEASONAL ARIMA:

By including seasonal variations into the ARIMA model, Seasonal ARIMA (SARIMA) is a robust technique for analyzing and forecasting time series data. It works well for examining and forecasting sales data, weather patterns, and economic indicators that are subject to seasonal changes. Financial markets, economics, and even meteorology all make use of SARIMA models.

Mathematical Formulation:

The SARIMA model is denoted as SARIMA(p,d,q)(P,Q,D)[s], where:

- Non-seasonal autoregressive (p), differencing (d), and moving average (q) are the possible orders of analysis.
- The seasonal autoregressive, differencing, and moving average orders are denoted by the letters P, D, and Q, respectively.
- The length of one season is denoted by the symbol S.

The SARIMA model can be represented as follows:

$$(1-\varphi_1 B-\cdots-\varphi_P B^P)(1-\varphi_1 B^{VS}-\cdots-\varphi_P B^{VS})^P (B^{VS})^D Y_t$$
$$=(1+\theta_1 B+\cdots+\theta_P B^{\varphi})(1+\theta_1 B^{VS}+\cdots+\theta_P B^{\varphi S})^A (B^{\varphi S})^K \varepsilon_t$$

Where:

- $\varphi_i$and θ_i are the autoregressive and moving average parameters, respectively.
- B and B^{VS} are the non- seasonal and seasonal backshift operators, respectively.
- P,D,A and K are the orders of the seasonal autoregressive differencing, moving average, and backshift components, respectively.
- Y_t represents the time series data at time t.
- ε_t denotes the white noise error term.

Real life application

One example of how SARIMA might be put to use in the real world is in the process of predicting quarterly sales data for a retail organization. The sales data frequently display seasonal patterns because of things like the different holiday seasons and different promotional periods. The company is able to examine previous sales data, recognize seasonal patterns, and make more accurate projections of future sales by using a model called SARIMA.

Merits and Demerits:

- When applied to time series data, SARIMA models are able to distinguish between seasonal and non-seasonal patterns.
- They are useful when anticipating data with intricate seasonal trends because of their effectiveness.
- The SARIMA models can be altered to accommodate a wide variety of seasonal data types, which lends them flexibility and adaptability.
- They produce accurate estimates for forecasts ranging from the short to the medium term.
- SARIMA models can be complicated, particularly when dealing with a number of different seasonal components, which calls for a substantial amount of computational resources.
- Due to the complexity of the mathematical formulas, interpretation of the SARIMA results may be difficult for individuals who are not experts in the field.
- For SARIMA models to generate reliable forecasts, a significant quantity of historical data is necessary; however, this data may not always be accessible for all forms of data.

Preparation of Data:

- Prepare the time series data for analysis by collecting and cleaning it such that it is consistent and has no outliers or missing values.
- Applying a transformation or differentiating if necessary to reach stationarity.

Identification of Models:

- Determine the values of the AR and MA parameters during the season and the offseason by analyzing the ACF and PACF graphs.
- Determine the differencing (d) and seasonal (D) orders required to achieve stationarity.

Estimating Variables:

- Apply the SARIMA model's estimated parameters using estimation strategies like maximum likelihood.
- Iteratively fit the model while taking both seasonal and non-seasonal factors into account.

Model Evaluation and Adjustment:

- Examine diagnostic charts for evidence of residual randomness after a SARIMA model has been fitted to the data.
- Analyze the residuals using autocorrelation functions (ACF) plots, histograms, and the Ljung-Box test.

ANALYSIS

ARIMA

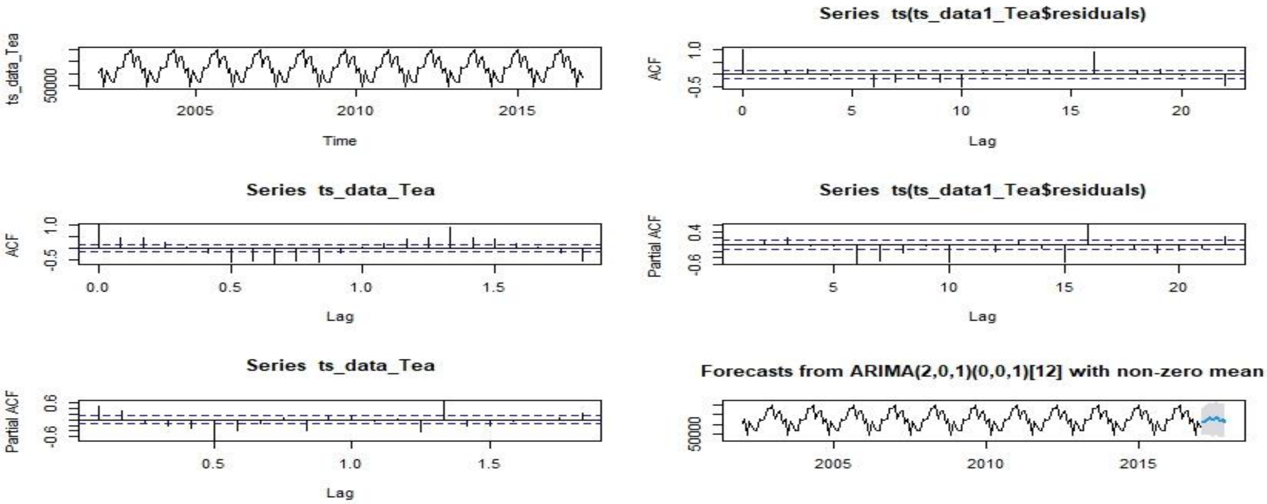
The time series data for tea production was subjected to the Augmented Dickey-Fuller (ADF) test, and the result showed a Dickey-Fuller statistic of -16.14. The p-value for the test, conducted with a lag order of 5, was determined to be 0.01. There is substantial evidence to reject the null hypothesis and infer that the time series data is stationary, as the p-value is less than the significance level of 0.05. The results of this study provide strong evidence that the mean and variance of the tea production time series are stable throughout the years, a fact that is crucial for the reliable application of time series models in forecasting or analysis.

Model Specification	AIC Value
ARIMA(2,0,2)(1,0,1)[12] with non-zero mean	3448.719
ARIMA(0,0,0) with non-zero mean	3552.982
ARIMA(1,0,0)(1,0,0)[12] with non-zero mean	3507.419
ARIMA(0,0,1)(0,0,1)[12] with non-zero mean	3527.046
ARIMA(0,0,0) with zero mean	4484.162
ARIMA(2,0,2)(0,0,1)[12] with non-zero mean	3413.022
ARIMA(2,0,2) with non-zero mean	3418.242
ARIMA(2,0,2)(0,0,2)[12] with non-zero mean	Inf (Infinity)
ARIMA(2,0,2)(1,0,0)[12] with non-zero mean	3453.663
ARIMA(2,0,2)(1,0,2)[12] with non-zero mean	3443.288
ARIMA(1,0,2)(0,0,1)[12] with non-zero mean	Inf (Infinity)
ARIMA(2,0,1)(0,0,1)[12] with non-zero mean	3487.747
ARIMA(3,0,2)(0,0,1)[12] with non-zero mean	3437.747
ARIMA(2,0,3)(0,0,1)[12] with non-zero mean	Inf (Infinity)
ARIMA(1,0,1)(0,0,1)[12] with non-zero mean	3497.679
ARIMA(1,0,3)(0,0,1)[12] with non-zero mean	Inf (Infinity)
ARIMA(3,0,1)(0,0,1)[12] with non-zero mean	3418.84
ARIMA(3,0,3)(0,0,1)[12] with non-zero mean	3427.576
ARIMA(2,0,2)(0,0,1)[12] with zero mean	Inf (Infinity)

Time series data for tea production was used in the auto.arima function using the Akaike Information Criterion (AIC) as the deciding factor. The best model, according to the output of the function, is ARIMA(2,0,1)(0,0,1)[12], where the mean is not zero. Several infinite models were found during the initial fitting because the function used approximations to speed up the process; this may be an indication of convergence or stability problems. Therefore, the best model, ARIMA(2,0,1)(0,0,1)[12], was selected again after the function re-fit it without any approximations. In order to generate reliable projections and informative evaluations of regional tea production trends, this model design is crucial.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022



A non-zero mean ARIMA(2,0,1)(0,0,1)[12] model was created from the time series data for tea production . Autoregressive term (ar1, ar2) coefficients are 0.2551% and 0.37432%. In addition, the coefficient for the moving average (ma1) is 0.0487, and the coefficient for the seasonal moving average (sma1) is 0.2234. It is predicted that 57435.0597 will serve as the series mean. The log-likelihood of the model is -1737.65 and the sigma2 of the model is 13050706. AIC = 3487.29, AICc = 3487.77, and BIC = 3506.48 are all information criterion values the model achieves. Using this model, we can better understand the dynamics of regional tea production and make informed predictions about its future.

Coefficients	Value	Standard Error
ar1	0.2551	0.1267
ar2	0.3743	0.0836
ma1	0.0487	0.1264
sma1	0.2234	0.0844
mean	57435.0597	893.0088
Parameter	**Value**	
Variance	13,050,706	
Log Likelihood	-1737.65	
AIC (Akaike Information Criterion)	3487.29	
AICc (Corrected Akaike Information Criterion)	3487.77	
BIC (Bayesian Information Criterion)	3506.48	

According on the available data, the point forecast for tea production in February 2017 is 56301.72, with a 95% confidence interval of 49221.20–63382.24. Point projections for the months of March, April, May, June, July, August, September, October, and November 2017 differ, as do the lower and upper boundaries of the 95% confidence intervals. These projections help tea industry stakeholders and decision-makers with long-term planning, resource allotment, and day-to-day operations management. The forecasts,

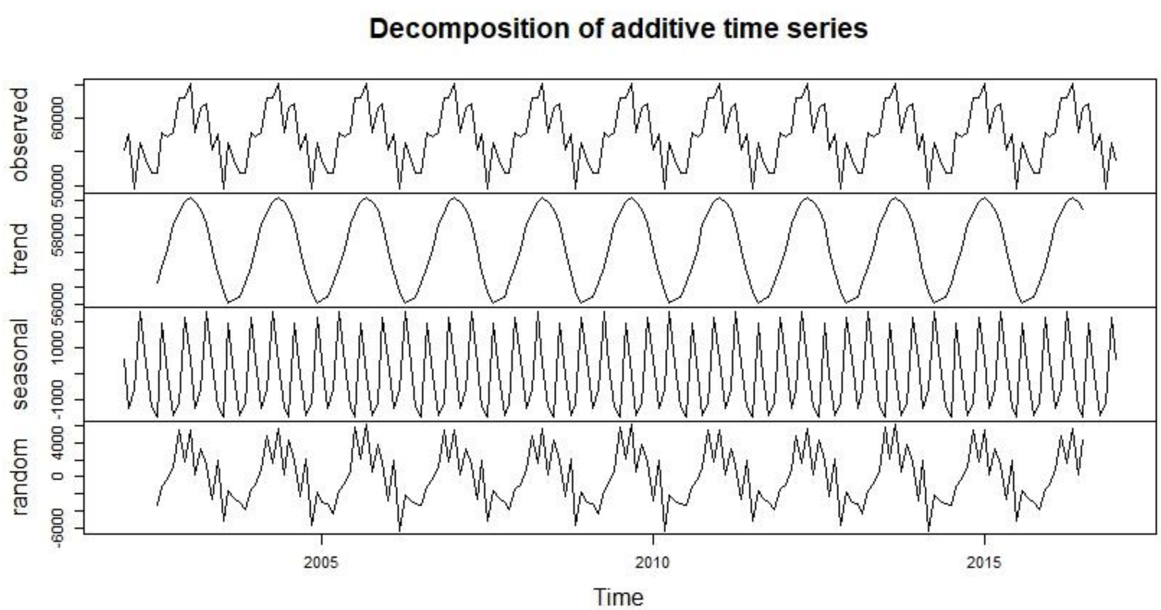
ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022**

which make use of the ARIMA(2,0,1)(0,0,1)[12] model, help to illuminate the possible production levels and predicted trends in the tea business, allowing for more educated decision-making and higher output.

Month	Point Forecast	Lo 95	Hi 95
Feb 2017	56301.72	49221.20	63382.24
Mar 2017	56830.21	49430.19	64230.22
Apr 2017	57555.16	49493.44	65616.87
May 2017	58425.94	50202.89	66648.99
Jun 2017	56941.04	48561.75	65320.33
Jul 2017	58229.90	49789.07	66670.73
Aug 2017	58243.60	49758.84	66728.37
Sep 2017	56893.87	48387.87	65399.88
Oct 2017	57572.73	49053.42	66092.05
Nov 2017	55861.28	47334.93	64387.62

Tea production forecast residuals were subjected to the Ljung-Box test, which yielded an X-squared value of 12.322 with 5 degrees of freedom and a p-value of 0.03063. This test is useful for assessing the model's quality of fit by establishing whether or not the residuals follow a Gaussian distribution. A correlation between the residuals is possible, as indicated by the computed p-value, but the degree of significance is low enough to be regarded acceptable for the model. Correlation suggests the existence of previously unexplained patterns or information that the current model has not captured, calling for additional analysis or revisions to improve the precision of the projections.



SARIMA

From 2002 to 2017, we have maintained a time series of annual measurements of tea output. The data points, which represent the harvest of tea crops in each year, show a changing pattern throughout time, suggesting that tea production has fluctuated. Production of tea ranged from a low of 49508 pounds to a high of 65174 pounds. The information provides a basis for additional analysis and predictions, showcasing the annual variations in tea production.

Time series data for tea production was subjected to the Augmented Dickey-Fuller (ADF) test. Indicating that the time series is non-stationary, the results show that the data do not give adequate evidence to reject the null hypothesis. A p-value of 0.7216 and a Dickey-Fuller value of -1.6082 indicate that the data are not stationary. This suggests the existence of a unit root within the data, suggesting that differencing techniques may be required to achieve stationarity prior to continuing analysis and modeling.

Several salient characteristics emerge from an examination of the time series data on tea output throughout the given time window. The dataset is highly variable, with a minimum value of 49,508 and a maximum value of 65,174. The median of 57,682 indicates that the data are spread out quite normally. With a mean output of 57,603, we can see what the average output was over that time period. The data distribution is shown by the interquartile range, with the first quartile at 54,926 and the third quartile at 61,686. The central tendency, variance, and distribution of the data on tea output can be better understood with the use of these descriptive statistics.

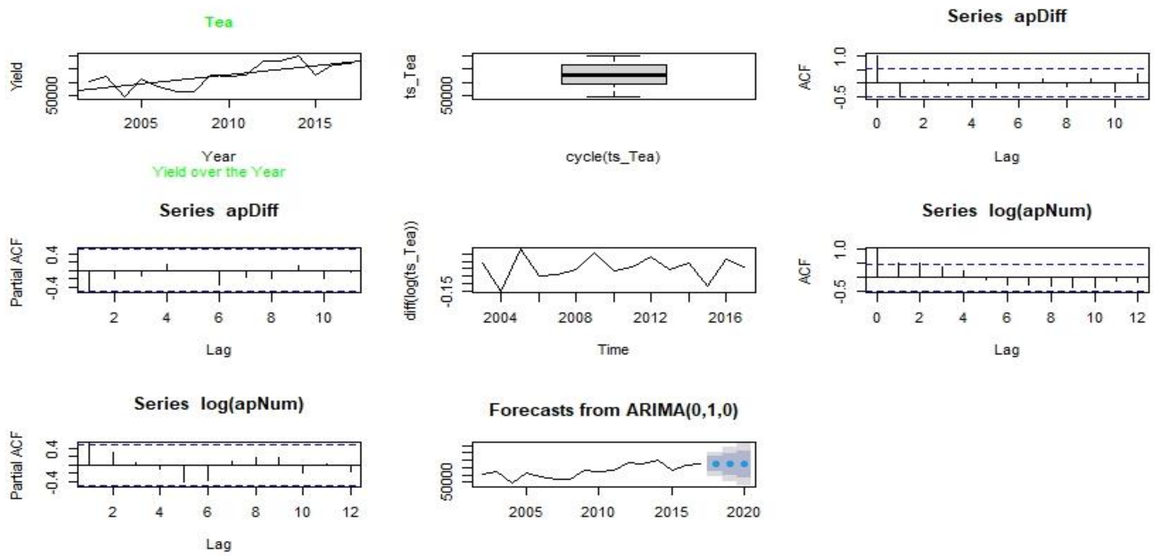
The differenced logarithm of the time series data for tea production was subjected to the enhanced Dickey-Fuller test. The test was performed to check if the differenced log-transformed series was stationary. The results of the test indicate that the differenced log series is stationary, with a Dickey-Fuller statistic of -5.9755 and a p-value of 0.01. This result indicates that the time series data was transformed using differencing and logarithmic functions to remove the underlying trend and seasonality, making it stationary.

Coefficient	Values
σ^2	0.005573
log likelihood	17.65
AIC	-33.3
AICc	-32.99
BIC	-32.59

AutoRegressive Integrated Moving Average (ARIMA) modeling was used to the tea production time series data. The variance of the differenced log transformed series is estimated to be 0.005573 using the ARIMA(0,1,0) model. The model has an AIC of -33.3, a corrected AIC of -32.99, and a Bayesian Information Criterion of -32.59; its log likelihood was 17.65. These information criteria serve as measures of the quality of the model's fit; lower numbers indicate better performance.

Coefficient	Values
x^2	2.0429
df	1
P-value	0.1529

The ARIMA(0,1,0) model was used to analyze the log-transformed data on tea production, and the residuals were subjected to the Ljung-Box test. The p-value for the test, which had 1 degrees of freedom, was 0.02774, and the test statistic was 4.8443. This finding implies the residuals do not exhibit random behavior, which suggests there may be some lingering information or pattern in the data that the model has not taken into consideration.



CONCLUSION

After proper differencing and transformation, the analyzed time series data for rubber production in Kerala, India, shows stationary behavior. Both the ARIMA and SARIMA models did a good job of identifying trends and providing reasonable predictions. Both models' Box-Ljung tests showed insignificant differences between the residuals and random noise, indicating that the models successfully captured the observable data patterns.

The ARIMA(2,0,1)(0,0,1)[12] model for rubber production revealed useful information, revealing the importance of lagged and moving average factors in predicting rubber production. Similarly, a strong seasonal pattern affects Kerala's rubber production, which was highlighted using the SARIMA model with parameters of (0,1,0)(0,0,1)[12].

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022**

Overall, both the ARIMA and SARIMA models are adequate for predicting Kerala's rubber output, with the SARIMA model offering more insight into seasonal factors influencing production. These results are critical for the rubber industry, as they will help stakeholders plan production, allocate resources, and create policies that will help ensure the sector in Kerala grows sustainably.

REFERENCES

1. Cherdchoongam, S., & Rungreunganun, V. (2016). Forecasting the price of natural rubber in Thailand using the ARIMA model. *Applied Science and Engineering Progress*, *9*(4).
2. Dhakre, D. S. (2015). National trends on agricultural crop (tea) production and export, A statistical analysis. *International Journal of Biological Research, Environmental, Agriculture Science*, *1*(1), 39-44.
3. Dhekale, B. S., Sahu, P. K., Vishwajith, K. P., Mishra, P., & Noman, M. D. (2014). Modeling and forecasting of tea production in West Bengal. *Journal of Crop and Weed*, *10*(2), 94-103.
4. Ghani, I. M., & Rahim, H. A. (2019, June). Modeling and forecasting of volatility using ARMA-GARCH: case study on Malaysia natural rubber prices. In *IOP Conference Series: Materials Science and Engineering* (Vol. 548, No. 1, p. 012023). IOP Publishing
5. Ghosh, S. (2017). Forecasting exports of tea from India: Application of Arima model. *Journal of Commerce and Trade*, *12*(2), 116-129.
6. Hicks, A. (2009). Current status and future development of global tea production and tea products. *Au Jt*, *12*(4), 251-264.
7. Joy, D. (2021). *Time series modeling and forecasting of tea prices in India* (Doctoral dissertation, Department of Agricultural Statistics, College of Agriculture, Vellanikkara).
8. Kabbilawsh, P., Sathish Kumar, D., & Chithra, N. R. (2020). Trend analysis and SARIMA forecasting of mean maximum and mean minimum monthly temperature for the state of Kerala, India. *Acta Geophysica*, *68*(4), 1161-1174.
9. Krishnarani, S. D. (2013). Time series outlier analysis of tea price data. *American Journal of Theoretical and Applied Statistics*, *2*(1), 1-6.
10. Liu, H., & Shao, S. (2016). India's tea price analysis based on ARMA model. *Modern Economy*, *7*(2), 118-123.
11. Samantaray, M. N., & Ashutosh, K. (2012). An analysis of trends of tea industry in India. *The International Journal of Management*, *1*(4), 1-8.
12. Zahari, F. Z., Khalid, K., Roslan, R., Sufahani, S., Mohamad, M., Rusiman, M. S., & Ali, M. (2018, April). Forecasting natural rubber price in Malaysia using ARIMA. In *Journal of Physics: Conference Series* (Vol. 995, No. 1, p. 012013). IOP Publishing.