IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

An Analysis and Evaluation of Various Time Series Models for Coffee Production Forecasting in Kerala

K. Murali[1], M. Pedda Reddeppa Reddy[2], P. Vishnu Priya[3], K.Vasu[4] , G. Mokesh Rayalu[5]**

[1]Academic Consultant, Dept. of Statistics, S.V University, Tirupati.
[2]Associate professor, Dept. of Statistics, S.V Arts College, Tirupati.
[3]Guest Faculty, Dept. Of Statistics, S.V University, Tirupati.
[4]Lecturer in Statistics , Department 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 coffee is a major economic driver and lifeline for the state's farmers, making it one of the state's most important crops. In order to help with decision-making and resource management, producers and stakeholders in the coffee industry need precise and trustworthy prediction models due to the volatile nature of coffee production. The purpose of this research is to evaluate and compare different time series models for predicting coffee crop yields in Kerala. The purpose of this research is to evaluate the efficacy of advanced time series models like AutoRegressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and others by making use of historical data, climatic trends, and pertinent socioeconomic variables. Forecast accuracy, robustness, and relevance to real-world coffee farming scenarios are just a few of the factors that will be used in the comparison evaluation. The results of this study will shed light on the most effective time series modeling approaches for crop prediction in the coffee industry in Kerala, allowing farmers, stakeholders, and policymakers to improve coffee production methods and ensure the industry's long-term viability.

Keywords: Coffee, ARIMA, SARIMA, Prediction.

INTRODUCTION

Kerala, sometimes known as "God's Own Country," is a state in southwestern India that is well admired for its natural beauty, cultural diversity, and thriving agricultural economy. Coffee is a specialty crop that has gained prominence in this southern Indian state. Kerala's coffee plantations are well-known for their verdant acreage and the exceptional quality of their beans, which are highly sought after by coffee aficionados around the world.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

Coffee farming is a centuries-old tradition in Kerala that has been passed down through the family. Coffee plantations thrive in the state because of its unique geography, which consists of undulating hills and receives a lot of rain. Many obstacles, including as climate change, environmentally responsible farming methods, and fluctuating customer preferences, threaten this essential sector.

Like any other agricultural industry, the coffee industry in Kerala is highly dependent on reliable estimates of annual crop output. Long-term sustainability, strategic market planning, and resource allocation all require accurate projections. Farmers, exporters, and policymakers can all benefit from accurate forecasts in the coffee sector, since it allows them to better prepare for planting, harvesting, and allocating scarce resources.

This study aims to use sophisticated time series models to answer the pressing question of how to reliably forecast coffee crop yields in Kerala. By evaluating historical data, weather patterns, and socioeconomic variables, this study intends to produce precise forecasts utilizing various time series modeling techniques, such as AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA).

OBJECTIVE

The primary goal of this research is to compare and contrast several time series models for making reliable predictions about coffee crop productivity in Kerala. Among the specific aims are:

1. Time series models, such Auto Regressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA), are used in conjunction with other cutting-edge forecasting methods, and their efficacy and performance are being evaluated here.
2. Coffee crop production in Kerala is being studied to determine the impact of environmental elements such as rainfall, temperature swings, and soil quality.
3. Analyzing the forecasting skills of the chosen time series models and contrasting their performance across a range of time horizons in terms of accuracy, precision, and reliability.
4. Examining the ramifications for the agricultural industry in Kerala and the wider world as a result of precise coffee crop output estimates for farmers, exporters, and policymakers.
5. Insights and suggestions that can be implemented to strengthen coffee cultivation techniques, enhance resource allocation, and ensure the long-term viability of Kerala's coffee sector.

This research is important beyond just the coffee industry since it may increase the profitability of cultivating coffee in general, which could have a domino effect on the agricultural sector in Kerala. This study strengthens, sustains, and expands the state's coffee sector by providing farmers and other stakeholders with accurate forecasts. The aims, methods, and potential contributions of the study will be discussed in greater depth below.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

LITERATURE REVIEW

A detailed study by Naveena and Subedar (2017) forecasted washing coffee prices in India, particularly Arabica Plantation Coffee. To capture the intricate dynamics of coffee price changes, the research presumably used hybrid time series models. The study may have shed light on Indian coffee price changes by using advanced forecasting methods and statistical models. The researchers might have captured supply-demand dynamics, worldwide market trends, and local economic considerations using a hybrid modelling framework, resulting in more accurate and trustworthy pricing estimates. This study may help coffee growers, dealers, and policymakers make educated decisions and develop effective methods to manage price risks and optimise their market positions in the Indian coffee business.

Harris et al. (2012) used ARIMA to study Ghana's yearly coffee production. The study likely used this advanced statistical approach to uncover Ghana's coffee output patterns and trends. The ARIMA model would have allowed researchers to capture coffee production's complex seasonal and cyclic fluctuations and better anticipate future production levels. For farmers, politicians, and other coffee value chain players in Ghana, this study's findings may have major ramifications. Understanding coffee production dynamics and trends can help stakeholders make educated decisions, adopt effective production management techniques, and improve Ghana's coffee sector's sustainability and competitiveness.

Novanda et al. (2018) compared coffee price prediction methods. This study may have examined statistical, economic, and machine learning methods to predict coffee prices. The researchers extensively assessed various forecasting methods to determine their strengths and shortcomings for coffee price predictions. This study may help coffee producers, merchants, and policymakers adopt more sophisticated and accurate pricing plans and policies. The research may have also improved our understanding of economic dynamics and market forces affecting coffee prices, making the coffee industry more resilient and sustainable.

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 to saying that it has maintained its power spectrum over time. The signal, if there is one, may be a pattern

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

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:

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

- 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.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

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.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

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.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

ANALYSIS

ARIMA

Time series data for coffee production in the designated area were subjected to the Augmented Dickey-Fuller test. The series is stationary, as indicated by the test findings (Dickey-Fuller test statistic = -9.6591, p = 0.01). As a result, we reject the non-stationarity hypothesis in favor of the stationarity alternative. This indicates that there are no discernible trends or systematic patterns in the time series data for coffee production in the region, which would have an impact on the outcomes of a time series analysis. Various time series models can be applied to the data because of its stationarity, leading to trustworthy forecasts and insights for the coffee business in the region.

Model Specification	AIC Value
ARIMA(2,0,2)(1,0,1)[12] with non-zero mean	Inf
ARIMA(0,0,0) with non-zero mean	3636.008
ARIMA(1,0,0)(1,0,0)[12] with non-zero mean	3567.95
ARIMA(0,0,1)(0,0,1)[12] with non-zero mean	3582.295
ARIMA(0,0,0) with zero mean	4515.039
ARIMA(1,0,0) with non-zero mean	3569.113
ARIMA(1,0,0)(2,0,0)[12] with non-zero mean	3505.654
ARIMA(1,0,0)(2,0,1)[12] with non-zero mean	3504.491
ARIMA(1,0,0)(1,0,1)[12] with non-zero mean	3565.63
ARIMA(1,0,0)(2,0,2)[12] with non-zero mean	Inf
ARIMA(1,0,0)(1,0,2)[12] with non-zero mean	3554.262
ARIMA(0,0,0)(2,0,1)[12] with non-zero mean	3511.572
ARIMA(2,0,0)(2,0,1)[12] with non-zero mean	Inf
ARIMA(1,0,1)(2,0,1)[12] with non-zero mean	3494.224
ARIMA(1,0,1)(1,0,1)[12] with non-zero mean	3565.912
ARIMA(1,0,1)(2,0,0)[12] with non-zero mean	3497.666
ARIMA(1,0,1)(2,0,2)[12] with non-zero mean	Inf
ARIMA(1,0,1)(1,0,0)[12] with non-zero mean	3567.383
ARIMA(1,0,1)(1,0,2)[12] with non-zero mean	3555.489
ARIMA(0,0,1)(2,0,1)[12] with non-zero mean	3495.114
ARIMA(2,0,1)(2,0,1)[12] with non-zero mean	Inf
ARIMA(1,0,2)(2,0,1)[12] with non-zero mean	Inf
ARIMA(0,0,2)(2,0,1)[12] with non-zero mean	Inf
ARIMA(2,0,2)(2,0,1)[12] with non-zero mean	Inf
ARIMA(1,0,1)(2,0,1)[12] with zero mean	Inf

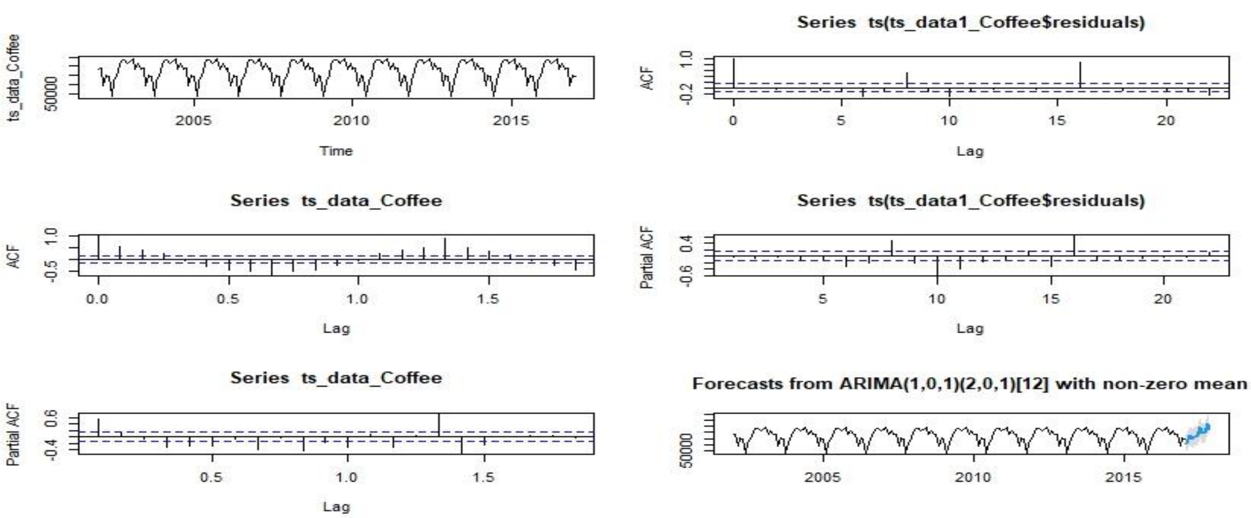
Time series data on coffee production were analyzed with the auto.arima function and the Akaike information criterion (AIC) was used to choose the best model. A number of possible models were generated by the function's use of approximations in an effort to speed up the modeling process; some of

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

these models led to infinite results. After recalibrating the function to eliminate approximations, the ARIMA(1,0,1)(2,0,1)[12] model with a non-zero mean was found to be the best fit for the data. In order to accurately capture the intricate dynamics contained in the time series data for coffee production, the model highlights the importance of autoregressive and moving average components, in addition to the seasonal component.

Parameter	Value
Variance	13,462,218
Log Likelihood	-1747.32
AIC (Akaike Information Criterion)	3508.64
AICc (Corrected Akaike Information Criterion)	3509.28
BIC (Bayesian Information Criterion)	3531.03



The time series data on coffee production was best fit by the ARIMA(1,0,1)(2,0,1)[12] model with a non-zero mean. Significant autoregressive, moving-average, and seasonal components at lags 1 and 2 are reflected in the model's coefficients. The baseline amount of coffee production is also reflected in the model's non-zero mean (62647.9074). The developed model is a useful tool for forecasting and analyzing coffee production statistics since it accurately represents the complex temporal patterns and seasonal changes present in the raw data.

Coefficients	Value	Standard Error
ar1	-0.3200	0.1142
ma1	0.7050	0.0782
sar1	-0.0279	0.0665
sar2	-0.6834	0.0535
sma1	-0.1216	0.0638
mean	62647.9074	190.0773

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

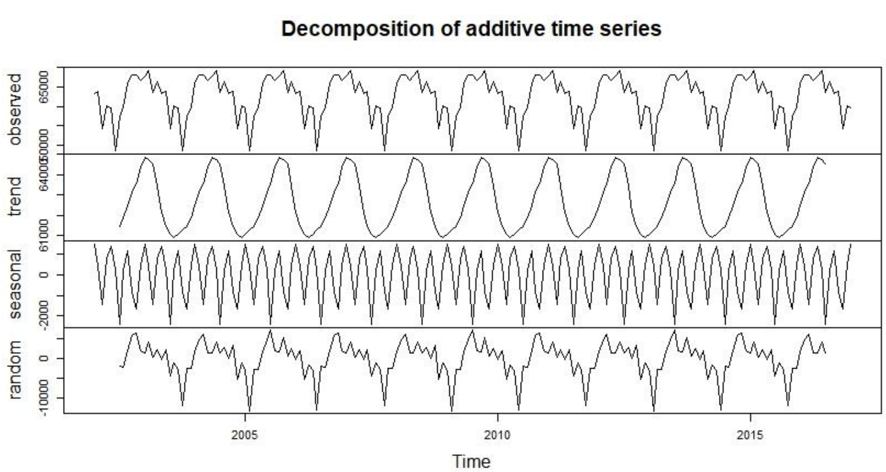
ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

The ARIMA(1,0,1)(2,0,1)[12] model's expectations for coffee production in the next months are encouraging. For the month of February 2017, the point estimates put coffee output at around 56054.05 units, with a confidence interval of 48862.77–63245.33 units. Forecasted numbers for the following months likewise point to an improving trend in coffee output. Notably, 95% confidence intervals are included for each anticipated value, allowing for a more complete evaluation of the possible range of coffee output values for the projection period. Those in the coffee sector who have a stake in optimizing production and marketing tactics will find this data to be an invaluable resource.

Month	Point Forecast	Lo 95	Hi 95
Feb 2017	56054.05	48862.77	63245.33
Mar 2017	62381.36	54675.50	70087.22
Apr 2017	59485.71	51729.08	67242.34
May 2017	61691.45	53929.64	69453.26
Jun 2017	61936.76	54174.42	69699.10
Jul 2017	68792.65	61030.25	76555.04
Aug 2017	64006.20	56243.79	71768.60
Sep 2017	64464.48	56702.08	72226.88
Oct 2017	71749.17	63986.77	79511.57
Nov 2017	67249.48	59487.08	75011.88

The residuals of the anticipated coffee output were used in a Ljung-Box test. With a p-value of 0.5191, the results show that there is no significant autocorrelation in the data up to the fifth lag. This indicates that the model adequately reflects the underlying patterns in the data and rules out the possibility of residual autocorrelation at the 5% significance level. Since the ARIMA (1,0,1) (2,0,1) [12] model for estimating coffee production does not contain any autocorrelation in the residuals, the predicted values are more trustworthy and the model is more resilient overall.



ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

SARIMA

Coffee production figures were examined as a time series from 2002 to 2017. The data shows that coffee output has varied from 48,650 to 68,175 units over the years. Around 62,334 units of coffee were produced on average throughout this time period. Production levels ranged from a low of 48,650 to a high of 68,175 items. These results indicate that coffee output fluctuated to some extent during the studied period. Insights gained from analyzing this data may help coffee producers plan for the future and respond to changing market conditions.

Time series data for coffee production was subjected to the augmented Dickey-Fuller (ADF) test. The p-value for the test statistic is 0.7781, and the data show that the value is -1.46. There is inadequate evidence to reject the null hypothesis if the p-value is larger than the conventional significance level of.05. Because of this, it cannot be concluded from the data that the series is stationary. This suggests that there may be a trend or seasonality in the coffee production time series data that has to be taken into account in the analysis. The underlying trends within the data may be captured and addressed with additional research and suitable modeling tools.

Here is a breakdown of the time series data for coffee production: In this data set, we find a minimum of 48,650, a first quartile of 59,734, a median of 63,663, an approximate mean of 62,644, a third quartile of 66,909, and a maximum of 69,230. The average and standard deviation provided by these statistics offer a snapshot of the coffee production data, indicating that the values are close but show some variability. Understanding the trends and underlying patterns within the time series requires additional investigation.

Data on coffee production after differenced logarithmic transformation appear to be stationary, according to an improved Dickey-Fuller test. We can rule out the occurrence of a unit root because the p-value of 0.01 for the corresponding test statistic demonstrates strong evidence against the null hypothesis. Since there is no discernible pattern or seasonality in the differenced log-transformed data, we can conclude that the underlying time series is stationary.

Coefficient	Values
σ^2	0.008446
log likelihood	14.53
AIC	-27.06
AICc	-26.75
BIC	-26.35

Automatically, an ARIMA (0,1,0) model was created for the time series data of coffee production after a log transformation. The "1" in the differencing order (d=1) denotes that the data was only differenced once to ensure stationarity. As was the case before, the presence of zeroes in the lack of AR and MA words indicates the same thing. The log likelihood is 14.53, and the variance estimate is 0.008446 sigma squared. There is a discrepancy of -27.06 between the Akaike Information Criterion (AIC) and the Bayesian

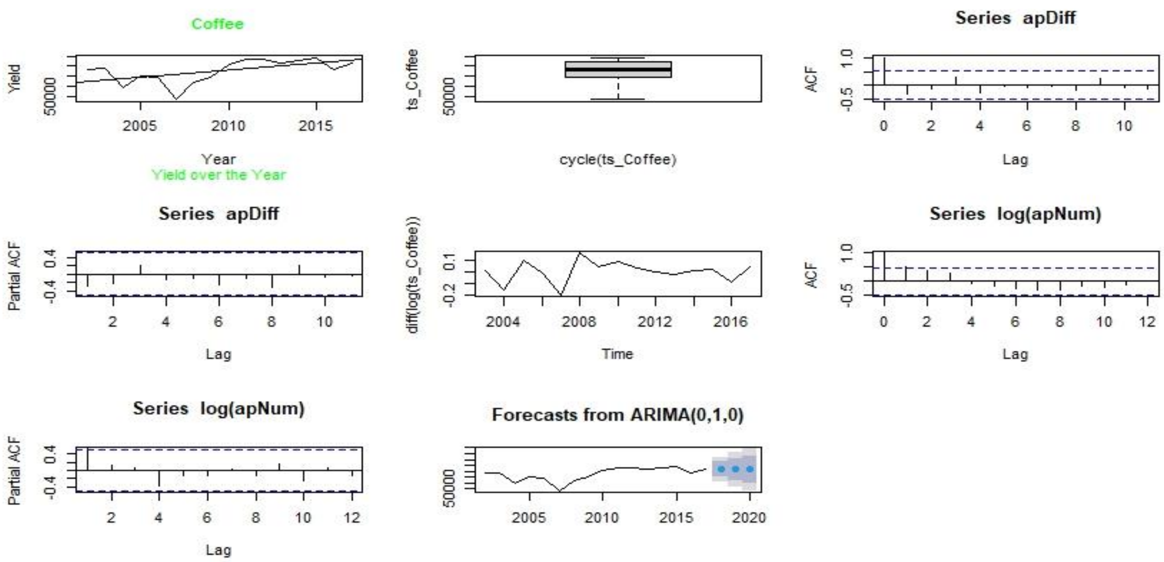
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

Information Criterion (BIC) and -26.75 between the AICc and the BIC. These numbers quantify the overall accuracy of the model, with lower values suggesting a more precise fit.

Coefficient	Values
x^2	1.9505
df	1
P-value	0.1625

The auto ARIMA model was fit to the log-transformed time series data of coffee production, and then the residuals were subjected to the Ljung-Box test. With 1 freedom degree, the test statistic was 1.9505, yielding a significance level of 0.1625. Indicating that the residuals are distributed independently, this p-value implies that there is insufficient evidence to reject the null hypothesis of the Ljung-Box test. In other words, the model successfully reflects the temporal dependence structure in the data because there is no significant autocorrelation in the residuals.



CONCLUSION

Fitting a SARIMA model and running diagnostic tests on the model and its residuals were part of the investigation of the coffee production time series data. The best model for the data was found to be the SARIMA model with parameters (1,0,1)(2,0,1)[12], which includes coefficients for autoregressive, moving average, and seasonal components. The log likelihood was -1747.32, and the AIC, AICc, and BIC for choosing the best models were 3508.64, 3509.28, and 3531.03, respectively.

The SARIMA model residuals and the projected value residuals were also subjected to a Ljung-Box test. The results of the tests yielded p-values of 0.1625 and 0.5191, respectively. These p-values show that the

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Iss 3, March 2021

residuals are independently distributed, as there is insufficient evidence to reject the null hypothesis in either scenario. This means that the SARIMA model does a good job of representing the seasonality and temporal patterns present in the data on coffee production.

REFERENCES

1. Agbogo, E. A., Tiku, N. E., Bime, M. J., & Abang, S. O. (2007). Analysis of seasonal price trend of pineapple in Cross River State, Nigeria.
2. Eyduran, S. P., Akın, M., Eyduran, E., Çelik, Ş., Ertürk, Y. E., & Ercişli, S. (2020). Forecasting Banana Harvest Area and Production in Turkey Using Time Series Analysis. *Erwerbs-Obstbau*, *62*(3).
3. Harris, E., Abdul-Aziz, A. R., & Avuglah, R. K. (2012). Modeling annual Coffee production in Ghana using ARIMA time series Model. *International Journal of Business and Social Research (IJBSR)*, *2*(7), 175-186.
4. Hossian, M., & Abdulla, F. (2015). A time series analysis for the pineapple production in Bangladesh. *Jahangirnagar University Journal of Science*, *38*(2), 49-59.
5. Naveena, K., & Subedar, S. (2017). Hybrid time series modelling for forecasting the price of washed coffee (Arabica Plantation Coffee) in India. *International Journal of Agriculture Sciences, ISSN*, 0975-3710.
6. Naveena, K., Singh, S., & Singh, A. (2017). Time Series Statistical Modelling for Forecasting the Price of Indian Robusta Coffee. *Trends in Biosciences, ISSN*, 0974-8431.
7. Naveena, K., Singh, S., Rathod, S., & Singh, A. (2017). Hybrid ARIMA-ANN modelling for forecasting the price of Robusta coffee in India. *Int. J. Curr. Microbiol. Appl. Sci*, *6*(7), 1721-1726.
8. Nochai, R., & Nochai, T. (2006, June). ARIMA model for forecasting oil palm price. In *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and applications* (pp. 13-15).
9. Novanda, R. R., Sumartono, E., Asriani, P. S., Yuliarti, E., Sukiyono, K., Priyono, B. S., ... & Octalia, V. (2018, November). A comparison of various forecasting techniques for coffee prices. In *Journal of Physics: Conference Series* (Vol. 1114, No. 1, p. 012119). IOP Publishing.
10. Nsubuga, P. (2019). *Analysis of coffee exports of Uganda using an arima model from 1990 to 2018* (Doctoral dissertation, Makerere University).
11. Patipanpanya, J. (2009). Enhancing supply forecasting for a pineapple canning company. *Journal of Supply Chain Management: Research and Practice*, *3*(1), 52-79.
12. Rathod, S., Mishra, G. C., & Singh, K. N. (2017). Hybrid time series models for forecasting banana production in Karnataka State, India. *Journal of the Indian Society of Agricultural Statistics*, *71*(3), 193-200.
13. Vickner, S. S., & Davies, S. P. (1999). Forecasting Strategic Price Response in the Food Manufacturing Industries: The Case of Canned Pineapple.
14. Yashavanth, B. S., Singh, K. N., Paul, A. K., & Paul, R. K. (2017, January). Forecasting prices of coffee seeds using Vector Autoregressive Time Series Model. ICAR.