SLIIT INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SCIENCES & HUMANITIES
SICASH 2023

Modeling Annual Coffee Production in Sri Lanka

[1] Jithmini, D., [2] Paboda, K. R., [3] Samaraweera, M. P. I. N., [4]Welaramba, W. D. K. H, [5]Peiris, T. S. G.

[1,2,3,4,5]Department of Mathematics and Statistics, Faculty of Humanities and Sciences,
Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka.

Corresponding author - hs21923692@my.sliit.lk

ARTICLE INFO

Article History:
Received: 10 September 2023
Accepted: 01 November 2023

Keywords:
Annual production; Autoregressive (AR); A toregressive Integrated Moving Average(ARMA); Coffee; Moving Average (MA)

Citation:

Jithmini, D., Paboda, K. R., Samaraweera, M. P. I. N., Welaramba, W. D. K. H, & Peiris, T. S. G. (2023). Modeling Annual Coffee Production in Sri Lanka . Proceedings of SLIIT International Conference on Advancements in Sciences and Humanities, 1-2 December, Colombo, pages 280-284.

ABSTRACT

Coffee production is a source of revenue for the economic sector in Sri Lanka. During 1988 to 2020, the mean annual coffee production is 7672 metric tons with a coefficient of variation is 30%. The advanced knowledge of annual coffee production has many advantages. However, past studies found that no model has been developed to model annual coffee production in Sri Lanka. In this study, an ARIMA (1,2,0) model was identified as the best fitted model to forecast the annual coffee production. The model was trained using data from 1988 to 2020 and validated using data in 2021. The best-fitted model was selected by comparing different statistical indicators such as Akaike Information Criteria, Schwarz Criteria, Log-likelihood Criteria, and volatility of the three parsimonious models. It was found that the errors of the best fitted model were white noise. The percentage errors for the forecast values for the training and validation data sets were within ± 10. The predicted annual production for 2022, 2023, 2024, and 2025 are 6987 MT, 6221 MT, 7209 MT, and 6664 MT, respectively. This is the first empirical study to develop a statistical model to predict annual coffee production in Sri Lanka. The model can be improved by using external variables

https://doi.org/10.54389/ETGI7523

as explanatory variables and considering dummy variables to capture the structural breaks.

1. INTRODUCTION

1.1. COFFEE CULTIVATION IN SRI LANKA

Coffee cultivation was introduced to Sri Lanka by Dutch settlers in the 17th century. Initially, it was grown in low-lying areas in the southwest of the island, but later, it spread to other parts of the country. However, coffee production in Sri Lanka experienced a decline after the 1870 coffee rust epidemic. With the decline of coffee production, tea became the primary agricultural crop, contributing significantly to the country's economy (Coffee, 2019). Nevertheless, coffee production in Sri Lanka has continued to be an important economic activity, contributing to both local and foreign markets. The country's primary coffee-producing regions are Kandy, Dikoya, Rattota, and Rikillagaskada, located in the south-central region of the island. These areas are characterized by high altitudes, misty hills, fertile soil, and excellent water resources. Such a favorable combination of natural conditions and terrain creates microclimates that aptly suit the production of excellent coffee (Coffee production in Sri Lanka, 2023) .

Sri Lanka mainly produces Arabica coffee, which is of high quality and is sought after for its mild flavor and delicate aroma. The production of coffee in Sri Lanka is based heavily on smallholder farms as most of the plantations are under private ownership (MDF, 2022). The farmers often use eco-friendly farming methods, including crop rotation, using natural mulches such as banana leaves and composting, and integrating cattle grazing. These practices preserve the natural ecosystem and promote plant growth, enhancing the overall quality of the coffee.

Like many other coffee-producing countries, Sri Lanka faces challenges that threaten coffee crop quality and quantity. One of the most significant threats is coffee leaf rust, also known as Hemileia vastatrix, which can significantly reduce crop yields (Coffee Rust, 2022). These factors make the prediction of annual coffee production difficult.

1.2. COFFEE PRODUCTION IN SRI LANKA

Sri Lanka was once the third-largest producer of coffee in the world. Sri Lanka is the 111th largest exporter of coffee in the world and coffee was the 395th most exported product in Sri Lanka. (Coffee in Sri Lanka, 2021). Though Sri Lanka coffee production fluctuated substantially in recent years, it tended to decrease through the 1972 - 2021 period ending at 5,306 mt in 2021 (Coffee Rust, 2022). Various time series models have been developed to predict coffee production in different countries (Binuomote, 2018). However, no statistical models have been developed to forecast annual coffee production in Sri Lanka.

1.3. OBJECTIVE OF THE STUDY

On view of the above, the objective of this study is to develop an univariate time series model to forecast the annual coffee production in Sri Lanka and validate the model.

2. MATERIALS AND METHODS

Auto-Regressive Integrated Moving Average (ARIMA) model is used. This method was first introduced by Box and Jenkins (1976) and until now it has become the most popular model for forecasting various products (Hyndman, 2001) . The ARMA (p, q) model can be represented by:

$$Y_t = \mu + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + e_t - \theta_1 e_1 - \theta_2 e_2 - \cdots \theta_q e_{t-q}$$

where (i=1, 2, p) and (j=1,2…., q) are the coefficient AR part and MA part respectively. {et} is white noise. The ARMA (p, q) model can only be made if the time series is stationary. If the series is nonstationary, d^{th} difference of the series: is considered. It is denoted by ARIMA(p,d,q).

2.1. SECONDARY DATA

The annual coffee production data from 1988 to 2021 was obtained from the Food and Agriculture Organization of the United Nations database. The statistical analysis was performed using EViews 12 software.

3. RESULTS AND DISCUSSION

3.1. TEMPORAL VARIABILITY OF THE ORIGINAL SERIES

The time series plot in Figure 1 shows the temporal variability of coffee production in Sri Lanka from 1988 to 2020.

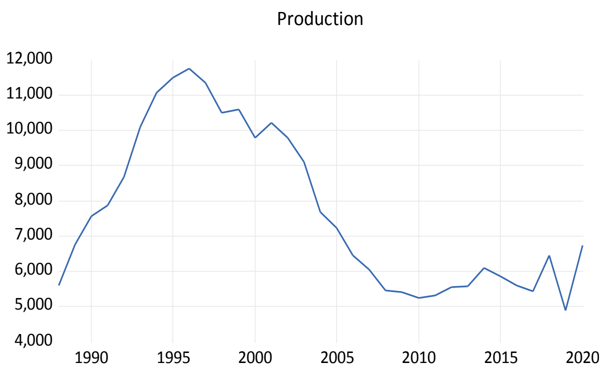


Figure 1. Time series plot of the yearly coff production

Figure 1 indicates that the yearly coffee production of Sri Lanka varies from 4887 (minimum) during the year 2019 to 11760 (maximum) during the year 1996 with a mean of 7672 and SD is 2250. There is a gradual decrease in Coffee Production from 1996 to 2020. Figure 1 also illustrates that the original observed series was nonstationary. This was confirmed by the Augmented Dickey-Fuller (ADF) test (ADF Test statistic = -1.709, p = 0.4168). Furthermore, it was found the first six autocorrelations are significantly different from zero ($p < 0.05$) and the first few autocorrelations are gradually declining.

It was also found that the first order difference of series is not stationary (ADF Test statistic = -2.253, p = 0. 193). Then the second-order difference was considered. As the ADF was significant (ADF Test statistic =-2.9640, p = 0.000), it can be concluded that the second-order difference series is stationary using the ADF test (Table 1). The correlogram of the stationary series is shown in Figure 2.

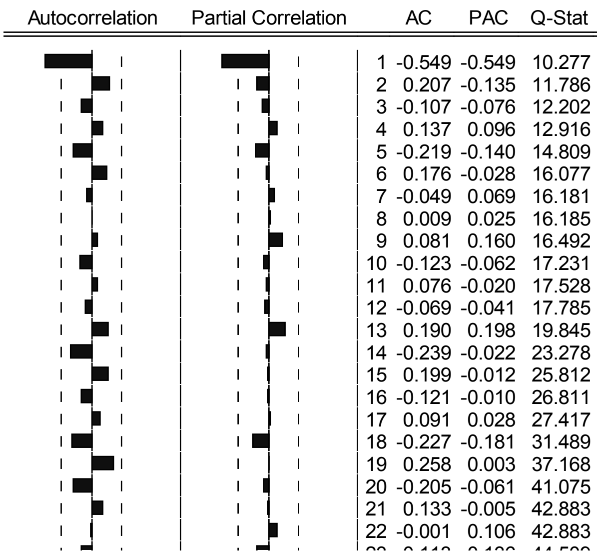


Figure 2. Correlogram of the second order differenced series

It can be seen that in ACF and PACF of the correlogram (Figure 2) were statistically significant only at lag 1. Thus, by comparing theoretical ACF and partial auto correlation function of AR(1) and MA(1), ARIMA (1,2,0), ARIMA (0,2,1), and ARIMA (1,2,1) were considered the parsimonious models.

3.2. SELECT ONE OF THE BEST-FITTED MODEL

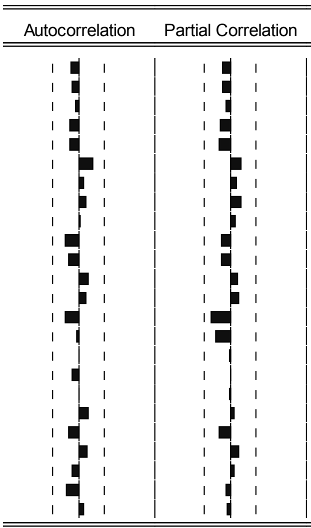
To select the best fitted model out of above three parsimonious models, various statistical indicators were compared (Table 1).

Table 1. Comparison of the selected ARIMA models

Parameters and Indicators	Model		
	ARIMA (1,2,0)	**ARIMA (0,2,1)**	**ARIMA (1,2,1)**
Parameter– AR (1)	significant	not applicable	significant
Parameter– MA (1)	not applicable	significant	not significant
σ^2_Volatility	457828.6	551940.9	450804.8
AIC	16.107	16.275	16.151
SBIC	16.246	16.413	16.336
Log-likelihood	-246.66	-249.26	-246.35

Results in Table 1 show that the MA parameter in ARIMA (1,2,1) is not significant and thus it was not an appropriate model to compare with other models. Of the remaining two models, the ARIMA (1,2,0) model has the lowest AIC, lowest BIC, and highest log-likelihood. Furthermore, the σ^2_volatility in ARIMA (1,2,0) was significantly less than that in ARIMA (0,2,1). Therefore, ARIMA (1,2,0) was chosen as the best-fitted model.

3.3. DIAGNOSTIC FOR BEST-FITTED MODEL



	AC	PAC	Q-Stat	Prob
1	-0.105	-0.105	0.3736	
2	-0.090	-0.102	0.6605	0.416
3	-0.037	-0.059	0.7098	0.701
4	-0.112	-0.136	1.1833	0.757
5	-0.109	-0.155	1.6493	0.800
6	0.204	0.149	3.3593	0.645
7	0.074	0.085	3.5927	0.732
8	0.101	0.142	4.0483	0.774
9	0.024	0.073	4.0750	0.850
10	-0.185	-0.123	5.7507	0.765
11	-0.130	-0.108	6.6188	0.761
12	0.135	0.093	7.5949	0.749
13	0.095	0.110	8.1045	0.777
14	-0.187	-0.257	10.211	0.677
15	-0.023	-0.191	10.244	0.744
16	0.007	-0.015	10.247	0.804
17	-0.085	0.002	10.777	0.823
18	0.008	-0.014	10.782	0.868
19	0.127	0.052	12.161	0.839
20	-0.136	-0.155	13.880	0.791
21	0.123	0.116	15.420	0.752
22	-0.088	0.049	16.294	0.753
23	-0.159	-0.054	19.518	0.613
24	0.068	-0.046	20.195	0.630

Figure 3. Correlogram of the residuals of the ARIMA model (1,2,0)

The correlogram of residuals from the selected model is shown in Figure 3. All the spikes are now within the significance limits in both ACF and PACF as all p values are greater than 0.05. Therefore, it can be concluded that the residuals are random. It was found that the error series having constant variance as the plot of residuals vs predicted values showed a random nature. Furthermore, from Jarque-Bera test, it was found that errors do not significantly deviate from the normality ($p > 0.05$). The properties of coefficients of the fitted model are shown in Table 2.

Table 2. Details of the parameters of ARIMA (1,2,0)

Dependent Variable: D(PRODUCTION,2)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 06/20/23 Time: 21:21
Sample: 1990 2020
Included observations: 31
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-23.91730	69.28588	-0.345197	0.7325
AR(1)	-0.849251	0.125657	-6.758506	0.0000
SIGMASQ	457828.6	151779.5	3.016406	0.0054

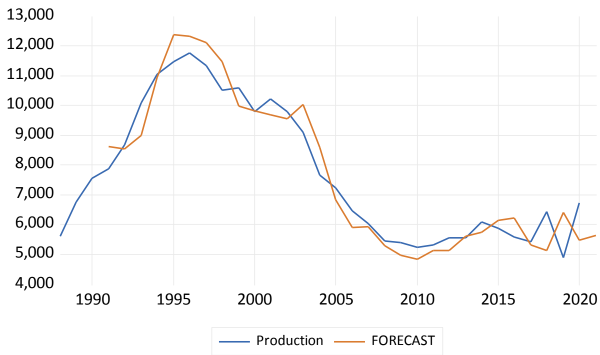
R-squared	0.507283	Mean dependent var	21.87097
Adjusted R-squared	0.472089	S.D. dependent var	979.8803
S.E. of regression	711.9562	Akaike info criterion	16.10688
Sum squared resid	14192688	Schwarz criterion	16.24565
Log likelihood	-246.6567	Hannan-Quinn criter.	16.15212
F-statistic	14.41389	Durbin-Watson stat	2.095374
Prob(F-statistic)	0.000050		
Inverted AR Roots	-.85		

The best fitted model can be written as .

$$Y_t = -23.91730 - 0.849252 * Y_{t-1} + Y_{t-2} - 0.849251 * Y_{t-3}.$$

Figure 4 indicates that the trend betw actuals and forecast for the training set is not much deviated. The correlation between actuals and predicted for the training set is strong positive significant($r = 0.960$, $p< 0.05$).

Figure 4. Comparison between actuals vs predicted for the training set.



3.4. SHORT TERM PREDICTION

The annual coffee production for the years 2021 to 2025 was predicted using the best-fitted model and the results are shown in Table 4:

Table 4: Forecast values from 2021-2025

Year	Forecast (MT)
2021	5635
2022	6987
2023	6221
2024	7209
2025	6664

According to the United Nations database, the actual coffee production in 2021 was 5306 MT. Our forecast value is 5635 MT and thus the percentage error is as low as –6%. Although according to the Department of Export Agriculture's (DEA) statistics, Sri Lanka's coffee production at the end of 2022 has increased compared that with 2021 (Attygalle, 2023). In fact, our model also showed an increasing trend in 2022 with respect to 2021. The percentage increase is 23.99%

4. CONCLUSIONS

Based on annual coffee production from 1988 to 2020, the identified best-fitted model was found as ARIMA (1,2,0). As the correlation between actuals and predicted for the training set is 0.960 ($p< 0.05$), the above model can be recommended to predict short-term annual coffee production in Sri Lanka. The predicted values for the years 2022, 2023, 2024, and 2025 are 6987 MT, 6221 MT, 7209 MT, and 6664 MT, respectively. The forecast values indicate that there will be an increase in the production of coffee from 2021 to 2025. It can be concluded that local coffee production will increase in the next few years. By implementing this model, Sri Lanka can significantly enhance forward contacts with various countries. In developing the model, the structural changes during 1988-2020 were not considered. Thus, it is suggested that the model to be improved by including dummy variables to capture the effect of structural changes during that period.

References

Attygalle, R. (2023). *Brewing a special cup with Sri Lankan coffee*. Daily FT: https://www.ft.lk/agriculture/Brewing-a-special-cup-with-Sri-Lankan-coffee/31-748432#:~:text=The%20production%20which%20was%20at,and%20Robusta%20(Coffea%20canephora).

Augmented Dickey-Fuller Test. (2022, 11 30). *Wikipedia* https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%93Fuller_test

Binuomote, S. L. (2018). Modelling Annual Yield of Coffee in Nigeria Using ARIMA Time Series Model (2018 – 2050). *Academic Research Journals*: https://www.academicresearchjournals.org/IJEBM/PDF/2018/October/Binuomote%20et%20al.pdf

Coffee. (2019). Department of Export Agriculture https://dea.gov.lk/coffee/#:~:text=Arabica%20type%20coffee%20prefers%20wet,%2C%20Kurunegala%2C%20Kandy%20and%20Matale

Coffee in Sri Lanka. (2021). OEC https://oec.world/en/profile/bilateral-product/coffee/reporter/lka

Coffee production in Sri Lanka. (2023). *Wikipedia* https://en.wikipedia.org/wiki/Coffee_production_in_Sri_Lanka

Coffee Rust. (2022). *BioProdection Portal* https://bioprotectionportal.com/resources/managing-pests/2022-11-09-coffee-rust-symptoms-causes-cycle-and-solutions/

Hyndman, R. J. (2001). *Box-Jenkins modelling.*

Lanka, E. S. (n.d.). *Sri Lanka Export Development Board*. Ceylon Coffee: https://www.srilankabusiness.com/food-and-beverages/coffee-from-sri-lanka.html

MDF. (2022). *Investment prospectus : Sri Lanka's Coffee.* Market Development Facility: https://marketdevelopmentfacility.org/wp-content/uploads/2022/09/Coffee-Prospectus.pdf

Nations, F. a. (n.d.). *FAOSTAT*. https://www.fao.org/faostat/en/#data/QCL