J Krishi Vigyan 2020, 9 (1) : 150-155 DOI : 10.5958/2349-4433.2020.00152.X

Forecasting Models for Economic Factors Affecting Banana Cultivation in Kerala

T Unnikrishnan* and V G Sunil**

Sree C. Achutha Menon Government College, Thrissur (Kerala)

ABSTRACT

An attempt was made to construct prediction models for the area, production and price of Banana in Kerala using ARIMA (Auto-Regressive Integrated Moving Average) model. The yearly secondary data on area, production and price of Banana collected from the Directorate of Economics and Statistics for the period from 1952-53 to 2018-19 were used for the analysis. The main objective of the study included assessment of trend and growth rates of area, production and price and identification of the best ARIMA model for prediction. Excellent parsimonious forecasting equations could be generated using the ARIMA technique. The validity of the models was tested using standard statistical techniques. The forecasting power of ARIMA model was used to forecast for six leading years, and results showed a good agreement between actual and predicted values.

Key Words: ACF, PACF, ARIMA, Banana, Kerala, Agriculture.

INTRODUCTION

Banana is an important fruit crop of many tropical and subtropical regions of India. It is cultivated in an area of 8,30,000 ha in India, and total annual production is around 2,97,79,000t. As a staple food, Banana contributes to the food security of millions of people. Bananas can be eaten fresh, cooked or processed into numerous value-added products, depending on the variety. India is the second-largest producer of banana in the world and is the second most important fruit crop in India next to mango. Bananas are cultivated commercially under tropical and subtropical conditions in all the states of India. The chief Banana growing states in India are Tamil Nadu, Andhra Pradesh, Maharashtra, Gujarat, Kerala, Karnataka (Saxena *et al*, 2017). Krishnamoorti and Hanif (2017) reported that the Arka banana special application through soil application 250 ml solution (%) on 45days after planting, followed by foliar application 0.5% on 5,6,7 and at shooting on hands recorded significantly highest yield (45.23 t/ha) over other two micronutrient application.

Kerala is one of the important Banana cultivating states in India. The total area of Banana in Kerala is 62,108 ha and an annual production 5,65,829t. Banana cultivation in Kerala is polyclonal, with a number of varieties under cultivation. It is cultivated in all most all districts of Kerala. Important varieties are Nendran, Palayankodan, Robusta, Rasthali etc. The system of banana cultivation is also diverse ranging from annually planted crops to semi perennial rainfed cultivation.

Forecasting models are used to provide an aid to decision-making and in planning for the future effectively and efficiently. These models are mostly developed using the past data to predict the future with the help of identifying the trends and patterns within the data. In other words, Statistical forecasting is the likelihood estimation of an event taking place in the future based on available data.

Corresponding Author's Email:: unniagstat@gmail.com

*Assistant Professor

**Assistant Professor, Department of Agricultural Extension, Communication Centre, Kerala Agricultural University, Vellanikkara, Thrissur, Kerala

ARIMA model is one of the widely used time series models. The popularity of the ARIMA model is due to its statistical properties as well as the well-known Box–Jenkins methodology (Box and Jenkins, 1970) in the model building process. ARIMA models (Box *et al*, 1994) have been utilised for predicting crop yield or any other agricultural production. Rathod *et al* (2011), Narayanaswamy *et al* (2012), and Pardhi *et al* (2016) applied regression analysis to study the effect of agricultural inputs and weather parameters on agricultural and horticultural crops. Sarika *et al* (2011) used, ARIMA model for modeling and forecasting India's pigeon-pea production. Suresh *et al* (2011) applied this model for forecasting sugarcane area, production and productivity in Tamil Nadu state of India. Naveena *et al* (2014) forecasted coconut production of India using ARIMA methodology. Considering the above mentioned facts, a study was conducted to model and forecast the area, production and productivity of Banana in Kerala using ARIMA (Auto Regressive Integrated Moving Average).

Saeed (2000) found that the best model for forecasting of wheat production in Pakistan was ARIMA(2, 2, 1). Forecasting sugarcane production in India was done using ARIMA(2, 1, 0) model by Mandal(2004). Carpio, CEBS (2002) explained the production response of cotton in India, Pakistan, and Australia using ARIMA(1,0,0) model. Price of oil palm was predicted efficiently using ARIMA (2,1,0) model (Nochai et al (2006)). ARIMA model for forecasting wholesale price of oil palm was ARIMA (1,0,1), and pure oil price of oil palm was ARIMA (3,0,0). Sen (2000) found that time series modelling of Black Banana price could be done using ARIMA (1,0,0). Unnikrishnan and Ajitha (2010) explained the advantage of ARIMA and cointegration models for the forecasting of Coconut and Rubber. Analytical studies in the ARIMA modelling of Banana in Kerala are not available, and hence a detailed study in this respect was essential to have an insight to the present scenario as well as future changes in the development of agriculture in Kerala.

MATERIALS AND METHODS

The principal objective of developing an ARIMA model for a variable is to generate post sample period forecasts for that variable. Its strength lies in the fact that the method is suitable for any time series with any pattern of change, and it does not require the forecaster to choose apriori the value of any parameter. Its limitations include its requirement of a long time series and short-term prediction. To develop advanced statistical models to predict the area, production and price of Banana in Kerala, secondary data from 1978-79 to 2018-19, which were collected from the publications of Planning Board and various Economic Reviews, were used. Forecasting models were developed for each crop using ARIMA models.

Auto-Regressive Integrated Moving Average (ARIMA) models

A time series X_t is an ARIMA (p,d,q) process if there exist polynomials Φ and Θ of degrees, p and q respectively and a white noise series Z_t such that the time series $D^d X_t$ is stationary and $\Phi(B)D^d X_t = \Theta(B)Z_t$ almost surely on the underlying probability space, where B denote the backshift operator $B(X_t) = X_{t-1}$. When d=0, $X_t = [\Theta(B)/\Phi(B)]Z_t$, is a stationary ARMA(p,q) process and in this case, all the zeroes of the polynomial $\Phi(z)$ lie outside the unit disc $|z| < 1$. When it has one or more values equal to one but no value inside the unit circle, it is nonstationary but integrated. In this case, the time series is an ARIMA process. In general when the function $[\Theta(B)/\Phi(B)]$ is well defined and analytic in the region $\{z \in C \mid \Phi(z)^{1}0\}$, if Φ has no roots on the circle $\{z \mid |z|=1\}$. Since it has p different roots there is an annulus $\{z/r<|z|<R\}$ with $r<1<R$ on which it has no root. On this annulus $\Phi(B)/\Theta(B)$ is analytic and it has a Laurent's series expansion $\Gamma(z)= \Sigma G_j B^j$ This series is uniformly and absolutely convergent on every compact subset of the annulus and the coefficients are uniquely determined by the value of X_t on the annulus. Hence the random variables X_t and Z_t are defined on a probability space $(\Omega, \mathcal{U}, P)$

and satisfying $\Phi(B)D^dX_t(\omega) = \Theta(B)Z_t(\omega)$ for almost every $\omega \varepsilon \Omega$.

Then the equation $\Phi(B)D^dX_t = \Theta(B)Z_t$ takes the form

$X_t = f_1X_{t-1}+f_2X_{t-2}+...+f_pX_{t-p} + Z_t+q_1Z_{t-1}+q_2Z_{t-2} +...+q_qZ_{t-q}$

For a stationary time series, the auto covariance and auto correlation at a lag kÎZ are defined by $g_X(k) = cov(X_{t+k}, X_t)$ and $r_X(k)) = \rho(X_{t+k}, X_t) = g_X(k))/ g_X(0))$, where $g_X(0)= Var(X_t)$ and $r_X(0) = 1$. The partial auto correlation at a lag k is defined as the correlation between $X_k -P_{k-1}(X_k)$ and $X_0 -P_{k-1}(X_0)$, where P_k is the projection of the vector yÎR^k on the subspace spanned by $(X_1,X_2,\ldots X_k)$ in R^k which is a linear combination $y=Sb_jx_j$ such that ||y-y^|| is minimal. It is the correlation due to intermediate values $X_1,X_2,\ldots X_{k-1}$ removed.

ARIMA model building

Identification

Simply saying, a time series is stationary if the autocorrelation function dies out fairly quickly. When autocorrelations drops off exponentially to zero, it is an autoregressive model whose order is determined by the number of partial autocorrelations which are significantly different from zero. On the other hand, if partial autocorrelations drop off exponentially to zero, it is a moving average model whose order is determined by the number of autocorrelations which are significantly different from zero. When both autocorrelation and partial autocorrelation move exponentially to zero, it is an ARMA model. The final models are achieved after going through the stages repeatedly.

Estimation

The precise estimates of the parameters of the model are obtained by the method of ordinary least squares as advocated by Box and Jenkins iterative procedure for finding the estimate through SPSS.

Diagnostic Checking

Different models can be obtained for various combinations of AR and MA individually and collectively. The best model is obtained with the following diagnostics like Coefficient of Determination (R^2), Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Portmonteau tests - Box Pierce or Ljung-Box Q-tests, Percentage Forecast Inaccuracy (PFI), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) etc. The smaller the values of MAE, RMSE and MAPE, the better the model is considered to be.

RESULTS AND DISCUSSION

ARIMA Model for the area, production and productivity of Banana

The exponential decay of significant spikes in ACF could be observed indicating that the series were nonstationary and should be transformed into a stationary one. The data corresponding to the area were made stationary after differencing the log-transformed data. Production and productivity of Banana have become stationary in the first difference alone. Augmented Dickey Fuller Test was also done, and the same result was observed.

Area under Banana

The best model identified for the prediction of the area under cultivation of Banana was ARIMA(0,1,0) with minimum Normalised Bayesian Information Criteria (BIC) and with highly significant coefficients. The model has a R^2 value of 96.6% with MAPE of 5.468, showing a good fit of the data.

Table 1. Parameter estimates for ARIMA (0,1,0) in the case of the area of Banana.

	Log transformed difference	**Estimate**	**SE**	**T**	**Sig.**
Constant	1	0.34	0.11	2.998	0.005

Table 2. Forecast for the area under Banana in the next six years using ARIMA (0,1,0).

2019-20	2020-21	2021-22	2022-23.	2023-24.	2024-25
46559	42422	39988	38998	39348	41074

The final model could be written in the form

$$A_t = e^{0.034}*(A_{t-1})^2/A_{t-2} \text{ ……..........…........……… (1)}$$

where A_t denotes the area of Banana in the year 't'.

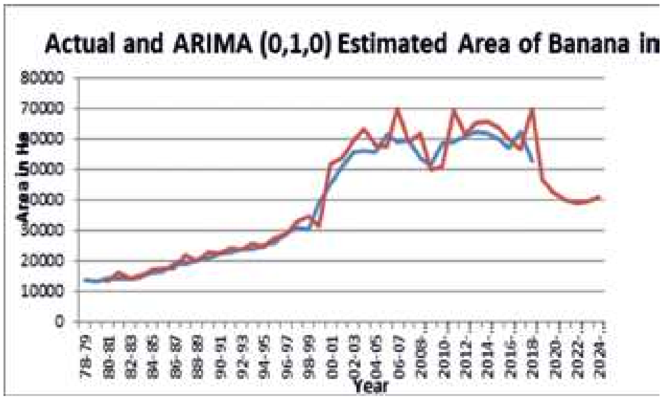

Fig1. Actual(Blue) and Estimated(red) area under banana cultivation in Kerala

The forecast for the area under Banana in the next six years could be given as in table 2. The forecasts show that there is a chance for a reduction in area under Banana in the state if new plans are not coming under the agricultural sector. The reasons may be due to the use of the area for non-commercial purpose, productivity loss due to continuous cultivation, and high labour cost in the state.

Banana recorded high positive growth rates in area during 1998-99 and continued up to 2006-07. After it the cultivation of Banana sustained in that level. This may be due to the decrease of the area under areca nut and coconut due to price fall. However, the rate of increase has shown a decreasing trend now.

Production of Banana

The best model identified for the prediction of banana production was ARIMA(0,1,0) with minimum Normalised Bayesian Information Criteria (BIC) and with highly significant coefficients. The model has an R^2 value of 89.5% with MAPE of6.983 showing a good fit of the data.

The final model could be written in the form

$$Y_t = Y_{t-1} + 6313.875 \text{ .... (2)}$$

where Y_t denotes the production of Banana in the year 't'.

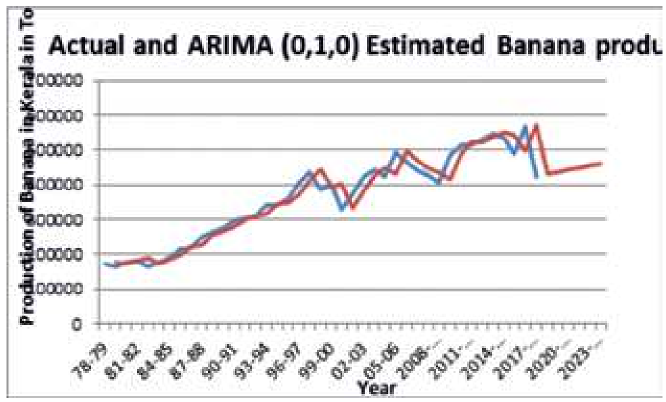

Fig2. Actual(Blue) and Estimated(red) Banana produced in Kerala

Though disease and drought have wrecked the farmers with production and productivity dropping to alarming levels, the chart showed an increased production of Banana in the state. Even if the 2018-19 figures is the lowest in the rent figures, the forecasts show an increase in production in the coming years.

Table 3. Parameters estimated for ARIMA (0,1,0) in the case of production of Banana

	Difference	Estimate	SE	T	Sig.
Constant	1	6313.875	6208.350	1.017	0.315

Table 4. Forecast for banana production in the next six years using ARIMA (0,1,0)

2019-20	2020-21	2021-22	2022-23.	2023-24.	2024-25
430362	436676	442990	449304	455617	461931

Table 5. Parameters estimated for ARIMA (0,1,0) in the case of productivity of Banana

	Differenced after Log Transform	Estimate	SE	T	Sig.
Constant	1	0.076	0.015	5.087	0.0001

Table 6. Forecast for price of Banana in the next six years using ARIMA (0,1,0)

2019-20	2020-21	2021-22	2022-23.	2023-24.	2024-25
4591	4978	5397	5851	6344	6878

Price of Banana

The final model with an $R^2 = 0.98$ and MAPE = 7.551% could be written in the form

$$P_t = 1.078963.e^{(P_{t-1})} \quad ....(3)$$

where P_t denotes the price of Banana in the year 't'.

The R^2 indicates that 98% of the variation in price could be predicted through the model.

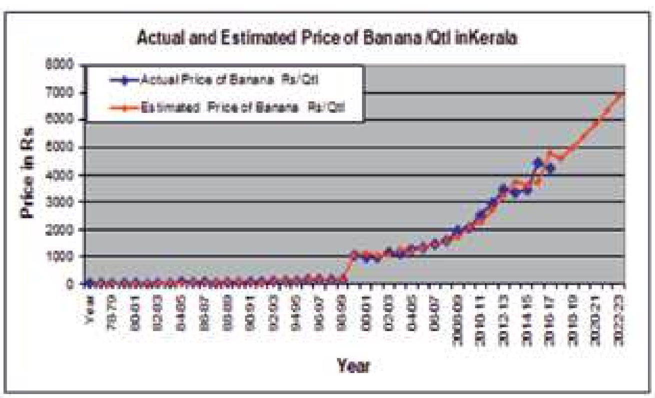

Fig3. Actual(Blue) and Estimated(red) price of banana cultivation in Kerala

Since the price shows an increasing trend in the coming years, it gives a good hope to the farmers of banana cultivation. The cause for this increase may be the decrease in the area of cultivation and less production increase the demand.

The residual ACF and PACF showed that the errors are white noise series and hence the models are perfect for forecasting the time series data of area, production and price of Banana in Kerala.

CONCLUSION

The forecasting power of ARIMA model was used to forecast for five leading years, and results showed a good agreement between actual and predicted values. Since the area under cultivation is showing a decreasing trend and production shows an increasing trend, it shows only the real interested farmers are doing banana cultivation in recent years. The low productivity may be expected due to the continuous cultivation of Banana in the same plot and hence lack of soil fertility. The forecast shows an increase in price, which may be by a decrease in area under cultivation will surely motivate Banana cultivating farmers in Kerala.

REFERENCES

Box G E P and Jenkins G M (1970). *Time Series Analysis: Forecasting and Control*, SanFrancisco: Holden-Day.

Box G E P, Jenkins G M and Reinsel G C (1994) *Time Series Analysis, Forecasting and Control*. 3rd edition, Prentice Hall, Englewood Clifs.

Carpio CEBS (2002) *Production Response Of Cotton In India, Pakistan, And Australia* [online], Msc Thesis In Agricultural And Applied Economics, Texas Tech University Available:http://etd.lib.ttu.edu/theses/available/etd-07312008-31295018555515 /unrestricted/31295018555515.pdf [27-02-09].

Engle R F and Granger C W J (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55(2): 251-76.

Krishnamoorthy V and Hanif Noorjahan A KA (2017). Influence of Micronutrients on Growth and Yield of Banana. *J Krishi Vigyan* **5**(2) : 87-89.

Mandal B N (2004). *Forecasting Sugarcane Production In India With ARIMA Model* (online), New Delhi [19-02-09].

Narayanaswamy T, Surendra H S and Rathod S (2012). Fitting of statistical models for growth patterns of root and shoot morphological traits in Sesame (*Seasamum indicum* L.). *Environ Ecol* **30(4)**: 1362-1365.

Naveena K, Rathod S, Shukla G and Yogish K J (2014). Forecasting of coconut production in India: a suitable time series model. *Int J Agril Eng* 7(1):190-193.

Nochai R and Nochai T (2006). *ARIMA Model for Forecasting Oil Palm Price.* 2nd IMT GT Regional Conference on Mathematics, Statistics and Applications,[online] Univ. Sains Malaysia, 13-15 June, 2006.

Pardhi R, Singh R, Rathod S and Singh P K (2016). Effect of price of other seasonal fruits on mango price in Uttar Pradesh. *Econ Affairs* **61**(4):1-5.

Rathod S, Singh, K N, Paul R K, Meher R K, Mishra G C, Gurung B, Ray M and Sinha K (2017). An Improved ARFIMA Model using Maximum Overlap Discrete Wavelet Transform (MODWT) and ANN for Forecasting Agricultural Commodity Price. *J Ind Soc Agric Stat* **71**(2): 103– 111.

Saeed N (2000). *Forecasting of Wheat Production in Pakistan using Arima Models. Int J Agri Biol* 2(4):352-353.

Sarika, Iquebal M A and Chattopadhyay C (2011). Modelling and forecasting of pigeon pea (*Cajanus cajan*) production using autoregressive integrated moving average methodology. *Indian J Agric Sci* **81**(6), 520-523.

Saxena M, Rathore, R P S, Gupta R P, Bhargav H and Thakur B (2017). Horticultural Statistics at a Glance 2017, Horticulture Statistics division, Department of Agriculture, cooperation & farmers welfare, Ministry of agriculture & farmers welfare, Government of India.

Sen L K, Shitan M and Hussain H (2000). *Time Series Modelling and Forecasting of Sarawak Black Banana Price*, Malaysia, November 2000.

Siji S and Nandini P V (2017). Chemical and nutrient composition of selected banana varieties of Kerala. *Int J Adv Eng, Manage and Sci* **3**(4): 401-404.

Suresh K K, Krishna, Priya S R (2011). Forecasting sugarcane yield of Tamil Nadu using ARIMA Models. *Sugar Tech* **13**(1):23-26.

Unnikrishnan T and Ajitha T K (2010). *Application of ARIMA models and Co-integration Techniques in Kerala Agriculture*, Proceedings of the International Seminar on Applied Statistics, Maharajas College, Ernakulam.

Unnikrishnann T (2009). *Changing Scenario of Kerala Agriculture – An overview*. M.Sc. Thesis, Kerala Agricultural University.

Weiss D (1992). International price relationships in the light of cointegration models: an introduction to the method from the example of the 'law of one price' for agricultural raw materials. Agrarokonomische Diskussionsbeitrage Universitat Giessen. (16): 39p.

Received on 28/08/2020 *Accepted on 05/11/2020*