SARIMA Model for Natural Rubber production in India

P. Arumugam*, V. Anithakumari #

* Reader Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli-627012, India
# Research Scholar Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli-627012, India

***Abstract*— In this paper an attempt is made to forecast the production of natural rubber in India by using monthly data for the period from January 1991 to December 2012. The comparative study of four different types of univariate time series models such as Linear Trend Equation, Additive Decomposition Model, Winter Seasonal Exponential Smoothing Model and Seasonal ARIMA Models were discussed. Mean Absolute Error (MAE), Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) were used as the selection criteria to determine the best forecasting model. Seasonal ARIMA (2, 1, 2) $(1, 1, 1)_{12}$ model for identification, parameter, estimation, diagnostic checking and forecasting future production. This study revealed that the time series data were influenced by a positive linear trend factor. It is therefore suggested that Additive decomposition could be used for forecasting the natural Rubber Production in India. The forecasting method for production of natural rubber production in India, as shown in this paper, can be a very useful tool for the Indian Rubber Industry Professionals and policy makers in India..**

***Keywords* — Decomposition, Exponential Smoothing, MAE, MAPE, SARIMA, Model Selection Criteria.**

I. INTRODUCTION

Rubber is an amorphous, elastic material obtained from the latex or sap or various tropical plants. There are two types of rubber-natural and synthetic. India is the fourth largest producer of natural rubber in the world and it is the fifth largest consumer. Though India is one of the leading producers of rubber it still imports rubber from other countries. Rubber-producing areas in India are divided into two zones-traditional and non-traditional. Under traditional zone we have Kanyakumari in Tamil Nadu and some districts of Kerala whereas under non-traditional zone. Rubber is produced in coastal regions of Karnataka, Goa, Andhra Pradesh, Orissa, some areas of Kankan regions of Maharashtra, Tripura, and Andaman and Nicobar Islands. Almost 95 per cent of India's total production of natural rubber comes from Kerala. Further, Kerala and Tamil Nadu together occupy 86 per cent of the growing areas of natural rubber. It is almost 90 per cent of the plantations in India that are fragmented and held by small and marginal cultivators. A few of the plantations are held by cooperatives and companies. Major corporations include A.V.Thomas Group and Harrison Malayalam. The main route of sales is through dealers and though the small growers sell through rubber-marketing cooperatives. Major consumption of rubber in India is in automotive tyre sector followed by bicycle tyres and tubes, foot wears, belts 50 per cent of rubber production is consumed in the tyre Industry whereas almost 15 per cent is for bicycle tyres and tubes, hoses, camelback, and latex products, etc. Slightly more than Kerala is the leading consumer of rubber followed by Punjab and Maharashtra.

The Indian rubber industry plays a very important role in the economy. The rubber plantation produces over 63 million tones of natural rubber in India and is expected to increase further in near future. The prospects of growth are further enhanced by a boom in vehicle industry and opening of Indian economy, resulting in improved standard of living of the people and rapid growth in industrialization. Natural rubber accounts for about 50 per cent of the production cost of a tyre manufacturer. There is a trend towards increase in the price of natural rubber which is reducing the profit margin of tyre manufacturers. This has resulted in their shifting to synthetic rubber as an alternative raw material. The ratio of natural rubber to synthetic rubber, 2000-07, has changed from 79:21 to 74:26. As per some of the tyre manufacturers, the use of synthetic rubber in tyre manufacturing may have to be increased to 30 per cent, if the price of natural rubber continues to (www.industrialrubbergoods. com). Domestic rubber prices depend upon the rate of import duty and the trend of international rubber prices. International rubber prices depend upon the production of rubber in Thailand and Malaysia, which accounts for over 80 per cent of global natural rubber production, and also the demand in the USA and China, the major markets for rubber.

II. REVIEW OF LITERATURE

There is hardly any literature available on forecasting production of natural rubber. However, a lot of work has been done on forecasting using various techniques. A number of studies are available where forecasting exercise is carried out by extrapolating long historical data using auto-regressive integrated moving average (ARIMA) method. The work of Chinye and Mesike [5] make use of ARIMA model for short term forecasting of Nigerian natural rubber export. Mad nasir shamsudin and Fatimahmohd [6] developed ARIMA model for short term forecasting the natural rubber prices. The monthly data for April 1991 to December 2000 were used to develop the model whereas forecasts were obtained for the period January 2001 to December 2001. The accuracy of expost forecast was measured through mean absolute percentage error (MAPE) and Thiel's U-statistic. Estimated ARIMA model for tea production using monthly data for

India from January 1979 through July 1991 and use the same to forecast tea production in India for the next twelve months.

Badmus and Ariyo [6][7] have conducted a forecasting study for inflation, industrial output, and exchange rate for India. The analysis was based on linear models, ARIMA, and bivariate transfer functions and restricted VAR. On the basis of root mean square error as a measure of forecasting accuracy, it is found that bivariate models do better than ARIMA for weekly data while for monthly data, ARIMA does a better job. The overall results support the use of a restricted VAR based on root mean square error. Badmus M.A and Ariyo.O.S [7] forecasting maize production using univariate time series models using ARIMA model and forecast the production in future year.

III. METHODOLOGIES

A. Trend Method

One of the best ways of obtaining trend values is the method of least squares. With this method a straight line trend is obtained. This line is called the line of the best fit. A procedure for estimating the parameter of any linear model, the method of least squares, can be illustrated simply by employing it to fit a straight line to a set of data points. In the trend method of forecasting, we shall use depersonalized data on rubber production as presented in this method, we treat

$$T = \alpha + \beta_t$$

$$\beta = \frac{n\Sigma ty - (\Sigma y)(\Sigma t)}{n\Sigma t^2 - (\Sigma t)^2}$$

$$\alpha = \left(\frac{\Sigma y}{n}\right) - b\left(\frac{\Sigma t}{n}\right)$$

Where

T= Trend, α= Intercept, β= Slope, t= Time period which take a value of 1 for January 1991, 2 for February 1991, and so on, and lastly 264 for December 2012.The equation to the model the trend is T=60596+0.015t Various trend equations were estimated using ordinary least square method. In this paper only the result for linear trend. In both the trend equations, Using linear trend equation, forecast for January 2013,February 2013 and so on for December 2013 are obtained by substituting t= 265, t=266, . . . t=276 respectively. It is pertinent to remind that these forecasts are the de-seasonal zed forecast of production of natural rubber.

An estimate of the Mean absolute Error (MAE) was found to be 0.489.Mean absolute Percentage Error (MAPE) was to be 0.088. This was done to examine the accuracy of the forecast for those periods for which actual production was available. The forecast were extended for the period January 2013 to December 2013 in Table1. A comparison of the model selection criteria in Table2 indicates that the accuracy of forecast obtained through trend equation is more than that through linear trend equation.

B. Additive Decomposition Model

The additive decomposition method is appropriate for modelling time series data containing trend, seasonal and error components, if we can assume the following. We have an additive model $X = T_t + S_t + C_t + e_t$, the error terms are random, and the seasonal components for any one season are the same in each year.

1) The quarterly rubber production for the year 1991 through 2012. Clearly, production of the rubber is seasonal. There also appears to be an upward trend in the data. We assume that we have an additive model and use this data to explain and demonstrate the following steps in the additive decomposition method. For the actual time series, compute a cantered moving average of length L(where L is the number of season in a year). To accomplish this, centered moving averages (two period moving averages of the initial moving averages) are computed.

Centred moving Average $(CMA) = Trend + Cycle$

2) Subtract the CMA_t from the data. The difference is equal to $S_t + e_t$.

3) Remove the error (e_t) component from $S_t + e_t$ by computing the average for each of the seasons. That is four estimates for the seasonal components.

Q1 = -186682, Q2= -183508, Q3 = - 178251, Q4= -176663

4) These average seasonal estimates should add up zero. If they do not we must adjust them so that they will be final adjustment consists of subtracting a constant $\left(\sum S_n / L\right)$ from each estimate.

$$(\Sigma S_n / L) = -181276$$

The final seasonal estimates are

$S_1 = 5406$, $S_2 = 2232$, S3 = 3025, S4 = 4614

5) Deseasonlize the data by subtracting from it their proper seasonal estimate $d_i = x_i - S_n$

d_1 = 1st actual production – S1 = 1 st deseasonlized value

6) Perform the proper regression analysis on the deseasonlized data to obtain the appropriate model for the data in a linear model. The equation to the model and trend is T_t = 60596+0.015t.An estimate or forecast for any time period can be found by adding together the estimate for the various components. The procedure was to add the appropriate estimate for the various components together. Thus the forecast constructing for the first quarter of 2013 January should be Table 3.

$$X_{265} = T_{265} + S_{265}$$
$$T_{265} = 60861$$
$$S_{265} = S_1$$
$$X_{265} = 66267$$

C. Winter Seasonal Exponential Smoothing Method

Winter Seasonal Exponential Smoothing is discussed modelling with Decomposition method, for modelling seasonal data. This method is efficient when the seasonal patterns are constant year after year and the necessary computations for updating with new data are not a problem. However, if the trend or seasonal patterns are included into the time series forecasts. The winter seasonal exponential smoothing technique employs the smoothing process three times.

1) to estimate the average level (level)of the series
2) to estimate the slope component (slope) of the series
3) to estimate the seasonal component (season) of the series Winter Additive Seasonal Exponential smoothing technique forecast a time series that has a linear trend and additive seasonal variation. The initial estimates of the parameters that are updated are usually obtained from the additive decomposition model. The winter methodology uses the following steps:

- The additive model containing linear trend is represented as
$$X_t = T_t + S_t + e_t$$
Where $T = a + b(t)$
- The basic concept of in exponential smoothing is Estimate = constant *(actual data) +(1-constant)*(old Estimate)
- The final forecast value is given by Forecast=(level estimate)+(slopeestimate)+(seasonal Estimate)

The level, slope and seasonal estimate have been smoothed a one step ahead forecast is obtained with the following equation

$$X_{t+1}(t) = [a_t + b_t] + S_{t+1}(t+1-L)$$

Where $X_{t+1}(t)$ =the forecast for the next time period t+1

a_t = the smoothed estimate for the level at time period t

b_t = the smoothed estimate for the slope at time t

$S_{t+1}(t+1-L)$ = the smoothed estimate for the t-1 season made at time t+1 We used additive decomposition to model the data for quarterly production. Using these result as initial estimates for the winter additive method and values of 0.01,0.02 and 0.05 for α, β and γ ,then length value L=4 we can smooth the estimate and obtain one period ahead forecast for each of the 264 time period.

Forecast for period 1.

$$X_1(0) = [a_0 + b_0] + S_t$$

Where a_0, b_0 and S_1 are the initial estimates from the decomposition analysis

Updated estimates in period 1

Level to be used in the second time period forecast

$$a_1 = \alpha(x_1 - S_1(0)) + (1-\alpha)[a_0 + b_0]$$

Slope to be used in the second time period forecast

$$b_1 = \beta(a_1 - a_0) + (1-\beta)b_0$$

Seasonal components to be used in the second time period forecast

$$S_2(t) = \gamma(x_i - a_i) + (1-\gamma)S_2(0)$$

The one step ahead forecast for the first quarter of the 2013 year (265 time period) can be calculated as

$$X_{t+1}(t) = [a_t + b_t] + S_{t+1}(t+1-L)$$
$$X_{264-1}(264) = [a_{264} + b_{264}] + S_{264-1}(264+1-4)$$

The procedure was to add the appropriate estimate for the various components together. Thus the forecast constructing for the Additive Winter seasonal Exponential smoothing factor of 2013 January should be Table 3.

D. Seasonal ARIMA Model

One of time series models which is popular and mostly used is ARIMA model. Based on assumption that the component is deterministic and independent from other non seasonal components. The non seasonal components may be stochastic and correlated with non seasonal components. We get the following well known Box Jenkins multiplicative seasonal ARIMA model. ARIMA (p, d ,q)(P,D,Q)$_s$ seasonal ARIMA model are defined by seven parameter ARIMA (p ,d ,q)(P,D,Q)$_s$.

$$\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B_p\right)\left(1-\beta_1 B^s-\beta_2 B^{2s}-\ldots-\beta_p B^{ps}\right)(1-B)^d\left(1-B^s\right)^D$$
$$Y_t = C + \left(1-\Psi_1 B-\Psi_2 B^2-\ldots-\Psi_q B^q\right)\left(1-\theta_1 B^s-\theta_2 B^{2s}-\ldots-\theta_Q B^{Qs}\right)\varepsilon_t$$

Where

$AR(p)$ - Autoregressive part of order p ,

$$\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B_p\right)$$

$AR_s(p)$ - Seasonal Autoregressive part of order p ,

$$\left(1-\beta_1 B^s-\beta_2 B^{2s}-\ldots-\beta_p B^{ps}\right)$$

$I(d)$ – differencing order d, $(1-B)^d$

$I_s(d)$ I_s(d) – seasonal differencing order d , $(1-B^s)^D$

$MA(q)$ - Moving average part of order q ,

$$\left(1-\Psi_1 B-\Psi_2 B^2-\ldots-\Psi_q B^q\right)_t$$

$MA_s(q)$ - Moving average part of order q ,

$$\left(1-\theta_1 B^s-\theta_2 B^{2s}-\ldots-\theta_Q B^{Qs}\right)\varepsilon_t$$

S is the period of the seasonal pattern appearing, i.e s=12,ε_t is a error term. The rubber production data in a seasonal time series .Let as consider the ARIMA (2,1,2) $(1,1,1)_{12}$ model

$$\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B_p\right)\left(1-\beta_1 B^s-\beta_2 B^{2s}-\ldots-\beta_p B^{ps}\right)\left(1-B\right)^d\left(1-B^s\right)^D$$

That is, s=12

$$\left(1-\phi_1 B-\phi_2 B^2\right)\left(1-\beta_1 B^{12}\right)\left(1-B\right)\left(1-B^{12}\right)Y_t=c+\left(1-\Psi_1 B-\Psi_2 B^2\right)\left(1-\theta_2 B^{24}\right)$$

The fitted ARIMA Model is given by

$$(1-0.521)(1-0.350)^{12}(1-B)(1-B)^{12}Y_t=-2349.524e_t$$

E. Phillips perron Test:

The Phillips perron test is a unit root test. That is it is used in time series analysis to test the null hypothesis that a time series is integrated of order 1. It builds on the Dickey Fuller test of the null hypothesis. $\delta=0$ in $\Delta Y_t=\delta Y_{t-1}+U_t$ where Δ is the first difference operator. The Dickey Fuller test involves fitting the regression models.

$$\Delta Y_t=\rho Y_{t-1}+(\text{cons}\tan t, timetrend)+U_t$$

The Phillips perren test involves fitting the regression.

$$Y_i=\alpha+QY_{i+1}+\varepsilon_i$$

Where we may exclude the constant or include a trend term. There are two statistics Z_p & Z_r

$$Z_p=n(\hat{\rho}_{n-1})-\frac{1}{2}n^2\hat{\sigma}^2/s^2 n\left(\lambda^2 n-\hat{\gamma}_{0,n}\right)$$

$$Z_r=Squrt\left(\left(\hat{\gamma}_{0,n}/\lambda^2 n\right)\left(\hat{\rho}_{n-1}/\hat{\sigma}\right)\right)-\frac{1}{2}\left(\hat{\lambda}^2 n-\hat{\gamma}_{0,n}\right)\left(\left(\frac{1}{\lambda_n}\right)\left(\frac{n\hat{\sigma}}{s_n}\right)\right)$$

$$\hat{\gamma}_{i,n}=\frac{1}{n}\Sigma\sum_{i=j+1}^{n}\hat{u}_i\hat{u}_{t-j}$$

$$\hat{\lambda}^2 n=\hat{\gamma}_{0,n}+2\sum_{j=1}^{n}\left(\frac{1-j}{q+1}\right)\hat{\gamma}_{j,n}$$

$$s^2 n=\frac{1}{n-k}\sum_{j=1}^{q}\hat{u}^2 i$$

Where u_i is the OLS residuals, k is the number of covariate in the regression, q is the number of newly-west lag to use in calculating $\lambda^{\wedge 2}n$ and $\hat{\sigma}$ is the OLS standard error of $\hat{\rho}$. Time series Dickey – Fuller test involve fitting the regression model is given by

$$\gamma=\alpha+\rho Y_{t-1}+\gamma_t+\varepsilon_t$$

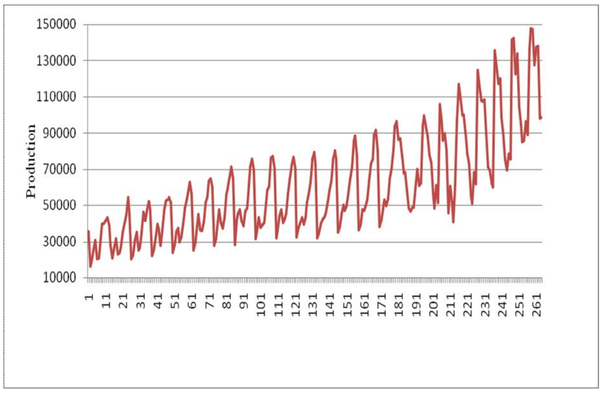

Fig. 1 Time series plot of actual rubber production in India

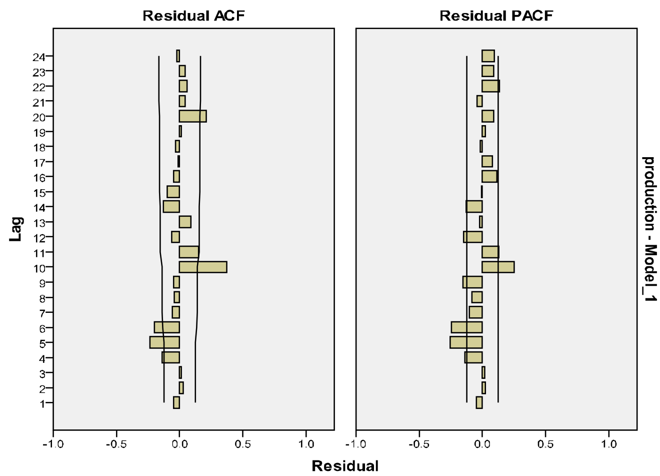

Fig. 2 The plot of Indian rubber production in residuals of ACF and PACF by Seasonal ARIMA (2 1 2) $(1\ 1\ 1)_{12}$ Method

TABLE 1
ACTUAL PRODUCTION AND FORECAST PRODUCTION BY SEASONAL ARIMA (2, 1, 2) $(1, 1, 1)_{12}$ METHOD

Fit Statistic	Mean
Stationary R-squared	.492
R-squared	.843
RMSE	11308.326
MAPE	9.725
MAE	6299.201
Normalized BIC	18.821

TABLE 2
FOUR MODEL SELECTION CRITERIA

Criteria	Trend	Additive Decomposition	Winter Exponential Smoothing	Seasonal ARIMA(2,1,2)$(1,1,1)_{12}$
MAE	0.488	0.000175	154.721	6299.201
MAPE	0.087	0.00031	0.0904	9.725
RMSE	1466	4020	28456	11308.326

TABLE 3
FORECASTING THE FOUR METHOD FOR 2013 JANUARY TO 2013 DECEMBER

Forecast	Time	Trend Method	Additive Decomposition	Winter Seasonal	Seasonal ARIMA (2,1,2) $(1,1,1)_{12}$
201301	265	60597	66006	60857	90069
201302	266	60597	62832	60858	91323
201303	267	60597	63625	60859	99681
201304	268	60597	65214	60860	111749
201305	269	60597	66006	60861	124079
201306	270	60597	62832	60862	133946
201307	271	60597	63625	60863	139795
201308	272	60597	65214	60864	141317
201309	273	60597	66006	60865	139237
201310	274	60597	62832	60866	134920
201311	275	60597	63625	60867	129916
201312	276	60597	65214	60868	125578

IV. RESULT AND DISCUSIONS

The final results of forecasting data obtained both approaches are compared with the test data. Subsequently, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAE) and Root Mean Square Error are working to evaluate the accuracy of the both models. Table2 report that the error of Additive decomposition method is smaller. It is implied that decomposition method a higher performance than Trend Model, Winter Exponential Smoothing and Seasonal ARIMA $(2,1,2)(1,1,1)_{12}$. The result of the study imply that Additive Decomposition forecasting method is a better alternative approach for predicting natural rubber production. Table 2 predict the future year of both models.

V. CONCLUSION

In this paper, we have presented efficient techniques to accurately predict time series data of natural rubber production. We have presented efficient techniques to accurately predict time series data of natural rubber production. The time series forecast based on comparison of four models across the two periods ahead in the forecast horizon. The accuracy of forecast of production of natural rubber as obtained by various methods is shown in Table 2 by presenting the absolute percentage error month wise for the period from January 2013 to December 2013. The mean absolute percentage as obtained by various methods is also presented. The mean absolute percentage errors are four models from Table2. Moreover, the Additive decomposition model developed for this study could be modified in term of learning rule, different training techniques, different of equation of decomposition method. Forecasting the future production of natural rubber through the most accurate univariate time series model can help the Indian government as well as the production is natural rubber industry to perform better strategic planning and also to help them in maximizing revenue and minimizing the natural rubber production.

Acknowledgment

The authors would like to thank the reference for given that very helpful comments and suggestions, their insight and comments led to a better presentation of the idea expressed in this paper. This research work was supported by the university Grant Commission, New Delhi.

REFERENCES

[1] Box, G.E.P,and G.M. Jenkins and G.C.Reinsel ,*Time series analysis Forecasting and* control, 4 th edition ,John Wiley and sons, Inc., New Jersey,1998.

[2] Bambang Widjanarko Otok and Suhartono, "Development of Rainfall Forecasting Model in Indonesia by using ASTAR, Transfer Function and ARIMA Methods," *European Journal of Scientific Research*, Vol.38 No. 3(2009),pp.386-395

[3] Walialak Atthir awong and Porntip Chatchaipun ,"Time series Analysis for Natural Field Latex Prices Prediction," *King Mongkut's Institute of Technology Ladkrabang*, Bangkok 10520, Thailand.

[4] A.A,Khin, Zainalabidin M.and Mad.Nasir.S," Comparative Forecasting Models Accuracy of Short- term Natural Rubber Prices," *Trends in Agricultural Economics4(1):1-17,2011*,ISSN 1994-7933/DOI: 10.3923/tae.2011.1.17.

[5] Chinye S. Mesike," Short term forecasting of Nigerian natural rubber export, "*Wudpecker journals of Agricultural Research* , Vol. 1(10), pp. 396-400,(2012).

[6] Mad nasir shamsudin and Fatimah mohd arshad,"Composite Model for Short Term Forecasting for Natural Rubber Prices" *Pertanika 13(2),283-288(1990),* 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia.

[7] Badmus M.A and Ariyo. O.S,"Forecasting cultivated Areas and production of Maize in Nigerian using ARIMA Model". *Asian Journal of Agricultural Science*, Vol.3 (3), Pp: 171-176, (2011).