RESEARCH PAPER International Journal of Agricultural Engineering | Volume 7 | Issue 1 | April, 2014 | 190–193

Forecasting of coconut production in India: A suitable time series model

K. NAVEENA, SANTOSHA RATHOD, GARIMA SHUKLA AND **K.J. YOGISH**

Received : 10.02.2014; **Revised :** 01.03.2014; **Accepted :** 14.03.2014

See end of the Paper for authors' affiliation

Correspondence to :

K. NAVEENA
Department of Farm Engineering, Institute of Agriculture Science, Banaras Hindu University, VARANASI (U.P.) INDIA

ABSTRACT : Coconut is an important and versatile tree crop with diverse end-uses, supporting livelihood of many farm households in the primary sector, grown in many states of India. The present study is an attempt to find an appropriate model to forecast the coconut production in India. Time series data for a period of 61 years from 1951-2012 were used. The best model has been selected based on the minimum root mean square error values. It has been found that ARIMA (1, 1, 1) model was found to be as an appropriate model to forecast the production. Further production of coconut in the next coming years would be also concluded in this study. The finding of this study will be helpful for farmers, coconut based industries and policy makers.

KEY WORDS : Coconut production, ARIMA, RMSE. Forecast, Ljung and Box Q statistic

HOW TO CITE THIS PAPER : Naveena, K., Rathod, Santosha, Shukla, Garima and Yogish, K.J. (2014). Forecasting of coconut production in India: A suitable time series model. *Internat. J. Agric. Engg.*, **7**(1) : 190-193.

Coconut (*Cocos nucifera* L.) is an important and versatile tree crop with diverse end-uses, supporting livelihood of many farm households in the primary sector, grown in many states of India. The classics of India have rightly eulogized coconut tree as "*Kalpavriksha*" or "the tree of heaven" which means "the tree that gives all that is necessary for living", owing to the multifarious uses of various parts of the palm and its products in our daily life. Coconut is mainly grown in tropical and sub-tropical countries. It forms an integral component of the socioeconomic and cultural lives of nearly 80 million people of the world in 92 countries. Major area located in Asia and the Pacific region. South East Asia is regarded as the place of origin of coconut. It is mainly cultivated in Philippines, Indonesia, India, Sri Lanka, Papua New Guinea, Thailand, Malaysia and Fiji.

The world annual production of coconut is 51 billion nuts from an area of 11.9 million hectares. More than 75 per cent of this is contributed by the four major player's *viz.*, India, Indonesia, Philippines and Sri Lanka. India had attained the top position in the production of coconut with an area of 1.895million hectares and a production of 14006.5 million nuts in the year 2011-12. The productivity is 6869 nuts per ha.

Despite India being a dominant producer of coconut she faces several challenges both at domestic and international front. Domestically the cost of production is too high and hence, in global market, Indian coconut products are price-wise less competitive. Another challenge is the threat of import of coconut products, a factor which causes decline in domestic production, so it is necessary to find the future of coconut production if existing trend continue in India. Several attempts have been made in the past to develop yield forecast models for various commodities. Jambhulkar (2013) studied the forecasting of rice production in Punjab using ARIMA Model. Boken (2000) have studied the forecasting of spring wheat yield.

METHODOLOGY

To forecast Indian coconut production, ARIMA model was employed. The yearly data from 1950-51 to 2011-12 on Indian coconut production collected from coconut board were used to forecast the future eight year production. The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. Box and Jenkins (1968) have studied ARIMA models extensively. The optimal forecast of future values of a time-series are determined by the stochastic model for that series. In its general form, the ARIMA model expressed as follows:

ARIMA (p, d, q) (P, D, Q) $_s$

p = order of non-seasonal auto regressive (AR), d = order of non-seasonal difference, q = order of non-seasonal moving average (MA), P = order of seasonal auto regressive (SAR) , D = order of seasonal difference, Q = order of seasonal moving average (MA),s = length of the season.

Auto regressive process (p, 0, 0):

If the observation $Y_{t\ \text{depends}}$ on its previous observation and error term e_t is called auto regressive process (AR process)

$$Y_t = \mu_1 + \text{¯}_1 Y_{t-1} + \text{¯}_2 Y_{t-2} + \ldots + \text{¯}_p Y_{t-p} + e_t \quad (2)$$

The values of auto regressive co-efficient restricted to lie between –1and + 1.

Moving average process (0, 0, q):

If the observation Y_t depends on the error term e_t and on one or more previous error terms (e_t's) then we have moving average (MA) process.

$$Y_t = (\mu_1 + e_{t-}\ \text{¯}_1 e_{t-1} + \text{¯}_2 e_{t-2} - \ldots + \text{¯}_p e_{t-q}) \quad (3)$$

where,

θ_i = i[th] moving average parameter, I = 1, 2… q, q= order moving average.

The values of the co-efficient are restricted to lies between –1 to 1.

Mixtures: ARIMA process:

If the non-stationary is added to a mixed ARMA process, then the general ARIMA(p,d,q) is implied.

$$(1 - B)^d (1 - c_p B^p)\, Y_t = C + (1 + \text{¯}_q B^q)\, e_t \quad (4)$$

where B is the backshift operator.

q And ϕ are the parameters, C is the intercept, e_t is the error term.

Diagnostic checking of the model:

After estimating the parameters of a tentatively identified ARIMA model, it is necessary to go for diagnostic check to confirm that the model is adequate. Ljung and Box (1978) 'Q' statistic was used to know whether the auto correlations for those residuals are significantly different from zero. This diagnostic checking helps us to spot the differences in the model, so that the model could be subjected to modification, if need to be.

Forecasting:

The ultimate test for any model is whether it can predict future events accurately or not. If the model is :

$$(1 - \text{¤}B)\, Y_t = (1 - \text{¯}_B)\, e_t \quad (5)$$

The above model gives the forecasting equation is

$$Y_t = \text{¤}Y_{t-1} + e_t - \text{¯}e_{t-1} \quad (6)$$

Root mean square error:

The accuracy of forecasts for both ex-ante and ex-post is tested using the following test (Makridakis and Hibbon, 1979).

$$\text{RMSE} = \left[\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)\right]^{1/2} \quad (7)$$

where, Y_t is actual price $\hat{Y}_t$ is the forecasted price.

Lesser the RMSE means the forecasted values from the model are closely matching the actual values. The model with accurate forecasts can be selected by using this criterion.

■ RESULTS AND DISCUSSION

For fitting the ARIMA Model the three stages of modeling as suggested by Box and Jenkins namely, identification, estimation and diagnostic checking was undertaken. Identification was done after examining the auto correlation function and the partial autocorrelation function. After that estimation of the model was done by the least square method. In the Diagnostic checking phase the model Ljung and Box Q Statistic was performed.

Fig. 1 shows that the series is not stationary because the mean of the time series is increasing with the increase in time. The auto correlation function of the time series in Fig. 2 shows that the series is not stationary because auto correlation co-efficients does not cut off to statistical insignificance fairly quickly. All the first 15 autocorrelations are significantly different from zero at about the 5% level: all the first 15 spikes in the acf extend beyond the square brackets. To make the series stationary it was first differenced.

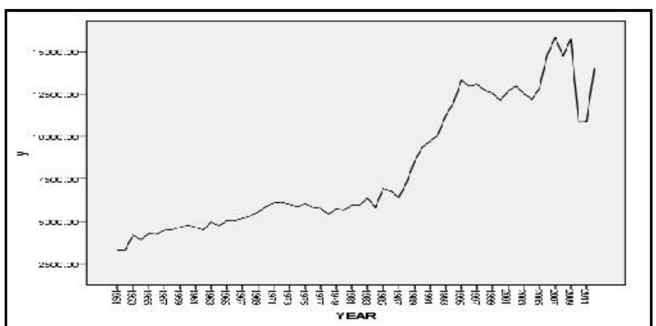

Fig. 1 : The time plot of coconut production in India

Fig. 3 shows the time plot of the differenced series and it clearly depicts that the series has now become mean stationary.

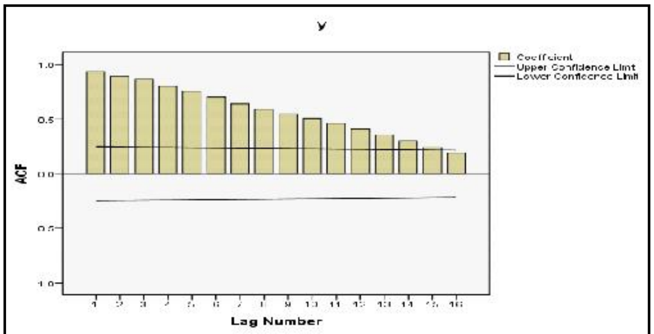

Fig. 2 : Autocorrelations at different lags of coconut production in India

The observations seem to fluctuate around a fixed mean, and the variance seems to be varying over time. However, the judgment about stationarity of the mean was withheld until the estimated acf and perhaps some estimated AR co-efficients were examined.

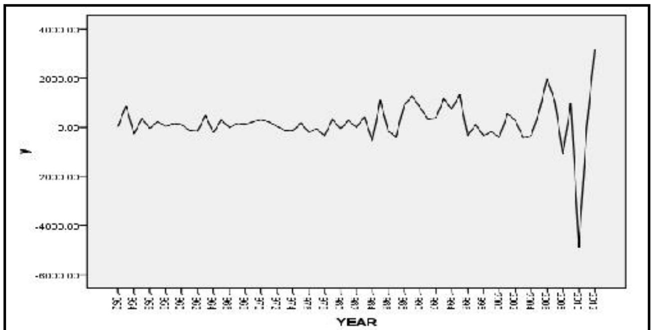

Fig. 3 : The time plot of the differenced series of coconut production in India

In Fig. 4 and 5 autocorrelation function and partial autocorrelation function of the differenced series are shown. The autocorrelations decay to statistical insignificance rather quickly. It was concluded that the mean of the series is stationary.

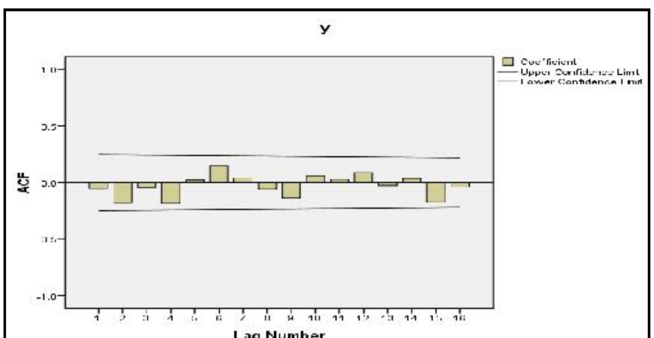

Fig. 4 : Autocorrelations at different lags of differenced time series of coconut production in India

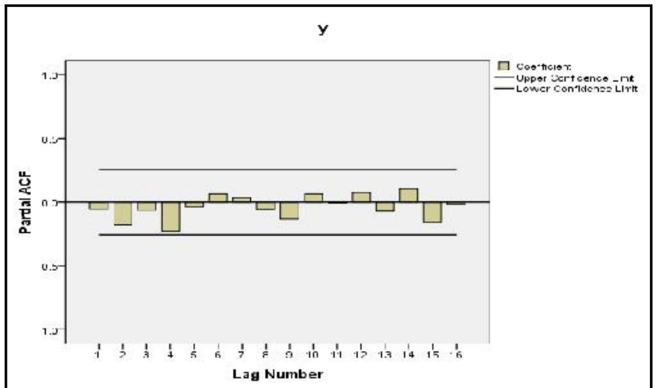

Fig. 5 : Partial autocorrelations at different lags of differenced time series

Once the time series has become stationary ARIMA model was estimated. After going through these stages ARIMA (1,1,1) model was found to be the best among the family of ARIMA models (Shafaqat Mehmood *et al.*, 2013.). ARIMA Model parameters and model fit statistics are given in Table 1.

Table 1 : ARIMA Model parameters and model fit statistics for coconut production in India

ARIMA model parameters		Estimate	SE	t-value	Sig (p-value)	Model fit statistics	
AR	Lag 1	0.832	0.141	5.886	.000	RMSE	907.86
Difference		1					
MA	Lag 1	1.000	31.429	0.032	0.975		

In Table 1 ARIMA (1,1,1) was found to be the best model the auto regressive (AR) co-efficient at lag 1 was found to be statistically significant at 1% level of significance, respectively with an estimate of 0.832. At the Diagnostic Checking Stage residual were examined and the autocorrelation co-efficients were found to be non-significant. Table 2 showing satisfactory model fitting.

Table 2 : Ljung-Box Q statistics for diagnostic checking

Model	Ljung-Box Q(18)		
	Statistics	DF	Sig.
Coconut-Model_1	9.378	16	0.897 ns

Forecasting:

Seven years' forecasts of coconut production are estimated using the best selected model and are presented in Table 3. Prediction intervals of forecast are also presented. Also graphical explanation of observed, fitted and forecasted value presented in Fig. 6.

Conclusion:

ARIMA (1,1,1) is the most appropriate model to forecast

Table 3 : Coconut production forecast in india

Model	2013	2014	2015	2016	2017	2018	2019	2020
Forecast	14100	14300	14500	14600	14800	15000	15100	15300
UCL	16000	16700	17200	17500	17900	18100	18400	18600
LCL	12300	11900	11700	11700	11700	11800	11900	12000

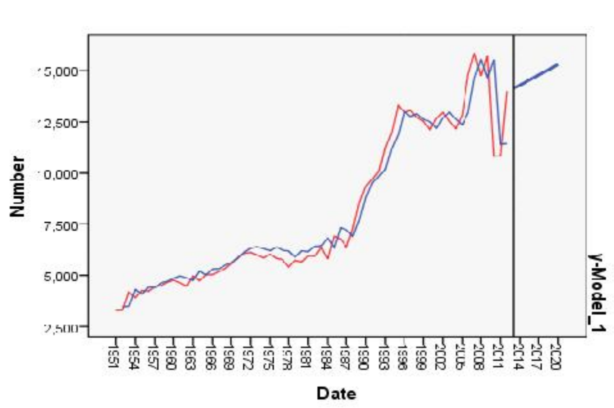

Fig. 6 : Observed and fitted time series for coconut production in India

production of coconut in India. Forecast indicates that production of coconut in the year 2020 is 1200 million nuts (8.51%) more than present production. It may because of increase in area, good management practice and increase in consumer demand. The finding of this study will be helpful for farmers, coconut based industries and policy makers.

Authors' affiliations:
SANTOSHA RATHOD, GARIMA SHUKLA AND **K.J. YOGISH,** Department of Farm Engineering, Institute of Agriculture Science, Banaras Hindu University, VARANASI (U.P.) INDIA

■ REFERENCES

Box, G.E.P. and Jenkins, G.M. (1968). Some recent advances in forecasting and control. *Appl. Statistics*, **17** : 91-109.

Boken, V.K. (2000). Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies. *Agron. J.*, **92**(6):1047-1053.

Jambhulkar, N.N. (2013). Forecasting of Rice Production in Punjab using ARIMA Model. *Internat. J. Scientific Res.*, **2** (8) : 1-2.

Ljung and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrica*, **65** : 297-304.

Makridakis, S. and Hibbon, M. (1979). Accuracy of forecasting: *An empirical investigation. J. Royal Statistical Society of Australia*, **142**(120): 97-145.

Shafaqat Mehmoodan Zahid Ahmad (2013). Time series model to forecast area of mangoes from Pakistan: An Application of Univariate Arima Model. *Academy of Contemporary Res. J.*, **7** (1): 1-7.