Journal of Pharmacognosy and Phytochemistry 2020; 9(2): 201-207

Journal of Pharmacognosy and Phytochemistry

Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
www.phytojournal.com
JPP 2020; 9(2): 201-207
Received: 16-01-2020
Accepted: 20-02-2020

Chinmayee Patra
Department of Agricultural Statistics, Institute of Agriculture Visva - Bharati, Shantiniketan West Bengal, India

Akhilesh Kumar Gupta
Department of Agricultural Statistics, Institute of Agriculture Visva - Bharati, Shantiniketan West Bengal, India

Ravi Ranjan Kumar
Department of Agricultural Statistics, Institute of Agriculture Visva - Bharati, Shantiniketan West Bengal, India

Corresponding Author
Chinmayee Patra
Department of Agricultural Statistics, Institute of Agriculture Visva - Bharati, Shantiniketan West Bengal, India

Forecasting of area, production and yield of jute crop in India using ARIMA model

Chinmayee Patra, Akhilesh Kumar Gupta and Ravi Ranjan Kumar

Abstract
The paper describes an empirical study on the trend of jute production, area & yield in India for the period starting from 1950 to 2017. As West Bengal contributes 80% of total jute production, so West Bengal is also taken for the study. For trend estimation Auto Regressive Integrated Moving Average (ARIMA) model was used. Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) were calculated for the data. Appropriate Box-Jenkins ARIMA model was fitted. Validity of the model was tested using Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE). For forecasting area, production and Yield ARIMA (2, 1, 1), (0, 1, 1) and (1, 1, 1) model for India and ARIMA (2, 1, 1), (1, 1, 2) & (2, 1, 1) for West Bengal were used respectively to forecast five leading years. The results also shows area forecast for the year 2022 to be about 778.4224 thousand hectare, production 10982.75 (In ' 000 Bales of 180 Kgs. Each) in India. Forecasted values of area to be 581.5554 thousand hectare and for production 8710.528(In ' 000 Bales of 180 Kgs. Each) in West Bengal.

Keywords: production, forecasting, ARIMA model

Introduction

Jute is an important traditional cash crop in India. It is known as golden fiber and 2nd important natural fibre crop in India next to cotton. In trade and industry, jute and mesta crop together known as raw jute as their uses are almost same. Raw jute plays an important role in the country's economy. Raw jute was originally considered as a source of raw material for packaging industries only. It plays a predominant role in the country's economy by generating employment, earning foreign exchange, solving many of the socio-economic problems, etc.

In India Jute is cultivated on 0.71 million hectares with an annual production of 9.98 million bales. India is the single largest producer of jute goods in the world, contributing about 60% of the global production. The domestic market continues to be the mainstay of industry consuming about 87% of the total production [1]. In India, It's grown on 0.686m.ha. Area and production is of 9.6 million bales (of 180kgs each). West Bengal is the most leading state of jute production in the country, followed by Bihar and Assam. In West Bengal, it is grown on 0.515m.ha. area, and production of 7.511million bales (of 180kgs each) [11]. About 90% of the jute farmers in India belong to the marginal and small categories of which almost 65% are with 1 hectare and below and 25% with 2 ha and below land holdings [1]. Jute is a bio-degradable crop grown mainly in the Ganges delta. West Bengal, Assam and Bihar are the major jute growing states in the country, which accounts for about 98 percent of the country's jute area and production (State of Indian Agriculture, 2016-2017). State of West Bengal occupies tops the list of jute production and contribute alone more than 80 per cent of the jute the country produces. Generally the crop (jute) is grown through-out the state except the hilly region of the north and the plateau area of the west. Murshidabad, West Dinajpur, Cooch Behar. Hugli. 24 Parganas (north and south), Nadia, Malda, Barddhaman, Jalpaiguri, Haora and Medinipur districts are the important producers. In West Bengal Corchorus capsularis (white jute) is grown in lowlands (Bills). A number of studies have been considered up by several authors in different forms to analyze the production behavior of jute with different objectives. Among the studies are works of importance [2-7].

Materials and Methods

Data collection

The time series data on area, production and yield of jute in West Bengal and India for the period of 72 years from 1950 to 2017 were collected from India Stat website for present study.

Analytical Method

Forecasting of area, production and productivity using ARIMA Model

The annual data on jute crop cultivated area, production and yield of India for the period from 1950-51 to 2016-17 were used for forecasting the future values using ARIMA models. The ARIMA methodology is also called as Box-Jenkins methodology. The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. The main objective in fitting this ARIMA model is to identify the stochastic process of the time series and predict the future values accurately. These methods have also been useful in many types of situation which involve the building of models for discrete time series and dynamic systems. But, this method was not good for lead times or for seasonal series with a large random component [8].

The first thing to note is that most time series are non-stationary and the ARIMA model refer only to a stationary time series. Since the ARIMA models refer only to a stationary time series, the first stage of Box-Jenkins model [9] is reducing non-stationary series to a stationary series by taking first order differences.

Box-jenkins auto regressive integrated moving average (ARIMA) models

Box-Jenkins methodology (Box and Jenkins of Time Series Analysis: Forecasting and Control) is used here for time series analysis which is technically known as the ARIMA methodology.

The ARIMA model includes

- The Autoregressive (AR) model.
- The Moving Average (MA) Model
- The ARMA Model

The Autoregressive (AR) Model

The Simplest form of the ARIMA model is called the autoregressive model. Let zt stand for the value of a stationary time series at time t, that is, a time series that has no trend, but fluctuates about a constant value referred to as the level of the series. (We deal with trends below.) By autoregressive, we assume that current zt values depend on past values from the same series. In symbols, at any t,

$$z_t = C + \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \dots\dots + \varphi_p z_{t-p} + \varepsilon_t$$

$$Z_t = C + \varepsilon_t + \sum_{i=1}^{p} \varphi_i z_{t-i}$$

Where C is the constant level, z_{t-1}, $zt_{-2\dots}$ z_{t-p} are past series values (lags), the ‘s are coefficients (similar to regression coefficients) to be estimated, and ε_t is a random variable with mean zero and constant variance. The ε_ts are assumed to be independent and represent random error. Some of the ‘s may be zero. If z_{t-p} is the furthest lag with a nonzero coefficient, the AR model is said to be of order p, denoted AR (p).

The moving average (MA) model

z_t can also be modeled as a linear combination of white noise stochastic error terms. We call this type of model a moving average (MA) model. If z_t is considered as a weighted average of the uncorrelated ε_t's, MA (q) moving average component of order q, which relates each z_t value to the residuals of the q previous z estimates may be expressed as

$$z_t = e_t - q_1 e_{t-1} - q_2 e_{t-2} - \dots - q_q e_{t-q}$$

The ARMA Model

The AR and MA models for stationary series to account for both past values and past shocks may be combined. Such a model is called an ARMA (p, q) model with p order AR terms and q order MA terms. Thus an ARMA (p, q) model is written as

$$z_t = C + \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \dots\dots + \varphi_p z_{t-p} + \varepsilon_t - q_1 e_{t-1} - q_2 e_{t-2} - \dots - q_q e_{t-q}$$

The main stages in setting up a Box-Jenkins forecasting model are Identification, Estimating the parameters, diagnostic checking and Forecasting.

In ARIMA modeling, the order of AR (p) is identified by partial autocorrelation function (PACF) while the order of MA (q) is identified by autocorrelation function (ACF) (Tsay, 2002). The order of ARIMA (p, d, q) is also identified by model selection criteria’s i.e. Schwarz Bayesian information criteria (SBIC) and Akaike’s Information Criteria (AIC). These criteria’s are further explained in model specification section.

Model specification

One of the important issues in time series forecasting is to specify model. Time series model is specified on the basis of some information criteria’s which includes AIC, BIC likelihood etc. Akaike’s (1973) introduced AIC criteria for model specification. AIC is mathematically defined as;

AIC= – 2log (maximum likelihood) + 2k

Where k = p+q+1 (if model includes intercept) otherwise k = p+q. model specified well if its AIC value is minimum as other fitted models (Tsay, 2005) [10].

Forecasting accuracy measuring techniques

After model selection, a next important step is to measure the accuracy to verify the reliability of forecasted value based selected model. Various tools are available in literature which includes Root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), mean error (ME) and mean percentage error (MPE). Further computation and literature of these accuracy measuring tools are given

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} |PE_t|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_t^2$$

$$PE_t = \left(\frac{Y_t - F_t}{Y_t}\right) \times 100$$

Where

Y_t the present is value for time t and F_t is the forecasted value for time t.

Results and Discussion

In this study, we used the data for jute crop cultivated area, production and yield for the period 1950-51 to 2016-17. As we have earlier stated that development of ARIMA model for any variable involves four steps: Identification, Estimation, Verification and Forecasting. Each of these four steps is now explained for jute crop cultivated area, production and yield as follows.

Model identification and validation

For forecasting jute crop area, production and yield, ARIMA model estimated only after transforming the variable under forecasting into a stationary series. The stationary series is the one whose values vary over time only around a constant mean and constant variance. There are several ways to ascertain this. The most common method is to check stationary through Augmented Dickey-Fuller test of the data. As result of Augmented Dickey-Fuller test the production, yield and area data for both India and west Bengal is not stationary, so one differencing is required to make the data stationary.

The next step is to identify the values of p and q. The Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) show that the order of p and q can. Basing on the p, d, q values from the ACF and PACF graph (figure 1,2,3,4,5 & 6) we entertained three tentative ARIMA models for area, production and yield and chose that model which has minimum AIC (Akaike Information Criterion), RMSE (Root Mean Square Error) & MAPE. The models and corresponding AIC, RMSE & MAPE values are given in Table 1.

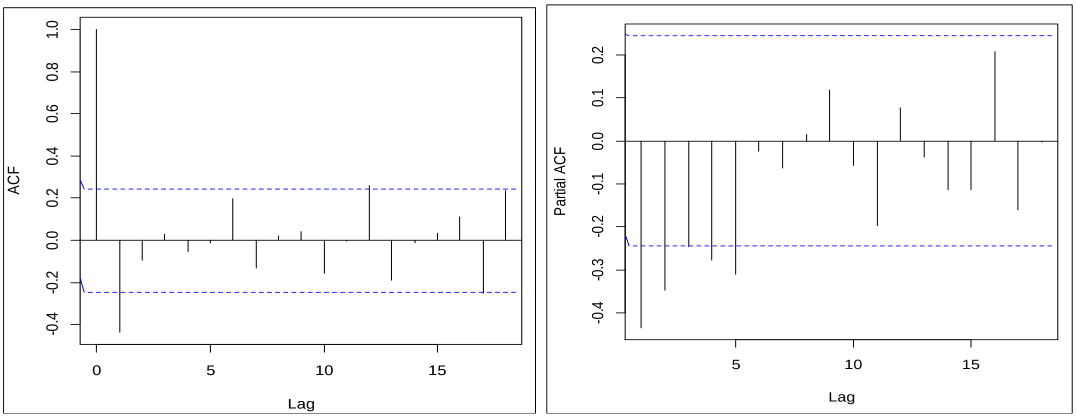


Fig 1: ACF & PCF graph for Production of India

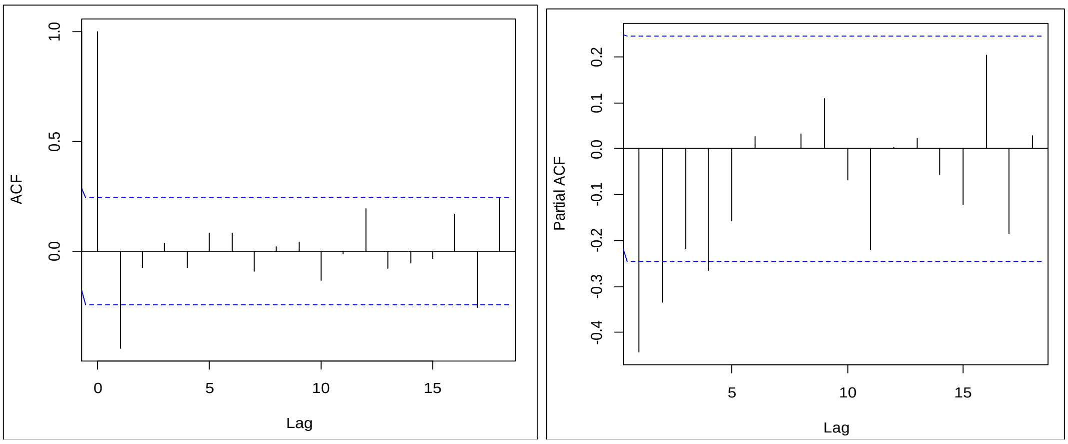


Fig 2: ACF & PCF graph for Production of West Bengal

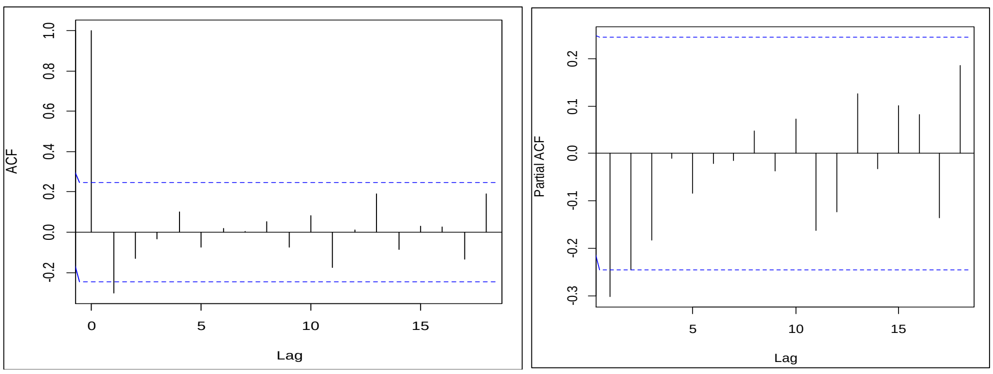


Fig 3: ACF & PCF graph for Yield of West Bengal

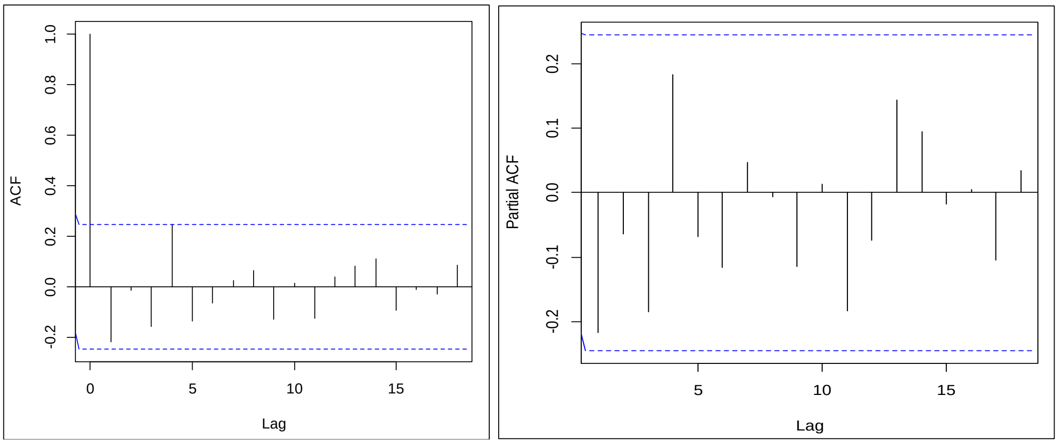


Fig 4: ACF & PCF graph for Yield of India

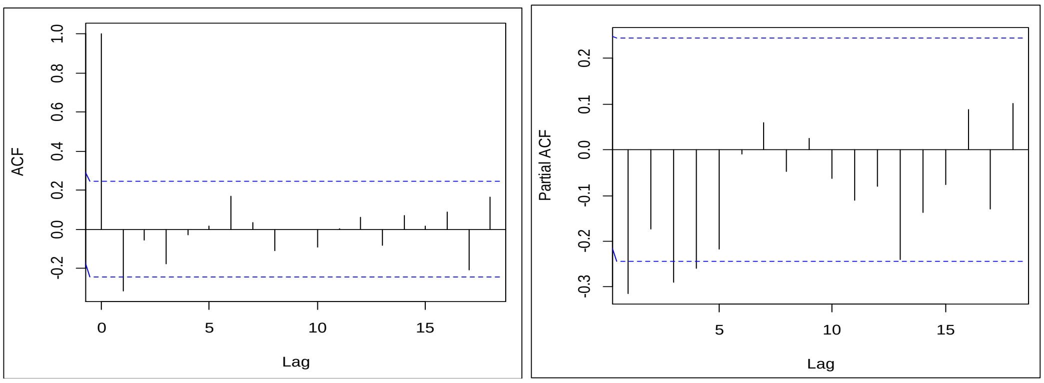


Fig 5: ACF & PCF graph for Area of West Bengal

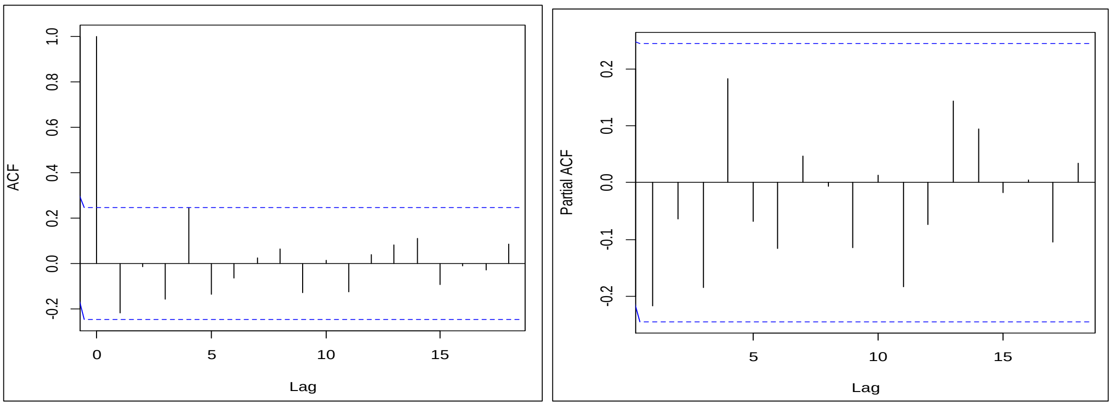


Fig 6: ACF & PCF graph for Area of India

Table 1: ARIMA models with AIC, MAPE & MAPE values

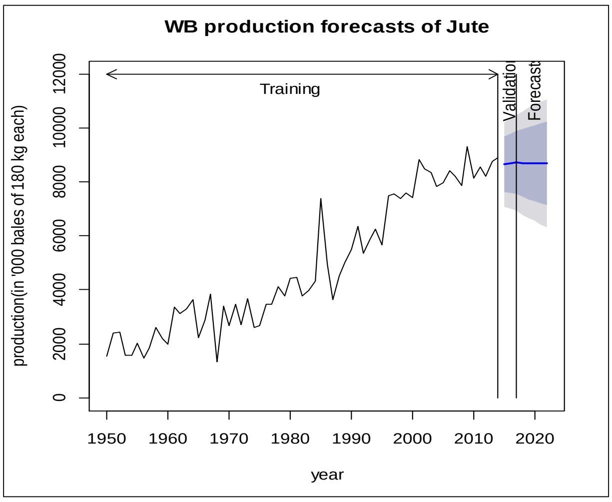
	ARIMA (p,d,q)	AIC	Rmse	Mape
Jute production in India	1,1,1	1080.56	312.6004	2.211118
	2,1,1	1079.64	312.2719	2.226247
	3,1,1	1082.24	316.4761	2.309968
Jute production in West Bengal	0,1,1	1045.60	749.1772	15.70118
	2,1,1	1035.38	739.4791	15.65608
	2,1,2	1042.17	756.0563	15.26028
Jute Yield in India	0,1,1	820.58	141.5322	6.994562
	1,1,1	822.57	141.519	7.00356
	0,1,0	820.87	144.127	7.169288
Jute Yield in West Bengal	1,1,1	826.05	143.6484	15.70118
	2,1,1	825.95	142.915	15.65608
	1,1,2	824.57	140.5157	15.26028
Jute Area in India	1,0,1	787.38	97.02904	9.687456
	0,1,1	777.49	100.2958	9.652849
	1,1,1	775.28	96.92478	9.192152
Jute Area in West Bengal	1,1,1	729.71	68.27193	11.63042
	2,1,1	713.30	68.20373	11.30175
	3,1,1	731.46	67.00694	11.65150

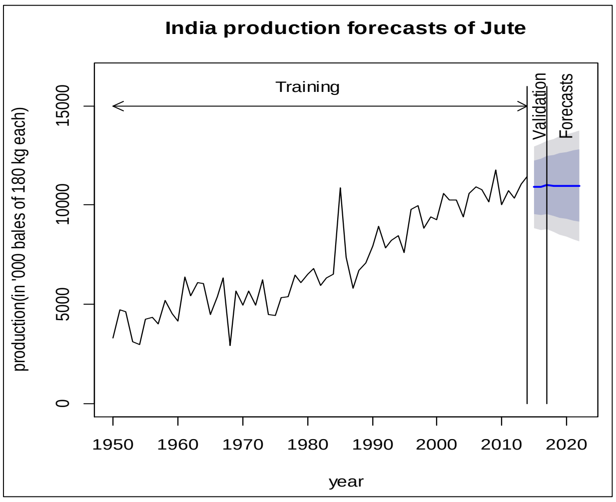
From this table, most suitable models are ARIMA (2, 1, 1) for jute production in India & ARIMA (2, 1, 1) for jute production in West Bengal has the lowest AIC and RMSE values, ARIMA (0,1,1) for jute yield in India & ARIMA (1,1,2) for jute yield in West Bengal has lowest AIC & MAPE values and ARIMA (1,1,1) for jute area in India & ARIMA (2,1,1) for jute area in West Bengal has lowest AIC & MAPE values selected for forecasting future values.

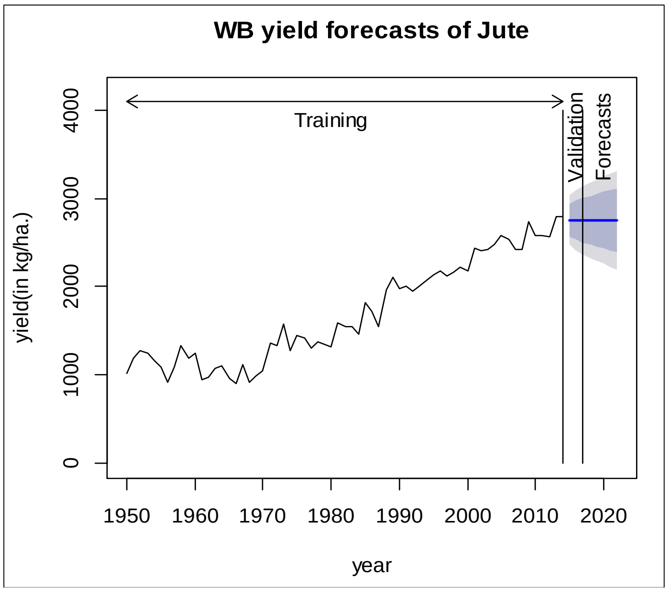
For the validation of model last 5years of data has taken. ARIMA (2,1,1) is the best model for forecasting the Production data in both the case of India and West Bengal, ARIMA (0,1,1) & ARIMA (1,1,2) is used for forecasting yield, ARIMA (1,1,1) & ARIMA (2,1,1) for forecasting area for India and West Bengal respectively as they have lowest MAPE.

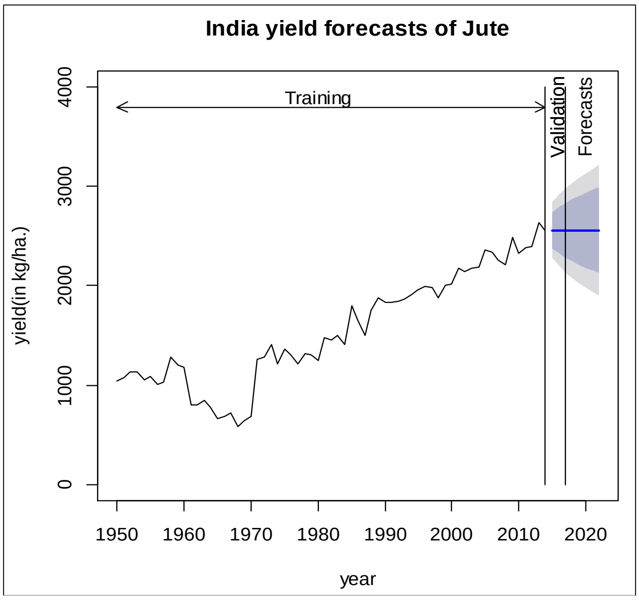
Table 2: Forecasted values of jute production, area and yield with 95% Confidence Level (CL)

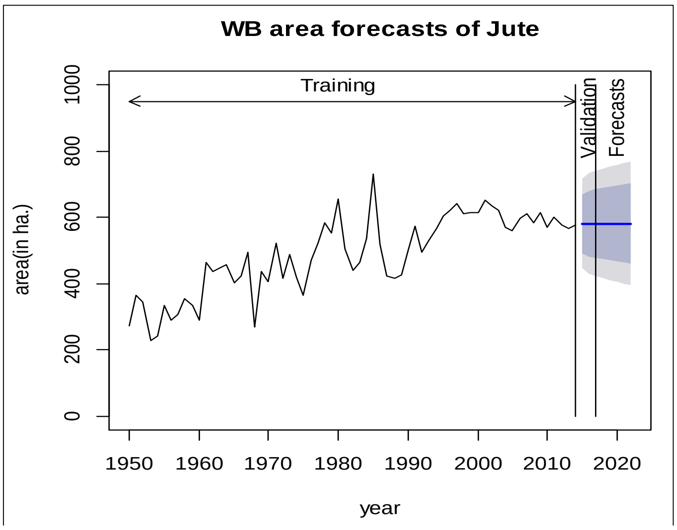
West Bengal production				India production		
year	Production	LCL	UCL	Production	LCL	UCL
2018	8711.696	6785.921	10637.47	10990.56	8626.574	13354.55
2019	8708.515	6657.285	10759.74	10978.3	8492.275	13464.32
2020	8710.867	6547.814	10873.92	10982.06	8390.774	13573.34
2021	8710.897	6439.323	10982.47	10983.55	8289.451	13677.65
2022	8710.528	6334.904	11086.15	10982.75	8188.181	13777.31
West Bengal yield				**India Yield**		
year	Yield(kg/ha)	LCL	UCL	Yield(kg/ha)	LCL	UCL
2018	2752.009	2327.319	3176.699	2558.026	2077.297	3038.755
2019	2732.004	2291.514	3212.494	2658.036	2026.910	3089.141
2020	2762.042	2258.307	3245.702	2698.025	1980.907	3135.145
2021	2769.00	2227.198	3276.809	2850.036	1938.309	3177.743
2022	2769.12	2197.833	3306.174	2876.17	1898.456	3217.596
West Bengal Area				**India Area**		
year	Area	LCL	UCL	Area	LCL	UCL
2018	581.5270	415.8793	747.1748	778.8048	572.53	985.0797
2019	581.5479	410.0867	753.0091	778.5397	572.2516	984.8277
2020	581.5535	404.5547	758.5523	778.4566	572.1673	984.7459
2021	581.5550	399.2088	763.9012	778.4306	572.1411	984.72
2022	581.5554	394.0199	769.0909	778.4224	572.133	984.7119

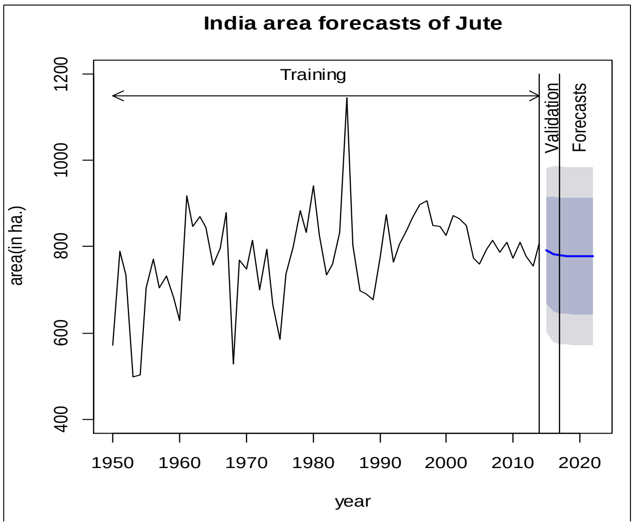












By using the suitable ARIMA models forecasted values for production, area and yield for both India and West Bengal are given in table 2. In 2022 the production of West Bengal is more so overall production of India is also more. In case of Yield West Bengal doesn't show drastically changes but in India yield of jute is more. The results also shows area forecast for the year 2022 to be about 778.4224 thousand hectare, production 10982.75 (In ' 000 Bales of 180 Kgs. Each) in India. Forecasted values of area to be 581.5554 thousand hectare and for production 8710.528(In ' 000 Bales of 180 Kgs. Each) in West Bengal.

Production & yield of India will increase in recent future as a result of increasing West Bengal production and yield of jute production, where as in case of yield the increment is not more in west Bengal but In India it has. This forecasting is a good comparison for jute production between India and its leading state. It is useful for planning different practices and recommended different strategies to improve jute production, also useful for Government for preparing budget & implementing new policies.

References

1. Kundu DK, Mazumdar SP, Ghosh D, Saha AR, Majumdar B, Ghorai AK, Behera MS. Long-term effects of fertilizer and manure application on soil quality and sustainability of jute-rice-wheat production system in Indo-Gangetic plain. Journal of Applied and Natural Science. 2016; 8(4):1793-800.
2. Sen SR. Growth and Instability in Indian Agriculture. Journal of the Indian Society of Agricultural Statistics. 1967; 21:832-833.
3. Narain D. Growth of productivity in Indian agriculture. Indian Journal of Agricultural Economics. 1977; 32(902-2018-1332):1-44.
4. Reddy VN. Statistical fitting of growth curves with illustrations from data on Indian economy. IIM, Kolkata, 1977.
5. Sawant SD. Investigation of the hypothesis of deceleration in Indian agriculture. Indian Journal of Agricultural Economics. 1983; 38(902-2018-1994):475-96.
6. Boyce JK. Agrarian impasse in Bengal: Institutional constraints to technological change, 1987.
7. De UK. Nature and Causes of Inter-District Variations in Yield of Rice in West Bengal, 1970-71 to 1994-95. Indian Journal of Agricultural Economics. 1999; 54(4):544-65.
8. Granger CW, Newbold P. Forecasting Transformed Variables. Nottingham University Forecasting Project. Note 6, 1970.
9. Box GE, Jenkins GM. Time series analysis: Forecasting and control San Francisco. Calif: Holden-Day, 1976.

10. Tsay RS. Analysis of financial time series. John Wiley & Sons; Sep 15, 2005.
11. Indiastat. https:\\www.indiastat.com/, 2018.