Int.J.Curr.Microbiol.App.Sci (2018) Special Issue-6: 1284-1293

EXCELLENT PUBLISHERS

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7692 Special Issue-6 pp. 1284-1293
Journal homepage: http://www.ijcmas.com

Original Research Article

Modelling and Forecasting of Sunn Hemp in India

P. Mishra[1*], R. P. Joshi[2], D. Ramesh[3], B. S. Dhekale[4] and P. K. Sahu[5]

[1]College of agriculture, JNKVV, Powarkheda, Hoshangabad (M.P.) 461110, India
[2](AICRP on small millets), College of agriculture, Rewa, JNKVV, 486001, India
[3]Department of Statistics & computer application, Acharya NG Ranga Agril. University, Agriculture College, Bapatla-522101, India
[4]Sher-e- Kashmir university of Agriculture scinces and technology, Kashmir, India
[5]Department of agriculture Statistics, Bidhan Chanda Krishi Vishwavidyalaya, Nadia–741 252, India
**Corresponding author*

Keywords

Production, Forecasting, Trend, Time series

A B S T R A C T

Sunn hemp (*Crotalaria juncea* L.) is most important distinctly named fibers of India. Present study describes the past and future trend of production and productivity of sunn hemp in India. The area of sunn hemp has been decreased due to expansion of area of rice in India, so Past research efforts have shown that the soft, lignified fibers produced in the stem of sunn hemp could be utilized in the manufacturing of pulp and paper, and more recent efforts have indicated that other potential products can be developed from these fibers like including newsprint, specialty papers, and as a component of commercial nursery potting media. From the forecast values obtained the regression model, it can be said that forecasted area will increases to some extent in future i.e. In 2011 area of Sunn hemp was 21.20 million ha. Up to the year 2020 it will be 4.5 million ha. In case of yield of Sunn hemp the ARIMA (1, 1, 5) model is best fitted., it can be said that forecasted yield will increases to some extent in future i.e. In 2011 yield of Sunn hemp was 658 kg/ha. Up to the year 2021 it will be 658.25 kg/ha.

Introduction

Most of the present-day sunn hemp production is located in India, Bangladesh, and Brazil, where it is grown as a green manure crop, a fodder crop, or for the bast fibers Past research efforts have shown that the soft, lignified fibers produced in the stem of sunn hemp could be utilized in the manufacturing of pulp and paper, and more recent efforts have indicated that other potential products can be developed from these fibers.. Sunn hemp is one of the earliest and most distinctly named fibers of India. Additional characteristics that enhance the value potential of sunn hemp as a non-wood, fiber crop are low nitrogen fertilization requirements, the ability to fix nitrogen and to grow in marginal soils, drought resistance, and resistance to root-knot nematodes. The bast fibers are utilized for the manufacture of cordage and high quality paper. Sunn hemp, a member of the legume family (Fabaceae), has great potential as an annually renewable, multi-purpose fiber crop. Time series (original or transformed) can be modeled using the simple moving averaging, simple exponential smoothing, and Box-Jenkins techniques (Box and Jenkins, 1976). There are two types of averaging techniques: (i) simple averaging, and (ii) moving averaging techniques. With simple averaging technique, the mean of all observations (i.e., yields in current and past years) in a series is used to forecast yield for the next year. In contrast, with moving averaging technique the analyst is more concerned with recent observations. Sahu and Mishra studied (2013) the production import –export and trade balance of total spices in India and China along with world using ARIMA model for time series data. Vishwjit *et al.*, (2014) forecasted pulses production in India using ARIMA model upto 2020**.**

Auto Regressive Integrated Moving Average (ARIMA) is the most general class of models for forecasting a time series. Appearance of lags of the forecast errors in the model is called "moving average". (ARIMA) model was introduced by Box and Jenkins in 1976 for forecasting variables. So it is necessary to study the future movement of Sunn hemp production by using past trends. (Wankhade *et al.*, 2010). Badmus and Ariyo (2011) studied forecasting the cultivated area and production of maize in Nigeria using Autoregressive Integrated Moving Average (ARIMA) model for time series data covering the period of 1970-2005 and forecasted area, production and yield for year 2020.

Materials and Methods

Data related to area, production and yield of Sunn hemp in India since 1960 to 2013 were collected from Agriculture at Glance, 2014. Statistical tools used to describe the above series are minimum, maximum, average, standard error, skewness, parametric trend models kurtosis, Box-Jenkins ARIMA modeling has been used to forecast series under consideration.

Descriptive statistics

To examine the nature of each series these have been subjected to get various statistics. Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. Descriptive statistics are typically distinguished from inferential statistics, with descriptive statistics we are simply describing what is or what the data shows. To get an overall movement of the time series data, trend equations are fitted. In this exercise different idea about the models like, polynomial, exponential, linear, compound etc are used for the purpose.

In this study also, we have tried different parametric models to describe the series under consideration, which are briefly given here under:

Linear Model

The equation of Linear model is given by $Y_t = b_0 + (b_1 t)$

Quadratic Model

The equation of Quadratic Model is given by $Y_t = b_0 + (b_1 t) + (b_2 t^2)$

The quadratic model can be used to model a series which "takes off" or a series which "dampens".

Compound Model

The equation of compound model is given by $Y_t = b_0 (b_1^t)$ Or $\ln (Y_t) = \ln (b_0) + t \ln (b_1)$

Cubic Model

The equation of cubic model is given by $Y_t = b_0 + (b_1t) + (b_2t^2) + (b_2t^3)$

Exponential Model

The equation of exponential model is given by $Y_t = b_0 e^{(b1t)}$ Or $\ln(Y_t) = \ln (b_0) + (b_1t)$

Logarithmic Model

The equation of logarithmic model is given by $Y_t = b_0 + b_1\ln (t)$

Growth Model

The equation of growth model is given by $\ln(Y_t) = b_0+b_1t\ Y_t$

Where Y_t is the value of the series at time t and b_0, b_1, b_2, b_3 are the parameters.

Among the competitive models, best model for each of the series is fixed on the basis of maximum R^2, minimum standard error and the significance of the coefficient. If, in any case the competitive models show equality in the above cases then, the model having minimum parameter is selected.

The best fitted models have also been presented in graphical form along with observed values.

ARIMA modeling

ARIMA models stands for Autoregressive Integrated Moving Average models. An ARIMA model is in-fact a combination of AR, MA models with integration.

Autoregressive model (AR)

The notation AR (*p*) refers to the autoregressive model of order *p*. The AR (*p*) model is written $X_t = c + \sum_{i=1}^{P} \alpha_i X_{t-i} + \mu_t$

Where $\alpha_1, \alpha_2 .... \alpha_p$ are the parameters of the model, *c* is a constant and μ_t is white noise i.e. $\mu_t \sim WN(0, \sigma^2)$. Sometimes the constant term is omitted for simplicity.

Moving Average model (MA)

The notation MA (*q*) refers to the moving average model of order *q*:

$$X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Where the $\theta_1,..., \theta_q$ are the parameters of the model, μ is the expectation of X_t (often assumed to equal 0), and the ε_t is the error term.

ARMA model

A time series $\{X_t\}$ is an ARMA (p, q) if $\{X_t\}$ is stationary and if for every t, $X_t - \phi_1 X_{t-1} - .... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}$ where, $\{Z_t\} \sim WN(0,\sigma^2)$ and the polynomials $(1 - \phi_1 Z - .... - \phi_p Z^p)$ and $(1 + \theta_1 Z + ... + \theta_q Z^q)$ have no common factors.

ARIMA model

A time series $\{X_t\}$ is an ARIMA (p,d,q) if $Y_t=(1-B)^d\ X_t$ is a causal ARMA(p,q) process. This means $\{X_t\}$ satisfies $\phi^*(B)X_t \equiv \phi(B)(1-B)^d X_t = \theta(B)Z_t$, where, $\{Zt\} \sim WN(0,\sigma^2)$; $\phi_{(z)}$ and $\theta_{(z)}$ are polynomials of degree p and q respectively and $\phi_{(z)} \neq 0$ for $|Z| \leq 1$. The polynomial $\phi^*(Z)$ has a zero of order d at z = 1. The process $\{X_t\}$ is stationary if and only if d = 0 and in that case it reduces to ARMA (p,q) process.

The stationarity requirement ensures that one can obtain useful estimates of the mean, variance and ACF from a sample. If a process has a mean that is changing in each time period, one could not obtain useful estimates since only one observation available per time period. This necessitates testing any observed series of data for stationarity. First the given data series are tested for stationarity through ADF and KPSS test. If the data are non-stationary, first order differencing was made to make data stationary. Given a set of time series data, one can calculate the mean, variance, autocorrelation function (ACF), and partial autocorrelation function (PACF) of the time series.

The calculation enables one to look at the estimated ACF and PACF which gives an idea about the correlation between observations, indicating the sub-group of models to be entertained. This process is done by looking at the cut-offs in the ACF and PACF. At the identification stage, one would try to match the estimated ACF and PACF with the theoretical ACF and PACF as a guide for tentative model selection, but the final decision is made once the model is estimated and diagnosed.

Results and Discussion

From the table1, In India since 1961 the area under sunn hemp has increased from 17.50 '000ha to 217'000ha.The average area under rice being 100 '000ha coupled with a positive skewness clearly indicate that the expansion in area has taken place during the later period under consideration. In fact the effect of green revolution is being reflected. Production figures show that in India has increased production from 18.10 *000 Bales of 180 Kgs. each* to 6150 *000 Bales of 180 Kgs*. The average yield under sunn hemp in India being 138.76 kg/ha.

Trends in production behavior of Sunnhemp

To workout the trends in area, production and Yield of Sunn hemp in India different parameter model like Linear, Logarithmic, Quadratic, Cubic, Compound, growth, and Exponential model where attempted to among the competitive models. The best model is selected on the basis of the maximum R^2 value, significance of the model and its coefficient. Production of Sunn hemp among the above mention model, Cubic model is found to be more attractive with maximum R^2 (0.524), coupled with significant coefficient values. From figure (1) it is also clear that the model has passed through most of the data points. Among the above mention model, Cubic model is found to be more attractive with maximum R^2 (0.782), coupled with significant coefficient values.

Box-Jenkins modeling and Forecasting

After the evaluation of trend of each and every series, our next task is to forecast the series for the year to come. For the purpose we adoptated the Box –Jenkins methodology and has been discussed in the material and method section.

Data for the period 1961- 2010 has been used for the model building, while data for years 2011-2013 are taken for model validation. The selected models are ARIMA (0, 1, 0), ARIMA (1, 1, 2) and ARIMA (1, 1, 5). These three models are again compared according to the minimum values of RMSE, MAE, MSE and MAPPE and maximum value of R^2 which are given in table 5. Hence, it can be concluded that ARIMA (1, 1, 2) is the best fitted model for forecasting the Area of Sunn hemp in India. Best fitted model used to forecast the series for the years to come.

Table.1 Descriptive Statistics

	Area (In ' 000 Hectare)	*Production (In ' 000 Bales of 180 Kgs. each)*	*Yield (In Kg./Hectare)*
Mean	100.10	50.66	138.76
Standard Error	8.24	2.85	23.07
Kurtosis	-1.09	-0.43	5.66
Skewness	0.30	0.15	2.69
Minimum	17.50	18.10	57.36
Maximum	217.00	101.40	659.86

Table.2 Trend models for area of Sunn hemp in India

Equation	Model Summary			Estimated parameters			
	R^2	F	Sig.	Constant	b1	b2	b_3
Linear	0.855	347.842	0	-117.864	0.062		
Logarithmic	0.855	349.007	0	-932.392	123.582		
Inverse	0.856	350.132	0	129.301	-2.45E+05		
Quadratic	0.855	347.842	0	-117.864	0.062	0	
Cubic	**0.980**	211.90	0	-117.864	0.000	0.014	-5.24
Compound	0.837	304.016	0	7.17E-10	1.012		
Power	0.839	307.065	0	4.13E-75	22.79		
S	0.84	310.132	0	24.523	-4.52E+04		
Growth	0.837	304.016	0	-21.056	0.012		
Exponential	0.837	304.016	0	7.17E-10	0.012		

Table.3 Trend models for production of Sunn hemp in India

Equation	Model summary			Estimated Parameter			
	R^2	F	Sig.	Constant	b1	b2	b3
Linear	0.364	374.691	0	-428.977	0.221		
Logarithmic	0.462	369.309	0	-3.31E+03	436.373		
Inverse	0.361	364	0	443.77	-8.63E+05		
Quadratic	0.366	380.145	0	-210.792	0	5.58E-05	
Cubic	**0.524**	385.671	0	65.11	0.003	0.183	1.565
Compound	0.326	742.051	0	1.10E-25	1.03		
Power	0.427	746.293	0	1.02E-195	59.401		
S	0.427	750.329	0	61.328	-1.18E+05		
Growth	0.478	742.051	0	-57.474	0.03		
Exponential	0.426	742.051	0	1.10E-25	0.03		

Table.4 Trend models for yield of Sunn hemp in India

Equation	Model summary			Estimated Parameter			
	R^2	F	Sig.	Constant	b1	b2	b3
Linear	0.753	445.892	0	-4.59E+04	23.818		
Logarithmic	0.672	439.994	0	-3.56E+05	4.71E+04		
Inverse	0.564	434.13	0	4.84E+04	-9.32E+07		
Quadratic	0.772	451.822	0	-2.23E+04	0	0.006	
Cubic	**0.782**	457.78	0	-28.77	0.029	1.749	28.35
Compound	0.708	582.264	0	1.36E-13	1.019		
Power	0.708	580.32	0	1.00E-118	36.73		
S	0.707	578.21	0	43.833	-7.27E+04		
Growth	0.708	582.264	0	-29.628	0.019		
Exponential	0.708	582.264	0	1.36E-13	0.019		

Table.5 Diagnostic tools and model selection for area of Sunn hemp in India of best fitted models

Area	R^2	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
ARIMA (0,1,0)	0.957	0.242	2.731	0.156	16.95	1.217	-2.769
ARIMA (1,1,2)	0.974	0.240	2.675	0.150	16	1.149	-2.382
ARIMA (1,1,5)	0.958	0.25	2.776	0.158	16.056	1.153	-2.295

Table.6 Diagnostic tools and model selection for production of Sunn hemp in India of best fitted models

Production	R^2	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
ARIMA (0,1,1)	0.919	1.201	10.231	0.818	**39.223**	3.908	**0.503**
ARIMA (1,1,4)	**0.931**	**1.147**	**9.948**	**0.779**	39.478	3.972	0.683
ARIMA (1,1,5)	0.925	1.205	10.256	0.789	48.893	**3.476**	0.851

Table.7 Diagnostic tools and model selection for yield of Sunn hemp in India of best fitted models

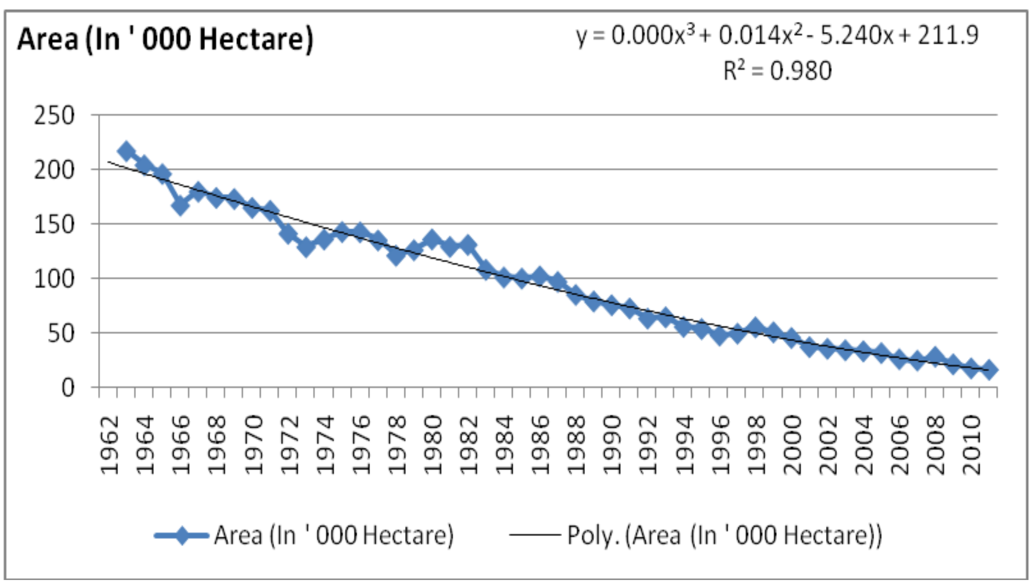
Yield	R^2	RMSE	MAPE	MAE	MaxAPE	MaxAE	Normalized BIC
ARIMA (0,1,2)	0.903	140.688	8.353	105.72	29.629	354.05	10.098
ARIMA (1,1,5)	**0.93**	121.598	**7.082**	**88.161**	**27.543**	320.905	**9.943**
ARIMA (0,1,4)	0.914	134.537	7.588	93.122	33.32	342.86	10.145
ARIMA (0,1,5)	0.918	**132.248**	7.625	95.329	31.099	**301.232**	10.179

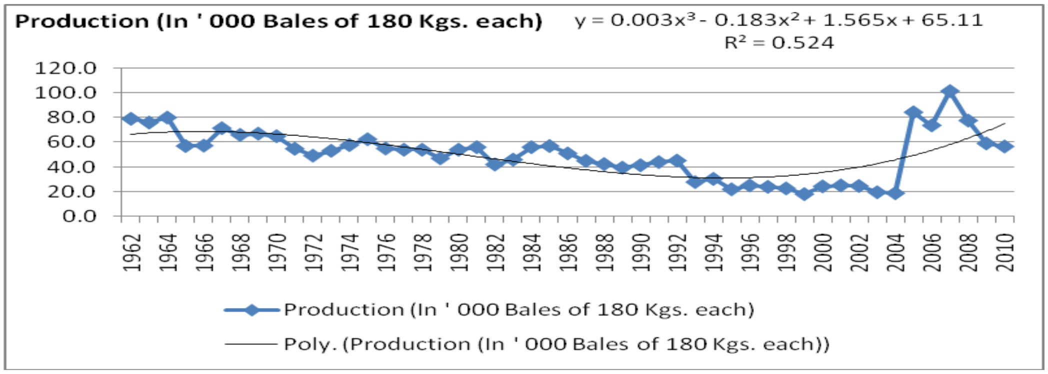
Table.8 Observed and predicated area, production and productivity of Sunn hemp in India

year	Obs_ared	Fore_area	Obs_prod	Fore_prod	Obs_yield	Fore_yield
2011	21.2	21.13	77.6	59.49	658.87	618.09
2012	17.5	18.81	59.1	53.57	607.89	677.67
2013	16.2	16.24	56.6	50.64	628.89	590.55
2014		6.04		52.09		638.1
2015		5.75		51.16		614.54
2016		5.5		49.88		656.91
2017		5.25		49.55		639.99
2018		5.05		49.22		671.82
2019		4.8		48.89		664.67
2020		4.5		48.4		658.25

Area (A) is in ' 000 ha, Production (In ' 000 Bales of 180 Kgs. each) and Yield (In Kg./Hectare)

Fig.1 Observed and expected trend values of area, production and yield of sunn hemp in India





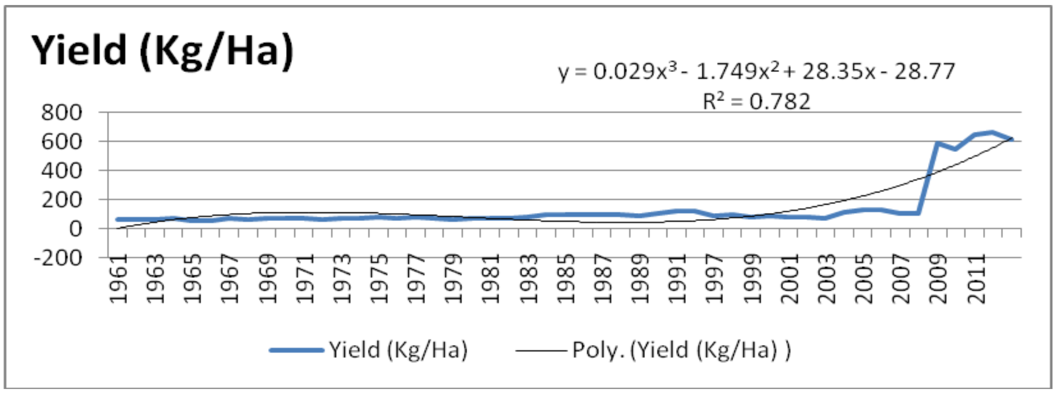
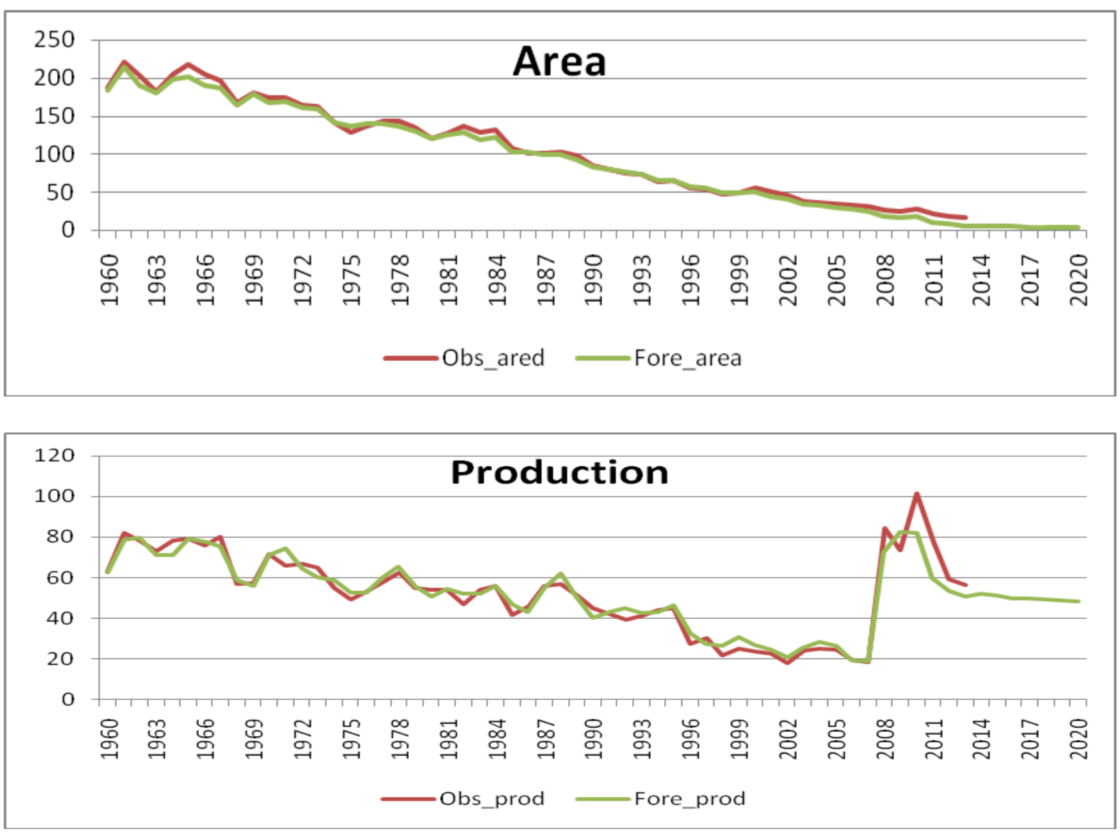
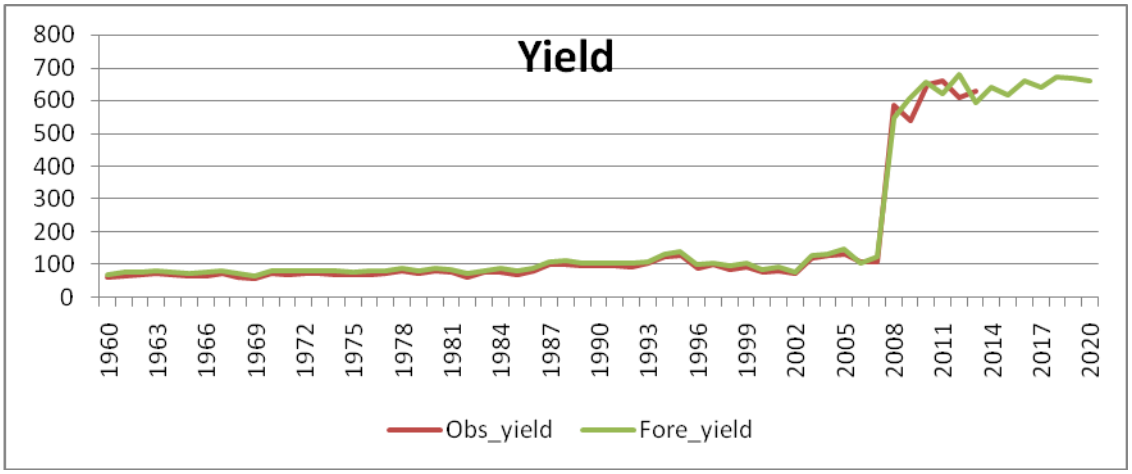


Fig.2 Observed and forecasted area, production and yield of Sunnhemp in India





Area (A) is in ' 000 ha, Production (In ' 000 Bales of 180 Kgs. each) and Yield (In Kg./Hectare)

Production behavior of Sunn hemp has been modeled with the help of Box- Jenkins's ARIMA modeling technique. Among the comparative models, best model has been selected on the basis of the above mentioned criteria and are presented in table-6. It is clear from the table the different types of model starting from (0-1-0) to (1-1-5) could be used to explain the series. Based on the model the predications for year 2020 have been made for the series. Forecasting with respective models suggest that there will be increased production for Sunn hemp. This justified that the selection of ARIMA (1, 1, 4) is the best model to represent the data generating process very precisely. From the above table 7 it is observed that the model ARIMA (1, 1, 4) is better than the other models because this model satisfied the all criterion except MSE and R^2. So the select ARIMA (1, 1, 4) is the best. Yield behavior of Sunn hemp has been modelled with the help of Box- Jenkins's ARIMA modeling technique. Among the comparative models, best model has been selected on the basis of the above mentioned criteria and are presented in table-7. Forecasting with respective models suggest that there will be increased yield for Sunn hemp. This justified that the selection of ARIMA (1, 1, 5) is the best model to represent the data generating process very precisely. From the above table 7 it is observed that the model ARIMA (1, 1, 5) is better than the other models because this model satisfied the all criterion except MSE and R^2. So the select ARIMA (1, 1, 5) is the best. About 20 percent of the produce was used to be exported to the overseas market till 1970s.But with the advent of Green Revolution, its area and production dropped down gradually, the reason being increased irrigation facilities, intensive cropping system with more profitable crops like vegetables, oilseed, pulses and paddy. Presently the annual production of fibre is around 18.8 thousand ton from 0.31 lakh hectare distributed in states. Orissa, Bihar, Madhya Pradesh, Rajasthan, West Bengal, Uttar Pradesh, Maharashtra, Tamil Nadu and Jharkhand. Even though India is the largest producer of sunnhemp fibre, followed by Bangladesh and Brazil, the demand for its fibre is showing sharp decline because of the competition rom cheaper synthetic fibre-based products in the trade. In area, production and productivity data none of the series is stationary; 1st differencing with original data makes all the series stationary. ARIMA model offers a good method for predicting the magnitude of any 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 a prior value of any parameter. In our study ARIMA (1, 1, 2) model is best suited for estimation of Sunn hemp area data. From the forecast values obtained the regression model, it can be said that forecasted area will increases to some extent in future i.e. In 2011 area of Sunn hemp was 21.20 million ha. Up to the year 2020 it will be 4.5 million ha. In case of yield of Sunn hemp the ARIMA (1, 1, 5) model is best fitted., it can be said that forecasted yield will increases to some extent in future i.e. In 2011 yield of Sunnhemp was 658 kg/ha. Up to the year 2021 it will be 658.25 kg/ha.

References

Badmus M.A and O.S. Ariyo 2011. "Forecasting Cultivated Areas and Production of maize in Nigerian using ARIMA Model" *Asian Journal of Agricultural Sciences* 3(3): 171-176,

Box, G.E.P. and Jenkins, G.M. 1976. Time Series Analysis: Forecasting and Control, Holden-Day, San Fransisco.

Sahu, P.K., and P. Mishra 2013. "Modeling and forecasting production behaviour and import- export of total spices in two most populous countries of the world" *Journal of Agriculture Research*, 51(4): 81-97.

Vishwajith K.P, B S Dhekale, P K Sahu, P. Mishra AND MD Noman 2014. "Time series modeling and forecasting of pulses production in India." *Journal of Crop and Weed, 10(2):147-154 (2014).*

Wankhade R., Mahalle S, Gajbhiye S., Bodade, V.M. 2010. "Use of the ARIMA Model for Forecasting Pigeon Pea Production in India*" International Review of Business and Finance*, ISSN 0976-5891, 2(1) (2010), pp. 97–102.