The Andhra Agric. J 62(3): 658-665, 2015

Forecasting of Area, Production and Producitivity of Jowar in Andhra Pradesh Using Growth Models

V Nireesha, V Srinivasa Rao, DV S Rao and G Raghunadha Reddy
Department of Agricultural Statistics, Agricultural College, Bapatla 522 101, Andhra Pradesh

ABSTRACT

This paper attempted to identify the trend of area, production and productivity of Jowar(Sorghum) in Andhra Pradesh through fitting different Linear, Non-Linear Growth models. Influence of weather parameters on area, production and productivity of Jowar crop by using Karl Pearson's correlation and Multiple Linear Regression Analysis was also studied. From the best fitted model forecasting of area, production and productivity was also done. Cubic model was identified as the best model for the observed data and forecasting was done for the Jowar area, production and productivity up to 2020 AD. It was observed that there was an increasing trend in the productivity, but both area and production was in the decreasing trend during the study period.

Key words : Adj.R^2 and MAPE, Karl Pearson's correlation coefficient, Linear and Non-Linear growth models, Multiple Linear regression, Theil's U-Statistic, R^2.

Jowar (Sorghum) is the fifth most important cereal crop after wheat, rice, maize and barley. The origin of sorghum is generally believed to be around the present day Ethiopia or East Central Africa. Sorghum was taken from East Africa to India during the first millennium. Major production share is contributed by the African countries, followed by USA accounting up to 17% of the world's total production. World trade in sorghum is dominated by the largest producer of the crop in the world i.e. U.S.A as most of the production in the country accounts for exports.

It is adapted to a wide range of environmental conditions but is particularly adapted to drought capable of growing well under contrasting climatic conditions. Cultivation of sorghum is mainly concentrated in peninsular and central India. Maharashtra, Karnataka, Andhra Pradesh, Madhya Pradesh, Gujarat, Rajasthan, Uttar Pradesh (the Bundelkhand region) and Tamil Nadu are the major Sorghum growing states.

In India the area under sorghum has decreased from 6245.1 thousand hectares in 2011 to 6214.4 thousand hectares in 2012 and production from 6006.5 thousand tonnes in 2011 to 5281.5 thousand tonnes. Thus area and sorghum production are estimated to have decreased by 30.7 thousand hectares and 725 thousand tonnes in 2012. Similarly the yield has marginally decreased from 962 kg ha^{-1} to 850 kg ha^{-1}.

Andhra Pradesh is one of the important sorghum growing states. The area under sorghum was 287 thousand hectares, production was 472 thousand tonnes and productivity was 1643 kg ha^{-1} during 2012 in Andhra Pradesh.

Growth models are useful in drawing inferences like the exact relationship between time and growth, the rate of growth at each point of time, the turning points in the growth, etc. These are used to describe the dependent variable with the help of multiple repressors. Srinivasa Rao and Srinivasulu (2006) measured the growth rates of turmeric and to estimate the future projections up to 2020 AD by using the linear, quadratic, exponential, logarithmic, compound growth models. Results revealed that turmeric production would increase considerably in Andhra Pradesh and India. Aparna *et al*. (2008) analyzed trend in growth rates of major vegetables in Visakhapatnam district with the help of compound growth by using exponential function. Growth models can provide a convenient data summary and be useful for prediction. Singh *et al*. (1998) analysed the trend in area, production and yield of major food grains, coarse cereals, pulses, oilseeds, sugarcane, and Sorghum at the state level in India. Compound growth rates of area,

Table 1. Growth models in Area, Production and Productivity of Sorghum in Andhra Pradesh.

Model	Linear	Logarithmic	Inverse	Quadratic	Cubic	Power	Exponential
				Area			
a	2938.19	3934.23	1191.20	3002.61	2623.10	8332.55	4129.74
b	-62.05	-853.85	2729.44	-69.94	20.22	-0.682	-0.053
c				0.164	-4.48		
d					0.065		
Theil's U-Statistic	94.84	87.21	76.86	94.90	**96.56**	71.02	89.16
R^2	0.959	0.756	0.259	0.960	**0.982**	0.637	0.808
ADJ R^2	0.958	0.751	0.242	0.959	**0.981**	0.628	0.804
MAPE	14.00	37.51	85.00	13.00	**7.00**	42.00	15.00
				Production			
a	1487.10	1850.62	801.51	1433.33	1214.17	2184.45	1666.03
b	-24.19	-324.42	1109.93	-17.61	34.45	-0.378	-0.029
c				-0.13	-2.82		
d					0.037		
Theil's U-Statistic	90.83	87.43	82.74	90.90	**91.72**	84.00	89.75
R^2	0.966	0.578	0.226	0.967	**0.972**	0.568	0.958
ADJ R^2	0.966	0.569	0.209	0.966	**0.972**	0.558	0.957
MAPE	15.07	23.75	37.93	15.60	**13.75**	24.41	15.87
				Productivity			
a	295.29	74.50	840.12	574.68	438.60	298.11	403.93
b	20.149	241.99	-648.63	-14.06	18.26	0.305	0.024
c				0.713	-0.954		
d					0.023		
Theil's U-Statistic	89.90	85.37	81.34	92.77	**93.30**	85.78	91.02
R^2	0.724	0.435	0.104	0.858	**0.878**	0.502	0.788
ADJ R^2	0.718	0.423	0.085	0.854	**0.875**	0.481	0.783
MAPE	16.89	24.68	32.55	13.50	**11.92**	20.15	13.42

production and yield were estimated by fitting log-linear functions using data for 1960-61 to 1992-93. The study revealed that for total food grains, as well as for all the individual grain crops, yield witnessed a higher growth rate as compared to acreage during 1972-73 to 1992-93. Sudha *et al.* (2013) analysed growth trends in area, production and productivity of maize from 1970 to 2008. And the future projections were estimated up to 2015AD by using growth models like linear, logarithmic, quadratic, etc., they concluded that maize production would increase considerably in Guntur district.

Singh *et al.* (2008) analysed the effect of rainfall and temperature effect on wheat yield in south western region of Punjab. The Maximum, minimum temperature and rainfall from December to March for each period of five years 1977-81 to 1997-2001 were analyzed. The results revealed that the temperatures during February and March were showed significant effect on wheat yield. The grain yield revealed positive correlation with minimum temperature but no trends observed for other parameters. Rankja *et al*. (2010) made an attempt to study the quantitative relationship between weather parameters and district level yield of cotton and to develop pre harvest forecast models for cotton yield. For this purpose 32 years data of weather parameters and crop yield was collected. The results showed that the 26 week crop period model was recommended for pre harvest forecast due to highest coefficient of determination (R^2) and lower forecast error. The analysis also revealed that time trend, maximum temperature, morning and evening relative humidity have significantly affected on crop yield.

MATERIAL AND METHODS

In this paper an attempt has been made for forecasting of area, production and productivity of Jowar (Sorghum) in Andhra Pradesh by using linear, non-linear growth models and influence of weather parameters on area, production and productivity by using Karl Pearson's Correlation coefficient and Multiple Linear Regression Analysis. 47 years time series data from 1966 to 2012 on area, production and productivity of Jowar were collected from the www.Indiastat.com and Ministry of Agriculture, Govt of India. The projections were also estimated up to 2020 AD.

For Forecasting area, production and productivity of Sorghum in Andhra Pradesh upto 2020 AD the data for a period of 47 years *i.e* from 1966 to 2012 was used and fitted Linear and Non-Linear growth models and best fitted model was selected based on model selection criteria. Among those models which were having highest Theil's U-Statistic (model accuracy), highest R^2, highest Adjusted R^2 and least MAPE values was selected as appropriate for the projections. The association between weather parameters on area, production and productivity also analysed by using correlation and regression analysis. The linear, non-linear growth models, correlation and regression analysis which were taken under consideration are as follows.

LINEAR AND NON-LINEAR GROWTH MODELS:

1. Linear function:

A linear model is one in which all the parameters appear linearly.

The linear equation is given by $Yt = a + bt$

2. Quadratic function:

This function is useful when there is a peak or a trough in the data of past periods.

Quadratic fit is given by $Yt = a + bt + ct^2$

3. Cubic function:

This function is useful when there is, two peaks or two troughs in the data of past periods.

Cubic fit or third degree curve is given by $Yt = a + bt + ct^2 + dt^3$

4. Exponential function:

If, when the values of t are arranged in an arithmetic series, the corresponding values of y form a geometric series, then the relation is of the exponential type.

The function of this type can be given by $Yt = ab^t$

5. Logarithmic function:

This model shows very rapid growth, followed by slower growth

The mathematical equation is given by

$Yt = a + b \ln(t)$

6. Inverse function:

Inverse curve shows a decreasing growth.

Inverse fit is given by $Yt = a + b/t$

7. Power function:

If, when the values of t are arranged in a geometric series, the corresponding values of y form a geometric series, the relation is of power type.

The fit is given by $Yt = at^b$

The fit is similar to exponential fit, but produces a forecast curve that increases or decreases at different rate.

In all the above functions,

Yt is dependent variable i.e., area / production / productivity

t is independent variable, time in years

a is an intercept

b,c are regression coefficients

Karl Pearson's Correlation:

Correlation coefficient is a measure of degree of linear association between two variables.

Correlation coefficient between *two variables X and Y*

$$r\ (X,Y) = \frac{Cov\ (X\ Y)}{\sqrt{Var\ (x).\ Var\ (y)}}$$

For test the significance of correlation coefficient (r) t-test is used.

Null Hypothesis (H0): ρ = 0

Alternate Hypothesis (H1): ρ ± 0

Test the Significance of correlation coefficient:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t\ (n\text{-}2)\ d.f$$

Multiple Linear Regression:

Prediction of an unknown dependent variable with the help of independent variables is called linear regression. Multiple Linear Regression Analysis is if more than one independent variable is to be used in the model, linear regression can be extended to multiple linear regression to accommodate several independent variables (Pearson 2011). It is of the form

$$Y_i = a + b1x1 + b2x2 + \ldots + bnxn$$

Where

a = intercept in the model,

x1, x2 , …… xn are independent variables,

b1, b2,……bn are regression coefficients.

MODEL SELECTION:

The choice of the linear and non-linear growth models equation amongst the available alternatives is very crucial. Many researchers uses Theil's U-Statistic, coefficient of multiple determination (R^2) or adjusted R^2 ($\bar{R}^2$) and least MAPE as the criterion of model selection.

Theil's U – Statistic:

This statistics allows a relative comparison of normal forecasting methods with naive approaches and also squares the errors involved so that the large errors are given much more weight than small errors. The positive characteristic that is given up in moving to Theil's U-statistic as a measure of accuracy is that of intuitive interpretation. The difficulty will become apparent as the computation of this statistic and its application are examined. Mathematically, it is defined as,

$$U = \sqrt{\sum_{t=1}^{n-1} \frac{\left(\frac{F_{t+1} - Y_{t+1}}{Y_t}\right)^2}{\left(\frac{Y_{t+1} - Y_t}{Y_t}\right)^2}} * 100$$

Where F t+1 is the forecasted value at time t +1, and where Y_t is the actual value at a given time period *t*, *n* is the number of data points . If U = 100 both models forecast with equal accuracy U < 100 model forecast is best.

Coefficient of Determination (R^2):

R^2 is a statistic that will give some information about the goodness of fit of a model. In regression, the R^2 i.e coefficient of determination will explains the variability of the model . An R^2 of 1.0 indicates that the regression line perfectly fits the data. It provides a measure of how well future outcomes are likely to be predicted by the model.

$$\mathbf{R^2} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

Adjusted R^2 ($\bar{R}^2$)

Adjusted R^2 is a modification of R^2 that adjusts for the number of explanatory terms in a model. The adjusted R^2 can be negative, and will always be less than or equal to R^2. The adjustedR^2 is defined as

$$\bar{R}^2 = 1 - \frac{(1-R^2)\ (n-1)}{n-p-1}$$

Where,

p is the number of parameters in the equation

n is the total number of observations

Mean Absolute Percent Error (MAPE): It is a measure of accuracy of a method for constructing fitted time series values in statistics, specifically in trend estimation. It is usually expresses accuracy as a percentage, and is defined by the formula:

$$MAPE = \left(\sum_{t=1}^{n}\left(\left|\frac{A_t - F_t}{A_t}\right|\right) \times 100\right) / n$$

Where,

At is the actual value and

Ft is the forecast value

Table2. Future Projections of Area, Production and Productivity of Sorghum in Andhra Pradesh.

YEAR	AREA ('000 ha)	PRODUCTION ('000 tonnes)	PRODUCTIVITY (kg ha^{-1})
2013	403.59	490.93	1677.91
2014	444.47	514.80	1767.00
2015	495.35	543.99	1860.99
2016	556.64	578.71	1960.01
2017	628.71	619.21	2064.21
2018	711.95	665.68	2173.73
2019	806.75	718.37	2288.69
2020	913.49	775.50	2409.25

Table 3. Correlation coefficient values of Area, Production and Productivity (Yield) of Jowar With Weather Parameters.

Weather Parameters	Area	Production	Yield	Test of significance (t) for Area	Test of significance (t) for production	Test of significance (t) for yield
Max temp(^{0}C)	0.18	0.05	-0.22	0.997	0.269	1.222
Min temp(^{0}C)	0.41	0.37	-0.22	2.419*	2.149*	1.234
Rainfall(mm)	-0.27	-0.37	0.04	1.530	2.153*	0.238
RH1(%)	-0.75	-0.68	0.65	6.116*	5.029*	4.547*
RH2(%)	-0.39	-0.39	0.34	2.290*	2.255*	1.967

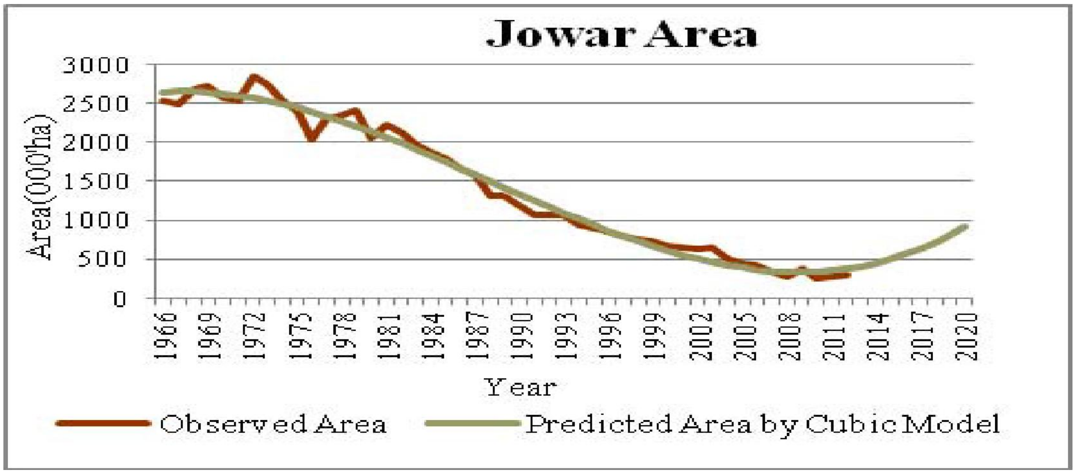
(* significant at 5% level) (t tab value = 2.048)

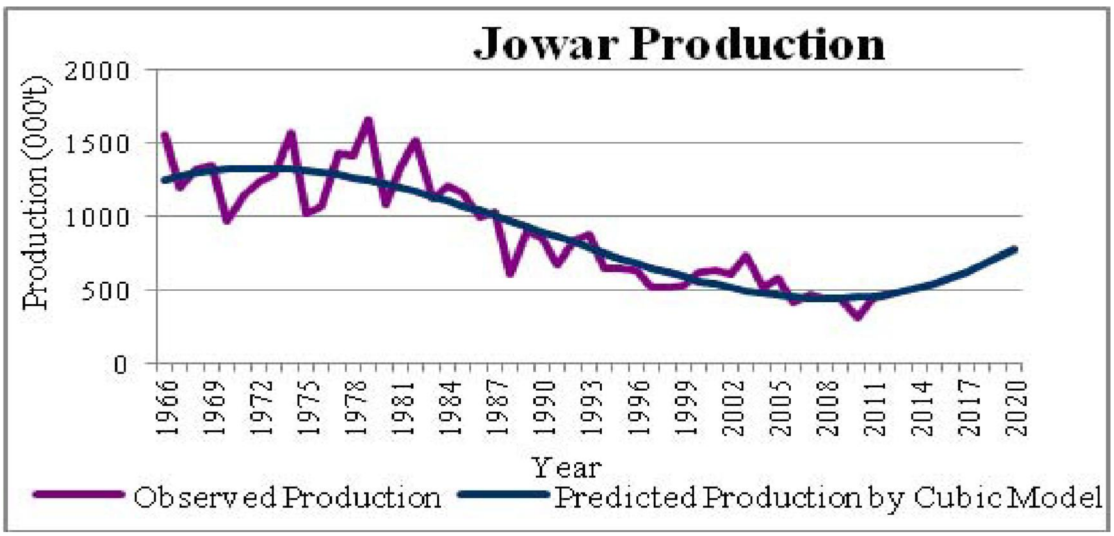
Table 4. Predicted models for Area, Production and Yield of Jowar on Weather Parameters.

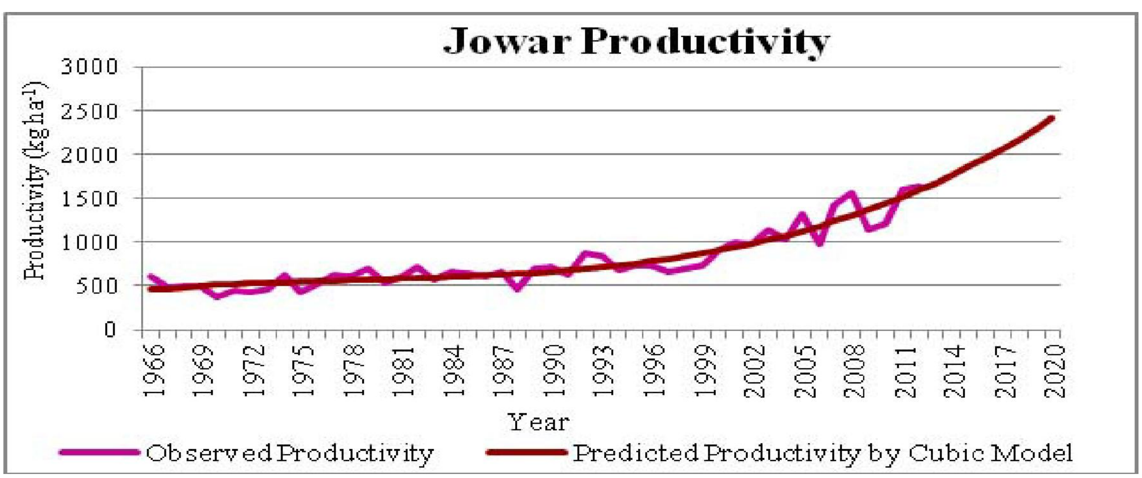
Dependent Variable	Model	R^2
Area	$\hat{Y}$ = 17543.5 + 5.43 (Max Temp) + 18.67 (Min Temp) + 2.19 (Rainfall) - 218.25**(RH1) + 1.31 (RH2)	0.615*
Production	$\hat{Y}$ = 7159.56 - 4.35 (Max temp) + 13.86 (Min temp) - 0.60 (Rainfall) - 76.75**(RH1) - 6.51 (RH2)	0.513*
Yield	$\hat{Y}$ = -12243.12 - 9.74 (Max temp) + 12.18 (Min temp) -5.73*(Rainfall) + 162.05**(RH1) + 9.86 (RH2)	0.578*

(**,*significant at 1% and 5% level respectively)

Observed and Projected Trend Charts For Area, Production And Productivity of Sorghum







RESULTS AND DISCUSSION

To understand the growth performance of Sorghum in Andhra Pradesh during the period from 1966 to 2012, time series data on area, production and productivity of Sorghum in Andhra Pradesh was analyzed by using Linear and Non-Linear growth models. The linear, logarithmic, inverse, quadratic, cubic, compound, power, exponential models were fitted and the parametric values are presented in table 1.

1. Area of Sorghum:

In Andhra Pradesh, the average area covered under Sorghum during the study period (1966 to 2012) was 1449 thousand hectares. Area under Sorghum in Andhra Pradesh exhibited a decreasing trend up to 2008 later it showed an increasing trend during the study period. The results obtained by fitting all the models were presented in Table1. In case of Sorghum area, cubic function had highest Theil's U-Statistic (96.56%) R^2 (0.982) and Adjusted R^2 (0.981) having least MAPE (7.00) values. Hence it was found to be suitable for future projections.

2. Production of Sorghum:

The average production under Sorghum was 906 thousand tonnes in the study period. Regarding the production of Sorghum in Andhra Pradesh, it has been observed that there was an increasing and decreasing trend. The results obtained by fitting the models were presented in Table 1. In case of Sorghum production the model selection criteria like Theil's U-Statistic (91.72%), R^2 (0.971), Adjusted R^2 (0.972) and MAPE are higher for cubic model. Hence cubic model was found to be the best fitted one for the projection of production under Sorghum in Andhra Pradesh.

3. Productivity of Sorghum:

The average productivity under Sorghum was 779 kg ha^{-1} in the study period. Regarding the productivity in Andhra Pradesh it has been observed that there was an increasing trend. The results obtained by fitting all the linear and non-linear growth models were presented in Table 1. In case of Sorghum productivity also cubic model had highest Theil's U-Statistic (93.30), R^2 (0.878), Adjusted R^2 (0.875) and MAPE (11.92) values. Hence cubic model was found to be the best fitted model for future projections.

Projections:

The future projections of area, production and productivity of Sorghum in Andhra Pradesh by 2020 AD were calculated and the results were presented in Table 2. Cubic function was found to be the best fitted one for projecting the area, production and productivity under Sorghum by 2020 AD as it had the highest Theil's U-Statistic, R^2, Adj $R^2$and MAPE values. The projected area, production and productivity under Sorghum by 2020 AD would be 913.49 thousand hectares, 775.50 thousand tonnes and 2409.25 kg ha^{-1} respectively. The projections of Sorghum area, production and productivity would be increasing by 2020 A.D.

From correlation analysis it was observed that Maximum Temperature and Minimum Temperature were positively correlated with area, production but on productivity these parameters were negatively correlated. Rainfall, morning relative humidity (RH1) and evening relative humidity (RH2) were negatively correlated with area, production but with yield positively correlated. Even though the yield was increased because of less area the production was decreased. Correlation coefficient values and test of significance values are presented in table 3.

From multiple linear regression analysis it was observed that among all the weather parameters Morning relative humidity (RH1) was negatively significant at 1% level on area and production. But rainfall was negatively significant at 5% level, morning relative humidity (RH1) was positively significant at 1% level. Because of these models lowest coefficient of determination (61.5%, 51.3%, and 57.8%) it can be concluded that other than weather parameters many factors are influencing on area, production and yield of Jowar. The predicted models are presented in table 4.

LITERATURE CITED:

Aparna B, Shareef S M, Raju V T and Rao V S 2008 Growth trends of major vegetables in Visakhapatnam. *The Andhra Agricultural Journal,* 55 (1): 68-69.

Rankja N J, Upadhyaya S, Pandya H R, Parmar H R and Varmora S L 2010 Estimation of cotton based on weather of Banaskantha district in Gujarat state. *Journal of Agrometeorology,* 12 (1): 47-52.

Rao V S and Srinivasulu R 2006 Growth comparisions of turmeric up to 2020 AD. *The Andhra Agricultural Journal,* 53 (1&2): 108-109.

Singh R K P and Ranjan K P 1998 Growth and instability in production of principal food grain crops: a case of backward economy. *Bangladesh Journal of Agricultural Economics,* 21 (1-2): 1-20.

Singh S S H, Hndl and Kaur P 2008 Effect of temperature and rainfall on wheat yield in south western of Punjab. *Journal of Agrometeorology,* 10 (1): 70-74.

Sudha CH K, Rao V S and Suresh CH 2013 Growth trends of maize crop in Guntur district of Andhra Pradesh. *International Journal of Agricultural Statistical Sciences,* 9 (1): 215-220.

(Received on 28.07.2014 and revised on 10.06.2015)