Research Article

Impact of monsoon and statistical forecasting of production of important oil seeds in Aurangabad division

Chhaya Sonar[1*], S L Sananse[2], A Y Tayde[3], Ankush Ubale[4]

Department of Statistics, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, INDIA.
Email: chhayajadhav@gmail.com

Abstract

Monthly Rainfall, dry spells at different stages and their influence on productivity is of interest to farmers and many researchers. The present study focuses on fore casting productivity of *Kharip* crops, cotton and soyabeanwhich are major crops in Aurangabad division. In Aurangabad division there are three districts namely Aurangabad, Jalna and Beed. *Kharip* crop's productionis affected by variation in rainfall, dry spells hence the parameters focused in this study are monthly rainfall (June to September), dry spells(number of days) and number of rainy days in a season (June to September). Correlation analysis indicates that there is significant impact of dry spells in the month of June and July on production of cotton crop while the production of soyabean is significantly correlated with dry spells in the months of July and August. Regression analysis is used to forecast the production of cotton and soyabean. The small magnitude of coefficient of determination indicates the non-linear relationship hence transformations are used to obtain linearity. The transformed data is used for further analysis.

Keywords: Rainfall, dry spells, oil seeds, forecasting, regression, Cp mallows test statistic.

[*]**Address for Correspondence:**
Dr. Chhaya Sonar, Department of Statistics, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, INDIA.
Email: chhayajadhav@gmail.com
Received Date: 03/05/2015 Accepted Date: 25/05/2015

Access this article online

Quick Response Code:

Website: www.statperson.com

DOI: 27 May 2015

INTRODUCTION

The success of crop establishment and growth largely depends on availability of adequate rainfall in country like India. Indian agriculture is largely affected by the rainfall occurring during monsoon (June to September) season. Rainfall is also primary factor affecting crop production in rain fed agriculture. It affects the productivity of many crops taken during *Kharip* season. The monsoon rainfall affects the total food grain of India (Krishna Kumar *et al.* 2004). Its impact on different crops are studied by various researchers. V. Prasanna (2014) focuses on the impact of monsoon rainfall on *Kharip* and *Rabi* food grain yield over India. P. E. Naill, M. Momami (2009) used ARIMA model to forecast the monthly rainfall. Sawa and Ibrahim (2011) forecast the yield of Millet and Sorghum by using bivariate correlation analysis and stepwise regression using dry spell parameter for semi-arid region of Nigeria. Mustapha (2012) also studied effect of dry spells on millet yield. Seghatoleslami, kafi, Majidi (2008) studied effect of deficit irrigation on yield. Peris and Peries (1993) studied the effect of bimonthly rainfall on coconut yield in Shri Lanka.

MATERIAL AND METHODS

The data of monthly rainfall(June to September), dry spells(June to September), monthly rainy days(June to September) along with productivity of crops soyabean and Cotton in kg/ha for three districts in Aurangabad division viz. Aurangabad, Jalna and Beed for period 2000-2001 2013-2014 was collected from meteorological observatory of Agricultural Department. The aim of this study is to forecast the production and to find out the parameters which largely affect the productivity of soyabeanandcottonwhich are major crops in Marathwada region of Maharashtra. The parameters studied here are

How to site this article: Chhaya Sonar, S L Sananse, A Y Tayde, Ankush Ubale. Impact of monsoon and statistical forecasting of production of important oil seeds in Aurangabad division. *International Journal of Statistika and Mathemtika* May to July 2015; 13(1): 38-40. http://www.statperson.com (accessed 30 May 2015).

monthly rainfall (June to September), number of rainy days in the month (June to September), dry spells(in number of days in the months June to September).A dry spell here is considered as more than four consecutive days having rainfall less than 2.5mm. The variables are defined as:

Yc: Productivity of Cotton Ys: Productivity of Soyabean

X_1: actual rain in the month of June X_7: Rainy days in the month of August

X_2: actual rain in the month of July X_8: Rainy days in the month of September

X_3: actual rain in the month of August X_9: Dry spell (in days) in the month of June

X_4: actual rain in the month of September X_{10}: Dry spell (in days) in the month of July

X_5: Rainy days in the month of June X_{11}: Dry spell (in days) in the month of August

X_6: Rainy days in the month of July X_{12}: Dry spell (in days) in the month of September

Correlation analysis is done to find out which variables affect the production of cotton and soyabean. Lack of fit test is conducted to check the independence of observations using Durbin-Watson statistic. Linear relationship between the dependent variable and independent variables should be tested. To fit the multiple linear regression model the linear relationship should be tested, there are number of ways to check linear relationship. The multiple linear regression model

$$y_c(or\ y_s) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{12} x_{12} + \varepsilon$$

Where $\beta_0, \beta_1, \beta_2, \ldots, \beta_{12}$ are unknown parameters that are to be estimated using the data considered in this study. $\beta_i$is also called as ith regression coefficients which measures the expected change in response $y_c(or\ y_s)$ per unit change in x_i when all the remaining variables i≠j = 1, 2, …, 12 are held constant and β_0 is the intercept of the regression plane. A statistical measure of how close the data are to the fitted regression lineis R^2. It is also known as the coefficient of determination, or the coefficient of multiple determination for multiple regression and is computed as

$$R^2 = 1 - \frac{SS_{Res.}}{SS_T}$$

Where $SS_T = \sum_{i=1}^{n}(y_i - \bar{y})^2\ and\ SS_{Res.} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

However R^2increases with increase in number of predictor variables hence adjusted R-squared is computed which compares the explanatory power of regression models that contain different numbers of predictors. The adjusted R-squared increases only if the new term improves the model more than would be expected by chance. The adjusted R-squared is defined as

$$R^2(adj) = 1 - \frac{SS_{Res.}/(n-p)}{SS_T/(n-1)}$$

Where n is number of observations and p is number of predictor variables. Examination of coefficient of determination may also help in detecting the linear relationship. The small magnitude of coefficient of determination indicates that there is non-linear relationship among response and predictor variables. In such cases non-linear regression analysis can be performed or the transformations such as log, square root, arcsine, inverse etc. can be used depending on the type of data to obtain linearity. Then the transformed data can be used for further analysis. Although there was no evidence of lack of fit the small value of the coefficient of determination is an indication of non-linear relationship. The non-linear relationship is converted into linear relationship by using log transformation for tur and cotton productivity. Then the analysis is carried out on the basis of transformed data.

RESULTS AND DISCUSSION

To determine which dry spell causes more effect or are critical to the productivity of selected crops soyabean and cotton; the bivariate correlation between productivity of cotton (y_c) and soyabean (y_s) with all twelve variables is obtained. SPSS software is used to test the significance of correlation. The results are listed in Table 1:

Table 1: Correlation between variables and response (y_c and y_s)

variables	Correlation with	
	y_c	y_s
X_1	0.091	0.249
X_2	0.356*	0.367*
X_3	-0.006	0.328*
X_4	0.240	-0.340*
X_5	0.238	0.190
X_6	0.337*	0.337*
X_7	0.095	0.307*
X_8	0.353*	-0.170
X_9	-0.436**	-0.188
X_{10}	-0.318*	-0.399**
X_{11}	-0.032	-0.314*
X_{12}	-0.135	0.120

**: correlation is significant at 0.01 level (2 tailed test) *: correlation is significant at 0.05 level (2 tailed test)

Table 1 indicates that Cotton production is significantly correlated with rainfall in the month of July, rainy days in the month of July and dry spells in the months of June and July. Production of soyabean is significantly correlated with rainfall in the month of June, July and August, also with number of rainy days in the month of July and August. It is also significantly correlated with dry spells in July and August. Regression model is fitted to the data of productivity from 2000-2001 to 2011-2012.

The two years data 2012-2013 and 2013-2014 was purposefully omitted which is used afterwards to validate the fitted model. Stepwise regression is used to select important predictors also the best subset of predictor variables is determined based on Mallow's Cp statisticusing MINITAB software. The fitted regression equation to forecast production of cotton is

$Yc = -0.244+0.410X_1+ 0.418X_2 + 0.230 X_3 + 0.094X_4- 0.495X_5 - 0.066X_6+0.167X_7+0.551X_8+ 0.0357X_9- 0.0435X_{10}+0.0438X_{11}-0.0618X_{12}$

The adequacy of the fitted model is checked using Coefficient of determination (R^2) and adjusted R^2. For the above model R^2= 69.9%which indicates that these twelve variables causes 70% variability in the production of cotton and other than these variables several factors affects the production of cotton crop that need to be considered for forecasting. The best subset of variables is selected using MINTAB; the best subset consists of five variables selected on basis of Cp mallows test statistic is X_1, X_2, X_3, X_5, and X_8. The regression equation using this subset of variables is

$Yc = -0.220 + 0.472 X_1+0.421X_2+0.273 X_3+0.741X_5- 0.650 X_8$

with R^2= 67.5%, R^2(adj) = 62.1%.

The regression equation to forecast production of Soyabeanis

$Ys = 1.47 + 0.812 X_1 + 0.425 X_2+ 0.109 X_3- 0.871 X_4 - 0.896 X_5 - 0.066 X_6 + 0.311 X_7$

$+ 1.13 X_8 - 0.0646 X_9- 0.107 X_{10} + 0.079 X_{11} + 0.212 X_{12}$

For the above model R^2= 68.2%. Which indicates that these twelve variables causes68.2% variability in the production of soyabean. This also indicates that there are other variables other than these 12 variables which affects the production of soyabean in Aurangabad division, they may be soil type, fertilizers used, precipitation, onset, the period of sowing, rainfall intensity etc. The best subset consists of five variables selected based on Cp mallows test statistic is X_1, X_2, X_3, X_5, and X_8 . The fitted regression equation using this subset of variables is:

$Ys = 2.02 + 0.526 X_1+ 0.326 X_2- 0.931 X_4- 0.540 X_5 + 0.374 X_7 + 1.22 X_8 + 0.166 X_{12}$

with R^2= 67.2%, R^2(adj) = 58.7%.which shows improvement in the model by reducing the insignificant variables.

CONCLUSIONS

Cotton production is significantly affected by dry spells in the month of June and July, while that of Soyabean is affected by dry spells in the month of July and August. In this study it is attempted to predict the productivity from the parameters of monsoon rainfall such as monthly rainfall, monthly rainy days and monthly dry spells in a season from June to September. Regression analysis results in coefficient of determination69.9% and 68.2%for cotton and soyabean production respectively, which may be due to the fact that it is simply harder to predict the production of oil seeds from the parameters under study. Regardless of the less R-squared, the significant coefficients of variables still represent the mean change in the response for one unit of change in the predictor while holding other predictors in the model constant.

REFERENCES

1. A.B. Mustapha (2012), Effect of Dryspell Mitigation with Supplemental Irrigation on Yield and Water Use Efficiency of Pearl Millet in Dry Sub-Humid Agro ecological Condition of Maiduguri, Nigeria, 2nd International Conference on Environment Science and Biotechnology.
2. V Prasanna (2014), Impact of monsoon rainfall on the total food grain yield over India, J. Earth Syst. Sci. 123, No. 5, July 2014, pp. 1129–1145
3. T S G Peiris and R R A Peries (1991-1993), effects of bimonthly rainfall on coconut yield in the low country intermediate zone of Sri Lanka, COCOS , 9 , 0 1 – 11
4. B.A. Sawa and A.A. Ibrahim(2011), Forecast Models for the Yield of Millet and Sorghum in the Semi Arid Region of Northern Nigeria Using Dry Spell Parameters, Asian Journal of Agricultural Sciences 3(3): 187-191, 2011
5. K. Krishna Kumar, K. Rupa Kumar, R. G. Ashrit, N. R. Deshpande, J. W. Hansen (2004) Climate impacts on Indian agriculture, International Journal of Climatology,Volume 24, Issue 11, pages 1375–1393
6. P.E. Naill M. Momani (2009), Time Series Analysis Model for Rainfall Data in Jordan: Case Study for Using Time Series Analysis, American Journal of Environmental Sciences 5 (5): 599-604
7. Seghatoleslami, M. J., Kafi, M., Majidi, E. (2008) Effect of drought stress at different growth stages on yield and water use efficiency of five proso millet (Panicummiliaceum L.) genotypes. Pak J Bot 40: 1427-1432.

Source of Support: None Declared
Conflict of Interest: None Declared