Multiple Regression Model For Optimization Of Yield Of Cotton In Rain Fed Zone Of Maharashtra, India

M. O. Wankhade, U. V. Kale

Abstract: This study relates to the optimization of yield of cotton with respect to variations in the monthly rainfall from June to September and the area under crop. Multiple regression model with interaction effects is fitted for the time series data of rainfall, area under cotton and yield of cotton from 1964 to 2018. The optimized solutions for yield of cotton with respect to various combination of monthly rainfall are suggested. Also an interaction effects between monthly rainfall, area under cotton and yield of cotton are studied.

Key Words: Multiple Regression, interactions, optimization, cotton yield, rain fed.

---

1 INTRODUCTION

Cotton is the most important cash crop of India and Maharashtra is the prominent cotton producing state of India especially Vidarbha and Marathwada regions. The renowned black cotton soil zone comprising of Vidarbha, Marathwada and Khandesh regions of Maharashtra produces about 75% of total cotton production of Maharashtra. The most of the area of Vidarbha and Marathwada being rain fed zone cotton is sown in the month of June-July and the temperature required for normal growth of cotton ranges from 20^0C to 40^0C. M. O. Wankhade, H. A. Bhosale[1] (2019), have made the comparison of linear, exponential growth curve and quadratic model for trend analysis of area, production and yield of cotton in Maharashtra and observed that quadratic trend model is the best fit as compared other models using the concept of accuracy measures viz. Mean Absolute Percent Error, Mean Absolute Deviation and Mean Square Deviation. Boreal Ecosystems Productivity Simulator for estimation of cotton yield in southern united states was proposed by Liming He, Georgy Mostovoy[2] 2019 and they found that 80 % of the data was explained by their model. V. Sellam and E. Poovammal[3] (2016) used Regression model for the analysis of the effects of various factors viz. area under crop, rainfall and prices of food, on the yield of crop and established the relationship between these factors using linear regression model. Quadratic, pure quadratic, linear, polynomial, interactions and stepwise linear regression model were adopted for estimation of yield of cotton, maize and wheat in India, by A. Shastry, H A Sanjay and E Bhanusree[4]. They studied the accuracy of the regression models on the use of R-square statistics and percentage error. S Bazgeer, Gh. Fadavi and S M Hossainy[5] (2014) tried to develop regression model for estimation of cotton yield in rain fed District Gharakhil of Iran. They proposed different regression models at two different stages viz. germination stage and squaring stage. Their study revealed that Exponential model are best at germination level while cubic regression model at squaring stage for estimation of yield of cotton. Patel, Amiksha A. & Kathiriya Dhaval R[6] (2017) studied effect of rainfall on cotton yield by data mining technique using Gaussian process algorithm in Sabarkantha agricultural region of Gujarat state, India. They observed that there is significant relationship between rainfall and yield of cotton. Raju Prasad Paswan, Shahin Ara Begum[7] (2013) reviewed artificial neural network and regression models for forecasting and prediction crop yields using different parameters. They observed that neural networks and statistical models together are more suitable for forecasts and prediction of yields. N. J. Rankja, S. M. Upadhyay, H. R. Pandya, B. A. Parmar and S. L.Varmora[8] developed pre-harvest forecast models for cotton yield considering various weather parameters in Banaskantha district of Gujarat. They suggested that temperature, humidity have significant effect on yield of cotton. Time series and cross-sectional statistical models were applied by David B. Lobell, Marshall B. Burke[9] (2010) for simulation of historical variability and prediction of the effect of climate changes on crop yield. Their study revealed that statistical models give accurate predictions of crop responses to climate changes. M. Sundar Rajan, M. Palanivel[10] compared various regression models to project the area under cotton, production and productivity cotton in India for 1951 to 2013. They stated that the third degree polynomial models are best for future trends of cotton in India as compared other regression models models, on the basis of R^2 statistics. Karim, M. R., M. A. Awal and M. Akter[11], (2010) focused on various deterministic models of time series viz. linear as well as nonlinear models for forecasting of wheat production in, Dmajpur, Rajshahi, and Rangpur districts of Bangladesh for the period 1971-72 to 2004-05.

2 METHODOLOGY

The monthly rainfall, particularly in June, July, August and September, has vital role for the growth and yield of cotton in rain fed zone of Maharashtra, especially Vidarbha and Marathwada regions. We focused to study the effect of monthly rainfall on cotton yield, interaction effects between monthly (June, July, August and September) rainfall and area under crop as well between monthly rainfall and yield of cotton by fitting Multiple Regression model. The same model is used for determining an optimum yield of cotton in Maharashtra using Minitab19 statistical software. The secondary data on monthly rainfall for the period 1964-2018 from Indian Institute of Tropical Meteorology, Earth System Science Organization Ministry of Earth Sciences Pune, India.

- *M. O. Wankhade, Dept. of Statistics,*
- *G. S. Sci., Arts & Comm. College, Khamgaon*
- *mail: mowankhade289@gmail.com*
- *. U. V. Kale Dept. of Statistics, Shri Shivaji Sci.*
- *College, Amravati, mail: uvkale@rediffmail.com*

The data relating to yield of cotton(in bales) and area under cotton (hectare) for the period 1964-2018 in Maharashtra is taken from cotton corporation and cotton advisory board of India.

2.1 Multiple Regression Model

Y_{ld}- Yield of cotton (Kg/hectare);
X_1- Rainfall in the month of June;
X_2- Rainfall in the month of July;
X_3- Rainfall in the month of August;
X_4- Rainfall in the month of September;
X_5- Area under cotton (in Lakh hectare)
The yield of cotton is function area under the crop, rainfall in June, July and August. Thus yield of cotton is expressed as;

$$Y_{ld} = f(X_1, X_2, X_3, X_4, X_5, X_1 * X_5, X_2 * X_5, X_3 * X_5) \quad ......(1)$$

Multiple regression model with interaction for the data under study is as;

$$Y_{ld} = \beta_0 + \sum_{i=1}^{5} \beta_i * X_i + \sum_{i=1}^{3} \gamma_{i5} * X_i X_5 \quad ....(2)$$

Where Y_{ld} is the dependent variable, Xi (i=1,2,3,4,5) the predictors and βi (i=0,1,2,3,4,5) are regression coefficients to be estimated.

	Rainfall				Cotton	
	JUN-AVG (in cm)	JUL-AVG (in cm)	AUG-AVG (in cm)	SEP-AVG (in cm)	Area Lakh hect.	Yield MS Kg/hect.
Year	X1	X2	X3	X4	X5	Yld
1964	138.55	240.15	274.3	208.55	28.24	80
1965	135.8	253	187.75	121.5	26.63	68
1966	66.55	350.45	149.7	204.5	26.11	74
1967	135.35	308.55	145.3	107.5	27.94	87
1968	89.35	254.75	136.2	220.25	27.17	88
1969	122	287.3	160	257.3	28.11	78
1970	283.1	184.35	419.75	214.45	28.12	31
1971	146.6	63.35	224.25	151.3	23.78	69
1972	120.3	103.75	158.55	65.35	25.31	75
1973	91.65	295.9	374.8	160.95	22.47	77
1974	102.85	177.85	219.75	107.45	25.02	112
1975	188.25	251.7	269.95	308.3	23.1	57
1976	99.9	300.55	218	113.65	21.2	67
1977	161.35	195.45	260.85	123.3	23.14	93
1978	217.35	287.15	237.15	65.55	25.09	90
1979	196.25	207.8	259.85	198.75	25.88	111
1980	224.25	194.65	349.65	95.8	26.67	81
1981	154.35	224.8	249.55	317.6	27.1	92
1982	86.4	264.7	134.25	185.3	26.48	103
1983	118.6	266.65	292.8	367.75	26.85	52
1984	66.45	191.15	144.6	87.35	26.85	95
1985	176	231.45	114.5	87.25	27.53	91
1986	160.6	219.75	278.3	99.7	26.92	78
1987	134.1	143.8	264.55	40.65	25.17	96
1988	176.25	355.4	316.25	362.6	26.27	107
1989	197.55	301.1	274.4	157.35	26.35	121
1990	269.35	230.45	447.95	139.65	27.3	93
1991	227.1	276.7	151.7	36.4	27.27	78
1992	203.3	132.4	363.4	146.5	24.8	141
1993	113.05	288.5	170.35	159.35	27.3	156
1994	167.15	316.05	219.95	155.7	27.6	98
1995	136.75	261.85	120.05	188.95	30.7	159
1996	46.85	225.7	229.5	229.45	30.85	182
1997	98.85	193.45	173.85	147.35	31.39	116
1998	164.25	252.1	279.1	255.95	31.99	141
1999	199.55	179.5	186.9	235.35	32.54	199
2000	253.65	285.4	280.5	55	30.77	101
2001	237.8	109.45	344.75	72.35	29.8	195
2002	304.05	88.1	285.8	111.85	28.01	158
2003	130.1	343.9	231.85	95.75	27.66	191
2004	110.75	199.05	153.2	126.4	28.4	311
2005	110.9	404.05	196	195.35	28.75	213
2006	154.25	244.95	342.45	188.95	31.07	274
2007	224.5	228.45	212.65	235.45	31.91	330
2008	108.3	179.8	234.65	272.5	31.33	336
2009	67.95	176.45	155.75	112.4	35.03	319
2010	142.1	345.75	310.05	189.65	39.42	379
2011	110.35	244.65	262	140.55	41.25	313.21
2012	75.15	262.6	183.1	180.6	41.46	332.13
2013	268	419.9	205.35	164.4	41.92	340.65
2014	62.9	227.05	197.15	155.75	41.9	324.58
2015	190.95	83.55	176.05	188.85	42.07	307.11
2016	163.35	386.35	110.4	313.55	38	395.92
2017	145.77	299.84	129.45	278.96	42.07	343.48
2018	99.78	195.89	174.51	110.02	41.19	334

Source:
1. Rainfall: Indian Institute of Tropical Meteorology, Earth System Science Organization Ministry of Earth Sciences Pune, India.
2. Area under crop and yield of cotton in Maharashtra: Cotton Advisory Board of India

3 Result and Discussion

Multiple regression model with first order interactions fitted by using Minitab19 software is;

$$Y_{ld} = -177 - 0.375 X_1 - 0.563 X_2 + 0.293 X_3 + 0.017 X_4 + 12.20 X_5 + 0.0067 X_1X_5 + 0.0182 X_2X_5 - 0.0091 X_3X_5 \quad ....(3)$$

ANOVA TABLE

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	8	446973	55871.6	15.6	0.000
Jun-AVG	1	705	704.7	0.20	0.660
JUL-AVG	1	4266	4266.3	1.19	0.281
AUG-AVG	1	580	580.1	0.16	0.689
SEP-Avg	1	85	85.1	0.02	0.878
Area MS	1	10724	10723.9	2.99	0.091
Jun-AVG*AreaMS	1	228	227.7	0.06	0.802
JUL-AVG*AreaMS	1	4399	4399.4	1.23	0.274
AUG-AVG*Area MS	1	490	489.9	0.14	0.713
Error	46	165025	3587.5		
Total	54	611998			

3.1 Factorial Plots for Yield MS

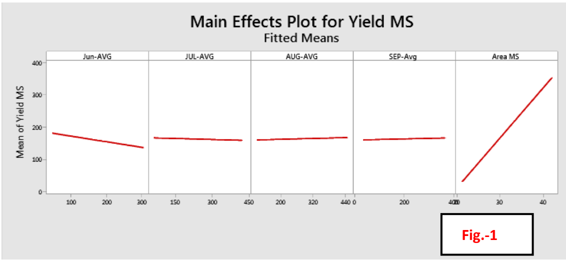

Fig.-1

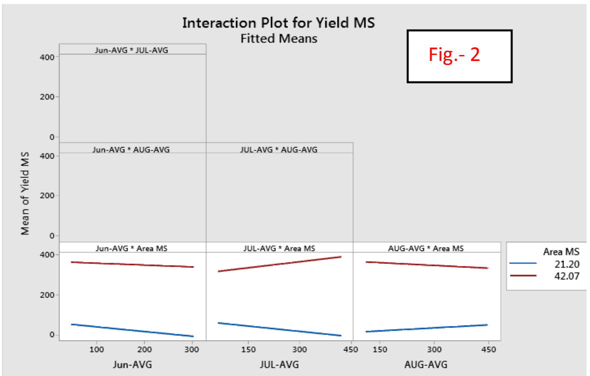

Fig.- 2

3.1.1 Parameters for Response Optimization Yield MS

Table 1

Response	Goal	Lower	Target
Yield MS	Maximum	31	395.92

Variable	Values	
	Min.	Max.
JUN-AVG	46.85	304.05
JUL-AVG	63.35	419.9
AUG-AVG	110.4	447.95
SEP-Avg	36.4	367.75
Area MS	21.2	42.07

3.1.2 Optimum Solutions-1

Table 2

Solution	Jun-AVG	JUL-AVG	AUG-AVG	SEP-AVG	Area MS	Yield MS Fit
1	46.850	419.900	110.400	367.75	42.07	413.437
2	46.850	419.900	110.400	36.400	42.07	407.653
3	237.482	419.900	110.400	367.75	42.07	395.920
4	237.482	419.900	110.400	367.75	42.07	395.920
5	46.850	416.599	112.156	367.75	42.07	412.609
6	46.850	416.599	112.156	36.702	42.07	406.831
7	237.482	419.900	110.400	367.75	42.07	395.920
8	237.482	419.900	110.400	367.75	42.07	395.920
9	237.482	419.900	110.400	367.75	42.07	395.920
10	46.850	333.327	110.400	367.75	42.07	395.920
11	46.850	333.327	110.400	367.75	42.07	395.920
12	46.850	333.327	110.400	367.75	42.07	395.920
13	46.850	333.327	110.400	367.75	42.07	395.920
14	46.850	419.900	303.355	367.75	42.07	395.920
15	46.850	419.900	303.355	367.750	42.07	395.920

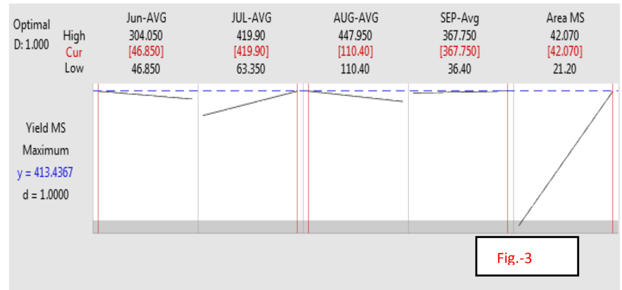

Fig.-3

3.1.3 Parameters for Response Optimization: Yield MS

Table 3

Response	Goal	Lower	Target
Yield MS	Maximum	31	395.92

Variable	Values	
	Min.	Max.
JUN-AVG	46.85	304.05
JUL-AVG	63.35	419.9
AUG-AVG	110.4	447.95
SEP-Avg	36.4	367.75
Area MS	21.2	37.2

3.1.4 Optimum Solutions-2

Table 4

Solution	Jun-AVG	JUL-AVG	AUG-AVG	SEP-AVG	Area MS	Yield MS Fit
1	46.850	419.900	110.400	367.75	37.200	320.200
2	46.850	419.900	110.400	367.75	37.200	320.200
3	175.45	250.800	110.400	367.75	37.200	284.928
4	46.850	63.368	110.400	367.75	37.200	279.641
5	46.850	63.350	349.150	367.75	37.200	268.573
6	46.850	63.350	389.675	367.75	37.200	266.694
7	46.850	63.350	110.401	367.75	35.622	259.662
8	46.850	63.350	110.401	36.400	35.992	258.569
9	46.850	419.895	447.950	367.75	34.173	255.926
10	238.084	63.350	110.400	367.75	37.200	255.794
11	46.850	419.892	447.950	36.401	34.523	255.762
12	241.403	63.350	110.400	367.75	37.200	255.381

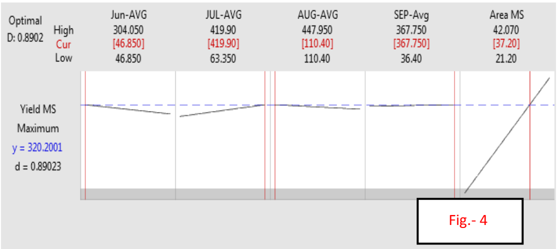

Fig.- 4

As the rainfall increases from 46.8 cm in June (Fig.-3, Fig.-4) the yield of cotton shows the downward trend. An increase in the rainfall in July shows an increase in the yield of cotton while increase in the rainfall in August results in decrease in the yield of cotton. There is no significant effect of rainfall in September on the yield of cotton. There is significant main and interaction effect (Fig.-1, Fig.-2) between the area under cotton, rainfall in June, July and August. An optimum solution for yield of cotton (Table-2, Table-4) under various parameters (Table-1, Table-3) are presented.

4 CONCLUSIONS:

We observed that multiple regression model with interaction effects is effective method to study of effect of rainfall in the various months from June, July, August September, and area under crop on the yield of cotton. The model showed that there is significant main and interaction effect of rainfall in June, July and August on the yield of cotton. Also the area under crop significantly effects the yield of cotton. We suggest that multiple regression model may be used for determining the optimum combinations of monthly rainfall, area under crop and yield of the crop.

REFERENCES:

[1] Wankhade, H. A. Bhosale (2019), "Trend Analysis With Respect To Area, Production And Yield of Cotton In Maharashtra", International Journal of Scientific & Technology Research Volume 8, Issue 09, September 2019

[2] Liming He, Georgy Mostovoy (2019),"Cotton Yield Estimate Using Sentinel-2 Data and an Ecosystem Model over the Southern US", Remote Sens. 2019, 11, 2000

[3] V. Sellam and E. Poovammal (2016), "Prediction of Crop Yield using Regression Analysis" Indian Journal of Science and Technology, Vol. 9(38), 10.17485/ijst/2016/v9i38/91714, October 2016.

[4] A. Shastry, H A Sanjay and E Bhanusree (2017) "Prediction of crop yield using regression Techniques", International Journal of Soft Computing 12(2), 96-102, 2017

[5] S Bazgeer, Gh. Fadavi and S M Hossainy (2014),"Statistical modeling for cotton yield estimation using agricultural indices, A case study of Gharakhil District in Mazandaran Province, Iran", Research Journal of Environmental Sciences, 8(2) 109-116, 2014.

[6] Patel, Amiksha A. & Kathiriya Dhaval R(2017), "Impact Of Rainfall On Cotton Crop Yield In Sabarkantha District Using Data Mining Techniques", AGRES, An International e. Journal (2017) Vol. 6, Issue 4:704-710

[7] Raju Prasad Paswan, Shahin Ara Begum (2013), "Regression and Neural Networks Models for Prediction of Crop Production", International Journal of Scientific & Engineering Research, Volume 4, Issue 9, September 2013

[8] N. J. Rankja, S. M. Upadhyay, H. R. Pandya, B. A. Parmar & S.L.Varmora (2010), "Estimation of cotton yield based on weather parameters of Banaskantha district in Gujarat state", Journal of Agrometeorology 12 (1), 47-52, June 2010

[9] David B. Lobell, Marshall B. Burke (2010), "On the use of statistical models to predict crop yield responses to climate change", Agricultural and Forest Meteorology xxx (2010) xxx–xxx.

[10] M. Sundar Rajan, M. Palanivel (2018), "Application of Regression Models for Area, Production and Productivity Growth Trends of Cotton Crop in India", International Journal of statistical Distributions and Applications, 2018; 4(1): 1-5.

[11] Karim, M. R., M. A. Awal and M. Akter, (2010), "Forecasting of wheat production in Bangladesh", Bangladesh Journal of Agriculture Research, 35: 17-28.

[12] Aiken, L. S., and S. G. West (1991). Multiple Regression: Testing and interpreting interactions. Newbury Park: Sage Publications

[13] Richard Williams, Interaction effects between continuous variables (Optional), University of Notre Dame, https://www3.nd.edu/~rwilliam/