International Journal of Chemical Studies

**P-ISSN: 2349–8528**
**E-ISSN: 2321–4902**
IJCS 2018; 6(6): 1111-1116

Received: 21-09-2018
Accepted: 24-10-2018

**KB Banakara**
PhD Scholar, Department of Agricultural Statistics, Navsari Agricultural University, Navsari, Gujarat, India

**Raj C Popat**
M. Sc Scholar, Department of Agricultural Statistics, Navsari Agricultural University, Navsari, Gujarat, India

**Amaresh**
M. Sc Scholar, Department of Agricultural Statistics, Applied Mathematics and Computer Sciences, University of Agricultural Sciences, Bengaluru, Karnataka, India

**HR Pandya**
Professor and Head, Department of Agricultural Statistics, Navsari Agricultural University, Navsari, Gujarat, India

**Correspondence**
**Amaresh**
M. Sc Scholar, Department of Agricultural Statistics, Applied Mathematics and Computer Sciences, University of Agricultural Sciences, Bengaluru, Karnataka, India

# Pre-harvest forecast of *Kharif* Rice using weather parameters in Bharuch district of Gujarat state

**KB Banakara, Raj C Popat, Amaresh and HR Pandya**

**Abstract**
Indian economy mainly depends on agriculture sector as it account 18 per cent of national GDP and it is most important occupation for most of the Indian families so; it is the soul of Indian economy. Rice is the most important staple food in India which play crucial role in daily requisite of diet. In the current study statistical crop modeling was engaged to provide forecast in advance. In this paper Multiple Linear Regression (MLR) Technique and Principal Components (PC) were derived for estimating average rice production for the Bharuch of district in south Gujarat. The weather indices were developed using correlation coefficient as weight to weekly weather parameters for the years from 1990 to 2012. The cross authentication of the developed forecast model were confirmed using data of the years 2013 to 2016. It was observed that value of *Adj.* $R^2$ has varied from 72.10 to 80.80 in different models. The study discovered that high value of *Adj.* $R^2$ was obtained in the model and which indicated that it was appropriate forecast model than other models. Based on the outcomes in Bharuch district, MLR techniques found to be better than PCA for pre harvest forecasting of rice crop yield.

**Keywords:** weather indices; MLR techniques; PCA; forecast

## 1. Introduction

Agriculture plays important role in Indian economy in which it is possible to cultivate large number of principal crops. Rice is the most important staple food in Asia. More than 90.00 per cent of the world's rice is grown and consumed in Asia, where 60.00 per cent of the world's population lives. India ranks second with 154.6 million tonnes of paddy (FAO, 2015). In the Gujarat state, rice occupies about 7.00 to 8.00 per cent of the gross cropped area of the state and accounts for around 14.00 per cent of the total food grain production. About 90.00 per cent of area under rice is confined to South and middle Gujarat.

The crop weather relationship has been studied by Fisher (1924) [6] and Hendricks and Scholl (1943) [10] studied crop weather relationship and pioneer to this research at Indian Agricultural Statistic Research Institute, New Delhi. They developed models which required small number of parameters to be estimated while taking care of distribution pattern of weather over the crop season. Agrawal *et al.* (1980) [2] and Jain *et al.* (1980) [2] modified this model by expressing effects of changes in weather parameters on yield in the particular week as second degree polynomial in respective correlation coefficients between yield and weather parameters. This model was further modified (Agrawal *et al.* 1986, 2011) [1] by explaining the effects of changes in weather parameters on yield in particular week using correlation as weight using linear function. Patel *et al.* (2007) [14], Chauhan *et al.* (2009) [4], Garde *et al.* (2012) [9], Mahdi *et al.* (2013), Singh *et al.* (2014) and Pandey *et al.* (2015) studied the relationship of weather parameters and rice crop yield in different regions of world. Varmola *et al*. (2004) [19], Agarwal *et al*. (2012) [3] Sisodia *et al*. (2014) [18] and Garde *et al*. (2015) [8] developed forecast models for Wheat crop in different regions of India. Similarly, for pigeon pea Kumar *et al.* (1999) [11] and Sarika *et al.* (2011) [16], for Sugarcane Priya and Suresh and for Ground nut Dhekale *et al.* (2014) [5] developed models.

The development of forecasting models for rice is very important, pre-harvest forecast needed in policy decision regarding export and import, food procurement and distribution, price policies and exercising several administrative measures for storage and marketing of agricultural commodities. Thus, the use of statistical models in forecasting food production and prices for agriculture hold great significance. Although no statistical model can help in forecasting the values exactly but by knowing even approximate values can help in formulating future plans.

## 2. Materials and Methods

The present study was carried out in the Bharuch district of South Gujarat. Considering the specific objectives of the study, *kharif* rice yield data were collected from the Directorate of Economics and Statistics, Government of Gujarat, Gandhinagar, Gujarat from 1990 to 2016 and the weekly weather data was collected from the Department of Agrometeorology, Agricultural University, Navsari. The maximum temperature ($X_1$), minimum temperature ($X_2$), Morning relative humidity ($X_3$), Evening relative humidity ($X_4$), and total rain fall ($X_5$) were considered for studying the effect on kharif rice grain yield. The weekly weather data related to Kharif rice crop season starting from a first fortnight before sowing to last of reproductive stage were utilized for the development of statistical models. Therefore for the each year weather data, from May-June (23rd standard meteorological week, SMW) to October (40st standard meteorological week, SMW) were utilized for *kharif* rice crop.

### 2.1 Statistical methodology

### 2.1.1 Multiple Linear Regression models (MLR)

The MLR models were developed using weather indices (Agrawal *et al.* 1986, 2011) [1], in this method, weekly data on weather variables of 17 weeks have been utilized for constructing weather indices (Weighted & un-weighted along with their interactions).

#### 2.1.1.1 Development of weather indices for yield forecasting method-1

$$Z_{ij} = \sum_{w=1}^{m} r_{iw}^{j} X_{iw} \quad \text{and} \quad Z_{ii',j} = \sum_{w=1}^{m} r_{ii'w}^{j} X_{i'w}$$

Where,

$J$ = 0, 1, 2 (where, '0' represents un-weighted indices, '1' represents weighted indices and '2' indicates weighted square indices)

$M$ = Week up to forecast ($m=20^{th}$)

$W$ = week number (1, 2, ...., $m$)

$r_{iw}$ = Correlation coefficient between adjusted crop yield and $i^{th}$ weather variable in $w^{th}$ week

$r_{ii'w}$ = Correlation coefficient between adjusted crop yield and the product of $i$ and $i'^{th}$ weather variable in $w^{th}$ week

$X_{iw}$ and $X_{i'w}$ are the $i$ and $i'^{th}$ weather variable in $w^{th}$ week respectively

The pre-harvest forecast models were obtained by applying the MLR techniques by taking predictors as appropriate un-weighted and weighted weather indices. Stepwise regression analysis was used for selecting significant variables (Draper and Smith 1981; Gomez and Gomez 1984). The regression model was as follows:

**Model-1**

The model was developed using 30 weather indices (15 unweighted and 15 weighted indices) as a independent variable and crop yield as dependent variable. The developed model is given as

$$Y = A_0 + \sum_{i=1}^{p}\sum_{j=0}^{1} a_{ij} Z_{ij} + \sum_{i \neq i'=1}^{p}\sum_{j=0}^{1} a_{ii'j} Z_{ii'j} + cT + \varepsilon$$

Where,

$Z_{ij}$ and $Z_{ii'j}$ are the weather indices

$i,i'$ = 1, 2, …$p$

$p$ = Number of weather variables under study

$Y$ = District total crop yield (kg/ha)

$T$ = Year number (trend parameter)

**$A_0$ is the intercept**

$a_{ij}$ and $a_{ii'j}$, $c$ are the regression coefficient $e$ is error term normally distributed with mean zero and constant variance

### 2.2 Principal component analysis

Principal Component Analysis (PCA) is a multivariate statistical technique which reduces data with large number of correlated variables into substantially smaller set of new variables through linear combination of the variables that accounts for total variation present in the original variables. The linear combination of variables are called principal components and estimated using correlation or covariance matrix. When the variables measured with different units, scale effects can influence the composition of derived components. In this case it is necessary to standardize the original variables. Therefore correlation matrix is considered to better as it does not require standardization.

**Method-1**

In this model 30 weather indices (unweighted and weighted, $j$=0, 1) were utilized to develop principal components. The number of components retained using scree plot and Kaiser's Criterion. The developed components were utilized for model development using regression analysis.

**Model-2**

$Y = \beta_0 + \beta_1 T + \beta_2 PC_1 + \beta_3 PC_2 + \beta_4 PC_3 + \ldots + \beta_{k+1} PC_k + \varepsilon$

Where,

Y is the rice crop yield

T is the trend variable

$\beta_0, \beta_1, \beta_2, \beta_3 \ldots \beta_{k+1}$ are the regression coefficients

$PC_1, PC_2, PC_3 \ldots PC_k$ are principal components

$\varepsilon$ is the error term

### 2.3 Comparison and validation of models

The comparisons and validation of models were done using following approaches

**1. Forecast error (%)**

The validation of the model using observed yield ($O_i$) and forecasted yield ($E_i$) was computed using below formula,

$$\text{Forecast Error} = \left[\frac{O_i - E_i}{O_i}\right] \times 100$$

**2. Coefficient of multiple determination (Adjusted $R^2$)**

The best fitted model among developed models were decided based on highest value of Adjusted $R^2$

$$R_{adj}^2 = 1 - \frac{SS_{res}/(n-p)}{SS_t/(n-1)}$$

Where,
$ss_{res}/(n-p)$ is the residual mean square
$ss_t/(n-1)$ is the total mean sum of square.

**3. Root mean square error (RMSE)**
The cross validation of the model were done using RMSE, for the year 2012 to 2016 using observed yield ($O_i$) and forecasted yield ($E_i$) was computed using below formula

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(O_i - E_i)^2\right]^{1/2}$$

## 3. Results and Discussion
### 3.1 Association between Rice Yield and Weather Parameters
The associations between rice yield and week wise weather parameters were studied by using Karl Pearson correlation coefficient (Table 1). The main aim was to know strength between rice yield and weekly weather parameters.
Positively significant correlations were observed between rice yield and some of the weekly weather parameters *viz.* maximum temperature ($42^{nd}$ SMW of the cropping season), Morning Relative Humidity (27 and $32^{nd}$ SMW) and Evening Relative Humidity (32 and $37^{th}$ SMW) and negatively significant correlation coefficient were observed between rice yield and some of the weekly weather parameters *viz.* minimum temperature ($27^{th}$ SMW). However, during cropping season of rice (*i.e.* $2^{nd}$ July to $14^{th}$ October) the other week wise correlation coefficient between the yield and weather parameters found to be non-significant. The value of '*r*' varies from -0.47 to 0.51, indicating that individual character does not explain more than 51.00 per cent variation in the yield. This suggests that simple regression using single weather parameter is not adequate to forecast the yield. It is necessary to utilize all weather parameters simultaneously. It is done by constructing un-weighted indices and weighted indices.

### 3.2 Multiple linear regression models (MLR)
Based on strategies followed in model 1, the obtained forecast model equations are given in Table 2**.** The Table 2 observed that the values of adjusted $R^2$ for different models were varied from 76.20 per cent (model $A_1$) to 80.80 per cent (model $A_6$). Based on highest value of adjusted $R^2$ model $A_6$ was selected as a best model among developed six models which found to be appropriate in the 40 SMW (ripening phase of rice crop) *i.e.* five weeks before the harvest of crop. The model showed 80.80 per cent variation accounted due to weather indices $Z_{150}$ *i.e.* interaction of maximum temperature and rainfall and trend variable *T.*

### 3.3 Principal component Analysis
Based on different methods principal components were developed. Number of components retained based on scree plot and Kaiser's criterion. Based on Kaiser's criterion 35, 37, 38, and 39 have seven principal components and 36 and 40 have six principal components (Table 3 & Figure 2). These principal components were utilized for the development of model by stepwise MLR analysis.
Based on strategies followed in model 2, the obtained forecast model equations are given in Table 4. The Table 4 observed that the values of adjusted $R^2$ for different models were varied from 72.10 per cent (model $B_1$, $B_2$ and $B_3$) to 79.00 per cent (model $B_6$). Based on highest value of adjusted $R^2$ model $B_6$ was selected as a best model among developed six models which found to be appropriate in the 40 (ripening phase of crop) SMW *i.e.* five weeks before the harvest of crop. The model showed 79.00 per cent variation accounted due to principal components $PC_1$ and trend variable *T.*

### 3.4 Comparison of MLR and PCA models
Comparison between MLR and PCA models was carried out by using Adj. $R^2$. The comparison of selected best fit models was done by forecast error and RMSE. The details of comparative study are given in Table 5 and the graphical representation is given in Figure 3.
It observed from Table 5 that, the value of adjusted $R^2$ for both models 79.00 to 80.80 and a value of RMSE 214.19 to 272.41, respectively. The forecast error per cent varies from 1.80 to 16.80. However, model $A_1$ was selected as best fit model based on highest Adj. $R^2$ and lower RMSE and Forecast Error. Therefore pre-harvest forecasting was done using model $A_1$ in the $40^{th}$ SMW *i.e.* five weeks before harvest of the rice crop. MLR model considered as best fit model for Bharuch district as compared to PCA.

**Table 1:** Week wise correlation coefficient between rice yield and weather parameters for Bharuch district

| SMW | $T_{max}(X_1)$ | $T_{min}(X_2)$ | MRH($X_3$) | ERH($X_4$) | RAINFALL ($X_5$) |
|---|---|---|---|---|---|
| 23 | -0.21 | -0.07 | -0.36 | -0.07 | -0.03 |
| 24 | 0.08 | -0.02 | 0.17 | -0.12 | -0.08 |
| 25 | 0.08 | -0.10 | 0.20 | -0.07 | -0.08 |
| 26 | -0.07 | -0.17 | 0.29 | 0.19 | 0.11 |
| 27 | -0.33 | -0.47* | 0.46* | 0.32 | 0.30 |
| 28 | -0.31 | -0.18 | 0.17 | -0.09 | -0.20 |
| 29 | 0.36 | 0.14 | -0.12 | -0.17 | -0.25 |
| 30 | 0.09 | -0.17 | 0.37 | 0.17 | 0.18 |
| 31 | -0.02 | -0.32 | 0.34 | 0.19 | 0.20 |
| 32 | -0.31 | -0.19 | 0.51* | 0.45* | 0.33 |
| 33 | -0.04 | -0.07 | 0.26 | 0.17 | -0.04 |
| 34 | 0.19 | 0.09 | -0.03 | -0.10 | 0.23 |
| 35 | -0.15 | -0.11 | 0.30 | 0.40 | 0.14 |
| 36 | 0.06 | 0.10 | 0.36 | 0.34 | -0.10 |
| 37 | -0.21 | 0.40 | 0.35 | 0.44* | 0.19 |
| 38 | -0.24 | 0.02 | 0.40 | 0.41 | 0.33 |
| 39 | -0.21 | -0.03 | 0.25 | 0.15 | -0.16 |

| | | | | | |
|---|---|---|---|---|---|
| 40 | -0.03 | 0.15 | 0.20 | 0.24 | 0.10 |
| 41 | -0.07 | -0.02 | 0.14 | -0.08 | -0.18 |
| 42 | 0.50* | -0.10 | -0.25 | -0.29 | -0.12 |

**Table 2:** Rice yield forecasting model-1 equations

| Model | WEEK | MODEL | Adj. $R^2$ |
|---|---|---|---|
| $A_1$ | 35 | $Y$=-3071.37+42.10$T$+3.42$Z_{30}$ | 76.20 |
| $A_2$ | 36 | $Y$=-3182.29+41.43$T$+3.27$Z_{30}$ | 76.60 |
| $A_3$ | 37 | $Y$=365.67+46.26$T$+0.01$Z_{150}$ | 76.70 |
| $A_4$ | 38 | $Y$=344.28+45.86$T$+0.01$Z_{150}$ | 78.30 |
| $A_5$ | 39 | $Y$=282.28+47.34$T$+0.01$Z_{150}$ | 80.50 |
| $A_6$ | 40 | $Y$=275.89+47.37$T$+0.01$Z_{150}$ | 80.80 |

**Table 3:** Per cent variation and cumulative variation of PCA, model-2

| SMW | 35 | | 36 | | 37 | |
|---|---|---|---|---|---|---|
| **PCs** | **% VAR** | **CUM VAR** | **% VAR** | **CUM VAR** | **% VAR** | **CUM VAR** |
| PC1 | 37.68 | 37.68 | 40.93 | 40.93 | 42.34 | 42.35 |
| PC2 | 24.15 | 61.83 | 20.90 | 61.82 | 20.18 | 62.53 |
| PC3 | 12.17 | 74.02 | 10.90 | 72.73 | 10.31 | 72.83 |
| PC4 | 8.32 | 82.33 | 7.88 | 80.61 | 7.94 | 80.77 |
| PC5 | 5.46 | 87.80 | 6.26 | 86.88 | 6.31 | 87.08 |
| PC6 | 4.34 | 92.13 | 5.75 | 92.62 | 5.76 | 92.84 |
| PC7 | 3.46 | 95.59 | | | 3.33 | 96.18 |
| SMW | 38 | | 39 | | 40 | |
| **PCs** | **% VAR** | **CUM VAR** | **% VAR** | **CUM VAR** | **% VAR** | **CUM VAR** |
| PC1 | 39.02 | 39.02 | 37.18 | 37.18 | 40.19 | 40.19 |
| PC2 | 23.25 | 62.27 | 25.47 | 62.64 | 25.38 | 65.58 |
| PC3 | 9.70 | 71.97 | 9.67 | 72.32 | 11.87 | 77.44 |
| PC4 | 8.76 | 80.73 | 9.31 | 81.63 | 8.45 | 85.90 |
| PC5 | 6.47 | 87.20 | 6.33 | 87.96 | 5.09 | 90.99 |
| PC6 | 5.31 | 92.52 | 5.42 | 93.38 | 4.02 | 95.00 |
| PC7 | 4.00 | 96.53 | 3.38 | 96.76 | | |

**Table 4:** Rice yield forecasting model-2 equations

| Model | WEEK | MODEL | Adj. $R^2$ |
|---|---|---|---|
| $B_1$ | 35 | $Y$=593.93+46.26$T$ | 72.10 |
| $B_2$ | 36 | $Y$=593.93+46.26$T$ | 72.10 |
| $B_3$ | 37 | $Y$=593.93+46.26$T$ | 72.10 |
| $B_4$ | 38 | $Y$=583.46+47.13$T$+83.38$PC_1$ | 76.40 |
| $B_5$ | 39 | $Y$=566.48+48.54$T$+99.68$PC_1$ | 78.70 |
| $B_6$ | 40 | $Y$=565.96+48.59$T$+101.48$PC_1$ | 79.00 |

**Table 5:** Comparison of MLR and PCA models

| MODEL | SMW | Year | Observed Yield | Forecast Yield | Forecast Error | RMSE | Adj. *R2* |
|---|---|---|---|---|---|---|---|
| $A_6$ | 40 | 2013 | 1890 | 1807 | 4.39 | 214.19 | 80.80 |
| | | 2014 | 2212 | 1840 | 16.80 | | |
| | | 2015 | 2083 | 1890 | 9.27 | | |
| | | 2016 | 1988 | 1952 | 1.80 | | |
| $B_6$ | 40 | 2013 | 1890 | 1573 | 16.77 | 272.41 | 79.00 |
| | | 2014 | 2212 | 1947 | 12.00 | | |
| | | 2015 | 2083 | 1782 | 14.46 | | |
| | | 2016 | 1988 | 1800 | 9.44 | | |

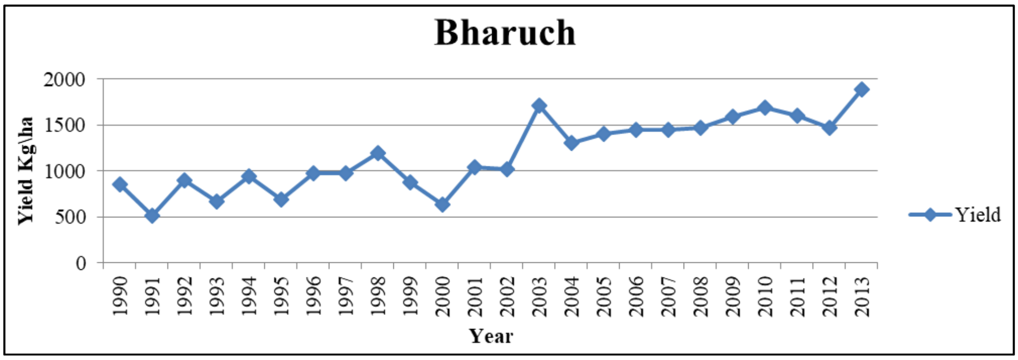
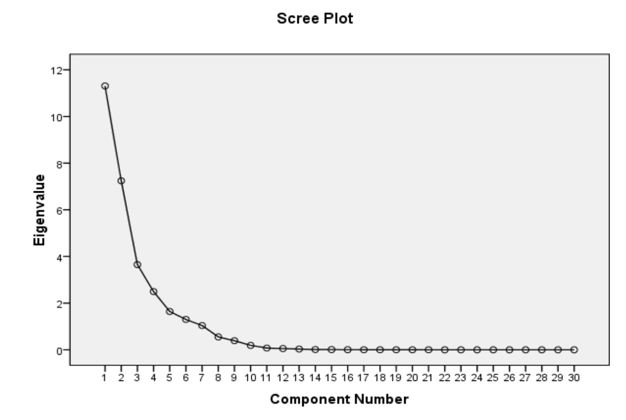
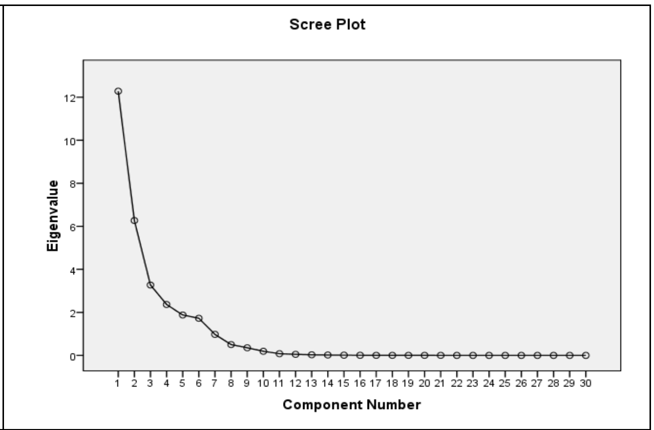


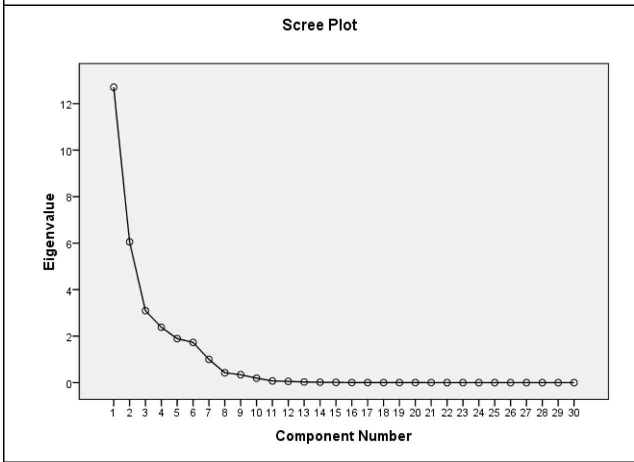

**Fig 1**: Trend of rice yield in Bharuch district



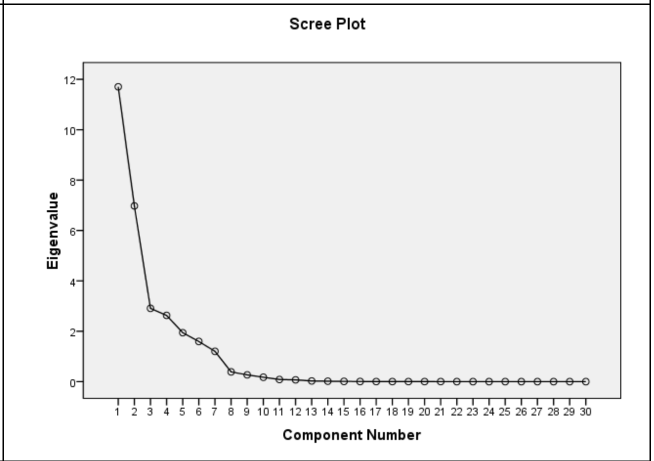

SMW 35



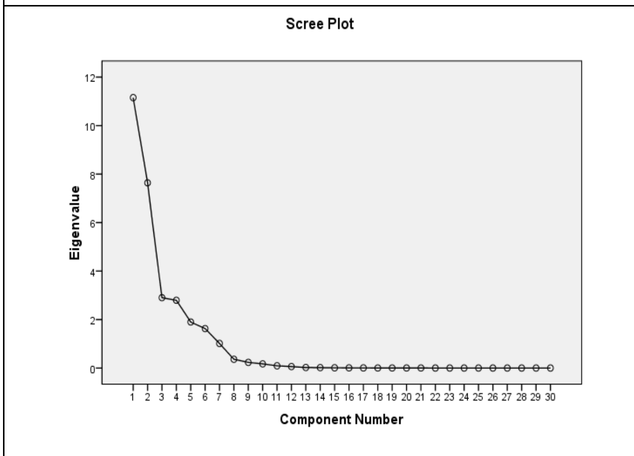

SMW 36



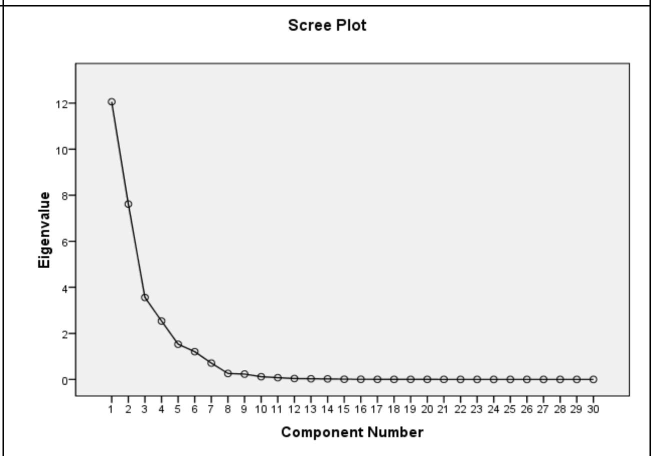

SMW 37




SMW 38




SMW 39




SMW 40

**Fig 2:** Scree plots of model-2

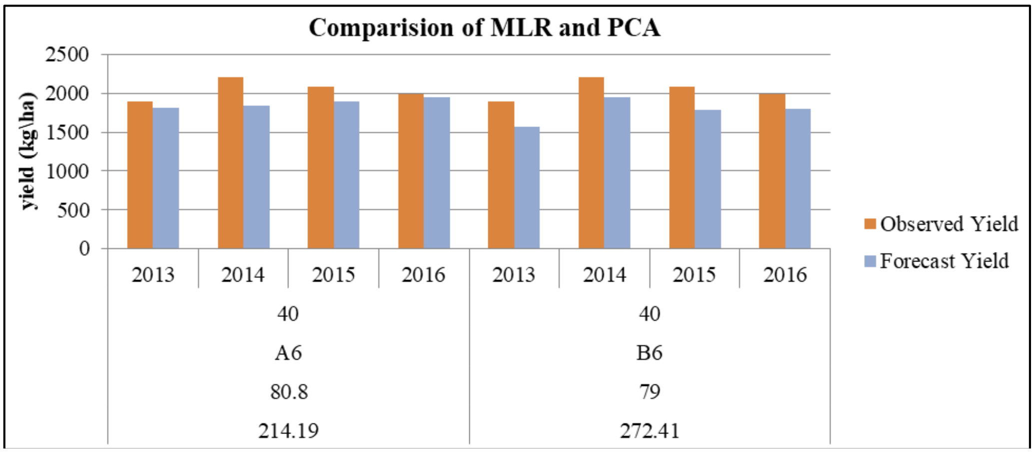



**Fig 3:** Graphical representation of observed yield and forecast yield for MLR and PCA

## 4. Conclusion

Using the forecast techniques like MLR and PCA, pre-harvest estimates of rice crop yield for Bharuch district could be computed successfully five weeks earlier, before the actual harvest *i.e.* during ripening phase of the crop period. It can be concluded from the results that there is a wide scope for using alternative approaches to develop predictors that could be used in forecasting models for reliable and dependable forecast. Therefore, it is important to continue research on these aspects for various other crops also on a continuous basis. This methodology can be applicable in many crops *viz.* rice, pulses, oil seeds, sugarcane etc. to develop pre-harvest forecasting models and these forecasts have significant value in agricultural planning and policy making.

## 5. References


1. Agrawal R Chandrahas, Kaustav A. Use of discriminant function analysis for forecasting crop yield. Mausam. 2011; 63(3):455-458.
2. Agrawal R, Jain RC, Jha MP. Modes for studying rice-weather relationship. Mausam. 1980; 37(1):67-70.
3. Agrawal R Chandrahas, Aditya K. Use of discriminant function analysis for forecasting crop yield. Mausam. 2012; 63(3):455-458.
4. Chauhan VS, Shekh AM, Dixit SK, Mishra AP, Kumar S. Yield prediction model of rice in Bulsar district of Gujarat. Journal of Agrometeorology. 2009; 11(2):162-168.
5. Dhekale BS, Mahdi S, Sawant PK. Forecast models for groundnut using meteorological variables in Kolhapur, Maharashtra. Journal of Agrometeorology. 2014; 16(2):238-239.
6. Fisher RA. The influence of rainfall on yield of wheat at Rothamsted. Philosophical Transaction of Royal Society of London, Series B. 1924; 213:89-142.
7. Food and Agriculture Organization. Rice market monitor, report. 2015; 18(2):2-6.
8. Garde YA, Dhekale BS, Singh S. Different approaches on pre harvest forecasting of wheat yield. Journal of Applied and Natural Science. 2015; **7**(2):839-843.
9. Garde YA, Shukla AK, Singh S. Pre-harvest forecasting of rice yield using weather indices in Pantnagar. International Journal of Agricultural Statistical Science. 2012; 8(1):233-241.
10. Hendricks WA, Scholl JC. Technique in measuring joint relationship. The joint effect of temperature and precipitation on corn yield. N. C. Staff Agricultural Experimental Techniques Bulletin, 1943, 74.
11. Kumar R, Gupta BRD, Athiyaman B, Singh KK, Shukla RK. Stepwise regression technique to predict Pigeon pea yield in Varanasi district. Journal of Agrometeorology. 1999; 1(2):183-186.
12. Mahdi SS, Lotus S, Singh G, Ahmad L, Singh KN, Dar LA, *et al*. Forecast of rice (*Oryza sativa* L.) yield based on climatic parameters in Srinagar district of Kashmir Valley. Journal of Agrometeorology. 2013; 15(1):89-90.
13. Pandey KK, Rai VN, Sisodia BVS, Singh SK. Effect of Weather Variables on Rice Crop in Eastern Uttar Pradesh, India. Plant Archives. 2015; 15(1):575-579.
14. Patel GB, Vaishnav PR, Patel JS, Dixit SK. Pre-harvest forecasting of rice (*Oryza Sativa* L.) yield based on weather variables and technological trend. Journal of Agrometeorology. 2007; 9(2):167-173.
15. Priya SRK, Suresh KK. A study on pre-harvest forecast of sugarcane yield using climatic variables. Statistics and Applications. 2009; 7&8(1&2):1-8.
16. Sarika, Iquebal MA, Chattopadhyay. Modelling and forecasting of pigeonpea (*Cajanus cajan*) production using autoregressive integrated moving average methodology. Indian Journal of Agricultural Sciences. 2011; 81(6):520-523.
17. Singh RS, Patel C, Yadav MK, Singh KK. Yield forecasting of rice and wheat crops for eastern Uttar Pradesh. Journal of Agrometeorology. 2014; 16(2):199-202.
18. Sisodia BVS, Yadav RR, Kumar S, Sharma MK. Forecasting of pre-harvest crop yield using discriminant function analysis of meteorological parameters. Journal of Agrometeorology. 2014; 16(1):121-125.
19. Varmora SL, Dixit SK, Patel JS, Bhatt HM. Forecasting of wheat yield on the basis of weather variables. Journal of Agrometeorology. 2004; 6(2):223-228.