

# Development of pre harvest forecasting model for Chhattisgarh plain Zone on Pigeonpea by Principal Component Analysis



**Gaind Lal\* and Pandey KK**

*Department of Agril Statistics, Indira Gandhi Agriculture University, India*

**Submission:** September 24, 2017; **Published:** October 02, 2017

**\*Corresponding author:** Gaind Lal, Department of Agril Statistics, Indira Gandhi Agriculture University, India, Email: [hirwani92@gmail.com](mailto:hirwani92@gmail.com)

## Abstract

The present investigation deals with the model have been also developed by Principal Component Analysis on districts level and zone level as well. Models fitted with 2 Principal Components (PC1 and PC2) and Time trend (T). Models are highly significant and R<sup>2</sup> value 69% for Pigeonpea and for CG plain Zone is also highly significant at 0.1% level of significance. The present investigation covers under the study of individual effect of weather variables, joint effect of weather variables forecasting model developed through stepwise regression technique following the concept of older studies by the different scientists in this direction for many crops. The principal component analysis has been also taken for the development of forecasting model. But in this section only Principal Component Analysis covered for the model development.

**Keywords:** Pigeonpea; Coefficient of determination; Principal component Analysis

## Introduction

Forecast of the crop production at suitable stages of crop period before the harvest are vital for rural economy. On the other hand, forecasts of crop yields are important for advance planning formulation and its implementation. Forecasting is also vital for crop procurement, distribution, price structure and import export decisions etc. These are useful to farmers to decide in advance their future prospects and course of action. Thus, reliable and timely pre-harvest forecasting of crop yield is very important. To meet such needs, crop forecasts under the prevalent system in India are being issued by the Directorate of Economics and Statistics, Ministry of Agriculture, New Delhi. The final estimates based on objective crop-cutting experiments are of limited utility as those become available quite latter after the crop harvest. The Statistical techniques employed for forecasting purposes should be able to provide objective, consistent and comprehensible forecasts of crop yield with reasonable precisions well in advance before the harvest. Several studies have been carried out to forecast crop yield using weather parameters etc. These are useful to farmers to decide in advance their future prospects and possible course of action. Thus, reliable and timely pre-harvest forecasting of crop yield is very important. In statistical model approach, one or several variables (representing weather or climate) are related to crop

responses such as yield and yield contributing characters. Therefore, there is a need to develop area specific forecasting models based on time series data to help the policy makers for taking effective decisions to counter adverse situations in food production. The forecasting of crop yield may be done by using three major objective methods

- a. Biometrical characteristics
- b. Weather variables and
- c. Agricultural inputs [1].

## Material and Methods

The present investigation covers under the study of individual effect of weather variables, joint effect of weather variables forecasting model developed through stepwise regression technique following the concept of older studies by the different scientists in this direction for many crops. The principal component analysis has been also taken for the development of forecasting model. Time series data on yield for 25 years (1990-91 to 2014-15) for 4 Districts, Seventeen year for 5 districts has been procured from Directorate of Agricultural, Govt. of Chhattisgarh. The weekly meteorological data (1990-91 to 2014-15) procured from IGV, Raipur for Chickpea and Pigeon pea crop period data.

PCA is a multivariate technique for data reduction. It is a mathematical function, which does not require user to specify the statistical model or assumption about distribution of original variables [2]. Let  $X_{ij}$  be the value of  $j$ th biometrical character ( $j=1, 2, \dots, p$ ) corresponding to  $i$ th varieties of experiment ( $i=1, 2, \dots, n$ ). The PCA for  $x_{ij}$ 's will be carried out. Let PC1, PC2, ..., PC $k$  be first  $K$  ( $K < P$ ) principal components explaining variability about more than 70 to 90 percent of the total variation in  $x_{ij}$ 's. Using these  $K$  principal components as regressed variables and varieties yield ( $y_i$ ) as regressed, the following linear multiple regression model for pre-harvest forecast of crop yield has been proposed.

$$Y_i = \beta_0 + \beta_1 PC1_i + \beta_2 PC2_i + \dots + \beta_k PCk_i + e_i, i = 1, 2, \dots, n.$$

Where  $Y_i$  is the crop yield of the  $i$ th plot;  $\beta_0, \beta_1, \beta_2, \beta_k$  are model parameters and  $e_i$  is error term assumed to follow independently normal distribution with mean 0 and variance  $\sigma^2$ . The aforesaid model is fitted with the data by least square technique.

**a) Measures for validation and comparison of the models:**

Two procedures have been used for the comparison and the validation of the developed models. These procedures are given below. Coefficient of determination ( $R^2$ ) and ( $R^2_{adj}$ ). The models were validated on the basis of ( $R^2$ ) and ( $R^2_{adj}$ ) which can be computed from the formula given by Drapper and Smith (1988). Adjusted  $R^2$  is given by the following formula,

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

**Table 1:** Pre-harvest Forecast Model through PCA for Plain zone on pigeonpea.

Variables	Coefficient	Standard error	P-Value	Significance ( $R^2$ )	$R^2$ (Adj)
Intercept	-60.02	457.6254		0.001	0.69***
PC1	-2244.8+	1169.402	0.06		
PC2	-4493.58*	1903.943	0.02		
T	-22.50***	4.60818	0.001		

\*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , + $P < 0.1$ .

**Table 2:** Validation of the model for Raipur district on pigeonpea.

Year	Actual rice yield(q/ha)	Predicted rice yield(q/ha)	Adjusted $R^2$ (%)	Percent deviation
2010	384.67	323.1569	0.69***	15.99
2011	416.22	397.9641		4.39
2012	546.11	310.4661		43.15
2013	481.62	544.3523		-13.02
2014	567.10	367.8191		35.14

\*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , + $P < 0.1$ .

Where  $ss_{res}/(n-p)$  is the residual mean square and  $sst/(n-1)$  is the total mean square. The total mean square is constant regardless of how many variables are in the model. On adding a regressed in the model Adjusted  $R^2$  increases only if the addition of the regressed reduces the residual mean square. It also penalizes for adding terms that are not helpful, so it is very important in evaluating and comparing the regression models.

**b) Percent Deviation:-**

The formula for computation of Percent Deviation of forecast yield from actual yield is given by [3,4]. This measures the deviation (in percentage) of forecast from the actual yield data. The formula for calculating the percent deviation of forecast is given below:

$$\text{Percentage deviation} = \frac{(\text{actual yield} - \text{Forecasted yield})}{\text{actual yield}} \times 100$$

**Results and Discussion**

The Principal Component PC1, PC2 and T fitted finally in the model in given below:  $Y = -60.02 - 2244.8 PC1 - 4493.58 PC2 - 22.50 T$ . The results given in Table 1 that all variables (PC1, PC2 and T), all the variables are found significant [5-8]. The coefficient of determination adjusted ( $R^2$ ) has been found to be 85 % which is significant at 0.01% level. Table 2 Indicates that the value of adjusted  $R^2$  good enough i.e. 85% (Figures 1 & 2). Deviation from yield indicates that the model is best fit and it has high power to pre-harvest forecast. It means model is very reliable for forecasting at zone level for pigeonpea [9-11].

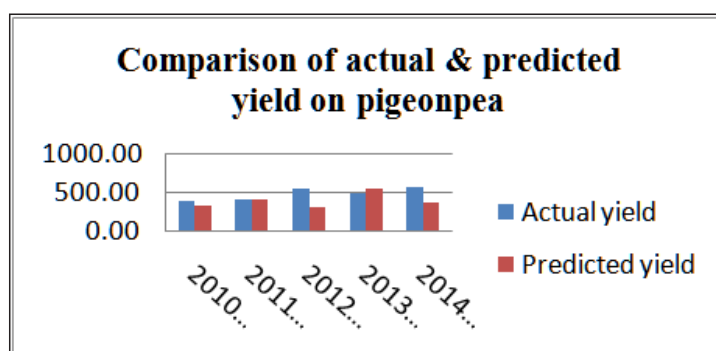


Figure 1: Comparison between actual and predicted yield on pigeonpea for PCA.

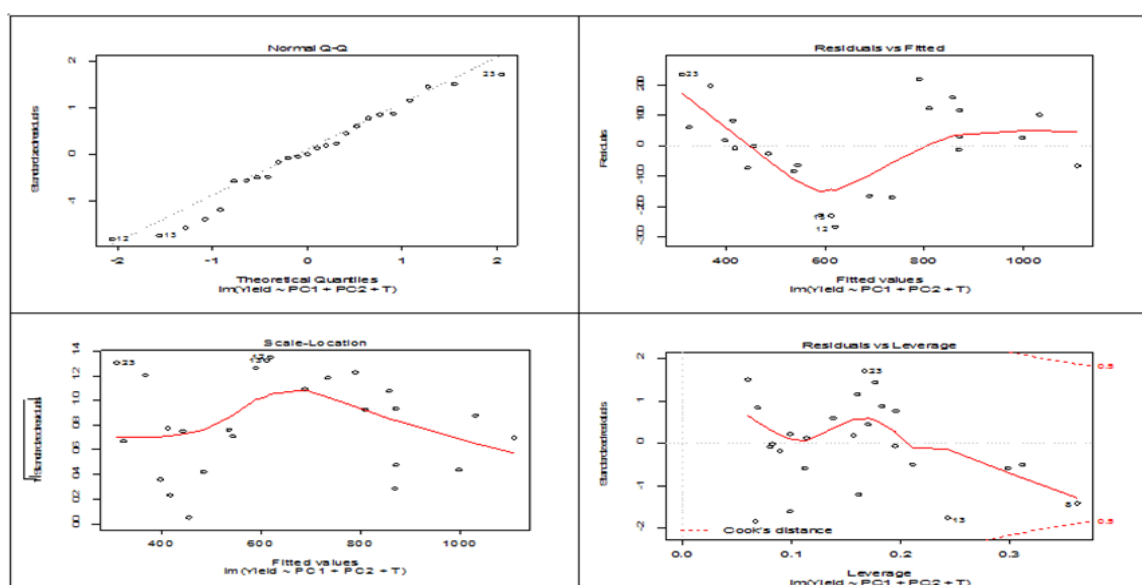


Figure 2 : Forecast model based on Principal Component analysis for Plain zone on pigeonpea.

## References

- Anderson TW (1984) An Introduction to Multivariate Statistical Analysis. John Wiley & Sons, New York.
- Azfar Mohd, Sisodia BVS, Rai VN, Devi Monika (2015) Pre-harvest forecasting models for rapeseed and mustard yield using principle component analysis of weather variable. *Mausam*, pp. 761-766.
- Agrawal R, Jain RC, Jha MP (1986) Models for studying rice crop weather relationship. *Mausam*, 37 (1): 67-70.
- Draper NR, Smith H (1988) Applied Regression Analysis. 3<sup>rd</sup> edn, John Wiley & Sons Inc, p.1-15.
- Agrawal R, Jain RC, Mehta SC (2001) Yield forecast based on weather variables and agricultural input on agrio climatic zone basis. *Ind J Agric Sci* 71(7): 487-490.
- Pandey KK Rai, VN Sisodia, BVS Bharti, AK Gairola, KC (2013) Pre-Harvest Forecast Models Based On Weather Variable And Weather Indices For Eastern UP. *Advances in Bioresearch*, pp. 118- 122.
- Pandey KK, Rai VN, Sisodia BVS, Singh SK (2015) Effect of weather variables on rice crop in eastern Uttar Pradesh, India. *Plant archives*, pp. 575-579.
- Kumar Neeraj, Pisal RR, shukla SP, Pandey KK (2016) Crop yield forecasting of paddy and sugarcane through modified hendrick and scholl technique for south Gujarat. *Mausam*, pp. 405-410.
- Pandey KK, D Maurya, Gupta G, Mishra SV (2016) Yield forecasting models based on weather parameters for eastern UP Vegetos. *An International Journal of Plant Research*.
- Pandey KK, Rai VN, Sisodia BVS, Pandey R, Pandey V (2016) Development of forecasting model on weather parameters for eastern UP. *Environ Life sci rel sci*, pp. 1370-1372.
- Pandey KK, Chandrakar CK, Singh SK, Maurya D, Gupta G (2017) Identification of most important weeks on minimum temperature for wheat crop. *International Journal of Current Microbiology and Applied Sciences*, pp. 788-794.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/IJESNR.2017.05.555667](https://doi.org/10.19080/IJESNR.2017.05.555667)

**Your next submission with Juniper Publishers  
will reach you the below assets**

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
**( Pdf, E-pub, Full Text, Audio)**
- Unceasing customer service

**Track the below URL for one-step submission**  
<https://juniperpublishers.com/online-submission.php>