Int.J.Curr.Microbiol.App.Sci (2018) *7*(7): 2000-2006

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 7 Number 07 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.707.236

Weather Forecast Models of Potato Yield Using Principal Componant Analysis for Sultanpur District of Eastern Uttar Pradesh, India

Snehdeep*, B.V.S. Sisodia, V.N. Rai and Sunil Kumar

Department of Agricultural Statistics Narendra Dev University of Agriculture and Technology, Kumarganj, Faizabad, U.P., India

**Corresponding author*

Keywords

pre-harvest forecast, Statistical model, Weather variables, Principal componant

Article Info

Accepted:
15 June 2018
Available Online:
10 July 2018

ABSTRACT

The present investigation entitled "Forecast Models of Potato Yield Using Principal Component Analysis for Sultanpur District of Eastern Uttar Pradesh." Time series data on yield of potato and weekly data from 40th SMW of the previous year to 6th SMW of the following year on five weather variables *viz.*, Minimum Temperature, Maximum Temperature, Relative humidity 08.30hrs, Relative humidity 17.30hrs, and Wind-Velocity covering the period from 1990-91 to 2011-12 have been utilized for development of pre-harvest forecast model. Statistical methodologies using multiple regression, principal component analysis for developing pre-harvest forecast model have been described. In both models (one based on regression and one from principal component) have been developed. The Model-Ist is based on step wise regression, and IInd based on principal component analysis. Models have been developed on the basis of adjR^2, RMSE and %SE, the best model obtained by the application of step-wise regression analysis of weekly weather data are Model-I_{st} for Sultanpur have further reduced the percentage standard error of the forecast yield to some extent. These models can be used to get the reliable forecast of potato yield two and half months before the harvest.

Introduction

Potato (*Solanum tuberosum* L.) is the most important vegetable crop of the India and known as "The king of vegetable". It is most important cash crop of Uttar Pradesh. Potato is nutritionally superior vegetable. Being a short duration crop, it produces more quantity of dry matter, edible energy and edible protein in lesser duration of time compared to cereals like rice and wheat. It is a native of tropical South America. India produced about 453.44 lakh tonnes of potato from 19.92 lakh hectares under the crop in the year 2012-13. The bulk of the produce come from state of Uttar Pradesh, West Bengal, Bihar and Punjab contributing 32, 26, 15 and 5% respectively in the year 2012-13. The area, production and potato yield at the national level increased during the period 1979-80 to 2010-11 by 172, 408 and 87% respectively. The heat sensitive potato crop is mostly confined to Indo-

Gangetic plains under irrigated conditions due to climate constraints. Small scattered area as rainfed crop are grown in hill during summers and in kharif season in plateau region, whereas winter season crop in the plateau region is irrigated. Usually the pre-harvest estimate of crop yield is obtained on the basis of visual observation which is not objective. There are two major objective approaches for forecasting crop yields one by using weather variables and the other by using weather variables and agriculture inputs jointly. These approaches can be used individually or in combination to give a composite model. Weather is one of the most important factors influencing crop growth. It may influence production directly through affecting the growth structural characteristics of crop such as plant population, numbers of tillers leaf area etc., and indirectly through its effect on incidence of pest and diseases. The effect of weather parameter at different stages of growth of crop may help in understanding their response in term of final yield and also provide a forecast of crop yields in advance before the harvest. The extent of weather influence on crop yields depends only on magnitude of weather parameters but also on their frequency distribution. Therefore, the knowledge of the frequency distribution of weather parameter is also essential while developing the pre-harvest model. Several studies have been carried out in past both in India and abroad on the crop weather relationship and forecasting crop yield, Fisher (1924) made first attempt to develop crop-weather relationship Hendrics and Scholl (1943) modified the Fisher's technique. Agarwal et al (1980) further modified the technique of Hendrics and Scholl (1943) by developing forecast model using weather indices for rice crop in Raipur district and Chhatisgarh such technique of Agarwal et al (1980) has been used by various author in the past for developing forecast yield of various crops in different region of the country. Notable among them are Sisodia et al (2014), Azfar et al (2014), Azfar et al (2015), Yadav et al (2016), R. R. Yadav *et al.*, (2014), etc.

Materials and Methods

This Chapter consists of the material used and the methodology employed for developing models to study the relationship between crop yield and weather variables, and to develop models for making pre-harvest forecast of yield. In order to facilitate systematic presentation, the chapter is divided into following sub-sections:

2.1 General information of the study area
2.2 Sources and description of data
2.3 Statistical methodology used for the development of models.

Description of the study area

Sultanpur is located at 26.27° N 82.07° E. It has an area of 1,713 square miles (4,437 km^2). The surface is generally level, being broken only by ravines in the neighborhood of the rivers. The central portion is highly cultivated, while in the south are widespread arid plains and swampy jhils or marshes. The principal river is the Gomti river, which passes through the centre of the district. According to the 2011 census, Sultanpur district has a population of 3,790,922. This gives it a ranking of 69th in India (out of a total of 640). The district has a population density of 855 inhabitants per square kilometre (2,210/sq miles). Its population growth rate over the decade 2001-2011 was 17.92%. Sultanpur has a sex ratio of 978 females for every 1000 males, and a literacy rate of 71.14%.

Yield data

Time series data on yield of potato for Sultanpur district of Uttar Pradesh for 22 years

(1990-91 to 2011-12) have been collected from the Bulletins of Directorate of Agricultural Statistics and Crop Insurance, Govt. of Uttar Pradesh.

Weather data

Weekly weather data for the same period on five weather variables viz., Minimum Temperature, Maximum Temperature, Relative Humidity at 8.30 and 17.30 hrs and Wind-Velocity have been used in the study. The weekly data on these weather variables have been obtained from the Department of meteorological centre Amausi Airport Lucknow. U.P. India.

Statistical tools used in the analysis

Keeping in view the objectives set out for the study, following statistical tools and methods have been used. The data are analyzed by using software like SPSS, and MS-EXCEL.

Development of the forecast model

This is based on the method given by Agrawal *et al.*, (1986) for developing forecast using weather indices. In this procedure, the entire 19 weeks data from 40th week to 52nd week of a year and 1st week to 6th week of the next year have been utilized for constructing weighted and un-weighted weather indices of weather variables along with their interactions. In all, 30 indices (15 weighted and 15 un-weighted) consisting of 5 weighted weather indices and 10 weighted interaction indices; 5 un-weighted indices and 10 un-weighted interaction indices have been obtained. Considering these 30 indices and trend variable (T) as regressors and yield as dependent variable, forecast has been developed. The fitted formula is

$$y = a_0 + \sum_{i=1}^{p}\sum_{j=0}^{1} a_{ij} Z_{ij} + \sum_{i\neq i'-1}^{p}\sum_{j=0}^{I} a_{ii'j} Z_{ii'j} + cT + \varepsilon$$

where

$$Z_{ij} = \sum_{w=1}^{n} r_{iw}^{j} X_{iw} \Big/ \sum_{w=1}^{n} r_{iw}^{j} \qquad j = 0,1$$

$$Z_{ii'j} = \sum_{w=n_1}^{n_2} r_{ii'w}^{j} X_{iw} X_{i'w} \Big/ \sum_{w=n_1}^{n_2} r_{ii'w}^{j}$$

y is the original crop yield, X_{iw} is the value of the i^{th} weather variable in w^{th} week, $r_{iw}/r_{ii'w}$ is correlation coefficient of yield adjusted for trend effect with i^{th} weather variable/product of i^{th} and i'^{th} weather variable in w^{th} week, n is the number of weeks considered in developing the weather indices, and p is number of weather variables used. a_0, a_{ij}, $a_{ii,j}$ and c are the parameters. ε is error term assumed to follow N (0, σ^2). The step-wise regression analysis was employed to develop the forecast.

Development of Statistical forecast models

Principal Component Analysis is more of a means to an end rather than an end in itself because this frequently serves as intermediate steps in much larger investigations by reducing the dimensionality of the problem and providing easier interpretation. It is a mathematical technique, which does not require user to specify the statistical model or assumption about distribution of original varieties. It may also be mentioned that principal components are artificial variables and often it is not possible to assign physical meaning to them. Further, since Principal Component Analysis transforms original set of variables to new set of uncorrelated variables, it is worth stressing that if original variables are uncorrelated, then there is no point in carrying out principal component analysis.

Let P_1, P_2, ... P_k be first k (k<p) principal components explaining about more then 90

per cent variability in original variables. We define the following forecast model based on first k principal components.

$$y_i = \beta_0 + \beta_1 P_{1i} + \beta_2 P_{2i} + \ldots\ldots\ldots \ \ldots\ldots\ldots + \beta_k P_{ki} + \beta_{k+1} T_i + e_i$$

Where y_i is crop yield during i_{th} year, β_i's are model parameters and e_i's are assumed to follow independently and normally distributed with mean 0 and variance σ^2.

Measures for validation and comparison of models

(i) Percentage Standard Error (%SE)

The % standard error of the (PSE) of the composite forecast of yield is computed as follows:

$$PSE = \frac{\sqrt{V(ycf)}}{ycf} \times 100$$

(ii) Percentage deviation

This measures the deviation (in percentage) of forecast from the actual yield. The formula for calculating the percentage deviation of forecast is given below

Percentage deviation =

$$\frac{\text{(actual yield- forecasted yield)}}{\text{(actual yield)}} \times 100$$

(iii) Percentage Standard Error of the forecast

Let $\hat{y}_f$ be forecast value of crop yield and x_0 be the selected value of X at which the forecast has been done. The variance of $\hat{y}_f$ as given in Draper and Smith (1998) is given by

$$V(\hat{y}_f) = \hat{\sigma}^2 X_0' (X'X)^{-1} X_0$$

where X′X is the dispersion matrix of the sum of square and cross products of regressors and $\hat{\sigma}^2$ is the estimated residual variance of the fitted. The standard error of $\hat{y}_f$ is given by

$$SE(\hat{y}_f) = \sqrt{V(\hat{y}_f)}$$

and, the % standard error (%SE) of $\hat{y}_f$ is given by

$$\%SE = \frac{SE(\hat{y}_f)}{\hat{y}_f} \times 100$$

(iv) Root Mean Square Error (RMSE)

It is also a measure for comparing two s. The formula of RMSE is given below

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

O_i and the E_i are the observed and forecasted value of the crop yield respectively and n is the number of years for which forecasting has been done.

Result and Discussion

This Chapter deals with results, salient finding and discussion of the study undertaken. Various pre-harvest forecast models as described in the preceding chapter have been developed. The results and findings and relevant discussion are presented as follows.

Pre-harvest forecast models using principal component analysis of weekly data of weather variables.

Results for Sultanpur district

Statistical models for pre harvest forecast of the potato yield in Sultanpur district of Eastern

Uttar Pradesh have been developed on the basis of weekly data on weather variables *viz.*, Minimum Temperature, Maximum Temperature, Relative humidity 08.30hrs, Relative humidity 17.30hrs, and Wind-Velocity using principal component. Following the two procedures and two different models have been developed. Sowing of potato starts generally from the first week of October in Sultanpur district. Therefore, weekly data on the weather variables have been considered from pre-sowing the 40^{nd} SMW of crop which fall during the first week of October. It has been proposed to make pre-harvest forecast of the potato yield at the stage of milking / dough, about two months before the harvest. Milking and dough stages generally start after about 130 days of sowing. Therefore, 6^{th} SMW of the next year (Feb.5-Feb.11) has been considered the week of pre-harvest forecast. Thus, in all 19 weeks data on the weather variables (40^{th} SMW of the previous year to 6^{th} SMW of the next year) have been utilized to develop the statistical models.

Comparison of the model

Based on these two forecast models, the forecast yields for the 2009-10, 2010-11 and 2011-12 have been computed and result are presented in Table-3.2.1. The values of R^2adj, percent deviation of forecast from actual yield, RMSE and %SE (CV) have also been computed for each model and are also presented in the Table-3.2.1.

Table.1 Comparison between actual and forecasted yield of different years of Sultanpur District

Model	Year	Actual yield	Forecast yield	Percent Deviation	RMSE	%SE	R^2	R^2adj.
I	2009-10	19.58	16.48	15.84	3.51	7.73	84.40	76.61
	2010-11	24.86	20.59	17.17		7.15		
	2011-12	24.35	21.34	12.36		7.08		
II	2009-10	19.58	14.85	24.12	5.51	8.44	80.56	68.20
	2010-11	24.86	17.45	29.81		6.05		
	2011-12	24.35	20.62	15.35		6.15		

Table.2 Best model from the application of discriminant function analysis of weekly weather data

Model-I	$\mathbf{Y=16.408-0.048Z_{21}-0.058Z_{121}+0.017Z_{140}-0.0006Z_{231}-0.0004Z_{451}+0.019T}$	**R^2=84.40**	**R^2adj=76.61**
Model-II	$Y=21.307-0.001P_1-0.004P_2-0.002P_3+0.001P_4-0.002P_5+0.00002P_6+0.261T$	R^2=80.56	R^2adj=68.20

It is evident from the results of the Table- 3.2.1 that coefficient of determination (R^2) has been found to be 84.40% for the Model-I with lesser percent standard error and minimum RMSE 3.51. Model-II is comparable with Model-I as it has R^2 80.56% and RMSE as 5.51. On the basis of the overall results of the Table- 3.2.1 it can be concluded that the Model-I, followed by Model-II are the most suitable model to forecast potato yield in Sultanpur district of Eastern Uttar Pradesh. Hence, a reliable forecast of potato yield about two and half months before the harvest can be obtained from the Model-I.

Summary and conclusion of the study are as follows:

As far as development of forecast models is concerned in both models based on the application of stepwise regression method and principle componant analysis. The best model obtained by the application of discriminant function analysis of weekly weather data have been given in Table 2.

References

Fisher, R. A. (1924). The influences of rainfall on the yield of wheat at Rothamsted. Philosophical Transaction of Royal Society of London, Series B, Vol. 213,pp. 89-142.

Hendricks, W.A. and Scholl, G.C. (1943). Technique in measuring joint relationship: The joint effects of temperature and precipitation on crop yield. N. Carolina Agric. Exp. Stat. Tech. Bull., pp. 74.

Jain, R.C. (1998). Forcasting of Crop Yields using Second Order Markov Chains. Jour. of the Ind. Sco. of Agril. Stats. Vol. LI (1): 61-72.

Pandey, K. K., V. N. Rai and B.V.S. Sisodia. (2014).Weather variable based rice yield forecasting modles for Faizabad district of eastern U.P. *Int. J. Agri. And Statistical Sciences, Vol. 10,No.2:381-385*

Mohd. Azfar, S B.V.S. Sisodia, V.N. Rai and Monika Devi.(2015). Pre harvest forecast models for rapeseed and mustard yield using principal component analysis of weather variables. *Mausam Vol.66,No.4:761-766.*

Yadav, R.R., B.V.S. Sisodia and Sunil Kumar, (2014): Application of principal component analysisin developing statistical models to forecast crop yield using weather variables. *Mausam, Vol.65(3):357-360.*

Sisodia, B. V. S., Yadav, R. R., Kumar, S. and Sharma, M. K. (2014). Forecasting of Pre- harvest crop yield using discriminant function analysis of meteorological parameter. *Journal of Agrometeorology*. Vol.16(1), pp.121-125.

Tripathi Ramesh (2013). A Study on Statistical Models for Pre-harvest forecasting of potato yield based on weather variables. M. Sc. (Agril.) thesis, N. D. U. A. & T. Kumarganj, Faizabad, Uttar Pradesh.

How to cite this article:

Snehdeep, B.V.S. Sisodia, V.N. Rai and Sunil Kumar. 2018. Weather Forecast Models of Potato Yield Using Principal Componant Analysis for Sultanpur District of Eastern Uttar Pradesh, India. *Int.J.Curr.Microbiol.App.Sci.* 7(07): 2000-2006.
doi: https://doi.org/10.20546/ijcmas.2018.707.236