International Journal of Humanities
and Social Sciences (IJHSS)
ISSN(P): 2319-393X; ISSN(E): 2319-3948
Vol. 7, Issue 4, Jun - Jul 2018; 47-50

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

WHEAT YIELD PREDICTION USING WEATHER BASED STATISTICAL MODEL IN NORTHERN ZONE OF HARYANA

Megha Goyal[1] & Urmil Verma[2]

[1]Teaching Associate, Department of Mathematics and Statistics, CCS Haryana Agricultural University, Hisar, India

[2]Department of Mathematics and Statistics, CCS Haryana Agricultural University, Hisar, India

ABSTRACT

The study was undertaken to investigate the impact of weather variables on crop productivity of wheat. Principal components of the weather parameters spread over the crop growth period were employed to forecast wheat yield(s) in northern zone (Ambala, Yamuna Nagar and Kurukshetra districts) of Haryana. Zonal weather models gave the desired predictive accuracy and provided a considerable improvement in the district-level wheat yield prediction. The results indicate the possibility of district-level wheat yield prediction, 4-5 weeks ahead of the harvest time, in Haryana.

KEYWORDS: *Eigen Values, Principal Component Scores, Percent Relative Deviation, Weather Parameters and Zonal Model*

Article History

Received: ***16 May 2018*** **|** ***Revised:*** ***25 May 2018*** **|** ***Accepted:*** ***07 Jun 2018***

INTRODUCTION

India has a well-established system for collecting agricultural statistics. The primary responsibility for collection of data regarding the area under different crops and production of crops is that of the State Governments. The official forecasts (advance estimates) of major cereal and commercial crops are issued by the Directorate of Economics and Statistics, New Delhi. However, the final estimates are given a few months after the actual harvest of the crop. This makes these statistics unusable for planning and management purposes. Hence, there is a considerable scope for improvement in the conventional system. Reliable yield forecasting within the growing season would enable improved planning and more efficient management of crop production, handling and marketing.

Various organizations in India and abroad are engaged in developing the methodology for pre-harvest forecasting of crop yield using different approaches. The most commonly used models are based on empirical statistical models. The processes determining crop growth are complex and final yield is governed by a number of factors. The main factors affecting crop yield are agricultural inputs and weather parameters. Use of these factors forms a broad category of models for forecasting crop yield. In many previous studies, yield forecasting models have incorporated a series of weather predictors {Bal (2004), Bazgeer *et al*. (2007), Esfandiary *et al*. (2009) and more specifically; temperature (Batts *et al.* (1997), Wheeler *et al.* (2000), Peng *et al.* (2004)) and rainfall (Mkhabela *et al.* (2005)). Azfar *et al.* (2015) have used principal component analysis for mustard yield forecasting.

The purpose of this study was to develop zonal yield models using time series data on weather parameters and crop yield by following principal component analysis for obtaining pre-harvest wheat yield forecasts in the northern zone of Haryana. The work was initiated to develop forecasting models on agro-climatic zone basis by combining the data of various districts within the zone so that a longer series could be obtained in a relatively shorter period and that provided the basis to use multivariate statistical analysis. The emphasis has been given to see the forecasting performance of the zonal yield model for district-level yield estimation during the period(s) of model development and model testing.

Study Region and Data Used

District-level wheat yield data of the past 34 years (1980-81 to 2013-14) were collected for northern zone in Haryana and the same were used to carry out linear time-trend analysis and then computing the district-level trend, based on yield TY = a + bt, where TY = Trend yield, a = Intercept, b = Slope and t = Year. The weather data viz., rainfall, maximum temperature and minimum temperature were collected from India Meteorological Department (IMD) and different meteorological observatories in Haryana state.

Computation of Weather Parameters

Weather data starting from the 1st fortnight of November to 1 month before harvest were utilized for the model building (crop growth period: 1st November to 15th April). The various fortnightly weather parameters were computed as follows:

Average Maximum Temperature (TMX) = $\dfrac{\sum_{i=1}^{15} \text{TMXi}}{15}$

Average Minimum Temperature (TMN) = $\dfrac{\sum_{j=1}^{15} \text{TMNj}}{15}$

Accumulated Rainfall (ARF) = $\sum_{k=1}^{15} \text{ARF}_K$

where TMX_i = ith day maximum temperature

TMN_j = jth day minimum temperature

ARF_K = kth day rainfall

i,j,k = daily meteorological data

Statistical Methodology

The use and interpretation of a multiple regression model often depend explicitly or implicitly on the assumption that the explanatory variables are not strongly interrelated. In most regression applications, the explanatory variables are not orthogonal. Usually the lack of orthogonality is not serious enough to affect the analysis. However, in some situations, the explanatory variables are so strongly interrelated that the regression results are ambiguous. Under such situation, the estimated values of the coefficients are very sensitive to slight changes in the data and to the addition or deletion of variables in the equation. To overcome the problem of multicollinearity observed among the weather data, the zonal wheat yield models were developed within the framework of Principal Component Analysis (PCA). Principal components P_i (i=1,2,….) were obtained as P = kX, where P and X are the column vectors of transformed and the original variables, respectively and k is the matrix with rows as the characteristic vectors of the correlation matrix R. The variance of P_i is the i^{th} characteristic root λ_i of the correlation matrix R; λ_s were obtained by solving the equation $|R- \lambda I| =0$. For each λ, the corresponding characteristic vector k was obtained by solving $|R-\lambda I|\ k=0$

RESULTS AND DISCUSSIONS

PCA offers a considerable improvement over least squares estimates in the presence of multicollinearity. The procedure consists of finding the eigen roots and eigenvectors of the correlation matrix of explanatory variables. The most frequently used convention is to retain the components whose eigenvalues are greater than one. Under this study, first ten eigenvalues of the correlation matrix of weather variables suggested ten factor solution. Eigenvalues and percent variance explained by different PCs are presented in Table 1.

Table 1: Eigen Values and Variance (%) Explained by Different Principal Components

Components	1	2	3	4	5	6	7	8	9	10
Eigenvalue (% variance explained)	4.91 (18.17)	3.48 (12.90)	2.67 (09.89)	2.33 (08.62)	2.07 (08.28)	1.94 (07.20)	1.73 (06.40)	1.36 (05.03)	1.24 (04.61)	1.06 (03.94)

Regression models via step-wise regression method (using SPSS) were fitted by taking yield as the dependent variable and principal components and trend yield as the regressors. The finally selected zonal wheat yield relationship on the basis of highest *adj-*R^2 and lowest standard error (SE) of estimate are as follows:

$$Yield_{est} = -2.93 + 1.08 *TY + 0.77 * PC_3 - 0.81 * PC_4 +0.85* PC_6$$

R^2= 0.880,   adj.R^2 =0.875 & SE= 2.64

where $Yield_{est}$ - Model predicted yield

TY - Linear time-trend based yield

PC_i - i^{th} principal component score (i = 1,2,…,10)

R^2 - Coefficient of determination

SE - Standard error of yield estimate

The above zonal yield model was used to obtain wheat yield forecasts at a district level in the northern zone of Haryana. Trend yield is an important parameter appearing in the model, indicating that most of the variability in yield is explained by TY, which is an indication of technological advancement, improvement in fertilizer/insecticide/pesticide/weedicide used and increased use of high yielding varieties. The model predicted yields along with observed yields and per cent relative deviations are as given below:

Table 2: District-Specific Wheat Yield Estimates Along with Percent Deviations from Observed Yield(s) Using Zonal Yield Model

Districts/ Years	Ambala			Kurukshetra			Yamunanagar		
	Obs. Yield (q/ha)	Fitted Yield (q/ha)	RD(%)	Obs. Yield (q/ha)	Fitted Yield (q/ha)	RD(%)	Obs. Yield (q/ha)	Fitted Yield (q/ha)	RD(%)
2010-11	36.08	41.12	-13.96	44.72	49.45	-10.58	45.57	41.21	9.57
2011-12	48.61	43.58	10.34	54.38	51.93	4.51	53.64	43.56	18.80
2012-13	42.06	42.30	-0.58	46.57	50.66	-8.77	43.39	42.15	2.85
2013-14	46.31	43.27	6.56	48.78	51.64	-5.85	45.31	43.00	5.09
Average absolute deviation	7.86			7.43			9.08		

Percent Relative Deviation (RD%) = 100 × [(observed (obs.) yield – fitted yield) / observed yield]

The results inferred that there is a considerable improvement in the district-level wheat yield forecasts on the basis of zonal models and also showed good agreement with the observed yield(s) in the northern zone of Haryana.

REFERENCES

1. *Azfar, M., Sisodia, B. V. S., Rai, V. N. and Devi, M.(2015). Pre-harvest forecast models for rapeseed and mustard yield using principal component analysis of weather variables, Mausam,* ***66****(4): 761-766.*
2. *Bal, S.K., Mukherjee, J., Mallick, K. and Hundal, S.S. (2004). Wheat yield forecasting models for Ludhiana district of Punjab state. J. Agrometeorology. 6: 161-65.*
3. *Batts, G.R., Morison, J. I. L., Ellis, R. H., Hadley, P. and Wheeler, T. R. (1997). Effects of CO2 and temperature on growth and yield of crops of winter wheat over four seasons, European Journal of Agronomy 7 (1997), 43-52.*
4. *Bazgeer, S., Gh. Kamali and A. Mortazavi (2007). Wheat yield prediction through agro meteorological indices for Hamedan, Iran. Biaban,* ***12****, 33-38.*
5. *Chhetri, B. I. N. O. Y., and S. K. Mahato. "Eco-friendly and alternative weed management practices in buckwheat (Fagopyru mesculentum Moench.) in relation to yield potential under rainfed condition." International J. Agricultural Science and Research 5.2 (2015): 165-172.*
6. *Esfandiary, F., G. Aghaie and A. D. Mehr (2009). Wheat yield prediction through agro meteorological indices for Ardebil district. World Academy of Science: Engineering and Technology,* ***49****, 32-35.*
7. *Mkhabela, M. S. and Mashinini, N.N. (2005). Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR, Agricultural and Forest Meteorology 129 (2005), 1-9.*
8. *Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X, Centeno, G. S., Khush G.S. and Cassman, K. G. (2004) Rice yields decline with higher night temperature from global warming, Proceedings of the National Academy of Sciences of the United States of America 101, 9971-9975.*
9. *Wheeler, T. R., Craufurd, P. Q., Ellis, R. H., Porter J. R. and Prasad, P. V. V. (2000)Temperature variability and the yield of annual crops, Agriculture, Ecosystems and Environment 82, 159-167.*