Int.J.Curr.Microbiol.App.Sci (2018) *7*(12): 2973-2978

EXCELLENT PUBLISHERS

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 7 Number 12 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.712.340

Development of Weather Based Wheat Yield Forecast Models in Haryana

Sanjeev, Puneet Verma and Urmil Verma*

Department of Mathematics & Statistics, CCS Haryana Agricultural University, Hisar-125004, India

**Corresponding author*

Keywords

Linear time trend, Eigen value, Eigen vector, Weather variables, Multicollinearity, Principal component score

Article Info

Accepted:
20 November 2018
Available Online:
10 December 2018

A B S T R A C T

Parameter estimation in statistical modelling plays a crucial role in the real world phenomena. Several alternative analyses may be required for the purpose. An attempt has been made in this paper to estimate the yield of wheat crop using principal components of the weather parameters spread over the crop growth period. Principal component analysis has been used for the purpose of developing zonal yield forecast models because of multicollinearity present among weather variables. The results indicate the possibility of district-level wheat yield prediction, 4-5 weeks ahead of the harvest time in Haryana, India. Zonal weather models had the desired predictive accuracy and provided considerable improvement in the district-level wheat yield estimates. The estimated yield(s) from the selected models indicated good agreement with State Department of Agriculture (DOA) wheat yield(s) in most of the districts.

Introduction

Crop yield is affected by technological change and weather variability. It can be assumed that the technological factors will increase crop yield smoothly through time and therefore, year or some other parameter of time can be used to study the overall effect of technology on yield. Weather variables affect the crop differently during different stages of development. This increases the number of variables in the model and thus, a technique based on relatively smaller number of manageable parameters and at the same time, taking care of entire weather distribution may solve the problem. Principal Component Analysis (PCA) in this direction was carried out for pre-harvest wheat yield estimation on agro-climatic zone basis in Haryana.

Some similar studies concerning this work viz., weather models developed by Mehta *et al.*, (2000), Agarwal *et al.*, (2001) and Ramasubramanian *et al.*, (2004) were successfully used for forecasting yields of various crops at district as well as agro-climatic zone level in different states of India. Hoogenboom (2000), Kandiannan *et al.*, (2002), Bazgeer *et al.*, (2007), Esfandiary *et* *al.*, (2009), Lobell and Burke (2010), Basso *et al.*, (2012) etc. have used a series of weather predictors for crop yield forecasting. Verma *et al.*, (2011 and 2016) and Goyal and Verma (2015) have used agromet/spectral indices for the development of crop yield models of different crops in Haryana (India). Azfar *et al.*, (2015) used principal component analysis for rapeseed and mustard yield forecast models for Faizabad district of U.P. (India).

Crop status and data description

Wheat is one of the most important cereal crops in India as it forms a major constituent of the staple diet of a large part of the population. India is the second largest producer among wheat growing countries of the World (Source: www.mapsofindia.com/indiaagriculture).
Haryana occupies the third position in wheat production among the various states in India (www.agricoop.nic.in/statistics). Haryana is self-sufficient in food grains production and also one of the top contributors of food grains to the central pool. Wheat occupies the foremost position followed by rice, not only regarding acreage and production but also in the versatility in adopting different soils and climatic conditions.

The Haryana state comprised of 22 districts is situated between 74^o 25' to 77^o 38' E longitude and 27^o 40' to 30^o 55' N latitude. The total geographical area of the state is 44212 sq. km. The DOA wheat yield data published by Bureau of Economics and Statistics, Haryana were compiled for the period 1980-1981 to 2012-2013 for Ambala, Kurukshetra, Rohtak, Karnal, Jind, Sonepat, Gurgaon, Faridabad, Mahendergarh, Hisar, Sirsa and Bhiwani districts, 1989-1990 to 2012-2013 for Yamunanagar, Panipat, Kaithal and Rewari, 1995-1996 to 2012-2013 for Panchkula, 1997-1998 to 2012-2013 for Jhajjar and Fatehabad districts and the same were used to carry out linear time-trend analysis and then computing the district-level trend based yield Tr = a + bt, where Tr = Trend yield, a = Intercept, b = Slope and t = Year. The weather data for the last 33 years (1980-1981 to 2012-2013) were collected from India Meteorological Department (IMD) and different meteorological observatories of Haryana.

Weather data starting from 1st fortnight of November to 1 month before harvest were utilized for the model building (crop growth period: 1st November to 15th April). Since the climatic data from adequate number of stations were not available; the districts having equable climatic conditions were grouped into four agro-climatic zones based on their physiography/soils and agro-climatic conditions in Haryana viz., zone-1: Ambala, Panchkula, Yamuna Nagar, Kurukshetra, zone-2: Karnal, Kaithal, Jind, Panipat, Sonipat, Rohtak, zone-3: Mahendergarh, Rewari, Jhajjar, Gurgoan, Faridabad and zone-4: Sirsa, Fatehabad, Hisar, Bhiwani.

PC method was used for extraction of factors which consists of finding the eigen values and eigen vectors. Principal components (i.e. P_i, i=1,2,….) were obtained as P = kX , where P and X are the column vectors of transformed and original variables and k is the matrix with rows as the characteristic vectors of the correlation matrix R. The variance of P_i is the ith 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 Discussion

PC method consists of finding the eigen roots and eigen vectors of the correlation matrix of explanatory variables.

Table.1 Eigen values and variance (%) explained by different principal components

Components	**Eigen value (% variance explained)**			
	Zone 1	**Zone 2**	**Zone 3**	**Zone 4**
1	4.91(18.17)	5.10(18.90)	4.48(16.59)	**4.52(16.75)**
2	3.48(12.90)	3.74(13.85)	3.41(12.62)	**3.45(12.79)**
3	2.67(09.89)	3.02(11.19)	2.73(10.12)	**2.77(10.25)**
4	2.33(08.62)	2.81(10.39)	2.33(08.65)	**2.59(09.61)**
5	2.24(08.28)	2.07(07.67)	2.11(07.81)	**2.09(07.74)**
6	1.94(07.20)	1.89(07.02)	2.04(07.56)	**1.79(06.63)**
7	1.73(06.40)	1.48(05.48)	1.71(06.34)	**1.62(05.99)**
8	1.36(05.03)	1.12(04.15)	1.28(04.74)	**1.34(04.96)**
9	1.24(04.61)	1.06(03.92)	1.17(04.34)	**1.10(04.08)**
10	**1.06(03.94)**	**0.90(03.32)**	**0.96(03.56)**	**1.07(03.96)**

Table.2 Zonal wheat yield models based on trend yield and PC scores

Model variable	**Zone1 Coefficients**	**Model variable**	**Zone 2 Coefficients**
Constant	***c*** ***-2.93***	***Constant***	***c*** ***0.31***
Tr	b_1 ***1.08***	***Tr***	b_1 ***0.99***
PC_3	b_2 ***0.77***	PC_1	b_2 ***-0.73***
PC_4	b_3 ***-0.81***	PC_2	b_3 ***0.81***
PC_6	b_4 **0.85**	PC_7	b_4 **-0.51**
R^2 = 0.880 adj.R^2 =0.875	**SE =2.64**	R^2 = 0.911 adj.R^2 =0.909	**SE= 2.14**

Model variable	Zone 3 Coefficients	Model variable	Zone 4 Coefficients
Constant	*C* *1.31*	*Constant*	*c* *-1.78*
Tr	b_1 *0.99*	*Tr*	b_1 *1.05*
PC_3	b_2 *0.84*	PC_6	b_2 *-1.10*
PC_6	b_3 *-0.50*	PC_{10}	b_3 *1.02*
R^2= 0.860 adj.R^2 =0.857	SE =2.55	R^2= 0.827 adj.R^2 =0.822	SE= 2.93

Zone 1: $Yield_{est} = \{c + b_1 \times Tr + b_2 \times PC_3 + b_3 \times PC_4 + b_4 \times PC_6\}$
Zone 2: $Yield_{est} = \{c + b_1 \times Tr + b_2 \times PC_1 + b_3 \times PC_2 + b_4 \times PC_7\}$
Zone 3: $Yield_{est} = \{c + b_1 \times Tr + b_2 \times PC_3 + b_3 \times PC_6\}$
Zone 4: $Yield_{est} = \{c + b_1 \times Tr + b_2 \times PC_6 + b_3 \times PC_{10}\}$

Where,
$Yield_{est}$ - Model predicted yield
c - Constant
b_i - Regression coefficient (i=1,2, 3, ...)
T_r - Linear time trend based yield
PC_i - i^{th} principal component score (i = 1,2,…,10)
SE - Standard error of yield estimate
R^2 - Coefficient of determination

Table.3 District-specific wheat yield estimates alongwith percent deviations from DOA yield(s) using fitted models

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%)
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

Districts/ Years	Rohtak			Karnal			Jind		
	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%)
2011-12	50.20	43.08	14.18	56.7	49.24	13.15	52.35	48.23	7.88
2012-13	37.38	44.38	-18.72	46.7	50.62	-8.39	42.70	49.69	-16.36

Districts/ Years	Sonipat			Panipat			Kaithal		
	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%)
2011-12	55.21	47.94	13.17	39.17	43.48	-11.00	54.51	47.06	13.66
2012-13	45.20	49.40	-9.29	42.47	45.24	-6.53	46.84	48.26	-3.04

Districts/ Years	Gurgaon			Faridabad			Mahendergarh		
	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%)
2011-12	49.62	43.82	11.69	48.37	45.67	5.58	46.11	45.68	0.93
2012-13	45.54	44.07	3.23	45.39	45.82	-0.94	47.77	45.74	4.25

Districts/ Years	Rewari			Jhajjar			Hisar		
	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%)
2011-12	50.02	44.33	11.38	48.58	41.33	14.92	50.98	49.92	2.07
2012-13	49.28	44.19	10.33	39.71	41.14	-3.59	42.73	48.48	-13.45

Districts/ Years	Sirsa			Bhiwani			Fatehabad		
	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%)
2011-12	53.57	49.56	7.49	43.06	43.67	-1.41	54.72	50.54	7.64
2012-13	48.42	48.10	0.67	40.55	42.20	-4.06	46.81	49.02	-4.72

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

The most frequently used convention is to retain the components whose eigen values are greater than one. Weather data starting from first fortnight of November to first fortnight of March *i.e.* one month before crop harvest were utilized for the model development. In this study, first ten eigen values of the correlation matrix of explanatory variables (weather parameters) suggested ten factor(s) solution (Table 1). However, the remaining components accounted for a smaller amount of total variation. Hence, those components were not considered to be of much practical significance. Eigen vectors being the weights were used to compute PC scores. Multiple linear regression models *via* step-wise regression method (Draper and Smith, 2003) were fitted by considering PC scores as regressors and wheat yield as dependent variable. The models finalized as shown in Table 2 were used to obtain the district-level wheat yield estimates (Table 3).

Trend yield is an important parameter appearing in all the models, indicating that most of the variability in yield is explained by *Tr*, which is an indication or technological advancement, improvement in fertilizer/ insecticide/pesticide/weedicide used and increased use of high yielding varieties. The predictive performance(s) of the fitted models were observed in terms of the percent deviations of wheat yield forecasts in relation to real time wheat yield(s). The estimated yield(s) from the selected models indicated good agreement with DOA wheat yield estimates in most of the districts.

References

Agarwal, R., Jain, R.C. and Mehta, S.C. (2001). Yield forecast based on weather variables and agricultural inputs on agroclimatic zone basis. *Indian J. of Agril. Sci.* 71(7), 487-490.

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.

Basso, B., C. Fiorentino, D. Cammarano, G. Cafiero and J. Dardanelli (2012). Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. *European Journal of Agronomy*, 41, 52– 65.

Bazgeer, S., Gh. Kamali and A. Mortazavi (2007). Wheat yield prediction through agrometeorological indices for Hamedan, Iran.*Biaban*, 12, 33-38.

Draper, N.R. and Smith, H. 2003. Applied Regression Analysis. 3rd edition, *John Wiley and Sons. New York.*

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.

Goyal, M. and Verma, U. (2015). Development of weather-spectral models for pre-harvest wheat yield prediction on agro-climatic zone basis in Haryana, *International J. of Agricultural and Statistical Sciences*, 11(1), 73-79.

Hoogenboom, G. (2000). Contribution of agrometeorology to the simulation of crop production and its applications. *Agricultural and Forest Meteorology*, 103, 137-157.

Kandiannan, K., K. K. Chandaragiri, N. Sankaran, T. N. Balasubramanian and C. Kailasam (2002). Crop-weather model for turmeric yield forecasting for Coimbatore District, Tamil Nadu, India. *Agricultural and Forest Meteorology*, 112, 133-137.

Lobell, D. B and M. Burke (2010). On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150, 1443–1452.

Mehta, S.C., Agarwal, R. and Singh, V.P.N. (2000). Strategies for composite forecast. *J. Ind. Soc. Agril. Statist.* 53(3), 262-272.

Ramasubramanian, V., Agrawal, R. and Bhar, L.M. (2004). Forecasting sugarcane yield using multiple Markov chains. (*IASRI, New Delhi publication*).

Verma, U., Dabas, D. S., Hooda, R. S., Kalubarme, M. H., Yadav, M. and Sharma, M. P. (2011). Remote Sensing Based wheat acerage and spectraltrend-agrometeoro-logical yield forecasting: Factor Analysis Approach. *Society of Statistics, Computers and Applications*, 9(1and2), 1-13.

Verma, U., Piepho, H. P., Goyal, A., Ogutu, J. O. and kalubarme, M. H. (2016). Role of climatic variables and crop condition term for mustard yield prediction in Haryana (India). *International J. of Agricultural and Statistical Sciences*, 12(1), 45-51.

How to cite this article:

Sanjeev, Puneet Verma and Urmil Verma. 2018. Development of Weather Based Wheat Yield Forecast Models in Haryana. *Int.J.Curr.Microbiol.App.Sci.* 7(12): 2973-2978.
doi: https://doi.org/10.20546/ijcmas.2018.712.340