



# Zonal trend-agrometeorological models for wheat yield estimation in Haryana

# M. Goyal and U. Verma<sup>\*</sup>

Department of Mathematics and Statistics, CCS Haryana Agricultural University, Hisar, Haryana, INDIA \*Corresponding author. Email: vermas21@hotmail.com

Received: August 15, 2015; Revised received: June 15, 2016; Accepted: August 07, 2016

**Abstract:** An attempt has been made to assess the impact of weather variables for district-level wheat yield estimation in Haryana. Fortnightly weather data and trend based yield were used for developing the zonal trendagrometeorological (agromet) models within the framework of multiple linear regression and discriminant function analyses. The district level wheat yield forecasts, percent deviations from the real time wheat yield (s) and root mean square error(s) at zonal level show a preference of using discriminant/weather scores as regressors in almost all the considered districts of the state. Zonal trend-agromet models provided considerable improvement in district-level wheat yield prediction moreover the yield estimates may be obtained 4-5 weeks in advance of the harvest time. The estimated yield(s) from the selected zonal models showed good agreement with State Department of Agriculture (DOA) wheat yields by showing less than 5 percent deviations in 9 districts and 6-11 percent deviations in the remaining 9 districts under consideration.

Keywords: Discriminant/weather score, DOA yield, Percent deviation, Root mean square error, Trend-agromet

## INTRODUCTION

Forecasting of crop yield is one of the most important aspects of agricultural statistics system. Crop yield models are abstract presentation of the interaction of the crop with its environment and can range from simple correlation of yield with a finite number of variables to the complex statistical models with predictive end. Forecasts can be formed in many different ways. The method chosen depends on the purpose and importance of the forecasts as well as the costs of alternative forecasting methods. An efficient crop forecasting infrastructure is pre-requisite for information system about food supply, especially export-import policies, procurement and price fixation. This is also useful to farmers to decide in advance their future prospects and course of action. The yield of any crop is affected by technological change and weather variability. It can be assumed that the technological factors will increase the yield smoothly through time and therefore, year or some other parameters of time can be used to study the overall effect of technology on crop yield. Weather variability both within and between seasons is the second and uncontrollable source of variability in yields.

Various statistical approaches viz., regression, timeseries and stochastic models are in vogue for arriving at crop forecasts. Every approach has its own advantages and limitations. Time series models have advantages in certain situations. They can be used more easily for forecasting purposes because the historical sequences of observations upon study variables are readily available. These successive observations are statistically dependent and the time series modelling is concerned with techniques for the analysis of such dependence. Climate change is a major concern today, and the researchers are engaged in understanding its impact on growth and yield of crops, and also identifying the suitable management options to sustain the productivity of crops under projected climate change scenario. Quantitative understanding of crop response to climate requires the development of statistical models for various characteristics of the crop by taking into account its time-series behavior along with the climatic factors. Winter crops are especially vulnerable to high temperatures during the reproductive stages and their differential responses to rising temperatures can have important consequences on crop yield. The key problem is how to incorporate the pertinent information into the forecasting process and subsequently into the decision making process. No satisfactory model which has universal validity exists today.

The official forecasts (advance estimates) of major cereal and commercial crops are issued by the Directorate of Economics and Statistics, Ministry of Agriculture, New Delhi. However, the final estimates are given a few months after the actual harvest of the crop. Thus, one of the limitations of crop cutting experiments is timeliness and quality of the statistics. Hence, there is a considerable scope of improvement in the conventional system.

The most commonly used models are empirical statistical models. In many previous studies, yield forecasting models have incorporated a series of

ISSN : 0974-9411 (Print), 2231-5209 (Online) All Rights Reserved © Applied and Natural Science Foundation www.jans.ansfoundation.org

weather predictors. Rai and Chandrahas (2000) and Ruíz and Ordonez (2003) have proposed the use of discriminant function analysis for crop yield forecasting. Chandran and Prajneshu (2004), Peng et al. (2004), Mkhabela et al. (2005), Bazgeer et al. (2007) and Esfandiary et al. (2009), Mehta et al. (2010) etc. have used weather data in the context of crop yield prediction. Discriminant function analysis has been used to transform several weather variables into a reduced set of discriminant/ weather scores for crop yield prediction by Chandrahas et al. (2010). Verma et al. (2011, 12) have used weather and remote sensing data for operational yield forecasting of different crops in Haryana under CAPE project. Vashisth et al. (2014) have applied different statistical procedures for comparing pre-harvest crop yield models.

The study serves the purpose of obtaining advance estimates or short-term forecasts of pre-harvest wheat yield, which has to be scientifically obtained by evolving reliable statistical methods of forecasting. 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 in the country. India is the second largest producer among wheat growing countries of the World (Source: www.mapsofIndia.com/ indiaagriculture/). Haryana occupies third place for wheat production among the various states in India (Source:www.agricoop.nic.in/statistics). Haryana is self-sufficient in food grains production and is one of the top contributors of food grains to the central pool. In view of the above points, an attempt has been made to obtain the zonal trend-agromet-yield models by incorporating the time/trend based yield along with weather variables using multiple linear regression and discriminant analysis.

The emphasis has been given to see the forecasting performance of the zonal models for district-level yield estimation in Haryana during the period(s) of model development and model testing. In subsequent sections, we first present the data used and methodology applied for the model building. Next, the crop yield estimates derived from the fitted models and accordingly the related discussion has been given.

#### **MATERIALS AND METHODS**

The Haryana state comprising of 21 districts is situated between 74° 25' to 77° 38' E longitude and 27°40' to 30°55' N latitude. The total geographical area of the state is 44212 sq. km. Time-series yield data for the past 30 years (i.e. 1978-79 to 2007-08) of wheat crop in various districts of Haryana published by Bureau of Economics and Statistics were used for computing linear yield trend i.e. T<sub>r</sub> = a+br, where  $T_r = Trend yield(q/ha)$ , a = Intercept, b = Slope and r = Year. The meteorological data for the same 30 years were collected from India Meteorological Department (IMD), Delhi

and different meteorological observatories in Harvana, India. The weather data of maximum temperature, minimum temperature and rainfall were used for the purpose. The total wheat growth period i.e. from 1<sup>st</sup> November to first fortnight of April was divided into 11 fortnights and the fortnightly weather parameters obtained may be expressed as:

 $\sum^{15} TMXi$ Average maximum temperature (TMX) = $\frac{\displaystyle\sum_{j=1}^{15} TMNj}{15}$ 

Average minimum temperature (TMN) =

 $\sum^{15} ARF_{K}$ 

15

Accumulated rainfall (ARF) = where  $TMX_i$  -= i<sup>th</sup> day maximum temperature,  $TMN_i = j^{\text{th}}$  day minimum temperature,  $ARF_K =$ k<sup>th</sup> day rainfall; i,j,k = daily meteorological data

The weather data were available only for the four different regions of the state. Since the climatic data from adequate number of stations were not available, hence, the districts having equable climatic conditions were grouped into different zones. Four agro-climatic zones are based on their physiography, soils and agro-climatic conditions in the state. These four zones and districts in each zone are as follows:

Zone-I: Ambala, Yamuna Nagar, Kurukshetra

Zone-II: Karnal, Kaithal, Jind, Panipat, Sonipat, Rohtak

Zone-III: Mahendergarh, Rewari, Jhajjar, Gurgoan, Faridabad and

Zone-IV:Sirsa, Fatehabad, Hisar, Bhiwani

Trend-agrometeorological wheat yield modeling: Yield is a complicated trait which is governed by a number of factors. The main factors affecting the crop yield are agricultural inputs and weather variables. The work was carried out to develop forecasting models on agro-climatic zone basis by combining the data of various districts within zones. Thus, a longer data series could be obtained in a relatively shorter period (i.e. inclusion of 30 years weather and yield data for each district(s) within zone) and that provided the basis to use multivariate statistical analyses. The focus was on the comparison of district- level yield estimates obtained under two different procedures by evaluating the forecasting performance of the zonal trend-agromet-yield models during the period of model development (1978-79 to 2004-05 i.e. 27 years ) and the model testing period (2005-06 to 2007-08 i.e. 3 post-sample years ). Multiple linear regression and discriminant function analyses were used to achieve the targeted objective.

The standard linear regression model considered may be written in the form  $Y=Xb+\varepsilon$ ; where Y is an  $(n \times l)$  vector of observations, X is an  $(n \times p)$  matrix of known form, **b** is a  $(p \times l)$  vector of parameters,  $\varepsilon$  is an  $(n \times 1)$  vector of errors with the assumptions E  $(\varepsilon)=0$  and  $V(\varepsilon)=I\sigma^2$ , so the elements of  $\varepsilon$  are uncorrelated. Regression models via stepwise regression analysis (Draper and Smith, 1981) were fitted using statistical software SPSS. The selected zonal yield models are given in Table-3.

Discriminant function analysis is a multivariate technique for identifying appropriate functions that discriminate best between the set of observations from two or more groups and classifies the future observations into one of the previously defined groups. The technique involves identifying linear/quadratic function(s) where the coefficients are determined in such a way that the variation between the groups gets maximized relative to the variation within the groups. The discriminant analysis is used to discriminate between various groups of objects when the dependent variable is categorical and the independent variables are metric. Thus, a crop year was categorized as congenial, normal or adverse to a crop using discriminant analysis which discriminated between three groups on the basis of 27 weather parameters ( i.e. 9 fortnights x 3 weather variables). At first, the data on yield were adjusted for the trend effect and then grouped into three categories. The weather parameters were used for the development of discriminant/weather scores. The contribution of each variable for discrimination in the group means was tested through  $\chi^2$  test for each fortnight of crop growth. Finally, the weather scores along with trend yield were used for the development of zonal trendagromet-yield models.

We considered a linear function of the form as  $Z = \sum_{i=1}^{p} \ell_i X_i$ . Here Z is total discriminant score for two groups, X<sub>i</sub> is the i-th weather variable used to discriminate the groups and  $\ell^i$  is the corresponding discriminant coefficient, p being the number of weather variables. The value of  $\ell^i$  can be obtained as  $\ell_i = (s_{ij})^{-1} d = C_{ij} d$ , where  $C_{ij} = (s_{ij})^{-1}$  and  $S_{ij}$  is discriminant dispersion matrix. Considering  $D^2 = \sum_{i=1}^{p} \ell_i d_i$  (Mahalanobis  $D^2$ -statistic ),  $d_i$  is the difference in means i.e  $\overline{X}_{1i} - \overline{X}_{2i}$  for the i-th variable,  $n_1$ ,  $n_2$  are the number of units in the two groups and  $\overline{X}_{1i}$ ,  $\overline{X}_{2i}$  are means of the ith variable for the two groups. The percent contribution of i-th weather variable may be computed as

$$C_i = {\ell_i \ d_i * 100} / D^2$$

For classifying the individual crop year into three groups, the distances between each individual observation and each group centroid in discriminant function were measured and then assigned to the group whose centroid in discriminant function was nearest.

Hair *et al.*, (1995) suggested that the contribution of each variable can be associated by knowing its potency value and is calculated as :

Potency value = (Square of discriminant coefficient) \* (relative eigen value)

The relative eigen value is the ratio of eigen value of the discriminant function to the sum of the eigen values for all significant discriminant functions. To assess the contribution of weather predictors, potency index is one of the measures to be employed. It represents the total discriminating effects across all the discriminant functions.

Extraction of discriminant/weather scores: Following discriminant function analysis, the thirty years wheat yield data (1978-79 to 2007-08) were classified into three groups namely congenial, normal or adverse on the basis of trend predicted yield and weather data. The first 27 years data from 1978-79 to 2004-05 were utilized for the model building and remaining three years were left for validation of the model. Since, the number of groups was three and the number of weather parameters was 33, therefore, only two discriminant functions (i.e.  $r = \min(p, k - 1)$ , where k is the number of groups and p is the parameters used) were obtained. Thus, DOA wheat yield data were classified into three groups and the same was used as categorical variate in discriminant function analysis (Bhuyan, 2005). The zone-wise percent contribution of each discriminant function alongwith chi-square values are presented in Table-1. The discriminant function coefficients were further used to obtain the discriminant scores for fortnightly weather database as has been given in Table-2. Discriminant analysis approach predicts the future observations qualitatively in different groups. Thus for quantitative forecasting, the zonal yield models shown in Table-3 were fitted by taking discriminant/ weather scores and trend yield as regressors and DOA crop yield as regressand in multiple linear regression analysis.

#### **RESULTS AND DISCUSSION**

The analysis was carried out to see the impact of weather parameters for pre-harvest wheat yield forecasting on agro-climatic zone basis in Haryana state. The developed zonal models are based on timeseries data of weather parameters from 1978-79 to 2004-05 and trend based yield as well, however, the data from 2005-06 to 2007-08 were used for validation of the models. Year/time variable was included to take care of variation between districts within zone as the weather data were not available for all the districts, though the zonal model utilized the same weather data in the adjoining districts under the zone. Data for the last one month of wheat crop season were excluded from the analysis, as the idea behind the study was to predict yield(s) about one month in advance of the

M. Goyal and U. Verma	. / J. Appl. & Nat.	. Sci. 8 (3): 1485 -	1492 (2016)
-----------------------	---------------------	----------------------	-------------

Zones	Discriminant Function	Eigenvalue	Percent Variation Explained	Wilk's Lambda	df	Chi- Square	Sig.
Zone-I	1	0.18	55.5	0.75	6	25.71	< 0.01
	2	0.14	44.5	0.88	2	11.54	< 0.01
Zone-II	1	11.44	82.7	0.02	54	529.56	< 0.01
	2	2.39	17.3	0.30	26	172.84	< 0.01
Zone-III	1	4.86	84.9	0.09	54	222.48	< 0.01
	2	0.87	15.1	0.54	26	58.05	< 0.01
Zone-IV	1	4.21	86.0	0.11	54	178.18	< 0.01
	2	0.69	14.0	0.59	26	42.82	0.02

Table 1. 7	Zone-wise a	contribution	of linear	discriminant	functions	and their significance	١.

Table 2. Discriminant scores based on fortnightly weather data for all the zones.

	Zoi	ne-I	Zon	e-II	-II Zone-III			Zone-IV		
Years –	DS <sub>1</sub>	DS <sub>2</sub>								
1978-79	0.36	0.94	5.67	1.26	3.58	0.99	3.28	1.21		
1979-80	0.81	2.12	5.86	1.76	3.58	0.99	3.20	1.20		
1980-81	1.16	1.73	6.03	1.85	3.58	0.99	3.36	1.22		
1981-82	-0.23	0.72	5.93	1.84	3.58	0.99	3.42	1.23		
1982-83	0.99	1.18	5.80	1.89	3.58	0.99	3.76	1.28		
1983-84	-0.27	0.06	6.01	1.93	3.58	0.99	3.49	1.24		
1984-85	-1.71	0.81	5.97	2.30	3.58	0.99	3.55	1.25		
1985-86	1.91	1.10	5.74	1.54	3.58	0.99	3.21	1.20		
1986-87	-0.40	2.12	5.14	0.75	3.58	0.99	2.05	-0.16		
1987-88	-0.59	0.04	4.08	0.46	2.18	-0.23	0.95	-1.51		
1988-89	1.55	0.10	4.45	-0.18	2.18	-0.23	0.95	-1.51		
1989-90	1.02	0.13	2.40	-1.51	0.42	-1.75	1.27	-1.46		
1990-91	2.06	-0.14	1.99	-2.43	0.42	-1.75	0.79	-1.54		
1991-92	1.96	-0.10	1.82	-2.17	0.42	-1.75	0.77	-1.54		
1992-93	-0.57	-1.09	1.54	-2.36	0.42	-1.75	0.98	-1.51		
1993-94	-0.34	-2.09	1.27	-2.67	0.42	-1.75	1.26	-1.47		
1994-95	-1.65	-1.45	1.46	-2.32	-0.63	-2.66	-0.26	-2.88		
1995-96	-0.30	0.00	1.15	-2.34	-1.17	-1.53	-0.13	-2.86		
1996-97	-0.27	-1.83	1.45	-2.30	-1.17	-1.53	-0.97	-1.40		
1997-98	-0.01	0.09	-0.71	-0.25	-1.50	-0.86	-2.03	0.43		
1998-99	0.85	-0.61	-0.95	-0.75	-1.50	-0.86	-2.12	0.42		
1999-00	0.85	-0.61	-1.67	0.20	-2.38	0.95	-2.07	0.42		
2000-01	0.34	0.25	-1.87	1.55	-2.38	0.95	-2.21	0.40		
2001-02	0.67	-0.03	-1.94	1.34	-2.38	0.95	-2.02	0.43		
2002-03	-0.41	0.00	-2.95	1.87	-2.38	0.95	-1.89	0.45		
2003-04	0.05	-0.91	-3.24	1.78	-2.82	1.85	-2.03	0.43		
2004-05	-2.09	2.39	-2.40	2.12	-2.82	1.85	-2.01	0.43		
2005-06	-0.88	-0.67	-5.04	1.35	-2.38	0.95	-2.05	0.43		
2006-07	-1.01	-0.19	-4.38	1.65	0.42	-1.75	-2.69	1.52		
2007-08	-1.43	0.22	-3.74	1.56	0.42	-1.75	-2.67	1.53		

actual harvest. The multiple linear regression and discriminant function analyses were used to obtain different zonal trend-agromet-yield equations. The best subsets of weather variables were selected using stepwise regression method in which all variables were first included in the model and eliminated one at a time with decisions at any particular step conditioned by the result of previous step. The best supported weather variables were retained in the model if they had the highest adjusted adj.  $R^2$  and low-

est standard error (SE) of estimate at a given step. The selected zonal trend-agromet-yield models are as follows: Zone-1 : (Ambala , Kurukshetra, Yamunanagar) Yield<sub>est.</sub> (model-1) = { $c_1 + (a_1 x T_r) + (a_2 x TMN_1) + (a_3 x TMN_8) + (a_4 x ARF_4)$ }

 $\begin{aligned} \text{Yield}_{est.} \quad (\text{model-2}) &= \{ c_2 + (b_1 x \ T_r) + (b_2 x \ DS_1 \ ) + (b_3 x \ DS_2 \ ) \} \end{aligned}$ 

 $\begin{array}{l} \text{Zone-II: (Rohtak, Karnal, Jind, Sonipat, Panipat, Kaithal )} \\ \text{Yield}_{est.} \quad (model-1) = \{c_1 + (a_1 \ x \ T_r) + (a_2 \ x \ TMN_3 \ ) + \\ (a_3 \ x \ TMX_4 \ ) + (a_4 \ x \ ARF_1) \ + (a_5 \ x \ ARF_7) \} \end{array}$ 

Table 3.   Selected zon	al trend-agromet wheat	yield models.		
Zone-I Model Variable	Coefficients		Zone-II Model Variable	Coefficients
		Model-1		
Constant	c <sub>1</sub>	-6.25	Constant	-8.93
T <sub>r</sub>	a <sub>1</sub>	1.05	Tr	1.05
$TMN_1$	$a_2$	0.90	$TMN_3$	-0.29
TMN <sub>8</sub>	a <sub>3</sub>	-0.72	$TMX_4$	0.45
$ARF_4$	$a_4$	0.03	$ARF_1$	-0.06
			ARF <sub>7</sub>	0.05
$R^2 = 0.905$ adj. $R^2 = 0.900$	SE= 2.23		$R^2 = 0.930$ adj. $R^2 = 0.928$	SE= 1.86
		Model-2		
Constant	c <sub>2</sub>	-0.38	Constant	2.33
$T_r$	<b>b</b> <sub>1</sub>	1.00	T <sub>r</sub>	0.93
DS <sub>1</sub>	b <sub>2</sub>	1.27	$DS_1$	-0.09
$DS_2$	b <sub>3</sub>	-0.62	$DS_2$	-0.38
$R^2 = 0.889$	SE= 2.40		$R^2 = 0.896$	SE= 2.26
$adj.R^2 = 0.885$			$adj.R^2 = 0.894$	
Zone-III			Zone-IV	
Model Variable	Coefficients		Model Variable	Coefficients
		Model-1		
Constant	c <sub>1</sub>	-6.03	Constant	15.33
T <sub>r</sub>	$a_1$	0.91	T <sub>r</sub>	1.08
$TMN_1$	$a_2$	0.52	$TMX_3$	-0.36
$TMN_4$	a <sub>3</sub>	0.56	$TMX_8$	-0.39
TMN <sub>9</sub>	$a_4$	-0.22	$ARF_1$	-0.35
ARF <sub>6</sub>	a <sub>5</sub>	0.05	$ARF_2$	-0.43
$R^2 = 0.906$	SE= 2.23		$R^2 = 0.864$	SE= 2.60
$adj.R^2 = 0.901$			$adj.R^2 = 0.857$	
		Model-2		
Constant	c <sub>2</sub>	5.66	Constant	3.51
T <sub>r</sub>	<b>b</b> <sub>1</sub>	0.82	Tr	0.90
$DS_1$	<b>b</b> <sub>2</sub>	-0.30	$DS_1$	-0.08
DS <sub>2</sub>	<b>b</b> <sub>3</sub>	-0.94	$DS_2$	-1.04
$R^2 = 0.897$	SE= 2.31		$R^2 = 0.877$	SE= 2.42
$adj.R^2 = 0.894$			$adj.R^2 = 0.872$	

M. Goyal and U. Verma. / J. Appl. & Nat. Sci. 8 (3): 1485 - 1492 (2016)

T	h	1.00
1.	U1	1.00
/	- 1	

Yield<sub>est.</sub> (model-2) = { $c_2 + (b_1 \times T_r) + (b_2 \times DS_1) + (b_3 \times DS_2) + (b_3$  $x DS_2$  )}

Zone-III: (Gurgaon, Mahendergarh, Jhajjar, Faridabad, Rewari)

Yield<sub>est.</sub> (model-1) = { $c_1 + (a_1 x T_r) + (a_2 x TMN_1) +$  $(a_3 \times TMN_4) + (a_4 \times TMN_9) + (a_5 \times ARF_6)$ 

Yield<sub>est.</sub> (model-2) = { $c_2 + (b_1 x T_r) + (b_2 x DS_1) + (b_3 x DS_2) + (b_3$  $x DS_2$  )

Zone-IV : (Hisar, Bhiwani, Sirsa, Fatehabad)

Yield<sub>est.</sub> (model-1) = { $c_1 + (a_1 x T_r) + (a_2 x TMX_3) +$  $(a_3 \times TMX_8) + (a_4 \times ARF_1) + (a_5 \times ARF_2)$ 

Yield<sub>est.</sub> (model-2) = { $c_2 + (b_1 x T_r) + (b_2 x DS_1) + (b_3 x DS_2) + (b_3$  $x DS_2$  )

Regression models: Model-1: Weather parameters and trend yield as regressors, Model-2: Discriminant/weather scores and trend yield as regressors,

Where, Yieldest. - Model predicted yield (q/ha) Tr -Trend yield (q/ha) TMX - Average maximum tempera-TMN - Average minimum temperature ARF ture Accumulated rainfall (1,2,3,...,11 refer to different  $DS_i - i^{th}$  discriminant score (i = 1,2) SE fortnights) Standard error of the estimate  $R^2$  - Coefficient of determination

The performance of the selected models were evaluated on the basis of percent relative deviations from real-time yields and root mean square error(s) (Tables- 4 & 5).

The district-level pre-harvest wheat yield estimates for 2005-06, 2006-07 and 2007-08 in Table-4 showed good agreement with real time wheat yield(s) for almost all the districts under consideration. Figures-1 to

**Table 4.** Zonal model based pre-harvest wheat yield estimates along with percent relative deviations for different districts of Haryana.

Forecast	Zone-I					Zone-III				
years	Ambala					Gurgaoi	1			
	DOA	Model-1		Model-2		DOA	Model-1		Model-2	
	Vield	Fitted	RD	Fitted	RD	Vield	Fitted	RD	Fitted	<b>RD</b> (%)
	$(\alpha/ha)$	Yield	(%)	Yield	(%)	(a/ha)	Yield	(%)	Yield	<b>RD</b> (70)
	(4,)	$(\alpha/ha)$	(,,,)	(a/ha)	(,,,,)	(4,)	(g/ha)	(,,,,)	(q/ha)	
2005-06	37.89	33.91	10.50	37.96	-0.18	34.95	34.92	0.09	36.45	-4.29
2006-07	38.06	39.12	-2.79	38.26	-0.53	39.66	37.96	4.29	37.08	6.51
2007-08	39.82	38.06	4.42	38.24	3.97	39.13	38.00	2.89	38.45	1.74
	Kurukshe	etra				Mahend	ergarh			
2005-06	45.82	42.63	6.96	45.95	-0.28	41.19	41.73	-1.31	42.54	-3.28
2006-07	46.72	46.99	-0.58	46.27	0.96	39.86	44.84	-12.49	43.24	-8.48
2007-08	47.72	47.78	-0.13	46.29	3.00	37.21	44.96	-20.83	44.67	-20.05
	Yamuna	nagar				Jhajjar				
2005-06	36.82	34.21	7.09	37.89	-2.91	37.10	36.89	0.57	38.21	-2.99
2006-07	41.55	38.37	7.65	38.03	8.47	39.17	39.73	-1.43	38.67	1.28
2007-08	37.71	38.97	-3.34	37.85	-0.37	39.23	39.58	-0.89	39.86	-1.61
	Zone-II									
	Rohtak					Faridaba	d			
2005-06	34.41	41.88	-21.71	46.78	-7.73	36.49	39.75	-8.93	40.77	-11.73
2006-07	39.58	42.80	-8.14	45.42	7.38	37.06	42.86	-15.65	41.47	-11.90
2007-08	35.10	41.08	-17.04	44.80	-3.13	46.46	42.97	7.51	42.88	7.71
	Karnal					Rewari				
2005-06	43.67	48.74	-11.61	45.85	-4.99	41.87	45.75	-9.27	46.13	-10.17
2006-07	44.23	47.62	-7.66	46.57	-5.29	39.86	39.45	1.03	38.42	3.61
2007-08	46.29	48.12	-3.95	47.40	-2.40	40.27	34.74	13.73	35.53	11.77
						Zone-IV	-			
	Jind					Hisar				
2005-06	40.18	47.23	-17.55	44.42	-10.55	37.04	38.61	-4.24	43.56	-17.60
2006-07	42.12	46.19	-9.66	45.21	-7.34	43.92	47.05	-7.13	43.07	1.94
2007-08	41.93	46.76	-11.52	46.10	-9.95	39.20	45.44	-15.92	43.65	-11.35
	Sonipat					Bhiwani				
2005-06	40.24	45.97	-14.24	43.22	-7.41	32.78	31.68	3.36	37.80	-15.31
2006-07	45.47	44.85	1.36	43.94	3.36	39.64	40.04	-1.01	37.24	6.05
2007-08	42.94	45.36	-5.64	44.77	-4.26	39.63	38.36	3.20	37.76	4.72
	Panipat				· · -	Sirsa				
2005-06	43.97	46.79	-6.41	44.00	-0.07	34.32	37.80	-10.14	42.89	-24.97
2006-07	44.77	45.48	-1.59	44.53	0.54	45.85	46.22	-0.81	42.38	7.57
2007-08	43.75	45.78	-4.64	45.17	-3.25	42.55	44.58	-4.77	42.94	-0.92
2005 05	Kaithal	16.21	10.04	12 5 1	2.74	Fatehab	ad	7.01	10	4.01
2005-06	41.97	46.31	-10.34	43.54	-3.74	40.72	37.54	7.81	42.68	-4.81
2006-07	43.73	44.92	-2.72	44.01	-0.64	46.32	45.67	1.40	41.93	9.48
2007-08	44.39	45.16	-1./5	44.58	-0.43	40.46	45./6	5.81	42.26	9.04

**Table 5.** Predictive accuracies of zonal trend-agromet models in terms of average absolute percent deviations of estimated wheat yields from real-time wheat yields based on two different statistical procedures.

Zones	Districts	Average absolute percent deviations of f	orecast yields of 2005-06, 06-07 and 07-08
		Model -1	Model -2
Zone-I	Ambala	5.90	1.56
	Kurukshetra	2.56	1.41
	Yamunanagar	6.03	3.92
Zone-II	Rohtak	15.63	6.08
	Karnal	7.74	4.23
	Jind	12.91	9.28
	Sonipat	7.08	5.01
	Panipat	4.21	1.28
	Kaithal	4.93	1.60
Zone-III	Gurgaon	2.42	4.18
	Mahendergarh	11.54	10.60
	Jhajjar	0.96	1.96
	Faridabad	10.70	10.44
	Rewari	8.01	8.52
Zone-IV	Hisar	9.09	10.30
	Bhiwani	2.52	8.70
	Sirsa	5.24	11.15
	Fatebabad	5.01	7 78



**Fig. 1**. Percent deviations of model(s) based wheat yields (2005-06) from real-time yields.



**Fig. 3.** *Percent deviations of model(s) based wheat yields* (2007-08) *from real-time yields.* 

3 show the percent deviations of models' based yield forecasts from DOA yields based on two different statistical procedures for three post-sample years separately. The predictive accuracies of the zonal wheat yield models in terms of average absolute percent deviations from DOA wheat yield(s) falling within acceptable limits and RMSEs at zonal level expressed in Tables-5 and 6 prefer the use of zonal models based on discriminant function analysis for almost all the districts falling in different zones of Haryana. Trend yield has been observed as a significant parameter appearing in all the models, which is an indication of technological advancement, improvement in fertilizer/insecticide/ pesticide/ weedicide use and increased use of high yielding varieties over time. The zonal trend-agromet-yield forecasting models developed for this empirical study allow prediction of district-level wheat yield well in advance of the actual harvest of the crop. The methods introduced in this paper may be applied to other crops/locations in the state.

### Conclusion

This study has demonstrated the utility of understanding and quantifying the relationships between wheat yield and weather variables. The relationships can be employed in studies that explore the impact of climate change on probable future wheat yields at regional scales. The predictive performance(s) of the zonal wheat yield models were assessed using multiple metrics, including the adj- $R^2$ , percent deviations of the forecast yields from the state DOA yields and RMSEs.



**Fig. 2.** Percent deviations of model(s) based wheat yields (2006-07) from real-time yields.

Table 6. RMSEs at zonal level in Haryana.

Zones	Model -1	Model -2
Zone-I	233.00	143.00
Zone-II	411.34	226.58
Zone-III	357.82	340.02
Zone-IV	271.38	431.94

The fitted trend-agromet wheat yield relationships were used to obtain the wheat yield forecasts of 2005-06, 2006-07 & 2007-08. The estimated yield(s) from the selected zonal models showed good agreement with the real-time wheat yields by showing less than 5 percent deviations in 9 districts and 6-11 percent deviations in the remaining 9 districts under consideration. Moreover, the fitted models may be used to provide reliable yield forecasts of wheat crop about one month in advance of the crop harvest while the state DOA yield estimates are obtained quite late after the actual harvest of the crop. Thus, the zonal models developed may be considered usable for district-level operational wheat yield forecasting in Haryana. Although this empirical analysis produced a model with adequate accuracy for pre-harvest forecasting purposes but it would also be worthwhile exploring if other summaries of the weather variables using alternative time windows besides the fortnight summaries we used, may improve the models' performance.

#### ACKNOWLEDGEMENTS

The weather data received from Haryana Space Applications Centre, Hisar and Department of Agrometeorology, CCS HAU, Hisar, India are gratefully acknowledged.

### REFERENCES

- Bazgeer, S., Kamali, G.H. and Mortazavi, A. (2007). Wheat yield prediction through agro-meteorological indices for Hamedan. *Iran. Biaban*, 12 : 33-38.
- Bhuyan, K.C. (2005). Multivariate analysis and its applications. 1<sup>st</sup> edition, New Central Book Agency, Kolkata, India.
- Chandrahas, Agrawal, R. and Walia, S.S. (2010). Use of discriminant function and principal component techniques for weather based crop yield forecast. IASRI/PR-09/2010, IASRI, New Delhi.
- Chandran, K.P. and Prajneshu (2004). Modelling effect of

sunshine and temperature on rice (*Oryza sativa*) tiller production using non parametric regression. *Ind. J. Agri. Sci.* 74(10) : 563-65.

- Draper, N. and Smith, H. (1981). Applied Regression Analysis. 2nd edition, Wiley, New Esfandiary, York., F., Aghaie, G. and Mehr, A. D. (2009). Wheat yield prediction through agro meteorological indices for Ardebil district. World Academy of Science. *Engineering and Technology*, 49 : 32-35.
- Hair, J.F., Anderson, J.R.E., Tatham, R.L., Black and William C. (1995). Multivariate data analysis with readings.Pretice Hall, New Jersey.
- Mehta, S.C., Pal, S. and Kumar, V. (2010). Weather based models for forecasting potatoyield in Uttar Pradesh. Project Report : IASRI/P.R.-01/2010, IASRI, New Delhi.
- Mkhabela, M.S., Mkhabela, M.S. and Mashinini, N.N. (2005). Early maize yield forecasting in the four agroecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. Agricultural and Forest Meteorology, 129 : 1-9.
- 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: 9971-75.1

- Rai, T. and Chanrahas (2000). Using the discriminant function of weather parameters for developing forecast model of rice crop. IASRI, Library Avenue, New Delhi.
- Ruíz, J.S. and Ordonez, Y.F. (2003). Prediction of corn yield in Mexico using vegetation indices from NOAA-AVHRR satellite images and degree-days. *Geocarto International*, 18(4): 33-42.
- Vashisth, A., Singh R. and Choudhary M. (2014). Crop yield forecast at different growth stages of wheat crop using statistical model under semi Arid region. *Journal of Agroecology and Natural Resources Management*, 1: 1-3.
- Verma, U., Dabas, D.S., Grewal, M.S., Singh, J.P., Hooda, R.S., Yadav, M., Kalubarme, M.H., Sharma, M.P. and Prawasi, R. (2011). Crop yield forecasting in Haryana: 1986 to 2010. Summary Report, pp.1-148. Department of Soil Science, CCS HAU, Hisar.
- Verma, U., Piepho, H.P., Ogutu, J.O., Kalubarme, M.H. and Goyal, M. (2012). Multi-level mixed modelling for weather-crop-yield relationships on agro-climatic zone basis in Haryana. Advances and Applications in Statistics, 28(1): 1-22.