Comparison of Different Regression Models to Predict Mustard yield in Central Punjab

Kavita Bhatt, K. K. Gill and Sandeep Singh Sandhu
School of Climate Change & Agricultural Meteorology
Punjab Agricultural University, Ludhiana
Email: kgill2002@gmail.com

ABSTRACT

An attempt was made to predict mustard yield over central Punjab (Ludhiana, Amritsar and Patiala) by regression models. Three statistical models have been developed for forecasting the yield of the mustard using the yield data and different weekly weather variables viz. maximum and minimum temperature, morning and evening relative humidity, sunshine hours, rainfall and number of rainy days. The sensitive period for mustard crop with respect to weather parameters were identified for different weather parameters by using correlation and selected windows with high correlation values were taken for further regression analysis. In the first Basic model, the multiple regression technique was used to predict the crop yield by using only weather parameters. The second model is Modified model, where technology trend was taken as one of the extra variable in multiple regressions. In the third model, multiple regression analysis was done using SPSS software. Regression equations were developed separately for all three models and were used to predict the mustard yield. The data for a period of (1974-2011) was used to develop the forecast model and further three years meteorological data (2012-14) was used to validate the models. For Ludhiana, among all the three models, Basic model explained up to 84 % variation in yield due to weather parameters while modified model explained highest i.e. 87% variation in mustard yield due to weather parameters. In case of Amritsar district, basic, modified and SPSS model explained 57, 59 and 62 % variation in mustard yield, respectively. In case of Patiala, the modified model was the best fit model to explain mustard yield as it explained up to 82 % variation in yield followed by SPSS (76 %) and basic model (49 %). The results reveal that modified model is best fit for Ludhiana and Patiala region, while SPSS model fits better for Amritsar region as far as mustard yield is concerned.

Key words: *Correlation, Multiple regression, SPSS, Technology trend, Yield forecasting.*

1. Introduction

Climate change is a challenge of 21st century for all of us for several reasons. Agricultural vulnerability is a major contributor in the overall vulnerability of a region to climate change (Hiremath and Shiyani 2012). It is likely to impact the yield of various crops threatening food security (FAO 2006). One important challenge is to reliably assess its quantitative impact on the yield of different crops. Weather has a very close relationship with the crops and it affects crop growth at different phenological stages. Therefore, large variation in yields from year to year and place-to-place can be attributed to the weather. An analysis of crop and weather relationships can also be of great help in prediction of growth and yield of crops.

Oilseeds are an important component of the agricultural economy, next to food grains, in terms of area, production and value. Among the oilseed crops, rapeseed and mustard occupy second position next to groundnut. At the' global level, India ranks second next (Yadava *et al*. 2001). In future, the demand for oilseeds production is likely to go up significantly due to increase in population and income. The importance of oilseed crops as a source of both fat and protein is well known. Mustard is highly sensitive to weather, showing quiet diverse pattern of growth and development of different sets of environmental conditions. Among all the weather parameters, temperature and solar radiation are the most dominant factors affecting the crop growth, dates of occurrence of different phenological stages and crop yield. A part of the decline and/or stagnation in mustard yields causing negative growth rate from 1997 was possibly due to unfavorable monsoon which created moisture stress (drought and excess rainfall) and temperature increases (Kumar 2005). High temperature during mustard crop establishment (mid September to early November), cold spell, fog and intermittent rains during crop growth also affect the crop adversely and cause considerable yield losses by physiological disorder along with appearance and proliferation of aphid pest, white rust, downy mildew and stem rot diseases. Since decades, attempts have been made to predict phenology and yield (Tripathi *et al*. 1999) of crops using thermal based indices. Statistical models can be used to evaluate the agricultural production risks as a function of climatic variability to assess regional crop yield potential across a wide range of conditions and to determine suitable planting dates and other management factors for increasing crop yield (Boote *et al.* 1996).

Under Punjab conditions, the growing season (October - April) coincides with a period of very low to high evaporation demand, abundant sunshine and moderate to high solar radiation. Mustard prefers moderate temperature between 25 and 28^{0}C with optimum around 20^{0}C and moderate rain of about 25-40 cm during growth period. Regression analysis can discover quantitative relationships between the related variables (George *et al*. 1978).The present studies proposed a statistical modeling methodology for mustard crop. The major objectives of the present study is to identify the sensitive period of mustard crop towards different weather parameters and to also to analyse the best fit model for the prediction of mustard yield for central Punjab.

2. Data and Method of Analysis

For developing forecasting models, long series of actual crop yield data, weather data technological parameters were used. By studying the relationship of yield with different weather elements, predictors are identified. The present attempt is based on the correlation regression technique. The study involves seven weather parameters (maximum and minimum temperature, morning and evening relative humidity, sunshine hours, precipitation and number of rainy days on weekly basis) for developing the three different statistical models for predicting the mustard yield for different districts of central Punjab.

2.1. Mustard yield data

In the present study, the yearly production (q) and area (ha) under mustard crop for the period 1973 to 2014 in respect of Ludhiana, Amritsar and Patiala has been collected from Statistical Abstract of Punjab. For each year the total production of mustard for different districts were divided by the total acreage to calculate the mustard productivity.

2.2. Weather data

Weekly data of different weather elements (maximum temperature, minimum temperature, rainfall, sunshine hour, number of rainy days, morning and evening relative humidity) for the years 1973 to 2014 were collected from the meteorological observatory installed at School of Climate Change and Agricultural Meteorology, PAU, Ludhiana for Ludhiana district. The other two districts representing central plain zone i.e. Amritsar and Patiala, the weather data have been collected from IMD, Chandigarh for the months covering life cycle of the crop (43rd week to 12th standard meteorological week) except the harvesting period, since the forecast is to be given before harvesting.

2.3. Yield forecast models

The correlation analysis has been carried out using Pearson correlation technique and the statistical model is developed using multi-regression method. All the three models were developed from a data series of 35-38 years (1973 to 2011 for Ludhiana, Amritsar and Patiala district) and the model has been verified with independent data for the years from 2012 to 2014, outside their sampling series. The performance of the model has been examined critically by computing percentage deviations of estimated productivity (from statistical abstracts) and forecast yield Figs.

2.3.1. Basic model (Model 1)

A basic model was developed by using weather parameters taking into account the sensitive period window for mustard crop. The correlations were worked out for sensitive periods and then multiple regression equations were developed by using G-stat. The average reported crop-yield was taken as dependent variable with weather parameters as independent variables. In the correlation and regression technique significant correlation between yield and the meteorological parameters were identified. The critical periods when weather parameters exert significant influence on yield were identified by analyzing the correlation coefficients for statistical and phenological significance. Out of all the periods, the sensitive periods of statistical and phenological significance were selected in terms of standard meteorological weeks (SMWs) for regression analysis.

2.3.2. Modified model (Model 2)

Secondly, a modified model was obtained by introducing an assumed technological trend in the basic model keeping other independent variables as constant. The development of modified model was intended to improve the accuracy of yield forecast, by superimposing the impact of agricultural technology in the form of linear time scale dummy variable.

2.3.2.1. Technology trend

Since 1950 crop yield in many meteorological sub-divisions registered a general upward trend due to advancement in the field of agricultural technology, like use of high yielding variety seeds, timely operations, better irrigation and drainage facilities, large scale use of fertilizers and pesticides

Table 1

Baseline data used in developing the model for different districts

Year	Tmax			Tmin			Rainfall			Technology
	Ldh.	Amt.	Pat.	Ldh.	Amt.	Pat.	Ldh.	Amt.	Pat.	
1973-74	28.0	26.8	27.7	9.7	10.8	14.4	0.0	0.0	0.0	1.0
1974-75	25.8	26.1	26.6	8.9	9.6	12.0	0.0	0.0	1.1	1.0
1975-76	23.7	23.5	24.5	8.3	10.3	13.4	0.0	7.2	2.1	1.0
1976-77	29.8	28.1	28.6	9.4	10.7	13.7	0.0	0.0	0.5	1.0
1977-78	23.2	22.9	23.5	9.1	9.7	11.4	8.2	42.5	31.4	1.5
1978-79	23.4	20.9	23.3	8.4	9.0	12.2	1.0	33.3	3.4	2.5
1979-80	25.3	24.3	25.5	9.9	10.9	12.2	5.2	6.1	13.2	3.5
1980-81	24.1	23.7	24.4	10.4	11.4	14.0	17.5	3.2	4.7	5.0
1981-82	21.8	19.7	21.4	9.1	10.2	12.9	1.9	52.7	50.6	5.0
1982-83	25.0	24.1	24.7	10.2	9.7	12.5	0.0	22.6	9.6	5.0
1983-84	28.1	27.0	27.7	9.7	12.5	15.4	0.0	0.0	0.0	5.0
1984-85	29.4	29.6	29.2	10.9	10.7	13.9	0.0	0.0	0.0	5.0
1985-86	24.5	25.8	26.0	9.8	10.2	12.8	0.0	0.0	0.0	7.0
1986-87	26.3	25.0	26.5	11.7	12.0	14.6	7.5	0.0	0.0	9.0
1987-88	25.8	26.1	26.4	11.3	10.5	13.7	0.0	5.9	12.0	9.0
1988-89	26.4	26.1	27.2	9.8	10.1	14.7	0.0	0.0	0.0	11.0
1989-90	25.6	24.7	25.0	10.3	10.9	15.1	0.0	22	12.6	12.0
1090-91	25.3	24.2	26.2	11.0	9.8	13.1	0.0	6.0	1.1	12.0
1991-92	25.0	25.7	26.8	10.3	9.3	14.8	0.0	0	3.1	14.0
1992-93	25.0	24.8	25.2	10.8	10.6	13.9	0.0	10.0	3.6	14.0
1993-94	27.4	24.8	28.2	11.1	10.5	16.6	0.0	11.8	0.0	14.0
1994-95	25.8	23.8	25.6	9.0	8.5	14.5	0.0	11.8	7.7	15.0
1995-96	26.0	25.9	26.3	12.5	12.7	16.2	2.6	0.0	0.0	16.0
1996-97	25.9	25.9	26.2	11.0	11.1	14.2	0.0	0	0.0	16.0
1997-98	24.7	22.2	—-	10.0	9.0	—-	0.4	4.8	—-	17.0
1998-99	28.0	26.4	28.4	11.0	10.4	12.0	0.0	0.0	0.0	17.0
1999-00	25.5	25.2	26.3	8.8	8.1	11.6	0.0	0.0	2.0	17.0
2000-01	28.7	27.5	28.1	11.4	11.2	13.9	0.0	2.0	0.0	17.0
2001-02	29.0	25.7	27.4	12.2	12.4	15.4	0.0	8.0	11.0	18.0
2002-03	27.0	24.2	—-	11.6	10.6	—-	0.0	26.0	—-	18.0
2003-04	30.9	29.2	31.0	13.1	12.7	17.1	0.0	0.0	0.0	18.0
2004-05	26.2	23.1	27.1	12.1	11.3	15.2	1.5	2.2	4.7	18.0
2005-06	26.3	26.0	27.4	12.9	11.5	13.2	0.0	8.2	3.2	19.0
2006-07	25.0	22.9	25.3	11.1	9.7	13.5	0.0	14.4	13.8	19.0
2007-08	30.0	28.6	30.7	12.5	11.1	15.4	1.3	0.0	0.0	19.0
2008-09	29.5	27.0	28.5	12.0	11.6	13.5	0.0	0.0	0.0	20.0
2009-10	30.5	27.1	—-	13.1	14.3	—-	9.4	0	—-	20.0
2010-11	28.2	24.5	27.2	11.5	12.2	15.5	0.0	7.7	7.1	20.0

Ldh. = Ludhiana, Amt. = Amritsar, Pat. = Patiala (— = Data not available) etc. All these have contributed to the sharp rise in crop yield as compared to the yield of earlier years. Basic data used in developing different models along with technology trend was compiled as table 1.

Regression equation for model 2 is given by:

$$Y_e = a_0 + \sum_{i=1}^{n} a_i x_i + \Sigma a_j x_j$$

Where,

- Y_e = Estimated yield, kg/ha
- a_0 = Regression constant
- a_i = Regression coefficients for meteorological predictor variables
- x_i = i th meteorological predictor variable i = 1, 2, ……n
- a_j = Regression coefficients for technological trend variables.
- x_j = j^{th} technological trend variable (1-20)

2.3.3. SPSS model (Model 3)

The third model is based on analyzing regression using SPSS software. SPSS (Statistical Product and Service Solutions) developed by (Norman *et al.,* 1968). Pearson's correlations between observed mustard yield and weather parameters and with combinations of weather parameters were computed. Sum of weather parameter and sum product of different weather parameter and correlation coefficient was derived. Multiple regressions between dependent variable (yield) and independent variables (time, sum and sum products for different weather parameters) were analysed using SPSS software. Regression equation was written using the regression formula.

Regression equation for models 1 and 3 is given by:

$$Y_e = a_0 + \sum_{i=1}^{n} a_i x_i$$

Where,

- Y_e = Estimated yield, kg/ha
- a_0 = Regression constant
- a_i = Regression coefficients for meteorological predictor variables
- x_i = i^{th} meteorological predictor variable i = 1, 2, ……n

3. Results and Discussion

For different districts of central Punjab, three different yield forecast equations have been developed using the three different models by using weekly data of weather variables.

Out of all the periods, the sensitive periods of statistical and phenological significance were selected for different districts in terms of standard meteorological weeks (SMWs) for regression analysis. For Ludhiana, Amritsar and patiala districts, effect of sensitive period on mustard yield was analysed and represented in table 2. The sensitive periods for mustard crop represent emergence, flowering, pod formation, pod filling and physiological maturity. Analysis showed maximum and minimum temperature showed positive effect on the mustard yield during the sensitive period, whereas, total rainfall showed negative effect on the mustard yield.

The final yield forecast functions by different models using important weather variables are presented below.

3.1. Model 1

In the basic model the weather data at critical periods were correlated with the ield and these correlations were used for regression analysis. For Ludhiana district the regression equation is as follows :

Yield=(-270.445)-9.22943Tmax(1)-5.06593 Tmax (10-12)+37.9737 Tmin (49-51)+57.1909 Tmin (8-12)-18.0487 Rainfall (48)-145.469 Rdays (48)-74.504Rd(4)+19.1242Rmax(51-5)-4.72503Rmin(50-51)-3.47254Rmin(1-3)-21.1428SSH(43-44)-64.6413SSH(52-2)

R^2= 84 %

The above analysis showed 84 % of variation in mustard yield is due to weather parameters. The per cent error ranged between -24.3 to 18.5 %. The forecasted yield and per cent error of 38 years, based on above regression equation is given in Table 3.

Regression equation developed for Amritsar district is as follows:

Yield=1916.88-23.0133Tmax(1-3)-22.4153Tmax(9-11)-48.474Tmin(52)+28.4689Tmin(10-12)-14.657Rf(2)-3.93488Rf(9)-48.5287Rdays(9-10)

R^2= 57 %

The above used model is explained 57% variation in mustard yield. The per cent error ranged between -22.4 to 20.5 %. The forecasted yield and per cent error of 38 years, based on above regression equation is given in Table 4.

For Patiala district, regression equation is as follows:

Yield=(-1365.29)+9.12496Tmax(5-6)+51.0079Tmax(9-12)+52.8261Tmin(6)+18.2056Tmin(11-12)-9.08652Rf(48)+6.35599Rf(9-10)-146.612Rdays(48)-98.8484Rdays(9-10)

R^2= 49 %

This model explains 49 % of variation in the mustard yield. The per cent error ranged from -27.2 to 24.7 %. The forecasted yield and percent error of 35 years, based on above regression equation is given in Table 5.

3.2. Model 2

In second model (Modified model), the weather parameters in critical periods along with technology trend variable were used through multiple regression analysis to obtain forecast equation. Improvement in agricultural technology over time necessitated the need to modify the basic model by introducing technological trend as an independent linear time scale dummy variable. The regression equation for Ludhiana is as follows:

Yield=173.59-15.8159Tmax(1)+3.25709Tmax(10-12)+27.9941Tmin(49-51)+24.0326Tmin(8-12)-12.9118Rf(48)-122.621Rdays(48)-56.8275Rdays(4)+13.0294Rmax(51-5)-3.75642Rmin(50-51)-4.86862Rmin(1-3)-3.46756SSH(43-44)-43.6361SSH(52-2)+16.0761TT

R^2= 87 %

The value of R^2 has increased to 87 % in the modified model. The error per cent for modified model ranged from –22.5 to 15.8 % during the last 38 years (Table 3).

For Amritsar regression equation with modified model is as follows:

Yield=1508.68-13.6366Tmax(1-3)-15.4116Tmax(9-11)
43.0674Tmin(52)+25.6446Tmin(10-12)-14.0706Rf(2)-2.99853Rf(9)-44.5188Rdays(9-10)+5.26525TT

R^2 = 59 %

Table 2

Sensitive periods and effect of weather variables on mustard yield in Punjab

Districts	Sensitive Period (SMWs)	Stage of mustard crop	Effect on mustard yield
	Maximum Temperature		
Ludhiana	10 to 12 (26.4^{0}C)	late reproductive stage	+ve
Amritsar	9 to 11 (25.1^{0}C)	late reproductive stage	+ve
Patiala	9 to 12 (26.5^{0}C)	late reproductive stage	+ve
	Minimum Temperature		
Ludhiana	8 to 12 (10.7^{0}C)	late reproductive stage	+ve
Amritsar	10 to 12 (10.7^{0}C)	late reproductive stage	+ve
Patiala	11 to 12 (14.0^{0}C)	late reproductive stage	+ve
	Rainfall		
Ludhiana	48 (1.4 mm)	Flowering stage	-ve
Amritsar	9 (9.4 mm)	late reproductive stage	-ve
Patiala	9 and 10 (6.6 mm)	late reproductive stage	-ve

SMW = Standard Meteorological Week

Table 3
Forecasted yield and error per cent of three different models from 1974 to 2011 for Ludhiana

Year	Actual yield yield (kg /ha)	Model 1		Model 2		Model 3	
		Forecasted yield(kg /ha)	% Error	Forecasted yield (kg /ha)	% Error	Forecasted yield (kg /ha)	% Error
1973-74	750	830	10.7	770	2.7	803	7.1
1974-75	780	794	1.7	765	-2.0	785	0.7
1975-76	632	695	9.9	663	5.0	723	14.5
1976-77	705	626	-11.1	645	-8.5	723	2.6
1977-78	412	446	8.3	477	15.8	466	13.0
1978-79	681	724	6.4	698	2.6	645	-5.4
1979-80	729	707	-3.0	685	-6.0	688	-5.6
1980-81	520	394	-24.3	405	-22.1	590	13.4
1981-82	670	757	13.0	704	5.0	671	0.1
1982-83	758	898	18.5	839	10.7	697	-8.1
1983-84	862	849	-1.5	862	0.0	914	6.0
1984-85	1066	1135	6.5	1057	-0.9	1022	-4.1
1985-86	1017	981	-3.5	965	-5.1	919	-9.6
1986-87	739	746	1.0	743	0.6	936	26.6
1987-88	1087	1205	10.9	1089	0.2	1031	-5.2
1988-89	1087	1126	3.6	1138	4.7	966	-11.2
1989-90	1030	1076	4.5	1108	7.6	1033	0.3
1090-91	1275	1123	-11.9	1128	-11.6	1164	-8.7
1991-92	1200	1055	-12.1	1068	-11.0	1116	-7.0
1992-93	1155	1192	3.2	1173	1.5	1079	-6.6
1993-94	1086	1074	-1.1	1081	-0.5	1086	0.0
1994-95	1265	988	-21.9	1073	-15.2	1125	-11.1
1995-96	1068	1141	6.8	1161	8.7	1055	-1.2
1996-97	882	903	2.4	1020	15.6	1052	19.3
1997-98	1053	1165	10.6	1181	12.1	1058	0.5
1998-99	965	1025	6.2	1062	10.1	998	3.4
1999-00	1198	1061	-11.4	1160	-3.2	1186	-1.0
2000-01	1403	1333	-5.0	1306	-6.9	1195	-14.8
2001-02	1364	1332	-2.4	1280	-6.2	1302	-4.5
2002-03	1197	1357	13.4	1316	10.0	1109	-7.3
2003-04	1348	1369	1.5	1335	-1.0	1279	-5.1
2004-05	1577	1237	-21.6	1253	-20.6	1257	-20.3
2005-06	1359	1242	-8.6	1269	-6.6	1276	-6.1
2006-07	1199	1254	4.6	1309	9.2	1381	15.2
2007-08	1106	1131	2.2	1191	7.7	1379	24.7
2008-09	1306	1315	0.7	1340	2.6	1471	12.6
2009-10	1235	1366	10.6	1353	9.6	1371	11.0
2010-11	1395	1392	-0.2	1410	1.1	1409	1.0

The modified model showed increase in the value of R^2 (59 %) as compared to model 1. The error per cent for modified model ranged from -22.6 to 19 .8% for the last 38 years (Table 4).

The regression equation developed is as follows:

Yield=1093.22-16.7405*Tmax(5-6)-11.3736*Tmax(9-12)+12.681*Tmin(6)-4.62477*Tmin(11-12)-16.6668*Rf(48)-7.83039*Rf(9-10)-65.369*Rd(48)+13.4364*Rd(9-10)+34.6695*TT

R^2 = 82 %

The modified model showed increase in the value of R^2 to 82 % as compared to 49 % in case of model 1. The error per cent for modified model ranged between -21.4 to 24.1 for the last 35 years.

3.3. Model 3: Regression using SPSS software

The multi-regression analysis using SPSS has been employed for the estimation of mustard yield in different districts of central Punjab. The regression expression for Ludhiana is as follows:

Regression equation = 814.0522+14.99609Time+20.96947Z471

R^2= 81 %

Table 4

Forecasted yield and error per cent of three different models from year 1974 to 2011for Amritsar

Year	Actual yield yield (kg /ha)	Model 1		Model 2		Model 3	
		Forecasted yield(kg /ha)	% Error	Forecasted yield (kg /ha)	% Error	Forecasted yield (kg /ha)	% Error
1973-74	1033	1049	1.6	993	-3.9	1108	7.3
1974-75	770	815	5.8	782	1.6	943	22.5
1975-76	563	625	11.0	680	20.8	694	23.3
1976-77	494	595	20.5	568	14.9	543	9.9
1977-78	449	495	10.2	486	8.2	521	16.1
1978-79	753	734	-2.5	715	-5.1	703	-6.7
1979-80	712	850	19.4	824	15.7	729	2.4
1980-81	811	769	-5.2	755	-6.9	804	-0.9
1981-82	993	777	-21.7	781	-21.4	894	-10.0
1982-83	954	929	-2.6	904	-5.2	881	-7.7
1983-84	1074	1116	3.9	1081	0.7	1057	-1.6
1984-85	1173	1050	-10.5	1027	-12.4	1022	-12.9
1985-86	1208	938	-22.4	935	-22.6	1123	-7.0
1986-87	1154	1173	1.6	1140	-1.2	1156	0.2
1987-88	1068	1036	-3.0	1012	-5.2	887	-16.9
1988-89	1068	1039	-2.7	1025	-4.1	1101	3.1
1989-90	952	975	2.4	988	3.8	928	-2.5
1090-91	863	828	-4.0	842	-2.5	1035	19.9
1991-92	1056	862	-18.3	890	-15.7	927	-12.2
1992-93	873	935	7.1	960	10.0	982	12.5
1993-94	1075	992	-7.7	1011	-5.9	863	-19.7
1994-95	942	805	-14.5	829	-12.0	888	-5.7
1995-96	970	907	-6.5	934	-3.7	1065	9.8
1996-97	954	1116	17.0	1132	18.7	1154	20.9
1997-98	762	880	15.4	913	19.8	873	14.6
1998-99	966	999	3.4	1006	4.1	818	-15.3
1999-00	1127	1138	1.0	1134	0.6	1044	-7.4
2000-01	1012	1120	10.6	1117	10.4	939	-7.2
2001-02	1026	963	-6.1	1002	-2.3	921	-10.2
2002-03	942	958	1.7	952	1.0	823	-12.6
2003-04	1037	1170	12.8	1196	15.3	1222	17.9
2004-05	1139	1142	0.3	1144	0.5	1177	3.3
2005-06	1051	1061	0.9	1099	4.5	928	-11.7
2006-07	884	929	5.1	961	8.7	912	3.2
2007-08	1091	854	-21.7	902	-17.4	1132	3.7
2008-09	1030	1118	8.6	1152	11.9	1080	4.8
2009-10	1554	1358	-12.6	1337	-14.0	1334	-14.2
2010-11	1054	1143	8.4	1136	7.8	1084	2.8

Here,

Z471 is the sum product of Rainy days * Sunshine hours

The regression equation showed that time and combination of rainy days and sunshine hours, plays an important role on mustard yield in Ludhiana area. The per cent error ranged between -20.3 % in year 2005 to 26.6% in year 1987. The forecasted yield and per cent error based on above regression equation is given in Table 3. The value of R^2 is 81 % indicating that weather variables are able to explain 81 % of variation in the mustard yield in Ludhiana region.

The regression equation for Amritsar district is as follows:

Regression equation =2943.866-3.59036Z10+31.34268Z21+8.07826Z31

R^2= 62%

Table 5

Forecasted yield and error per cent of three different models from year 1974 to 2011for Patiala

Year	Actual yield yield (kg /ha)	Model 1		Model 2		Model 3	
		Forecasted yield(kg /ha)	% Error	Forecasted yield (kg /ha)	% Error	Forecasted yield (kg /ha)	% Error
1973-74	408	352	-13.7	421	3.2	419	2.7
1974-75	709	843	19.0	557	-21.4	617	-13.0
1975-76	632	727	15.0	554	-12.3	524	-17.1
1976-77	345	339	-1.7	421	22.2	346	0.2
1977-78	259	303	17.1	240	-7.3	259	0.0
1978-79	682	552	-19.1	594	-12.8	567	-16.9
1979-80	542	477	-11.9	471	-13.0	495	-8.6
1980-81	374	404	8.0	403	7.7	363	-3.0
1981-82	345	303	-12.1	299	-13.2	315	-8.7
1982-83	524	578	10.4	647	23.5	513	-2.1
1983-84	448	496	10.6	531	18.5	504	12.6
1984-85	634	663	4.6	600	-5.4	740	16.7
1985-86	713	837	17.3	753	5.6	750	5.2
1986-87	907	931	2.6	776	-14.5	780	-14.0
1987-88	713	737	3.4	689	-3.4	746	4.6
1988-89	713	802	12.4	835	17.1	796	11.6
1989-90	734	817	11.2	870	18.5	757	3.2
1990-91	1047	959	-8.4	931	-11.1	821	-21.5
1991-92	1149	868	-24.5	993	-13.6	853	-25.8
1992-93	694	866	24.7	862	24.1	803	15.7
1993-94	962	1075	11.7	952	-1.0	952	-1.0
1994-95	1041	904	-13.2	973	-6.5	945	-9.2
1995-96	890	832	-6.5	928	4.3	972	9.2
1996-97	1274	1099	-13.8	1051	-17.5	984	-22.8
1998-99	842	888	5.4	1060	25.9	980	16.4
1999-00	1070	1048	-2.1	1091	2.0	1051	-1.8
2000-01	1109	808	-27.2	984	-11.2	1013	-8.7
2001-02	876	789	-9.9	992	13.3	1014	15.7
2003-04	1238	1037	-16.3	1035	-16.4	1022	-17.4
2004-05	961	1044	8.6	1114	16.0	1063	10.6
2005-06	1188	1045	-12.0	1043	-12.2	1114	-6.2
2006-07	1003	994	-0.9	1057	5.4	1141	13.7
2007-08	1121	916	-18.3	1097	-2.1	1164	3.9
2008-09	1052	953	-9.4	1140	8.3	1189	13.0
2010-11	1137	1002	-11.9	1096	-3.6	1163	2.3

Here,

Z10 is the sum of maximum temperature

Z21 is the sum product of minimum temperature

Z31 is the sum product of rainfall

The regression equation showed that time, minimum temperature and rainfall plays an important role in determining the mustard yield in Amritsar area. The per cent error ranged between -21.3 % in year 2013 to 23.3 % in year 1976. The forecasted yield and per cent error based on above regression equation is given in Table 4. The value of R^2 is 62% indicates that weather variables are able to explain 62 % of variation in the mustard yield at Amritsar region.

For Patiala, regression equation is as follows:

Regression equation = 503.6099+19.67265Time+0.410622Z231

R^2= 76%

Here,

Z231 is the sum product of minimum temperature* rainfall.

Table 6
Forecasted yield and error per cent of three different models from year 2012 to 2014 for Ludhiana, Amritsar and Patiala

Year	Actual yield	Model 1		Model 2		Model 3	
	yield (kg /ha)	Forecasted yield(kg /ha)	% Error	Forecasted yield (kg /ha)	% Error	Forecasted yield (kg /ha)	% Error
Ludhiana							
2011-12	1584	1265	-20.1	1227	-22.5	1355	-14.4
2012-13	1234	1330	7.8	1402	13.6	1399	13.4
2013-14	1289	1150	-10.8	1198	-7.0	1275	-1.1
Amritsar							
2011-12	1555	1267	-18.5	1281	-17.6	1286	-17.3
2012-13	1395	1265	-9.3	1283	-8.0	1097	-21.3
2013-14	1410	1132	-19.7	1142	-19.0	1158	-17.9
Patiala							
2012-13	1434	1286	-10.3	1159	-19.2	1113	-22.4
2013-14	1431	1125	-21.4	1182	-17.4	1161	-18.9

The regression equation showed that time and combination of minimum temperature and rainfall plays an important role on mustard yield in Patiala area. The per cent error ranged between -25.8 % in the year 1992 to 16.4 % in the year 1999. The forecasted yield and per cent error based on above regression equation is given in Table 5. The value of R^2 is 76 % indicating that weather variables are able to explain 76 % of variation in the mustard yield in Patiala region.

3.4. Validation of the models

The forecasted mustard yields for the year 2012, 2013 and 2014 were taken for validation of the models. For Ludhiana district, the forecast yield of three years was compared with actual yield obtained that year to calculate the error percentage. The results for Ludhiana district showed (Fig 1) that in the year 2012 and 2014, all the three models showed less forecasted yield than the actual yield, but in year 2013, model 1, 2 and 3 predicted higher forecasted yield than the actual yield. The comparisons between the actual and forecasted yield for the year 2012, 2013 and 2014 by the three different models for mustard crop is summarized as Fig 1. The error per cent for model 1 in year 2012,

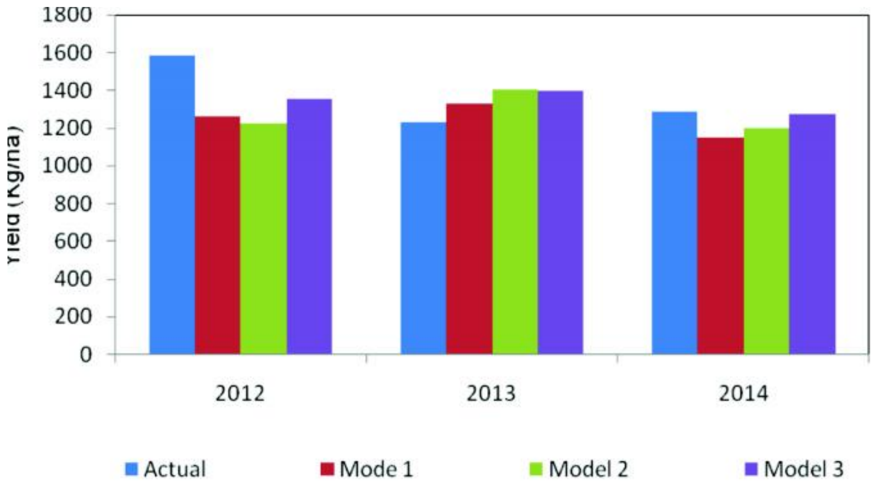

Fig. 1: Comparison between the actual and forecasted yield (by three different models) of mustard for the year 2012, 2013 and 2014 for Ludhiana district.

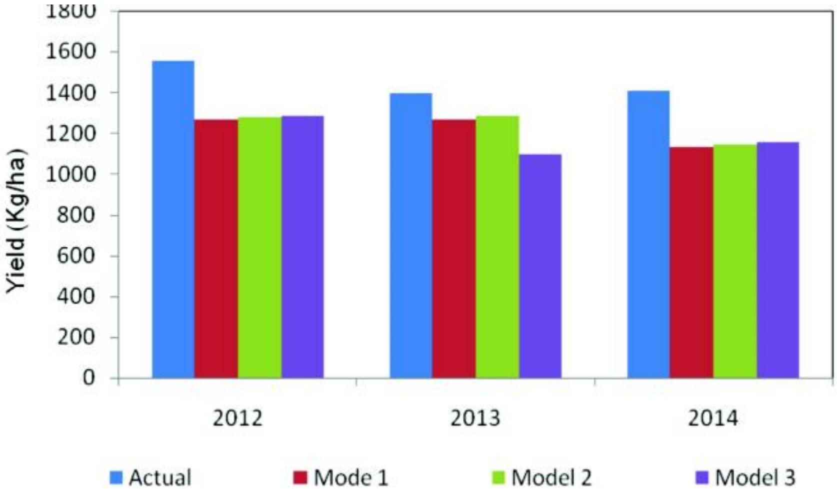

Fig. 2: Comparison between the actual and forecasted yield (by three different models) of mustard for year 2012, 2013 and 2014 for Amritsar.

2013 and 2014 were -20.1, 7.8 and -10.8 %, respectively. The error per cent of model 2 is more than the basic model for 2012 and 2013 and were -22.5, 13.6 and -7.0 % for the years 2012, 2013 and 2014, respectively. The error per cent for model 3 were -14.4 % (2012), 13.4 % (2013) and -1.1% (2014).

The forecasted yield for Amritsar district for three different years (2012, 2013 and 2014) was compared with actual yield obtained that year to calculate the error percentage. The results showed (Fig 2) that in the year 2012, 2013 and 2014, all the three models showed lesser yield than the actual yield. The comparisons between the actual and forecasted yield for the year 2012, 2013 and 2014 by the three different models for mustard crop is summarized as Fig 2. The error per cent for model 1 in year 2012 was -18.5 %, in 2013 it was -9.3 % and -19.7 % in the year 2014. The error per cent of model 2 was less than the basic model for all the three years and were -17.6, -8.0 and -19.0 % for the year 2012, 2013 and 2014, respectively. The error per cent for model 3 were -17.3, -21.3 and -17.9 % for the year 2012, 2013 and 2014, respectively.

For Patiala district, the forecasted yield only for the year 2013 and 2014 was compared with actual yield as the yield data for 2012 is not available. The results showed (Fig 3) that for both years, all three models (model 1, 2 and 3) forecasted lesser yield than the actual yield. The comparisons between the actual and forecasted yield for the year 2013 and

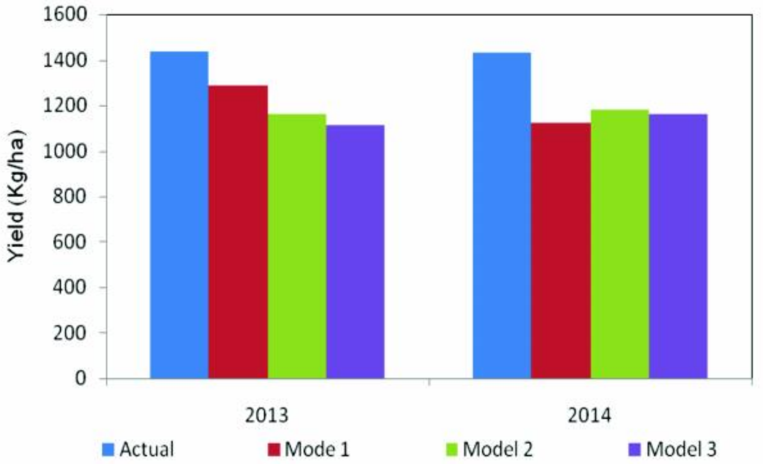

Fig. 3: Comparison between the actual and forecasted yield (by three different models) of mustard for the year 2013 and 2014 for Patiala district.

2014 by the three different models for mustard crop is summarized as Fig 3. Comparison of three statistical models (model 1, 2 and 3) for the prediction of mustard crop for Patiala showed that for the year 2013, model 1 showed less per cent error (-10.3%) than model 2 (-19.2%) and model 3 (-22.4). for the year 2014, per cent error calculated for model 1 was -21.4 %, model 2 was -17.4 % and -18.9 % for model 3.

4. Conclusion

Using the forecast models, pre-harvest estimates of mustard yield for Ludhiana, Amritsar and Patiala districts could be computed successfully very much in advance before the actual harvest. These three models predict the mustard yield very well and the error per cent of these models are below 27 %. As the data used for developing these model is of high degree of accuracy, its reliability is also high. The district government authorities also can make use of the forecast model developed using weather indices, for obtaining accurate pre-harvest estimates of mustard crop. The regression models are also useful to predict the yield of other crops all over the country but these are site specific. Till the final production of crops becomes known, decisions have to be made on the basis of informed predictions or scientific forecasts. The main beneficiaries are farmers, traders, exporters and importers.

References

1. Hiremath D. B. and Shiyani R.L., 2012: 'Adapting Gujarat to Climatic Vulnerabilities: The Road Ahead', Res.J.Recent Sci., Vol. 1, No. 5, pp. 38-45.

2. F.A.O., The State of Food insecurity in the World, FAO, Rome, Italy, 2006.

3. Yadava J.S., Singh A.K. and Chauhan J.S., 2001: (National Research Centre on Rapeseed Mustard, Bharatpur (India)). Stategic issues for doubling the productivity of rapeseed mustard in next decade in India. Annals of Biology (India). Vol. 17, No. 1, pp. 1 20.

4. Kumar A., 2005: 'Rapeseed-mustard in India: current status and future prospects. In: Winter School on Advances in Rapeseed-Mustard Research Technology for Sustainable Production of Oilseeds, National Centre on Rapeseed-Mustard, Sewar, Bharatpur', December 15 to January 04, 2005, pp. 278–288.

5. Tripathi P., Tomar S. K. and Singh A. K., 1999: 'Crop weather model to predict the growth and yield of mustard and wheat under mustard-wheat cropping system'. P. 285. In: Proc. Natl. Workshop on Dynamic Crop Simulation Modelling for Agromet Advisory Services, January 4-6, 1999, NCMRWF, New Delhi.

6. Boote K. J., Jones J. W., and Pickering N. B., 1996: 'Potential uses and limitations of crop models', J. Agron. Vol. 88, pp. 704-716.

7. Box George E.P., William G. H. and Hunter J. S., Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building, John Wiley and Sons, New York, 1978.