International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; SP-8(6): 110-115

https://www.mathsjournal.com
Received: 12-09-2023
Accepted: 14-10-2023

HV Gundaniya
School of Agriculture,
ITM University, Gwalior,
Madhya Pradesh, India

YA Garde
Department of Agricultural Statistics, NM College of Agriculture, Navsari Agricultural University, Navsari, Gujarat, India

PR Vekariya
Department of Agricultural Statistics, BA College of Agriculture, Anand Agricultural University, Anand, Gujarat, India

Corresponding Author:
HV Gundaniya
School of Agriculture,
ITM University, Gwalior,
Madhya Pradesh, India

Weather intervention based sugarcane yield forecasting models for enhancing farmers' income

HV Gundaniya, YA Garde and PR Vekariya

Abstract
Agriculture is one of the most important sector of an Indian Economy. Sugarcane is one of the most important commercial crops grown in India. Pre-harvest forecasts have significant value in agricultural planning and play important role in doubling the farmers' income. Therefore proper forecast of such important commercial crop is necessary for future planning, policy making and sustainable production. In the present investigation, relationship between yearly cane yield and weekly weather parameters were studied by using Karl-Pearson's correlation coefficient approach. The study revealed that all the weather variables were significantly correlated with cane yield in different weeks of cropping season. The sugarcane yield forecasting models were developed using 26 years of cane yield and corresponding weather data (1991-92 to 2016-17). The statistical tools of multiple linear regression (MLR) and discriminant function analysis were used for model development. The study showed that Model-1h, developed through MLR technique have high R^2 value (92.6%) and low value of RMSE (6.42) as compared to remaining models. Therefore, study concluded that multiple linear regression (MLR) was more reliable as compared to discriminant function analysis approach and which provide yield forecast well in advance of actual harvesting of the crop.

Keywords: MLR, discriminant function analysis, weather parameters, forecast models

Introduction

Pre-harvest forecasts have significant value in agricultural planning and play important role in doubling the farmers' income. Sugarcane is one of the most important commercial crops grown in India. In terms of production, India is the second largest producer of Sugarcane in World after Brazil. Sugarcane is main source of sugar in India and main source of raw material for white sugar, jiggery and Khandsari. Sugar industry is second largest agro based industry in India after textile industry. The Sugarcane cultivation and Sugar industry in India plays a vital role in socio-economic development of rural areas by mobilizing rural resources and generating higher income and employment opportunities.

The forecasting of crop yield may be done by using three major objective ways (i) biometrical characteristics (ii) weather variables (iii) agricultural inputs (Agrawal *et al.*, 2001) [3]. These approaches can be used on an individual basis or together to form a composite model. The most frequently used approach relies on crop weather relationship studies within which historical data of crop yield and weather parameters are utilized.

Similar work have been done by many scientists *viz.* Agrawal *et al.* (2012) [1] used discriminant function analysis for developing wheat yield forecast model for Kanpur. Sattar *et al.* studied relationship between Sugarcane yield and weather parameters using MLR at Samastipur. Garde *et al.* (2015) [7] developed pre harvest forecast models for wheat yield using MLR and discriminant function analysis in Varanasi. Kumar *et al.* (2016) [10] studied crop yield forecasting of paddy and sugarcane through modified Hendrick and Scholl technique for south Gujarat using weather variables. Biswas and Bhattacharya (2019) [5] used MLR and discriminant function analysis for developing rice yield forecast models for lower Gangetic Plain of India. Banakara *et al.* (2018) [4] studied forecasting of rice yield using discriminant function analysis and MLR in Surat district of Gujarat.

Timely and accurate yield forecast is essential for crop production, marketing, storage and transportation decisions as well as for managing the risk associated with these activities.

Crop yield forecasting is important for national food security. It is also needful in taking policy decisions related to food procurement and distribution, import-export policies, price policies with greater precision, (Garde *et al.*, 2012) [9]. Accurate early warning of crop failures reduce the undesirable effect of price rise. A sugarcane crop is sensitive to the climate, soil type, irrigation, fertilizers, insects, diseases, varieties, and the harvest period. Sugarcane crop should be harvested at right stage of maturity because both early and delayed harvesting results in loss of quality and quantity of product. As sugarcane is one of the most important commercial crops, therefore proper forecast of such important commercial crop is necessary for future planning and policy making.

Materials and Methods

The study utilized secondary yearly yield data of cane yield for 26 years (1991-92 to 2016-17) were collected from the annual reports of Main Sugarcane Research Station, NAU, Navsari and corresponding year weekly weather data were collected from Department of Agril. Engineering, NMCA, NAU, Navsari. Five different weather parameters were included in the study; namely Maximum Temperature (X_1), Minimum Temperature (X_2), Total Rainfall (X_3), Morning Relative Humidity (X_4) and Evening Relative Humidity (X_5). The statistical models were developed by using data from the year 1991-92 to 2012-13 and validation of the developed models were done by using remaining data (2013-14 to 2016-17).

Association between cane yield and weather parameters

The relationship between yearly cane yields with weather parameters were studied by using approach of Karl-Pearson's coefficient of correlation.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Where,

$I = 1, 2, ...N$

R is the correlation coefficient between cane yield and average of weekly weather parameters

Y_i is observed value of cane yield in i^{th} year

X_i is observed value of weekly weather parameter in i^{th} week

$\bar{X}$ and $\bar{Y}$ are the mean of X and Y variables respectively

Pre harvest forecast models based on weather indices:

In this method weekly data on weather variables of 45 weeks have been used for construction of weather indices (Agrawal *et al.*, 2007; Garde *et al.*, 2015) [7]. Weather indices were computed as unweighted indices and weighted indices. In weighted indices weight being correlation coefficient between yearly crop yield (detrended) and weather parameters for respective weeks. The forms of weather indices are given as below:

$$Z_{i,j} = \sum_{w=1}^{m} r_{ij}^{j} X_{iw} \quad \text{and} \quad Z_{i,i',j} = \sum_{w=1}^{m} r_{ii'j}^{j} X_{iw} X_{i'w} \quad , i, i' = 1, 2 \ldots p$$

Where,

P is the number of weather variables under study

$J = 0, 1$ (where, '0' represents un-weighed indices and '1' represents weighed indices).

M is the week up to forecast ($m<n$).

w is the week number (1, 2,….., m).

R_{iw} is correlation coefficient between yield (detrended) and i^{th} weather variable in w^{th} week.

$R_{ii'w}$ is correlation coefficient between yield (detrended) and the product of i^{th} and i'^{th} weather variable in w^{th} week.

Multiple Linear Regression

The forecast models were obtained by applying stepwise multiple linear regression approach by taking predictors as un-weighted and weighted weather indices for forecasting sugarcane yield well in advanced. The form of statistical forecast model using developed weather indices is given below: (Draper and Smith, 1981).

Model 1

$$Y = A_0 + \sum_{i=1}^{p} \sum_{j=0}^{1} a_{ij} Z_{ij} + \sum_{i \neq i'=1}^{p} \sum_{j=0}^{1} a_{ii'j} Z_{ii'j} + cT + e$$

Where,

Y is observed sugarcane yield.

A_0 is the intercept.

Z_{ij} and $Z_{ii'j}$ are the weather indices.

A_{ij} and $a_{ii'j}$ are the regression coefficients of Z_{ij} and $Z_{ii'j}$ weather indices.

P is the number of weather variables.

E is the error term.

C is the regression coefficients of trend variable.

T is the trend variable (year).

Discriminant function Analysis

The discriminant function is a statistical technique which discriminates between various groups of objects on the basis of characters which are considered to be relevant. The procedure start with crop years have been divided into three groups namely congenial, normal and adverse on the basis of crop yield adjusted for trend effect. In the present study four different methods were developed for developing discriminant score and further models were obtained as mentioned below.

Method-1: The method utilized fifteen weather indices (weighted) as described in previous section Two discriminant score have been generated from these weighted weather indices and categorical yield data. The forecast model was developed using two discriminant score along with trend variable through stepwise regression analysis (Agrawal *et al.*, 2012 [1]; Garde *et al.*, 2015) [7]. The form of the developed model is as follow.

Model 2

$Y = \beta_0 + \beta_1 ds_1 + \beta_2 ds_2 + \beta_3 T + \varepsilon$

Where

Y is sugarcane yield (t/ha).

β_i is regression coefficient, i =0,1,2,3

Ds_1 and ds_2 are discriminant scores,

T is the trend variable

ϵ is error term assumed to follow N(0, σ^2).

Method-2: In this method, total cropping season of sugarcane crop from 1st week to 45th week has been divided into four phases by considering development stages of Sugarcane crop. The each phase consist different number of weeks. The simple average of weather variables data in different weeks within the phase were obtained. For each phase average weather variables were used to extract discriminant score using discriminant function analysis. Total eight discriminant scores were obtained for all phases during the year. The forecasting model was fitted by taking sugarcane yield as the regress and and the discriminant scores and trend T as regress or (Agrawal *et al.*, 2012) [1]. The form of the developed model is as follow.

Model 3: $Y = \beta_0 + \sum_{l=1}^{2}\sum_{m=1}^{4}\beta_{lm} ds_{lm} + \beta_3 T + \varepsilon$

Where,

Y is the sugarcane crop yield.

β_0 is intercept of the model.

β_{lm}'s (l=1,2, m= 1, 2, 3,4) and β_3 are the regression coefficients.

Ds_{lm} is the l^{th} discriminant score in m^{th} phase.

T is the time trend variable (year).

ϵ is error term assumed to follow $N(0, \sigma^2)$.

Method-3: In this method, same procedure was followed as mentioned in Method-2, using weighted weather indices instead of average weather variable. The form of the developed model is as follow:

Model 4: $Y = \beta_0 + \sum_{l=1}^{2}\sum_{m=1}^{4}\beta_{lm} ds_{lm} + \beta_3 T + \varepsilon$

Where symbols are same as defined in Method-2.

Method-4: In this method, discriminant scores were computed using five weekly weather parameters data for the first week (37th SMW). In second week (38th SMW), five weather parameters along with the scores developed at first week (37th SMW) were utilized for development of score through discriminant function analysis. The two discriminant scores were obtained to the second week (38th SMW). The process was repeated for the successive week data till the time of forecast *i.e.* 9th week (45th SMW). At last 45th SMW two discriminant scores based on the data were obtained. These two developed scores and trend were utilized in developing forecast model using stepwise regression technique (Agrawal *et al.*, 2012 [1], Garde *et al.*, 2015) [7]. The form of the model is.

Model 5: $Y = \beta_0 + \beta_1 ds_1 + \beta_2 ds_2 + \beta_3 T + \varepsilon$

Where symbols are same as defined in Method-1.

Comparison and validation of models

The comparison and validation of the models were done by using following approaches.

1. Coefficient of multiple determination (R^2_{adj}): The developed statistical models were compared on the basis of adjusted coefficient of multiple determination (R^2_{adj}) as follows.

$$R^2_{adj} = 1 - \frac{SS_{res}/(n-p)}{SS_t/(n-1)}$$

Where,

$ss_{res}/(n\text{-}p)$ is the residual mean square.

$ss_t/(n\text{-}1)$ is the total mean sum of square.

2. Forecast Error (%): For the comparison of developed forecasting models, the Forecast Error percentage was calculated using following formula.

$$\text{Forecast Error (\%)} = \frac{O_i - E_i}{O_i} \times 100$$

Where, O_i and E_i are the observed and forecast value of crop yield respectively.

3. Root Mean Square Error (RMSE): In addition to Forecast errors (%), Root Mean Square Error (RMSE) was calculated as a measures of comparing developed models. The formula of RMSE is given below.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Oi - Ei)^2}$$

Where, O_i and E_i are the observed and forecast value of crop yield respectively and n is the number of years.

Results and Discussion

Association between cane yield and weather parameters

The significance of correlation coefficients were observed between yearly sugarcane yield and weekly average weather parameters *viz.* maximum temperature (42nd, 43rd and 49th SMW), minimum temperature (16th SMW), total rainfall (30th and 42nd SMW), morning relative humidity (2nd, 13th, 30th, 40th, 43rd and 50th SMW) and evening relative humidity (42nd and 43rd SMW) (Table 1). However, during cropping season of sugarcane the other week wise correlation coefficients between the yield and weekly average weather parameters turned out to be non-significant. It was observed that the morning relative humidity found significant at all growth stages which means morning relative humidity is indirectly increase crop yield. Similarly maximum temperature and rainfall also play important role in sugarcane yield.

Table 1: Week wise correlation coefficient between yield of sugarcane and weekly average weather parameters

SMW No.	Max Temp (℃)	Min Temp (℃)	Rainfall (mm)	RH-I (%)	RH-II (%)
W_i	X_1	X_2	X_3	X_4	X_5
1	-0.370	-0.020	-	0.034	0.093
2	0.125	0.208	-0.071	-0.527*	-0.213
3	0.149	0.387	-0.112	-0.083	-0.026
4	-0.102	0.195	-	-0.234	0.148
5	0.249	0.355	0.218	-0.136	0.074
6	-0.056	0.231	0.014	0.048	0.266
7	0.014	0.064	-0.373	0.319	-0.203
8	0.391	0.221	0.142	0.320	-0.173
9	0.233	0.196	-	-0.004	-0.042
10	0.008	0.251	-0.375	-0.293	-0.339

11	0.276	0.186	0.014	-0.364	-0.199
12	0.180	0.369	-0.375	0.214	-0.020
13	-0.165	0.362	-	0.552**	0.411
14	0.061	0.285	-0.138	0.043	0.054
15	0.248	0.164	0.173	-0.061	0.078
16	0.188	0.478*	0.014	-0.025	0.128
17	0.258	0.134	-0.095	0.170	-0.170
18	0.241	0.180	-0.138	0.173	-0.035
19	0.030	0.239	0.373	0.044	0.136
20	0.016	0.239	-0.274	0.347	0.045
21	0.253	0.303	0.022	0.291	0.264
22	0.288	0.037	0.062	0.095	0.091
23	-0.095	0.151	0.017	0.221	0.177
24	0.113	0.191	-0.101	0.089	-0.024
25	0.395	0.412	-0.398	0.011	-0.292
26	-0.050	-0.030	0.182	0.197	0.016
27	-0.137	-0.120	0.111	0.244	0.113
28	-0.139	0.018	-0.243	-0.007	-0.092
29	-0.010	0.157	-0.005	0.2069	-0.094
30	-0.417	-0.162	0.430*	0.570**	0.275
31	-0.411	-0.102	0.359	0.281	0.246
32	-0.356	-0.018	0.306	0.352	0.413
33	-0.276	0.227	-0.218	-0.018	-0.076
34	-0.090	0.003	-0.071	-0.043	-0.163
35	-0.201	0.006	0.372	-0.014	0.112
36	0.069	-0.023	0.074	0.391	0.031
37	0.214	0.343	0.014	-0.036	-0.112
38	0.129	0.376	0.084	0.276	0.179
39	-0.115	0.096	-0.071	0.252	0.104
40	-0.080	0.081	-0.341	-0.505*	-0.171
41	-0.035	-0.073	-0.321	0.027	-0.319
42	0.615**	-0.148	-0.452*	-0.311	-0.457*
43	0.456*	-0.251	-0.248	-0.493*	-0.559**
44	0.262	0.176	-0.238	-0.203	-0.039
45	-0.111	0.043	0.152	0.019	-0.060
46	-0.118	0.246	-0.062	0.387	0.070
47	0.179	0.055	0.243	-0.151	-0.051
48	0.132	0.410	0.012	0.107	0.255
49	0.473*	0.329	-0.138	0.361	0.259
50	0.391	0.245	-	0.531**	0.390
51	0.346	0.223	0.111	0.086	0.268
52	-0.145	-0.076	-	0.312	0.023

*Critical value at 5% level of significance is 0.423
** Critical value at 1% level of significance is 0.537

Pre harvest forecast models

Sugarcane yield forecasting models were developed by using stepwise MLR and discriminant function analysis. The forecast model equations were developed by using above discussed methods are presented in Table 2 along with their Adj. R^2. It is found that in all the models, trend variable T was significant. In Model-1 weather indices Z_{11} (Maximum Temperature), Z_{21} (Minimum Temperature) and Z_{41} (Morning Relative Humidity) were found significant. In Model-2 and Model-5 ds_1 found significant.in Model-3 ds_{13} and ds_{14} were found significant. In Model-4 ds_{11}, ds_{21}, ds_{12}, ds_{22}, ds_{13}, ds_{23}, ds_{14}, ds_{24} were found significant. The value of Adj. R^2 varied from 89.1 to 92.6 for different models. The highest value of Adj. R^2 was found in Model-1 (92.6) which considered that Model-1 is better than other models.

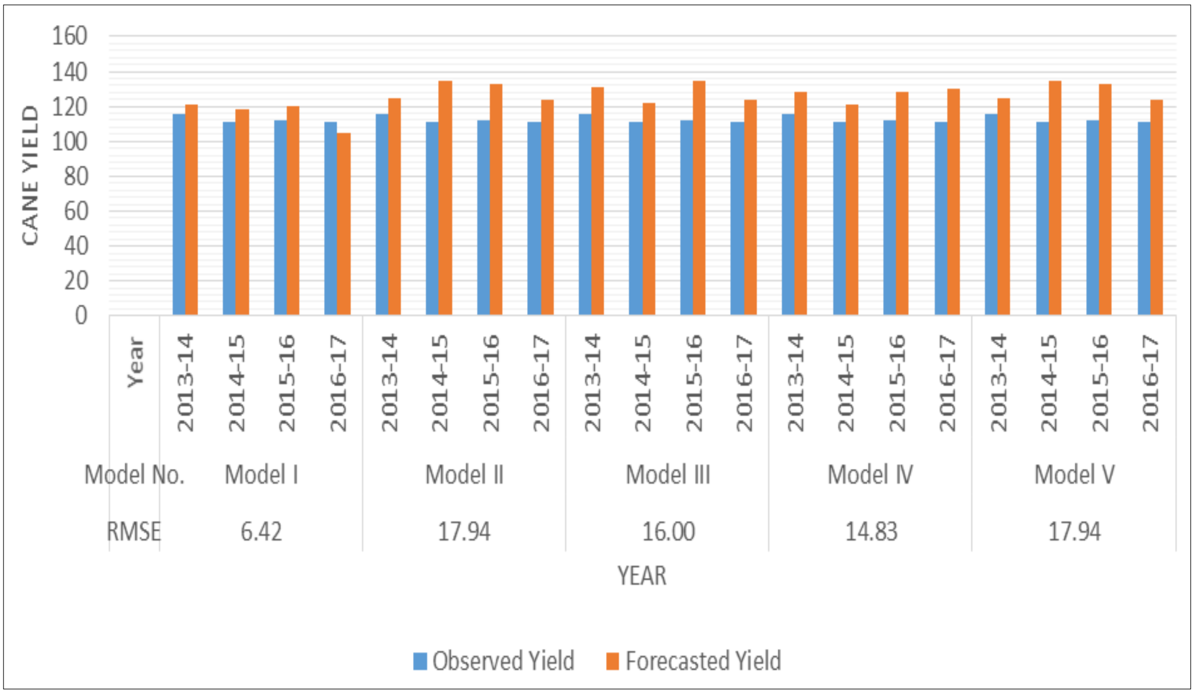
The developed models comparison was done by using RMSE and forecast error and results are presented in Table 3. The RMSE and percent deviation were calculated on the basis of observed and forecast yield (t/ha) for the year 2013 to 2016. The result revealed that the RMSE varied from 6.42 to 17.94. The result based on both the tables, Table 2 and Table 3 revealed that Model 1 have high value of Adj. R^2 and lowest value of RMSE. So it was concluded that Model 1 which is developed through MLR is the most suitable model for forecasting sugarcane yield. The model provide reliable forecast of sugarcane yield 5 weeks before harvest. The similar model results were observed by Kumar *et al.* (2014) [11] in Navsari and Bharuch districts of Gujarat and Priya and Suresh (2009) [12] in Coimbatore district.

Table 2: Sugarcane yield forecast models

Model No.	Forecast model equation	Adj. R^2
1	$Y = -47.320 + 1.394\ Z_{11} + 1.721\ Z_{21} + 0.459\ Z_{41} + 2.056\ T$	92.6
2	$Y = 69.911 + 4.930\ ds_1 + 2.343\ T$	91.7
3	$Y = 70.334 - 5.804\ ds_{11} - 2.139\ ds_{21} - 4.343\ ds_{12} + 0.524\ ds_{22} - 3.821\ ds_{13} + 6.789\ ds_{23} + 4.625\ ds_{14} + 6.210\ ds_{24} + 2.306\ T$	89.1
4	$Y = 66.876 + 5.234 ds_{13} + 3.343 ds_{14} + 2.606\ T$	90.9
5	$Y = 69.911 + 4.930\ ds_1 + 2.343\ T$	91.7

Table 3: Observed and forecast of sugarcane yield

Forecast year	Observed yield (t/ha)	Forecast yield (t/ha)				
		Model I	Model II	Model III	Model IV	Model V
2013-14	116	121	125	131	128	125
		(-3.78)	(-7.37)	(-12.25)	(-9.70)	(-7.37)
2014-15	111	118	135	122	121	135
		(-5.51)	(-21.25)	(-9.60)	(-8.19)	(-21.25)
2015-16	112	120	133	135	128	133
		(-7.89)	(-19.59)	(-20.76)	(-15.16)	(-19.59)
2016-17	111	105	124	124	130	124
		(4.97)	(-12.01)	(-11.80)	(-17.68)	(-12.01)
RMSE		6.42	17.94	16.00	14.83	17.94



*Figures in parenthesis are forecast error (%)

Fig 1: Graphical representation of observed yield and forecasted yield using MLR and discriminant function analysis

The comparison of observed yield and forecasted yield along with RMSE values are graphically presented in Fig 1.

Conclusion

Weather and climatic information plays a major role before and during the cropping season of the crop and if pre harvest forecast provided can be helpful in inspiring the farmer to organize and use their own resources in order to gather the benefits. Based on the results obtained in this research work it can be concluded that there is a wide scope for using alternative approaches to develop predictors that could be used in forecasting models for reliable and dependable forecast. The technique of discriminant function was found useful in classifying the crop year in to congenial, normal and adverse year with respect to crop yield further weather score developed through discriminant function analysis method. The study concluded that multiple linear regression (MLR) found better as compared to discriminant function analysis approach for pre-harvest sugarcane crop yield forecasting.

References

1. Agrawal RC, Aditya K. Use of discriminant function analysis for forecasting crop yield. Mausam. 2012;63(3):455-458.
2. Agrawal RC, Mehta SC. Weather based forecasting of crop yield, pest and diseases- IASRI models, Journal of Indian Society of Agricultural Statistics. 2007;61(2):255-263.
3. Agrawal R, Jain RC, Mehta SC. Yield forecast based on weather variables and agricultural inputs on agro-climatic zone basis. Indian Journal of Agricultural Science. 2001;71(7):487-490.
4. Banakara KB, Garde YA, Pisal RR, Bhatt BK. Pre-harvest forecasting of rice yield for effective crop planning decision in Surat district of South Gujarat, India. International Journal of Current Microbiology and Applied Sciences. 2018;7(6):3410-3422.
5. Biswas R, Bhattacharya B. Rice yield prediction in lower Gangetic Plain of India through multivariate approach and multiple linear regression analysis. Journal of Agro Meteorology. 2019;21(1):101-103.
6. Draper NR, Smith H. Applied Regression Analysis, Second edition, John Wiley and sons, New York; c1981.
7. Garde YA, Dhekale BS, Singh S. Different approaches on pre harvest forecasting of wheat yield. Journal of Applied and Natural Science. 2015;7(2):839-843.
8. Garde YA, Shukla AK, Singh S. Pre-harvest forecasting of rice yield using weather indices in Pantnagar (Uttarakhand). International Journal of Agricultural and Statistical Science. 2012;8(1):233-241.
9. Garde YA, Singh T, Singh S, Mishra GC. Weather based pre-harvest forecasting of wheat at Ghazipur (U.P.).

International Journal of Agricultural Sciences. 2012;8(2):325-328.
10. Kumar N, Pisal RR, Shukla SP, Pandey KK. Crop yield forecasting of paddy and sugarcane through modified Hendrick and Scholl technique for South Gujarat. Mausam. 2016;67(2):405-410.
11. Kumar N, Pisal RR, Shukla SP, Pandey SP. Crop yield forecasting of paddy, sugarcane and wheat through linear regression technique for south Gujarat. Mausam. 2014;65(3):361-364.
12. Priya KSR, Suresh KK. A study on pre-harvest forecast of sugarcane yield using climatic variables. Statistics and applications 2009-10. 2009;7-8(1&2):1-8.