Journal of Pharmacognosy and Phytochemistry

Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2019; 8(2): 2037-2042
Received: 19-01-2019
Accepted: 23-02-2019

VB Kihla
Department of Agricultural Statistics, N. M. College of Agriculture, NAU, Navsari, Gujarat, India

BK Bhatt
Department of Agricultural Statistics, ASPEE College of Horticulture and Forestry, NAU, Navsari, Gujarat, India

YA Garde
Department of Agricultural Statistics, College of Agriculture, NAU, Waghai, Gujarat, India

Correspondence
VB Kihla
Department of Agricultural Statistics, N. M. College of Agriculture, NAU, Navsari, Gujarat, India

Analyzing students enrollment of gender categories from different colleges of Navsari agricultural university

VB Kihla, BK Bhatt and YA Garde

Abstract
An attempt has been made to study the student performance of gender categories in two different colleges of Navsari Agricultural University *viz*, N. M. College of Agriculture (NMCA) and ASPEE College of Horticulture and Forestry (ACHF) and also combination of both the colleges for the period of 2004 to 2015. The analysis has been carried out on the performance of gender category students with the help of the multiple regression models for categorical performance in different colleges of Navsari Agricultural University. All fitted models were tested through ANOVA for their significance. NMCA showed the presence of multicollinearity in male category while it was absent in other models. The *VIF* value of variable mX_2 of NMCA was greater than 10 due to presence of multicollinearity, so the model was refitted after omitting this variable. The gender wise fitted models of all the colleges showed less variability between the colleges. Out of six fitted models, the model for female category of NMCA which was found best fitted multiple regression model as compared to all other models with R^2 value, 0.9872 and *RMSE* value 1.100. The study revealed that the fitted model of female category was better than the fitted model of male category for the all the colleges.

Keywords: students enrollment, multiple regression, multicollinearity, *VIF*, R^2, *RMSE*

Introduction
The agricultural education in India is facing one of the biggest challenges. It has to identify its role in equipping the human resources for enhancing agricultural productivity and sustainable use of natural resources. Agricultural colleges and universities were initially assigned to disseminating scientific knowledge and skills to the farming community and to train them to use such skills for better output. As a backup for such a mission, agricultural research was encouraged to adapt the scientific knowledge to suit the realities of rural societies. However, these initiatives could not keep pace with the fast changing scientific and technical improvements and gradually failed in their objective to cultivate the most modern skills and attitudes to both agricultural students as well as farmers (Sahu, 2012) [8].

According to DARE/ICAR annual report (2017–18) [1], the identified thrust areas continued to receive financial, technical and monitoring support from Agricultural Education Division, ICAR. The Agricultural Education Division is mandated for maintaining, upgrading quality and relevance of higher agricultural education through partnership with State Agricultural Universities (63 SAUs), Deemed-to-be-Universities (4 DUs) and Central Universities (4 CUs) with Agricultural faculties and Central Agricultural Universities (3 CAUs) under the National Agricultural Research and Education System (NARS) to address the challenges of agricultural growth and upgrading quality of higher agricultural education.

The higher education system in India grew rapidly after independence. According to All India Survey on Higher Education, (AISHE) 2017-18[1], there were 903 Universities and 39050 Colleges listed on AISHE web portal and out of them 882 Universities and 38061Colleges have responded during the survey. In which 500 General, 126 Technical, 70 Agriculture & Allied, 58 Medical, 22 Law, 13 Sanskrit and 10 Language Universities and rest 83 Universities are of other Categories.

During the year 2005-2006, the total enrollment in all courses and levels in regular stream had been 110.28 lakh, including 44.66 lakh girls students constituting40.50 % (Kumar *et. al*., 2008) [5]. Now a day's, the total enrollment in higher education has been estimated to be 36.6 million with 19.2 million boys and 17.4 million girls. Girls constitute 47.6% of the total enrolment. Gross Enrollment Ratio (GER) in Higher education in India is 25.8%, which is calculated for 18-23 years of age group. GER for male population is 26.3% and for females, it is 25.4%.For Scheduled Castes, it is 21.8% and for Scheduled Tribes, it is 15.9% as compared to the national GER of 25.8% (AISHE, 2017-18) [1].

Materials and Methods

The data for present investigation were collected from N. M. College of Agriculture and ASPEE College of Horticulture and Forestry of Navsari Agricultural University through the personal contact to the Principal or Dean of the colleges for the period of 2004 to 2015.

Multiple regression model

Multiple regression technique was used to estimate the performance of student enrollment (Rawlings *et al.*, 1998) [7].

$$\hat{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$$

Where,

$\hat{Y}$= Passed out students from gender category

X_1= No. of first year students enrollment in respective category

X_2= No. of second year students enrollment in respective category

X_3= No. of third year students enrollment in respective category

X_4 = No. of fourth year students enrollment in respective category

Here, Y is the dependent variable and X_i's are independent variables with b_i's as the partial regression coefficients of Y on X_i's. The estimates of coefficients (b_i's and a) were computed using the method of ordinary least squares.

Multiple coefficient of determination (R^2) was calculated as:

$$R^2 = \frac{\text{Regression sum of squares}}{\text{Total sum of squares}}$$

A higher coefficient is an indicator of a better goodness of fit for the observations.

Root mean square error (*RMSE*)

The *RMSE* is defined as:

$$RMSE = [\Sigma(Y-\hat{Y})^2/n-k\ ]^{1/2}$$

Where,

k = No. of explanatory variables

n = Total no. of observations

The model with smaller value of *RMSE* is the better model.

The ANOVA (F- test) was used to test the significance of the model and t-test was used to test significance of partial regression coefficients.

Case statistics and residual analysis

The regression equations for gender categories were developed and compared using R^2 and *RMSE*. Also diagnostic measures were employed to test certain assumptions to test validity of the conclusions to be drawn. These were Variance Inflation Factor (VIF), Studentized Residual (r_i), Leverages, Cook's distance *(D) etc*.

VIF is a measure of how much the variance of an estimated regression coefficient increases if the explanatory variables are correlated. The higher value of VIF is indicating the degree of collinearity. Hann (2002) [4]suggested that if the VIF is *>10* there is strong evidence that collinearity is affecting the regression coefficients and consequently they are poorly estimated.

$$VIF = 1/(1 - R_j^2)$$

Where, R_j^2 is the R^2 of X_i, regressed on all the other *X's* in the model.

Studentized residuals, denoted by r_i, i = 1, 2, 3, n are a standardized version of the ordinary residuals, which are useful for the detection of outliers. According to Rawlings *et, al,*. (1998) [7], larger value of r_i (greater than 2 or 3 in absolute magnitude) indicating possible problem in the model. Leverage values fall between 0 and 1. Marasinghe and Kennedy (2008) [6] consider a leverage value greater than *2p/n*, large and suggest to examine the corresponding observation. Large values for Cook's distance signify unusual observations. According to Cooks and Sanford (1982) [3] if Cook's D value is greater than 1 indicate the outlier, those *>4* are potentially serious outliers.

Results and Discussion

The multiple regression model was developed for enrollment in agricultural education of Navsari Agricultural University, using 12 years of enrolled students data. The multiple regression model was carried out from the year 2008 due to Navsari Agricultural University was bifurcated from Gujarat Agricultural University in the year of 2004, therefore 1st batch of NAU was pass out in the year 2008. The first year enrollments (%) of gender categories in different colleges are presented in following Table 1:

Table 1: Table showed that the enrollments of male students was higher as compared to the female students initially, but after 2007 enrollment of female students is continuously increasing in agricultural education.

Year	N. M. College of Agriculture			ASPEE College of Horticulture and Forestry			Combination of both the Colleges		
	Male	Female	Total %	Male	Female	Total %	Male	Female	Total %
2004	89.66	10.34	100	100.00	0.00	100	95.59	4.41	100
2005	89.09	10.91	100	84.21	15.79	100	87.10	12.90	100
2006	90.59	9.41	100	83.33	16.67	100	87.77	12.23	100
2007	95.00	5.00	100	77.55	22.45	100	88.37	11.63	100
2008	73.81	26.19	100	70.49	29.51	100	72.41	27.59	100
2009	74.03	25.97	100	75.36	24.64	100	74.66	25.34	100
2010	77.65	22.35	100	80.25	19.75	100	78.92	21.08	100
2011	73.00	27.00	100	71.26	28.74	100	72.19	27.81	100
2012	76.42	23.58	100	77.00	23.00	100	76.70	23.30	100
2013	50.41	49.59	100	58.33	41.67	100	54.15	45.85	100
2014	58.49	41.51	100	65.22	34.78	100	61.62	38.38	100
2015	78.57	21.43	100	65.14	34.86	100	71.16	28.84	100

Regression model was fitted for male students on different colleges using all the four, independent variables mX_1 (1st year enrollment male students), mX_2(2nd year enrollment male students), mX_3(3rd year enrollment male students) &mX_4(4th year enrollment male students) and mY (passed out male students) as dependent variable. Similarly, the regression model was fitted for female students on different colleges using all the four, independent variables fX_1(1st year enrollment Female students), fX_2(2nd year enrollment Female students), fX_3(3rd year enrollment Female students) & fX_4(4th year enrollment Female students) and fY (passed out Female students) as dependent variable. Following six models were fitted for different colleges.

Table 2: Fitted multiple regression model for the student enrollment for the gender category on different colleges

Category	Multiple regression model equation	R^2	*RMSE*	*P* value
N. M. College of Agriculture				
Male	$mY = 46.825 + 0.261mX_1 - 0.962mX_3^{**} + 0.686mX_4^{*}$	0.915	6.166	0.018
Female	$fY = -3.676 + 0.174fX_1^{*} - 0.095fX_2 + 0.138fX_3 + 0.641fX_4^{*}$	0.987	1.100	0.003
ASPEE College of Horticulture and Forestry				
Male	$mY = -4.927 + 0.181mX_1 - 0.224mX_2 + 0.407mX_3 + 0.541mX_4$	0.902	5.860	0.072
Female	$fY = -4.747 + 0.228fX_1 - 0.245fX_2 + 0.230fX_3 + 0.798fX_4$	0.952	2.298	0.025
Combination of both the colleges				
Male	$mY = -34.619 + 0.558mX_1 - 0.237mX_2 + 0.224mX_3 + 0.539mX_4$	0.910	9.860	0.063
Female	$fY = -6.362 + 0.143fX_1 - 0.025fX_2 + 0.227fX_3 + 0.432fX_4$	0.916	4.590	0.057

** Significant at 1 % level

* Significant at 5 % level

Table 3: VIF values of variables for different colleges

VIF values of variables	Different colleges		
	NMCA	ACHF	Combined colleges
mx_1	3.297	1.727	1.315
mx_2	12.584	2.734	1.804
mx_3	9.185	3.362	2.954
mx_4	9.304	2.730	2.329
fx_1	1.472	2.178	1.471
fx_2	3.475	4.640	4.164
fx_3	3.641	2.333	2.297
fx_4	7.201	5.730	6.380

The fitted regression equation, R^2, *RMSE*, *P*-value and *VIF* value are presented in Table 2 & 3. From the Table 2 it can be observed that the value of R^2 varies from 0.902 to 0.987 for all colleges. It was observed that male category of NMCA was significant at 5% level of significance with p-value 0.018. It was observed that value of *VIF* for male category of NMCA was 12.584 (mx_2) indicating presence of multicollinearity in the model. After omitting variable from the model, the refitted model showed that value of R^2 0.915 with *RMSE* 6.166.

The table 2 it was revealed that female category of NMCA found significant at 1% level of significance with *p*-value 0.003 and value of R^2 was 0.987 with *RMSE* 1.100. The fitted model was free from problem of multicollinearity, that value of *VIF* was less than 10 for all the variables (Table 3).

It was observed that male category of ACHF found non-significant with *p*-value 0.072 and female category of ACHF found significant at 5% level of significance with *p*-value 0.025 and value of $R^2$was 0.902 with *RMSE* 5.860 and value of R^2 was 0.952 with *RMSE* 2.298 respectively (Table 2). The fitted model was free from problem of multicollinearity, that value of *VIF* was less than 10 for all the variables (Table 3).

From the table 2 it was observed that male category of combined colleges found non-significant with *p*-value 0.063 and value of R^2 was 0.910 with *RMSE* 9.860 and female category of combined colleges found non-significant with *p*-value 0.057 and value of R^2 was 0.916 with *RMSE* 4.590. The fitted model was free from problem of multicollinearity, that value of *VIF* was less than 10 for all the variables (table 3).

The fit diagnostics for the fitted models of gender category for different colleges are presented in Fig. 1 to Fig. 6.

It was observed that Fig. 1(a) of normal quintile plot of the residuals doesn't seem to deviate from a normal distribution in any systematic manner. In Fig. 1(b) showed that despite one large value of standardized residual the fitted model does not seem to be deviated from the normality. This large value may create some problem on fitted model which is indicated in Fig. 1(h), as the fit mean and residual do not go parallel across this range. The Fig. 1(f) showed that all $Di < 1$, *i.e.* there is no problem in development of model. Similarly Fig. 1(e) showed three values are large. Leverage in Fig 1(c) indicated that one observation in the data set is an outlier which is influencing the result.

In Fig. 2(b), Fig. 3(b) and Fig. 5(b) showed two values are outside the ranges from -2 to +2 of residual plot. This large value may create some problem on fitted model which indicated in Fig. 2(h), Fig. 3(h) and Fig. 5(h) respectively as the fit mean and residual do not go parallel across this range. The Fig 2(f) showed that four observations and the Fig. 5(f) showed that one of the observation across the suggested range 0 to 1(Cook and Sanford 1982) [3]. In Fig. 3(f) showed the all *Di* values are below the one. Similarly Fig. 2(e) and Fig. 5(e) as one value and Fig. 3(e) as two values are too large. Leverage in Fig 2(c), Fig. 3(c) and Fig. 5(c) indicated that two observations in the data set is an outlier which is influencing the results.

In Fig. 4(b) and Fig. 6(b) showed three values are out of the range of -2 to +2 of residual plot. This large value may create some problem on fitted model which is indicated in Fig. 4(h) and Fig. 6(h), as the fit mean and residual do not go parallel across this range. The Fig. 4(f) and Fig. 6(f) revealed that three of the observation across the suggested range 0 to 1 (Cook and Sanford, 1982) [3]. Similarly Fig. 4(e) and Fig. 6(e) as one value is too large. Leverage in Fig. 4(c) and Fig. 6(c) indicated that three observations in the data set was an outlier which is influencing the result.

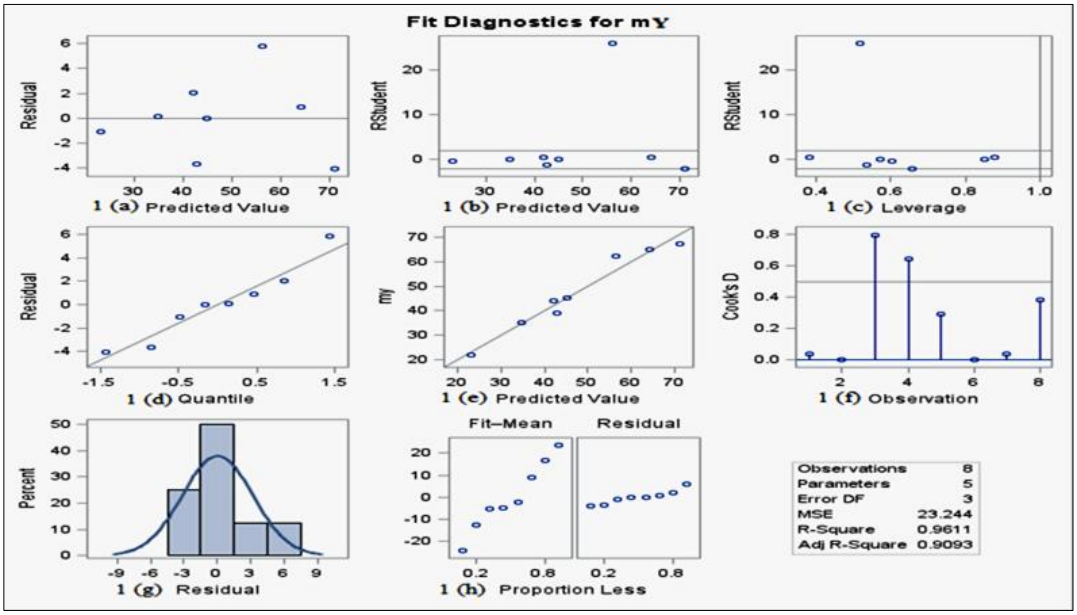

Fig 1: Fit diagnostics for male category of NMCA

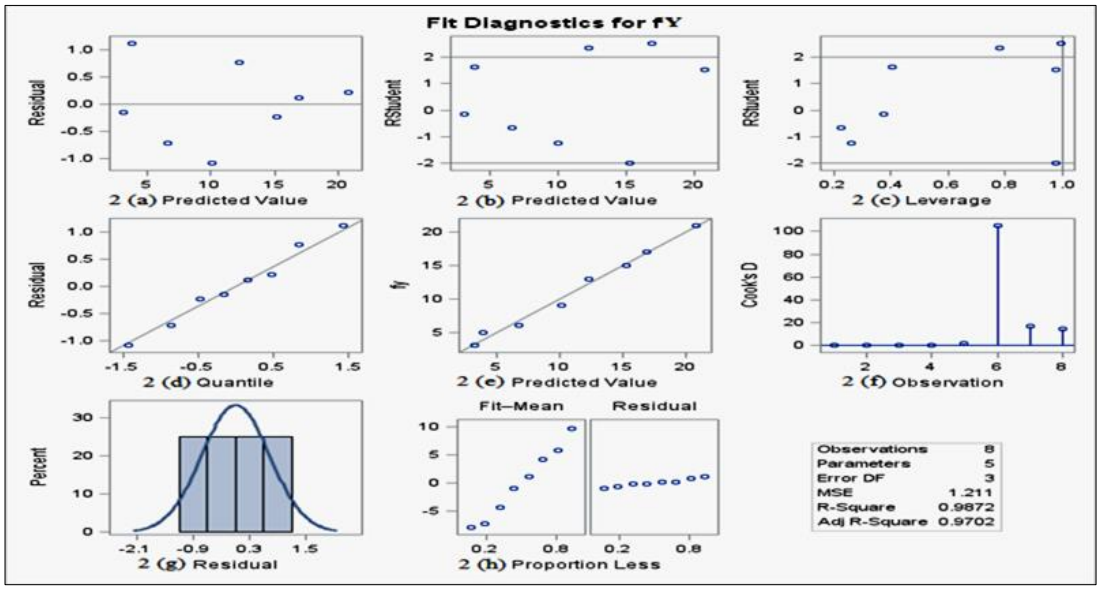

Fig 2: Fit diagnostics for female category of NMCA

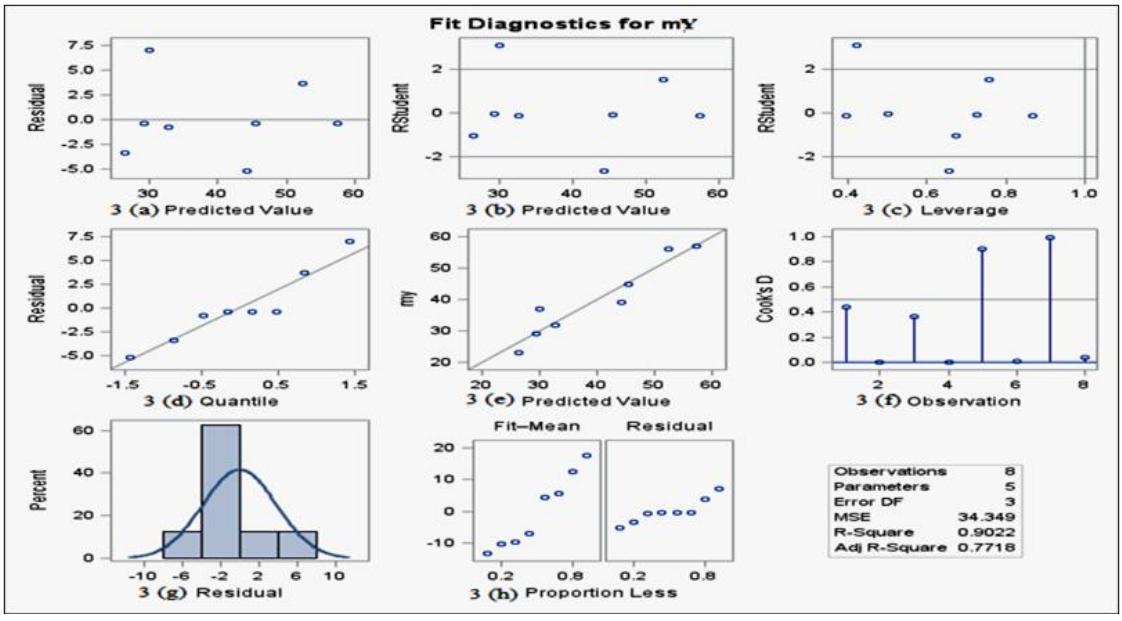

Fig 3: Fit diagnostics for male category of ACHF

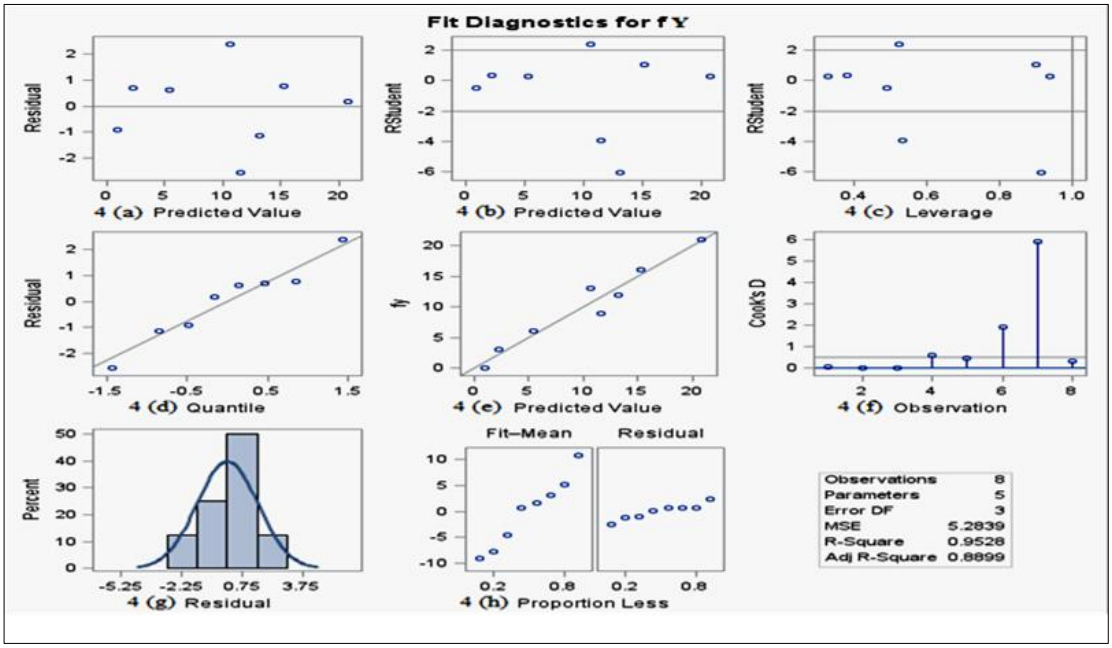

Fig 4: Fit diagnostics for female categoryof ACHF

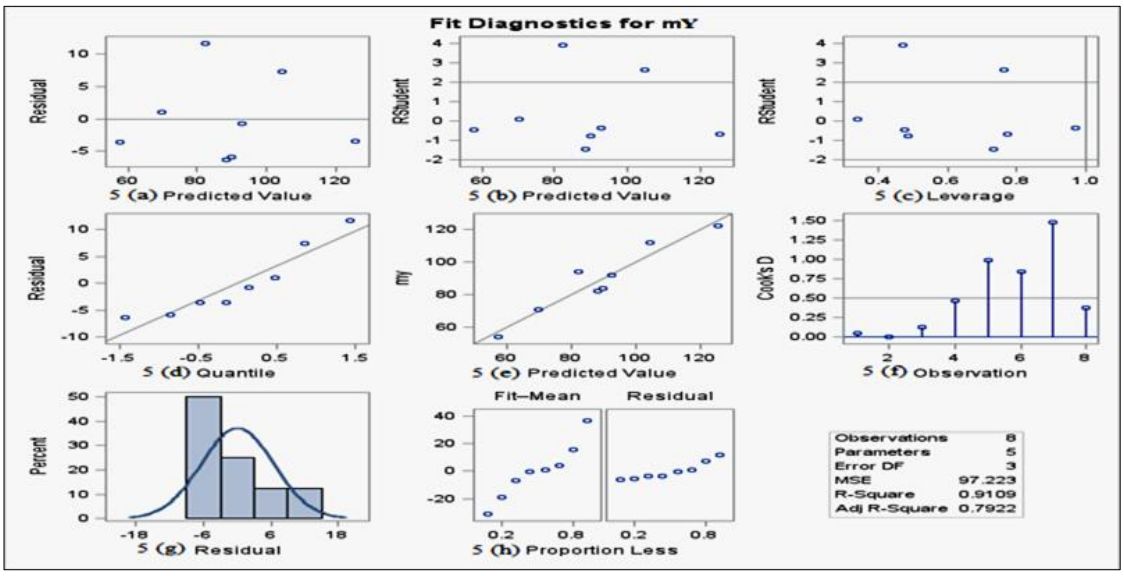

Fig 5: Fit diagnostics for male category of combined colleges

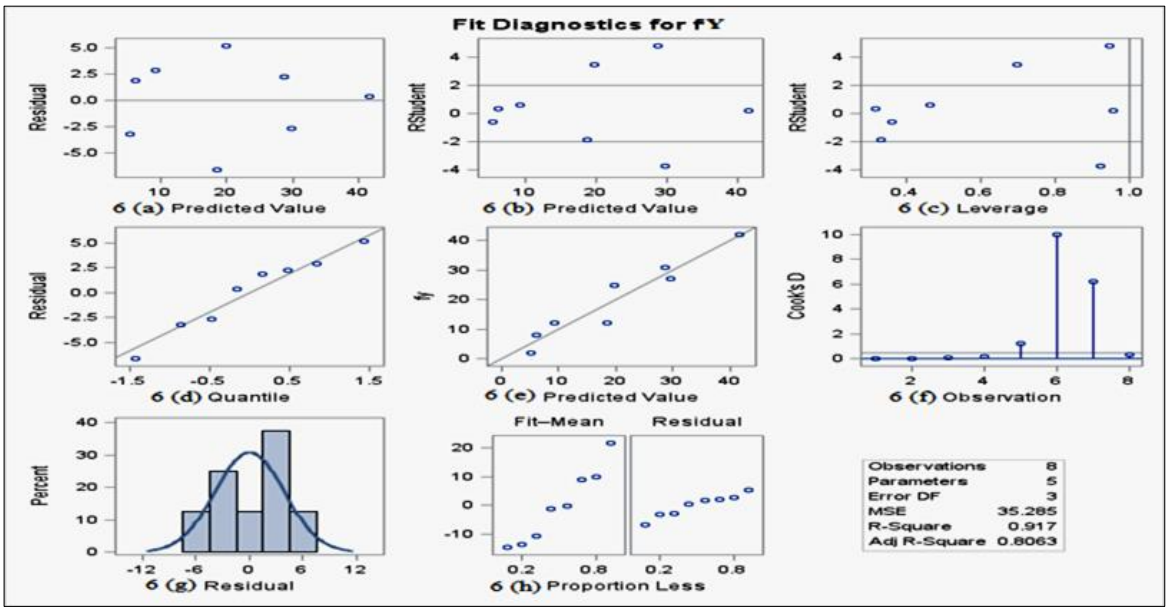

Fig 6: Fit diagnostics for female category of combined colleges

Conclusions

In the present study, the enrollments of male students were more as compared to the female students and also showed that the enrollment of female students is continuously increasing in agricultural education. The trend in enrollment of student in agriculture is necessary not only for forecast purpose of student performance but also useful for planning and effective implementation of educational management programmes and quality judgment for institute. More enrollments of students in agricultural education are better for future, because it is the backbone of Indian economy and also its helps in development of the country. The fitted models of gender category showed less variability which indicates the first year enrollments of both the colleges were similar. It was observed that variability is very less for all categories of both the colleges, because the value of R^2 for all the categories is above 0.90 which indicated the more variability has been explained by class variables. The root mean square error of all the models was found low. Out of six fitted models, the model of female category of N. M. College of Agriculture found best fitted models as compared to all developed models with highest value R^2 value (0.987) and lowest value of *RMSE* (1.100). The present study revealed that fitted models for female categories were found best for both agriculture as well as horticulture colleges as compared to male category.

References

1. Anonymus. All India Survey on Higher Education 2017-18 (AISHE-). Government of India, Ministry of Human resouce Development, 2017-18.
2. Anonymus. Annual report of DARE/ICAR. Agricultural Human Resource Development. 2017-18, 87-96.
3. Cooks RD, Weisberg S. Residual and influence in regression. New York: Chapman and Hall, 1982.
4. Hann CT. Statistical methods in hydrology. Iowa state university press, Ames, 2002, 378.
5. Kumar V, Kumar N, Kumar N. Pattern of enrollments at different education level. S and T Human Resources, India, 2008.
6. Marasinghe MG, Kennedy WJ. SAS for data analysis: intermediate statistical methods with 100 SAS program. Springer science and business media, New York, 2008.
7. Rawlings JO, Pantula SG, Dickey AD. Applied regression analysis: A research tool. Springer-Verlag New York Berlin Heidelberg, 1998.
8. Sahu KK. Development of statistical model for human resource generation in agricultural education in Chhattisgarh. M.Sc. (Agri.) Thesis, Dept. of Agricultural Statistics and Social Science, Indira Gandhi Agricultural University, Raipur, Chhattisgarh, 2012.