Prediction of 305 days lactation milk yield from fortnightly test milk yields in hill cattle under field conditions

R K Pundir*
ICAR- National Bureau of Animal Genetic Resources, Karnal-132001 (Haryana) India

ABSTRACT

The present investigation was undertaken to predict 305 days lactation milk yield from fortnightly test day milk yields using regression and principal component analysis approaches. Accuracy of different methods i.e. simple and multiple regression and principal component analysis using simple and cumulative fortnightly test milk yields was accessed by predicted value, standard error of predicted value, adjusted predicted value, studentized residual, deleted residual, studentized deleted residual, mahaboliny distance, cook's distance and centred leverage values. Test day milk yield were recorded on 688 cows during September, 2009 to June 2010 from 27 villages of Ukhimath block (Kedarnath valley), Rudarpryag district, Uttarakhand, India. Test day milk yields were recorded by fortnightly after 5- 10 days of calving. Daily milk yield (Test day) were computed as sum of morning and evening milk yield on test day. A total of 200 cows completed 20 test day recordings under the study. The 305 days lactation milk yield was estimated from the selected data set (200 cows) those had completed 20 test day yields records using Test Interval Method. Prediction equations were developed considering 305 days lactation milk yield as dependent variable and first 12 fortnightly test day yields as independent variables. All the analysis was carried out by SPSS and SAS programs. It was concluded that for prediction of 305 days lactation milk yield, principal components analysis method using fortnightly test day milk yields was best.

Keywords: Cattle, Lactation milk yield, Test day milk yield, Multiple Regression, Principal Component Analysis
***Corresponding author**: pundir.rakesh@Gmail.com

INTRODUCTION

In dairy cattle improvement programs, selection has been focused on lactation milk yield (Haile, 2006). Prediction of 305 days lactation milk yield from the test day yields with an appropriate accuracy helps in selection of animals at an early stage of life which increase the genetic gain through decreasing the generation interval and cost of recording. Similar studies conducted in cattle and buffaloes (Das and Sadana 1995; Saini et al. 2005) revealed high correlation between test day milk yields and 300 days lactation milk yield suggesting that test day milk yield can be used in prediction of lactation milk yield. The present investigation was undertaken to formulate principal component scores or synthetic variables and to predict lactation milk yield on the basis of these scores and to compare the efficiency of this criteria using early test milk yields with simple and multiple regression analysis in hill cattle of Uttarakhand.

MATERIALS AND METHODS

Test day milk yields were recorded from September, 2009 to June 2010 from 27 villages of Ukhimath block (Kedarnath valley), Rudarpryag district, Uttarakhand on 688 cows by 5 enumerators. Test day milk yields were recorded by fortnightly after 5 to 10 days of calving. Daily milk yield (Test day) were computed as sum of morning and evening milk yield on test day. A total of 200 cows completed 20 test day recordings during the study. The 305 days lactation milk yield was estimated from the selected data set (200 cows) those had completed 20 test day yields records using Test Interval Method (ICAR, 2003). Descriptive statistics in terms of fortnightly test day milk yield mean with number of observations, minimum, maximum and standards error and estimated lactation milk yield are presented in Table 1. Prediction equations were developed considering 305 days lactation milk yield as dependent variable and first 12 fortnightly test day yields as independent variable using simple and multiple regression and principal component analysis. All the 12 test day milk yields were used in principal component analysis to develop synthetic variables. The resulting eigen values and the percentage of variance explained by each principal component were subsequently used to decide those components to be used in regression analysis. Principal component with eigen value more than one were considered significant and used for prediction of 305 days lactation milk yield. The accuracy of prediction was determine by the coefficient of determination (R^2) values, standard error of predictive value, adjusted predictive value, deleted residual, studentised deleted residual and Mahalabony distances. The 305 days lactation milk yield of remaining cows was predicted by best developed prediction equations under the study as these cows did not complete all 20 test recordings. All the analysis was carried out by SPSS (2007) and SAS (2002).

RESULTS AND DISCUSSION

Hill cows of Uttarakhand

Cows were small in size with cylindrical type of body. Animals were well built and compact with strong legs. Black body colour predominates in the area. Skin was tight and eyes lids were black. The colours of hoof and muzzle were black in most of the cases. Dewlap and hump were small. Head was small and poll prominent. Horns were present. The colour of horn was black and length was small. Face was short and concave. Ears were small to moderate in length and horizontal in orientation. Neck was short in length and thin. Udder was small, touching to the body, not well developed and milk veins were not prominent. Tail was up to the hock with black, brown and white switch. Temperament was docile. Milk productivity of hill cows is low and daily milk yields ranged from 2.0 to 3.0 kg only. Cows were reared on extensive system of management i.e. grazing from morning to evening with some amount of feeds at home in the evening (Pundir et al. 2013).

Descriptive statistics

Descriptive statistics of different test day milk yield are given in Table 1. On the 20th fortnight there were only 200 cows available. The average test day milk yield ranged from 0.400 kg (20th fortnight) to 6.90 kg (Ist fortnight). The average test day milk yield was maximum as 1.579±0.391 kg on the second fortnight. The minimum average test milk yield was observed on last 20th fortnight as 0.632±0.018 kg. The 305 days lactation milk yield was estimated 603.98 kg using Test Interval Method (ICAR, 2003). The 305 days lactation milk yield was estimated from the test day milk yield which was the sum of AM and PM milk yield to avoid overestimation and underestimation when it was based on morning or evening milk yield as reported by Schaeffer and Rennie (1976). Gantner et al. (2009) also reported maximum accuracy for prediction of lactation milk yield from test day milk yield using both morning and evening test milk yield in place of any one of them. They also concluded that usages of simply doubling method gives overestimation and underestimation of daily yields when estimating based on morning or evening records, respectively.

Phenotypic correlations

The phenotypic correlations between 305 days lactation milk yield and different test day milk yields (T1 to T12) were significant, positive and ranged from 0.681 (T12) to 0.938 (T5). The estimates of phenotypic correlations between 305 days lactation milk yield with T4, T5 and T6 was almost similar (0.936 to 0.938). The

Table 1. Average fortnightly daily milk yield (kg) and estimated milk yield at different fortnight

Fortnightly milk yield	No. of obs. available	Mean	Minimum	Maximum	305 days milk yield*
T1	688	2.415	0.680	6.50	
T2	688	2.553	0.900	6.90	62.90
T3	688	2.575	0.900	6.80	101.15
T4	688	2.537	0.900	6.90	137.67
T5	686	2.342	0.700	6.10	172.24
T6	684	2.276	0.700	5.90	205.99
T7	680	2.235	0.650	5.90	238.84
T8	666	2.159	0.630	4.10	271.01
T9	636	2.145	0.600	4.03	302.66
T10	619	2.086	0.500	3.80	333.56
T11	577	2.043	0.450	3.40	364.08
T12	545	2.026	0.450	3.00	393.85
T13	506	1.951	0.430	0.25	422.57
T14	480	1.887	0.300	2.00	450.24
T15	438	1.811	0.300	2.00	477.24
T16	377	1.793	0.300	1.90	503.79
T17	347	1.756	0.250	1.90	529.81
T18	320	1.723	0.300	1.90	555.23
T19	295	1.675	0.250	1.60	579.83
T20	200	1.615	0.240	1.70	603.98

*Test interval method

amount of phenotypic correlation of 305 days lactation milk yield was declined after T5 to T12 records. Positive phenotypic correlations between 305 days lactation milk yield and different test day milk yields were obtained by Hatisa et al. (2007) in Sahiwal cattle and Chakaraborty et al. (2010) in Murrah buffaloes. The phenotypic correlations between 305 days lactation milk yield and cumulative fortnightly milk yield (first 2 to all 12) were in increasing order from 0.924 (first 2) to 0.982 (all 12). The increment in the amount of correlation between 305 days lactation milk yield and first 3 to first 4 and first 4 to first 5 cumulative test day milk yields was maximum as 0.009 and after first 5 it has showed declining trends. The high correlation coefficient between 305 days lactation milk yield and a particular test day milk yield indicated the high prediction accuracy.

Prediction of 305 days lactation milk yield from single independent variable

The constant, regression coefficient, F value and accuracy (R^2) for the estimation of 305 days lactation milk yield from different fortnightly test milk yield are given in Table 2. All the prediction equations had significant F values. The regression coefficients of 305 days lactation milk yield on different test day milk yields (from T1 to T12) showed increasing trends from 123.732±3.821 to 264.80±20.214, respectively based on single independent variable. The accuracy of prediction (R^2) was maximum with T5 as 0.880 and thereafter it was declined as the record of test day milk yield increased.

The prediction accuracy with T4 and T5 was almost similar (0.875). Deb and Gurnani (1994) reported similar estimates of R^2 for T5 and T6 records. The accuracy of different prediction equations was higher than the reports of Gokhale and Nagarcenkar (1979) and Singh and Rana (2008). The low R^2 values were observed at the terminal phase of test day milk yields (T10, T11 and T12). It indicated that the terminal part of lactation explained least of the total variance of the lactation milk yield as compared to rest of the test day milk yield. This could be due to the higher components of temporary

Table 2. Regression coefficients and constants for estimating 305 days lactation milk yield from different fortnightly milk yields

Fortnightly milk yield	R square	F value	Constant±SE	Coefficient±SE
Single independent trait				
T1	.841	1048	177.333±8.817	123.732±3.821
T2	.844	1070	168.553±8.967	129.224±3.950
T3	.814	865.383	159.812±10.195	136.244±4.631
T4	.875	1390.332	147.582±8.397	150.607±4.039
T5	.880	1446.516	124.309±8.787	173.043±4.550
T6	.876	1392.510	118.045±9.104	185.869±4.981
T7	.860	1212.578	120.311±9.682	189.779±5.450
T8	.781	706.007	112.514±12.858	210.880±7.937
T9	.719	505.920	93.533±15.882	231.013±10.271
T10	.657	379.029	101.169±17.880	245.00±12.584
T11	.582	275.420	102.288±20.763	259.437±15.633
T12	.464	171.606	107.595±25.621	264.80±20.214
Cumulative test day milk yields				
First2	.854	1156.48	169.61±8.60	64.06±1.88
First3	.862	1233.83	160.06±8.50	44.64±1.27
First4	.878	1420.22	153.21±8.17	34.94±0.93
First5	.894	1672.88	143.21±7.76	29.61±0.72
First6	.907	1934.52	134.74±7.46	25.98±0.59
First7	.916	2170.81	128.11±7.12	23.24±0.49
First8	.932	2711.40	116.92±6.58	21.60±0.41
First9	.943	3282.94	104.87±6.12	20.39±0.35
First10	.952	3942.05	93.72±5.84	19.45±0.31
First11	.959	4641.73	82.54±5.50	18.74±0.27
First12	.963	5207.46	70.51±5.35	18.20±2.52

Table 3. Regression coefficients and constants for estimating 305 days lactation milk yield from different first 12 fortnightly milk yield by stepwise procedure

Fortnightly milk yield	R square	F value	Constant±SE	Coefficient±SE
Test day milk yield				
T5	.880	1446.516	124.309±8.787	173.043±4.550
T5, T9	.929	1293.595	73.237±8.030	127.043±5.247
				91.112±7.750
T5, T9, T2	.947	1165.425	81.001±7.039	77.684±7.622
				83.608±6.794
				44.703±5.534
T5, T9, T2, T12	.956	1053.144	55.346±7.641	81.953±7.010
				54.316±7.790
				42.049±5.083
				54.524±8.734
T5, T9, T2, T12, T8	.961	957.800	55.360±7.187	71.386±6.905
				30.972±8.618
				40.884±4.786
				51.363±8.237
				39.100±7.602
T5, T9, T2, T12, T8, T3	.963	842.197	50.492±7.153	63.452±7.134
				33.333±8.428
				26.351±6.367
				55.809±8.137
				35.133±7.502
				22.924±6.836
T5, T9, T2, T12, T8, T3, T7	.965	745.772	51.489±7.052	52.767±8.085
				26.642±8.669
				27.105±6.275
				55.466±8.012
				34.064±7.397
				18.751±6.910
				22.300±8.360
T5, T9, T2, T12, T8,	.965	663.163	50.906±7.004	44.889±8.943
T3, T7, T11				23.786±8.721
				30.672±6.478
				43.786±9.875
				28.115±7.923
				17.768±6.874
				28.305±8.826
				21.201±10.632
Cumulative test day milk yields				
First12	.963	5207.462	70.513±5.350	18.201±0.252
First12 and First 4	.966	2799.537	58.449±6.017	21.858±.966
				-7.602±1.943

environmental effects in the later part of the lactation (Saini et al. 2005). The accuracy of different prediction equations based on simple regression suggested that test milk yield record at T5 is enough for prediction of 305 days lactation milk yield with a reasonable accuracy.

Prediction of 305 days lactation milk yield from cumulative test day milk yields i.e. first 2, first 3 etc. Indicated that R^2increased from first 2 to all 12 and ranged from 0.854 to 0.963. The increment in $R^2$was maximum from first 3 to first 4 and from first 4 to first 5 (0.016). The R^2 value with first 5 and single T5 was 0.894 and 0.880 showing slight improvement (0.006) when sum of all first five test milk yields considered for prediction of 305 days lactation milk yield. The increment in $R^2$was declined after first 8 to all 12, at the terminal phase of the lactation.

Prediction of 305 days lactation milk yield using stepwise regression procedure

The constant, regression coefficient, F value and accuracy (R^2) for the estimation of 305 days lactation milk yield from different fortnightly test milk yields using stepwise regression are given in Table 3. It was observed that when two variable T5 and T9 considered together for prediction of 305 days lactation milk yield, the R^2 value was 0.929 which gave maximum gain in accuracy as 0.049 as compared to single independent variable T5. The second maximum gain was 0.018 in R^2, when three variable T5, T9 and T2 considered together for the prediction of 305 days lactation milk yield which explained around 94.7 of total variance in the lactation milk yield. Increasing the variables (test day milk yields) more three did not add more than 0.009% of accuracy in the prediction of 305 days lactation milk yield. It was observed that there was no use of increasing the variable more than 2 (T5 and T9) because of minor improvement in R^2 value. It may be concluded that test milk yield at 2 and 9 fortnights may be pooled for prediction of 305 days lactation milk yield with an accuracy of 93%. The estimates of accuracy of predictions were comparable with those reported by Saini et al. (2005) and Singh and Rana (2008).

Cumulative test day milk yields were also used for prediction of 305 days milk yield. There were only two combinations resulted from stepwise regression procedure i.e. all 12 and all 12 and first 4. The corresponding coefficient of determination (R^2) was 0.963 and 0.966, respectively. It may be concluded that simple test day milk yield T5 and T9 together revealed better combination for prediction of 305 days lactation milk yield while considering the all traits and their time of recording or expression.

Prediction of 305 days lactation milk yield using principal component analysis (PCA) procedure

The estimated factors loading extracted by factor analysis, eigen values and variation explained by each factor are presented in table 4. The scree plot and factor components developed from different fortnightly milk yields are given in Figures 1&2, respectively. There were only two factors extracted with eigen values greater than 1 using all 12 test day milk yields independently.

Extraction Method: Principal Component Analysis.

Factor 1 and 2 accounted for 79.387% and 90.155% of total variation in the lactation milk yield, respectively. These two factors were used to predict 305 days lactation milk yield. Factor 1 and Factor 1&2 together gave the accuracy (R^2) of prediction as 91.4% and 96.6%, respectively (Table 5). When cumulative test day milk yields were used only one factor was developed which explained 98.793% of total variation in the 305 days lactation milk yield. This factor was used to predict 305 days lactation milk yield which explained 92.5% of total variation in lactation milk yield i.e. less than simple independent test day milk yield. It may be concluded that simple test day milk yields were better for prediction of 305 days milk yield as compare to cumulative test yields in the present study. Vaidya (2002) used PCA to predict first lactation milk yield and lifetime milk yield in crossbred cattle and observed that PCA were capable for prediction of lactation milk yield from test day milk yield with R^2 of 99% in the farmer herds. He also emphasized the

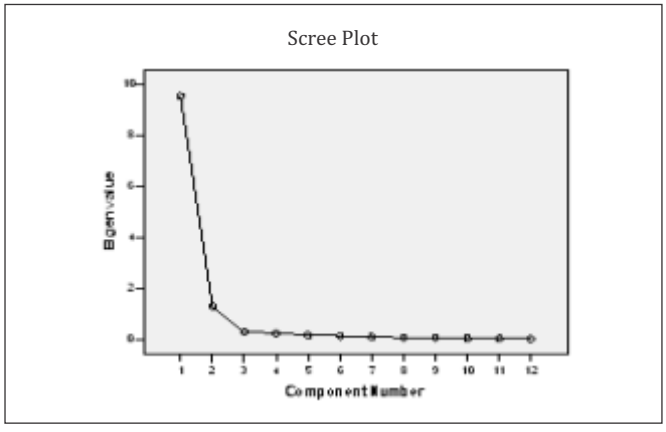

Figure 1: Scree plot for different factors developed from different fortnightly milk yields

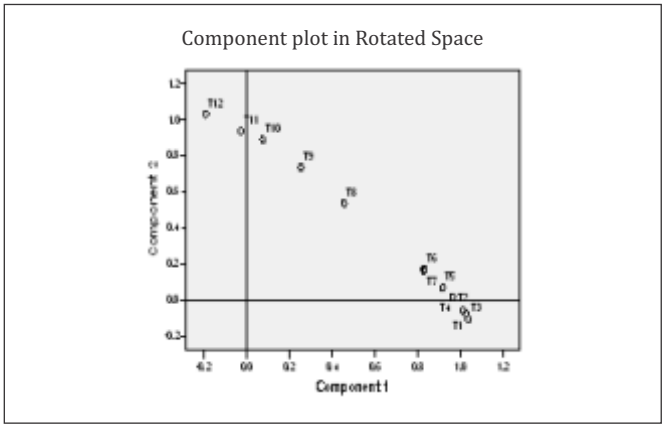

Figure 2: Factor components developed from different fortnightly milk yields

Table 4. Different component factors and variance explained developed factors from different fortnightly milk yields

Component	Eigen values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings(a)
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	9.526	79.387	79.387	9.526	79.387	79.387	8.942
2	1.292	10.768	90.155	1.292	10.768	90.155	7.432
3	.300	2.496	92.651				
4	.248	2.067	94.718				
5	.167	1.389	96.106				
6	.140	1.163	97.269				
7	.103	.860	98.129				
8	.069	.577	98.707				
9	.060	.498	99.205				
10	.043	.355	99.560				
11	.034	.282	99.842				
12	.019	.158	100.000				

Table 5. Regression coefficients and constants for estimating 305 days lactation milk yield from different synthetic factors developed by PCA

Factor	R square	F value	Constant±SE	Coefficient±SE
From fortnightly milk yield				
Factor1	.914	2108.454	427.116±3.139	144.503±3.147
Factor 1 and 2	.966	2791.051	427.116±1.983	110.992±2.776
				48.013±2.776
From Cumulative fortnightly milk yield				
Factor1	.925	2456.552	427.116±2.926	145.390±2.933

Table 6. Residuals Statistics by different procedures for estimation of 305 days lactation milk yield

Parameter	Regression method				Principal component analysis (PCA)			
	Fortnightly milk yield		Cumulative fortnightly milk yield		Fortnightly milk yield		Cumulative fortnightly milk yield	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Predicted Value	627.11	148.48	627.11	148.54	627.11	145.39	627.11	148.53
Standard Error of Predicted Value	5.14		3.19		3.93		3.21	
Adjusted Predicted Value	427.47	147.28	427.06	148.39	427.11	145.42	427.09	148.38
Studentized. Residual	-.006		.001		.000		.000	
Deleted Residual	-.357		.0480		-.0024		.024	
Studentized Deleted Residual	-.002		.005		.000		.004	
Mahal. Distance	7.96		1.99		.995		1.99	
Cook's Distance	.00		.006		.005		.005	
Cantered Leverage Value	.040		.010		.005		.010	

advantage of PCA in reducing the number of traits with a minor loss of accuracy as compared to multiple regressions. There are few limited studies those predict performance of dairy cattle using PCA from the body weights, age at conception and calving, service period etc. (Bhattacharya and Gandhi, 2005; Haile et al. 2008). Regression analysis and PCA was compared by Haile et al. (2008) and recommended that the parameter estimates generated with multiple regression were unreliable because of the multicollinearity among dependent variables. They suggested that PCA may be used to determine more reliable estimates by reducing the effects of multicollinearity.

The accuracy of prediction of 305 days lactation milk yield by different methods i.e. regression (T5), stepwise regression (T5 and T9) and PCA (all 12 cumulative and factor 1&2) was determine by the R2 values, standard error of predictive value, adjusted predictive value, deleted residual, studentised deleted residual and Mahalabony distances, Cook's Distance and Cantered Leverage Value (Table 6). Based on different statistical parameters it may be concluded that PCA with test day milk yields is the best method in the present study. Based on these prediction equations and available records (test day milk yield), 305 days lactation milk yield was predicted. The predicted lactation milk yield by different methods i.e. regression (T5), stepwise regression (T5 and T9) and PCA (all 12 cumulative and factor 1&2) was 657.21±5.31 kg (686), 667.59±6.23 kg (545), 668.15±6.15 kg (545) and 627.12±5.06 kg, (545), respectively. The lower standard error was with PCA with test day milk yields further suggested that this method is the best among the studied.

ACKNOWLEDGEMENT

Sincerely thanks are due to the Director, ICAR- National Bureau of Animal Genetic Resources (NBAGR), Karnal and Director and Staff, Animal Husbandry Department, Uttarakhand for providing necessary facilities and their support during the study.

REFERENCES

Bhattacharya, TK and Gandhi RS. 2005. Principal components versus multiple regression analysis to predict lifetime production of Karan Fries cattle. *Indian Journal of Animal Sciences.* 75 (11): 1317-20.

Chakaraborty D, Dhaka SS, Pander BL, Yadav AS and Dandapat A. 2010. Genetic studies of 305 days and test day milk yield records in Murrah buffaloes. *Indian Journal of Animal Sciences.* 80 (8): 729-32

Das G and Sadana DK. 2003. Predictability of lactation milk yield based on test day values in Murrah buffaloes. *Indian Journal of Animal Research.* 37 (2): 136-38

Gantner V, Jovanovac S, Kovac M, Malovrh S and Kompan D. 2006. Various approaches to daily milk yield prediction from alternative milk recording scheme. *Acta Agraria Kaposvariensis.* 10:105-112.

Gokhale SB and Nagarcenkar R.1979. Inheritance of part yield and their role in selection of buffaloes. *Tropical Animal Health and Production*. doi: 10.1007/BF02237887

Haile A. 2006. Genetic and economic analysis of zebu cattle and their crosses with Holstein Friesian in central Ethiopia, *Ph.D. thesis, NDRI, Deemed University, Karnal, India*.

Haile A, Joshi BK, Ayaleq W, Tegegne A, Singh A and Chakarvarti AK. 2008. Prediction of first lactation milk yield of Boran cattle and their crosses with Holstein-Friesian in central Ethiopia using multiple regression and principal component analysis. *Indian Journal of Animal Sciences.* 78 (1): 66-69.

Hatisa ED, Muasya WB and Kahi AK. 2007. Genetic and phenotypic parameters for test day milk yield of Sahiwal cattle in semi arid tropics. *Animal.* 1:185-92.

ICAR, 2003. International Agreement of Recording Practices. International Committee for Animal Recording ed., Interlaken, Switzerland.

Ptak E and Schaffer LR.1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livestock Production Science.* 34: 23-34

Pundir RK**,** Singh PK, Neelkant, Sharma D, Singh CV and Prakash B. 2013. Uttara- A new cattle germplasm from Uttarakhand hills. *Indian Journal of Animal Sciences.* 83 (1)**:** 51-58.

Saini T, Gahlot GC and Kachwaha RN. 2005. Prediction of 300 days lactation yield on the basis of test day milk yields in Rathi cows. *Indian Journal of Animal Sciences.* 75 (9): 1087-89.

SAS .2002. Statistical Analysis System for mixed models, SAS Institute Inc, Cary, NC, USA.

Schaeffer LR and Rennie JC. 1976. Am-pm testing for estimating lactation yields. *Canadian journal of Animal Sciences.* 56:9

Singh A and Rana JS. 2008. Prediction of 305 day milk yield based on test day values in Murrah Buffaloes. *Indian Journal of Animal Sciences.* 78(10): 1131-33

SPSS. 2007. Statistical Package for Social Sciences. SPSS Inc., 444 Michigan Avenue, Chicago, IL 60611.

Vaidya MS. 2002. Genetic evaluation of crossbred cows in Konkan region of Maharashtra state, *Ph.D. thesis, NDRI, Deemed University, Karnal, India.*