Research paper

Fitting random regression model to analyze test day milk yield in Sahiwal cattle

Manjari Pandey[1] and Raja K N[2]*

[1]Animal Genetics and Breeding division, ICAR-National Dairy Research Institute, Karnal, (Haryana) India
[2]ICAR- National Bureau of Animal Genetic Resources, Karnal, (Haryana) India

ABSTRACT

The present research was carried out for analyzing test day milk yield in Sahiwal cattle using random regression model (RRM). First lactation 3953 monthly test day milk yield (MTDMY) records over a period of 31 years (1986-2017) were collected on 392 cows from history cum pedigree sheets and daily milk yield register maintained at NDRI, Karnal, Haryana. The records were taken at an interval of 30 days, with first record on 6th day and last on 305th day. The test day milk yield data was modeled using RRM (WOMBAT) considering different order of Legendre polynomial (LP). 4th order and 6th order LP for the additive genetic effect and the permanent environmental effect respectively was found to be the best. The highest and lowest additive genetic variance was observed for 125th day milk yield (2.57 kg^2) and 35th day milk yield (1.02 kg2) respectively. Its magnitude was found to be lower in the beginning and towards the end of lactation and relatively higher in the mid-lactation. The permanent environmental variance was observed higher during initial test days. The residual variance was lower than the permanent environment variance for all the test day milk yields. The heritability of monthly test day milk yields was found to be highest (0.33 ± 0.04) for 125th day milk yield and the lowest (0.15 ± 0.03) for 6th and 35th day milk yield. The estimates of genetic correlations and permanent environment correlations between different test day milk yields ranged from 0.02 to 0.98 and from 0.27 to 0.98 respectively. The study highlights the usage of test day model analyzed through random regression model to be utilized for selecting animals at an early age.

Keywords: Genetic parameters, phenotypic parameters, Random regression analysis, Sahiwal, Test day model, WOMBAT
***Corresponding author:** mnpandey155@gmail.com
Manuscript received: 20.03.2019 ; accepted: 12.08.2019

INTRODUCTION

Livestock plays an important role in Indian economy and it provides employment to about 8.8% of the population. India is rich in livestock resources. Indigenous germ pool is the real asset of any country and among the various indigenous breeds Sahiwal is one of the best dairy breed which produces the most milk of all zebu cattle breeds. Good herds of pure Sahiwal are available around Fazilka and Abohar towns of Ferozepur district in Punjab, Sri Ganganagar district of Rajasthan and some pockets of U.P. and Chhattisgarh. India ranks first in milk production but productivity per animal is low due to low genetic potential for milk production. This can be improved by early selection and proper breeding strategies. In India mostly first lactation 305 days or less milk yield is used to select or cull the animals. It's use consumes time and money, leads to increased generation interval, decreased genetic gain per unit of time and also is based on less number of records. One solution to all this is the use of test day milk yield records for selecting animals (Kokate *et al.,* 2013; Gupta, 2013; Singh, 2014). Test day (TD) model is a statistical procedure which considers genetic and environmental effect directly on the test day basis (Swalve, 1995). TD increases the accuracy of sire evaluation as it uses large number of measurements per daughter than one lactation record in 305 days milk yield model and also in TD model more detailed fixed effects can be adjusted. The data on test day milk yield of dairy cows is a longitudinal data having correlation between tests on the same animal. It reduces residual variance by providing more information per sire by using all available test day of sire's daughters and by adjusting more environmental effects in genetic evaluation model. it is more feasible to maintain fortnightly, monthly, bimonthly milk yield records under field conditions where farmers are rarely involved in maintaining daily yield records.

Most of the developed countries are using monthly test day milk yield records using Random Regression Model (RRM) instead of 305 days milk yield for genetic evaluation of dairy animal as lactation curve of each animal is different. RRM is used for the analysis of longitudinal data of individual over time measured on trajectory. RRM allows the inclusion of random regression coefficients for the lactation curve for each cow (Henderson, 1982). In this the fixed regression explains the general shape of lactation curve for all animals belonging to the same sub class of age - season of calving and random regression describes the deviation which allows each cow to have differently shaped lactation curve and account for random genetic and permanent environmental effects.

MATERIALS AND METHODS

Source of data

Data of first lactation 3953 monthly test day milk yield (MTDMY) records were collected on 392 cows from history cum pedigree sheets and daily milk yield register maintained at NDRI, Karnal, Haryana and was drawn over a period of 31 years (1986-2017). Test day milk yield records were taken at an interval of 30 days, with first MTDMY recorded on 6th day and last MTDMY on 305th day.

Statistical Analysis

The present study was carried out with the objective of fitting random regression model to analyze genetic and phenotypic parameters of monthly test day milk yield in Sahiwal cattle.

For carrying out the analysis the season of calving was classification of season of calving was done as follows; S_1= winter (December to March), S_2= summer (April to June), S_3= rainy (July to August), S_4= autumn (September to November). For classification of period of calving, the total span of 31 years was divided into ten periods (P1 to P10) of three consecutive years with last period having 4 consecutive years. The age at first calving was coded as A1 ≤ 961 days to A10 ≥1417 days with a class interval of 80 days.

Random regression test day model was applied using WOMBAT software

The test day milk yield data was modeled considering different order of Legendre polynomial (LP) for the additive genetic effect (4th order) and the permanent environmental effect (6th order).The random regression model used can be represented as:

$$\mathbf{Y = Xb + Za + Wp + e}$$

where, Y is the vector of test day milk yields of cattle in different lactation; b is the vector of fixed effects (season, year, age groups); X is the incidence matrix relating test day milk yields to fixed effects, p is vector of permanent environmental random regression coefficients; a is vector of additive genetic random regression coefficients; Z and W are covariate matrices for 'a' and 'p' respectively and 'e' is the vector of random residual effects associated with Y. The assumption of this model was

$$\begin{bmatrix} a \\ p \\ e \end{bmatrix} \sim N(0, V) \text{ with } V = \begin{bmatrix} G \otimes A & 0 & 0 \\ 0 & I \otimes P & 0 \\ 0 & 0 & R \end{bmatrix}$$

where, G is Variance-Covariance matrix of additive genetic random regression coefficients; A is Additive genetic relationship matrix among the animals; x is Kronecker product function; P is Variance-Covariance matrix of permanent environment random regression coefficients; I is Identity matrix and R is diagonal matrix of residual variances.

Function in RRM

The Legendre polynomials (LP) have been used extensively in RR analysis. Kirkpatrick *et al. (1990)* proposed the use of covariance function for longitudinal data of this kind. To calculate LP, first define $P_{0(x)}$=1, and $P_{1(x)}$ = x. Then, in general, the *n+1* polynomial is described by the followings recursive equation:

$$p_{n+1}(x) = \frac{1}{n+1}((2n+1)xp_n(x) - np_{n-1}(x))$$

These quantities were normalized using

$$\emptyset_n(x) = (\frac{2n+1}{2})^{0.5} P_n(x)$$

where n is the order of the polynomials. Test day records in the interval 6 to 305 ti days were standardized to the interval -1 to +1 with the following formula:

$$\propto_i = -1 + 2\left(\frac{t_i - t_{min}}{t_{max} - t_{min}}\right)$$

Estimation of genetic and phenotypic parameters

Jamrozik and Schaeffer (1997) procedure was used to estimate the genetic and phenotypic parameters such as heritability of test day milk yields, genetic and phenotypic correlations between test day records at different days in milk (DIM) using the genetic (co)variances, permanent environmental (co)variances and homogeneous residual variances of test day milk yield

Variance-covariance component estimation: A mixed model analysis was carried out to obtain restricted maximum likelihood estimate of covariance components with software WOMBAT (Meyer, 2007).

Genetic variance of test day milk yield: The genetic variances of test day milk yields at different DIM was estimated using the covariates of Legendre polynomial function as follows: $\sigma^2 a_{(i)} = z_i G z'_i$

Permanent environmental variance of test day milk yield: $\sigma^2_{pe\ (i)} = z_i P z'_i$

Genetic covariance between test day milk yields: $\sigma_{a\ (ij)} = z_i G z'_j$

Permanent environmental covariance between test day milk yields: $\sigma_{pe(ij)} = z_i P z'_j$

where, $\sigma^2 a(i)$; $\sigma^2 pe\ (i)$ is genetic and permanent environment variance of test day milk yield on i^{th} DIM respectively; Z_i is coefficients of the covariate matrix corresponding to i^{th} DIM and G is Additive genetic variance –covariance matrix; P is permanent environmental variance–covariance matrix

Estimation of heritability of test day milk yields

The heritability of test day milk yield records at different DIM (6, 35…275 and 305 days) in first lactation was estimated as follows

$$h^2_{(i)} = \frac{\sigma^2_{a(i)}}{\sigma^2_{a(i)+}\ \sigma^2_{pe(i)} + \sigma^2_{e(i)}}$$

Where, $h_{(i)}^2$ is heritability of test day milk yield on i^{th} DIM; $\sigma^2 a_{(i)}$ genetic variance of test day milk yield on i^{th} DIM; $\sigma^2 a_{(i)}$ is permanent environmental variance of test day milk yield on ith DIM and $\sigma^2 a_{(i)}$ is residual variance of test day milk yield on ith DIM.

Genetic and Phenotypic Correlations

The genetic and phenotypic correlations among different monthly test day milk yields were calculated from the analysis of variance and covariance among test days milk yields.

Genetic correlation (r_g)

$$r_g(XY) = \frac{Cov\ a}{\sqrt{(\sigma_X^2)(\sigma_Y^2)}} \quad \text{with} \quad \text{S.E.}\left(r_g\right) = \frac{1-r_g{}^2}{\sqrt{2}} \sqrt{\frac{\text{S.E.}(h^2{}_x)\text{S.E.}(h^2{}_y)}{(h^2{}_x)(h^2{}_y)}}$$

where

$Cova\_{XY}$ is Genetic component of covariance between traits X and Y; $\sigma^2 x$ and $\sigma^2 y$ are Genetic components of variance for traits X and Y; h2x and h2y are the heritability estimates of the two traits x and y, respectively.

Similarly, using permanent environment variance-covariance estimate correlation between permanent environment effects was calculated

Phenotypic correlation (rp)

$$r_p(XY) = \frac{COV_a(XY) + COV_{pe}(XY) + COV_e(XY)}{\sqrt{\left[\sigma_a^2(X) + \sigma_{pe}^2(X) + \sigma_e^2(X)\right]\left[\sigma_a^2(Y) + \sigma_{pe}^2(Y) + \sigma_e^2(Y)\right]}} \quad \text{with} \quad SE\left(r_p\right) = \sqrt{\frac{\left[1 - r_p^2(XY)\right]}{\left[(N-2)\right]}}$$

where, COV_a (XY) is Genetic component of covariance between traits X and Y; COV_{py} (XY) is Permanent environment component of covariance between traits X and Y; COV_e (XY) is Error component of covariance between traits X and Y; $\sigma^2 e$ (X) and $\sigma^2 e$ (Y) = Error components of variance for traits X and Y; rp (XY) is Phenotypic correlation between the traits X and Y in the same individual; N−2 is Degree of freedom.

RESULTS AND DISCUSSION

Random Regression test day model (RR-TDM) was used on monthly test day milk yields (MTDYs) of first lactation of Sahiwal cattle to study the inheritance patterns of test day milk yields and to estimate the genetic correlations among test day milk yields. Legendre Polynomial (LP) was used separately to model the covariates for fixed and random regression coefficients in the random regression test day model.

Table 1: Estimates of covariate function parameters for fitting different models

	LP35	LP36	LP45	LP46
Log l	-4102.884	-4094.209	-4100.465	-4082.793
AIC	-8249.768	-8232.418	-8252.93	-8224.586
BIC	-8387.888	-8370.532	-8416.162	-8359.812

AIC: Akaike's information criterion, BIC: Bayesian information criterion

Random regression coefficients

Different orders of Legendre polynomial were used for modeling the covariates in the random regression test day model as given in Table 1. LP of order 4 and 6 were best fitted for additive genetic variance and permanent environment variance as these had lower log likelihood function. AIC and BIC values compared to other models.

The estimated variances (a_i, a_i) and co-variances (a_i, a_j) among the additive genetic random regression coefficients using Legendre polynomial in first lactation of Sahiwal cattle have been presented in the Table 2. Eigen values represent the amount of variation explained by the corresponding eigen function (Kirkpatrick *et al.,* 1990).In RR-TDM analysis, the first two eigen values 2.93 (81.16%), 0.58 (16.14) of the additive genetic covariance function accounted for at least 98% of the sum of all eigen values but the first three eigen values 6.88 (79.93%), 1.22 (14.12%), 0.33 (3.89%) permanent environment effect accounted at least 98% to total variation. Little variation was associated to other eigen values for additive genetic effects and permanent environment effects.

Additive genetic, permanent environment and residual Variances of test day milk yields

The variances (additive genetic, permanent environment and residual) of different test day milk yields (Table 3) were estimated using variance-covariances structure among random regression coefficients and covariate of the functions used in random regression test day model. The residual variance was assumed constant for all monthly test day milk yields. The highest additive genetic variance was observed for 125th day milk yield (2.57 kg^2) and the lowest was observed for 35th day milk yield (1.02 kg^2).The additive genetic variance increased up to 125th day and gradually decline thereafter towards the end of lactation. The magnitude of VA was found to be lower in the beginning and end of lactation and relatively higher in the mid-lactation. The permanent environmental variance was observed higher during initial test days. Ved Prakash (2015) reported highest individual permanent environment variance as well as phenotypic variance for TD3 (66th day).The residual variance was observed to be lower than the permanent environment variance for all the test day milk yields.

Table-2: Estimates of variances (diagonal, boldface type), co-variances (upper diagonal) and correlations (lower diagonal) between random regression coefficients in Sahiwal cattle

Additive genetic RRC					Permanent environment RRC						
	a0	a1	a2	a3		p0	p1	p2	p3	p4	p5
a0	2.66	0.11	-0.78	0.26	p0	6.75	0.29	-0.49	0.38	-0.62	0.13
a1	0.09	0.57	-0.01	-0.08	p1	0.11	1.04	-0.14	-0.31	0.16	-0.09
a2	-0.89	-0.03	0.29	-0.13	p2	-0.32	-0.23	0.35	0.24	-0.19	0.04
a3	0.53	-0.36	-0.82	0.09	p3	0.31	-0.63	0.08	0.23	-0.16	0.02
					p4	-0.49	0.32	-0.26	-0.69	0.21	-0.12
					p5	0.26	-0.24	0.15	0.38	-0.38	0.07

Table 3: Additive genetic, permanent environment and phenotypic variance (kg2) and temporary environment of monthly test day milk yields using Legendre polynomial

Additive genetic variance			Permanent Environmental variance		Phenotypic variance		Temporary Environmental variance
DIM	VA	S.E.	VEP	S.E.	VP	S.E.	VET
6th day	1.06	0.42	4.89	0.47	4.47	0.31	1.30
35th day	1.02	0.40	4.26	0.42	6.59	0.44	1.30
65th day	1.79	0.36	5.71	0.44	8.81	0.45	1.30
95th day	2.40	0.42	4.88	0.38	8.59	0.42	1.30
125th day	2.57	0.38	3.92	0.24	7.79	0.38	1.30
155th day	2.39	0.41	3.69	0.25	7.38	0.42	1.30
185th day	2.08	0.33	3.91	0.41	7.29	0.45	1.30
215th day	1.79	0.38	4.30	0.47	7.39	0.43	1.30
245th day	1.57	0.21	4.77	0.43	7.65	0.54	1.30
275th day	1.36	0.30	4.69	0.39	7.35	0.50	1.30
305th day	1.30	0.34	3.51	0.25	5.95	0.36	1.30

Heritability estimates of test day milk yields

The additive genetic, permanent environmental and residual variances of test day milk yields were utilized to estimate the heritability of test day milk yield using Legendre polynomial function for first lactation of Sahiwal cattle and presented in the Table 4. The h^2 of the monthly test day milk yields was found to be highest (0.33 ± 0.04) for 125th day milk yield and the lowest (0.15 ± 0.03) for 6^{th} and 35th day milk yield by RR-TDM. Danell (1982) reported that the heritability varied from 0.15 to 0.31

Table 4: Estimates of heritability and permanent environment variance ratio of monthly test day milk yields by random regression test day model

Test day	h2 ± S.E.
6th day	0.15 ± 0.03
35th day	0.15± 0.04
65th day	0.20 ± 0.03
95th day	0.28 ± 0.03
125th day	0.33 ± 0.02
155th day	0.32 ± 0.03
185th day	0.28 ± 0.05
215th day	0.24 ± 0.03
245th day	0.20 ± 0.03
275th day	0.18 ± 0.04
305th day	0.21 ± 0.03

Genetic and Permanent environment correlations between test day milk yields

The genetic and permanent environment correlations between different test day milk yields were estimated based on the genetic and permanent environmental variance and co-variances of test day milk yields using Legendre polynomial and are presented in the Table 5. The estimates of genetic correlations between different test day milk yields ranged 0.02 to 0.98. Genetic correlation between 6^{th} test day record with milk records of 125^{th} test day onwards were found to be negative and in 275^{th} and 305^{th} test day it was positive again. As the interval between test days increased the magnitudes of genetic correlations between test day milk yields decreased. Adjacent test day had higher correlations. Kettunen et al. (1998) also reported that the genetic correlation between consecutive test-day milk yields were higher (0.81 to 0.98) and decreasing when test-day interval increased in Finnish Ayrshire cows. The estimates of permanent environment correlations of all the monthly test day milk yields ranged from 0.27 to 0.98.

In conclusion, the present investigation was aimed at studying the genetic and phenotypic parameters of

Table 5: Genetic (below diagonals) and permanent environment correlations (above diagonals) among monthly test day milk yields estimated by Legendre polynomial using RR-TDM

Test Day	6th	35th	65th	95th	125th	155th	185th	215th	245th	275th	305th
6th	1	0.54	0.38	0.36	0.38	0.37	0.34	0.31	0.28	0.27	0.32
35th	0.62	1	0.97	0.92	0.78	0.61	0.51	0.52	0.57	0.59	0.43
65th	0.24	0.91	1	0.97	0.86	0.69	0.58	0.58	0.63	0.64	0.46
95th	0.07	0.82	0.98	1	0.95	0.83	0.73	0.71	0.72	0.70	0.54
125th	-0.02	0.75	0.95	0.98	1	0.96	0.89	0.85	0.81	0.77	0.64
155th	-0.07*	0.66	0.89	0.95	0.98	1	0.98	0.94	0.87	0.81	0.72
185th	-0.11*	0.57	0.79	0.88	0.94	0.98	1	0.98	0.92	0.86	0.77
215th	-0.15*	0.42	0.65	0.76	0.84	0.91	0.97	1	0.97	0.93	0.83
245th	-0.20	0.27	0.49	0.61	0.72	0.82	0.91	0.98	1	0.98	0.96
275th	0.02	0.12	0.36	0.49	0.61	0.73	0.85	0.94	0.97	1	0.92
305th	0.05	0.15	0.31	0.47	0.59	0.71	0.82	0.91	0.96	0.98	1

*All the additive genetic correlations and permanent environment correlations were highly significant (P≤.01) except marked with * (Non-significant)*

monthly test-day milk yield of Sahiwal cattle using random regression analysis. After fitting various orders of legendre polynomial for modeling the covariates LP of order 4 and 6 were found to be the best for additive genetic variance and permanent environment variance respectively. The heritability was found to be highest (0.33 ± 0.04) for 125th day milk yield. the higher additive genetic variance in the mid lactation test days as compared to early and late lactation test day milk records suggests that these can be used for predicting first lactation 305 day milk yield and also are more reliable for selecting animals at an early age even before the completion of first lactation. Part time records can be utilized more efficiently than full time records. The inclusion of random regression coefficients in the lactation curve of each cow via random regression analysis further enhances the efficiency of test day model. The study recommends the usage of test day model analyzed through random regression model to be utilized more for selection of animals and planning the breeding programs in India.

REFERENCES

Danell B. 1982. Studies on lactation yield and individual test-day-yields of Swedish dairy cows. IV. Extension of part lactation records for use in sire evaluation. Acta. Agric. Scand. **32:** 65-81.

Gupta, AK. 2013.Genetic evaluation of Sahiwal cattle using lactation curve models. Ph.D. Thesis, NDRI (Deemed University), Karnal, Haryana, India.

Henderson CR. 1982. Analysis of covariance in the mixed model: higher level, non homogeneous and random regression. Biometrices. 38: 623-640.

Jamrozik J and Schaeffer LR. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. J. Dairy Sci. 80(4): 762-770.

Kettunen A, Mantysaari EA, Strandén I, Pösö J and Lidauer M. 1998. Estimation of genetic parameters for first lactation test day milk production using random regression models. 6th WGCALP, Armidale, NSW2351, Australia, **23:** 307-310.

Kirkpatrick M, Lofsvold D and Bulmer M, 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics. 124: 979–993.

Kokate S, Singh A, Banu R, Gandhi RS, Chakravarty AK, Gupta AK and Sachdeva GK. 2013.Genetic and non-genetic factors affecting monthly test day milk yields in Karan Fries cattle. Indian J. Animal Sci.83

(4): 385–389.

Meyer K. 2007. WOMBAT – A tool for mixed model analyses in quantitative genetics by REML, J. Zhejiang Uni. SCIENCE B., 8:815–821.

Singh M. 2014. Genetic analysis of test day milk yields in Murrah buffaloes. M.V.Sc. Thesis, NDRI (Deemed University), Karnal, Haryana, India.

Swalve HH. 1995. The effect of test day models on the estimation of genetic parameters and breeding values for dairy yield traits. J. Dairy Sci.78: 929-938.

Ved Prakash. 2015. Genetic evaluation of Sahiwal cattle using test day milk yield models. Ph.D. Thesis, NDRI, (Deemed University), Karnal, India.