



## Random regression test-day milk yield models as a suitable alternative to the traditional 305-day lactation model for genetic evaluation of Sahiwal cattle

VED PRAKASH<sup>1</sup>, A K GUPTA<sup>2</sup>, M SINGH<sup>3</sup>, G S AMBHORE<sup>4</sup>, A SINGH<sup>5</sup> and R S GANDHI<sup>6</sup>

ICAR-National Dairy Research Institute, Karnal, Haryana 132 001 India

Received: 9 March 2016; Accepted: 29 August 2016

### ABSTRACT

In the present study, first three lactation 305-day milk yield variance components, genetic parameters and breeding value were estimated using test-day milk yield and 305-day actual milk yield data of Sahiwal cattle. The estimates were obtained using three methods viz. random regression model (RRM) with homogeneous residual variance (RRM-HOM), RRM with heterogeneous residual variance (RRM-HET) and univariate animal model. The additive genetic variance of 305-day milk yield estimated from RRM was higher compared to univariate animal model for all lactation. From RRM, it was possible to account for permanent environmental effects due to individual milk yield variations during lactation. The heritability estimates were low for first (0.072 to 0.079) and third lactation (0.087 to 0.112) 305-day milk yield from all three methods. For second lactation, low heritability estimate from univariate animal model (0.144) and moderate estimate from different RRM (0.206 to 0.219) were obtained. For all lactation, breeding value rank correlation was more than 0.78 between lactation model and random regression test-day model. The same bull was identified as top ranking bull from all three methods. It can be concluded that random regression test-day models can replace conventional 305-day lactation model for genetic evaluation as it resulted in higher additive genetic variance estimates, gave similar/higher heritability value and moderate to high rank correlation estimates for breeding values.

**Key words:** 305-day milk yield, Lactation model, Random regression, Sahiwal, Test-day milk yield

In India, sire evaluation is done mainly on the basis of 305-day or less first lactation milk yield (lactation model) at organized farms. This leads to increased generation interval, decreased genetic gain per unit of time and fewer numbers of daughters per sire due to smaller herd size. To overcome these constraints, the test-day milk yield records at monthly interval are suggested for sire and cow evaluation. Sire evaluation using a test-day model was reported to have higher accuracy, due to having a larger number of measurements per daughter than one lactation record in 305-day milk yield model, can directly account for short-term environmental factors specific to individual yields as well as individual variation of the shape of the lactation curve among cows. In 305-day milk yield models, only average lactation curve of cows are considered. Within the test-day model approach, the genetic component of the lactation curve can be modeled by fitting regression

coefficients for each animal, commonly referred to as random regression (RR) coefficients. Applying random regression models (RRM) to test-day records allows evaluation of cows for parameters associated with shape of the lactation curve simultaneously with production level. Use of orthogonal polynomial functions especially Legendre polynomials (LP) reduces correlation among estimated regression coefficients in comparison to other functions with the same number of parameters (Schaeffer 2004). The residual variance changes over time and this heterogeneity of residual variance across the lactations has been modeled by dividing the lactation length into different intervals, with an assumption of homogeneity within intervals and heterogeneity between intervals (Jaffrezic *et al.* 2000, Prakash *et al.* 2016).

Sahiwal is one of the best dairy cattle breed of India. There is no report of comparison of genetic evaluation of Sahiwal cattle using random regression test-day model and 305-day milk yield model. Hence, present study compared the conventional 305-day milk yield model and the random regression test-day model for the genetic evaluation of Sahiwal cattle in different lactations.

### MATERIALS AND METHODS

*Source of data:* Data were collected from the history sheets and daily milk record registers of Sahiwal cows

Present address: <sup>1</sup>Scientist (drvedagb@gmail.com), AGB Division, ICAR-CSWRI, Avikanagar, Rajasthan. <sup>2,5</sup>Principal Scientist (guptaak2009@gmail.com, avtar54@gmail.com), <sup>3</sup>PhD Scholar (manav21vet@gmail.com), Dairy Cattle Breeding Division. <sup>4</sup>(drgsambhore@gmail.com), Cattle Breeding Farm and Dangi Cow Research Station, Igatpuri, Nashik, Maharashtra. <sup>6</sup>Assistant Director General-AP&B (ravindersinghgandhi@gmail.com), ICAR, New Delhi.

Table 1. Basic statistics and data structure for different lactation in Sahiwal cattle

Particulars of data	First lactation	Second lactation	Third lactation
Number of animals	867	714	373
Number of animals with records	635	425	306
Number of sire with progeny in data	67	55	47
Number of dams	447	397	114
Number of test day records	5890	3884	2852
Mean±SE of 305-day milk yield (kg)	1889.19±25.05	2083.68±34.19	2244.28±40.09
C.V% of 305-day milk yield	33.47	33.83	31.20

maintained at ICAR-National Dairy Research Institute (NDRI), Karnal, India. The records on first to third lactation traits of Sahiwal cattle, spread over a period of 52 years (1961 to 2012), were collected for the present study. Detail of data structure and basic statistics of the data are given in Table 1. The geographical location of NDRI livestock farm is at an altitude of 250 meters above the mean sea level in the Indo-Gangetic alluvial plains on 29°42' N latitude and 72°02' E longitude. The climate of the farm is subtropical in nature. The detail of feeding, breeding and management practices followed at the farm was as described by Gupta (2013). The test-day milk yield records were collected from the daily milk yield recording registers of Sahiwal cows. Sahiwal cows were milked three times a day and daily milk yield was calculated by adding three milk yield records in a particular day. A total of ten individual monthly milk yields were collected at an interval of 30 days from sixth day to 276<sup>th</sup> day of lactation. Actual 305-day milk yield was calculated by adding the daily milk yield from sixth day of lactation to 305-day or less. The records of the animals with known pedigree and normal lactation as reported by Prakash *et al.* (2015) were considered for this study.

*Statistical analysis*

The variance components of 305-day milk yield of first, second and third lactation were estimated using univariate animal model and single trait random regression test-day animal model.

*305-day milk yield model:* For variance component estimation of actual 305-day lactation milk yield (lactation model), a univariate animal model was used. A mixed model analysis was carried out to obtain restricted maximum likelihood estimate of covariance components with software WOMBAT (Meyer 2007).

$$Y_{ijklmn} = \mu + S_i + P_j + A_k + D_l + a_m + e_{ijklmn}$$

Where,  $Y_{ijklmn}$ , 305-day or less milk yield;  $S_i$ , fixed effect of  $i^{th}$  season of calving;  $P_j$ , fixed effect of  $j^{th}$  year of calving;  $A_k$ , fixed effect of  $k^{th}$  age group class;  $D_l$ , fixed effect of  $l^{th}$  days open class;  $a_m$ , animal's random additive genetic effect

NID (0,  $\sigma_a^2$ ) and  $e_{ijklmn}$  is random residual effect NID (0,  $\sigma_e^2$ ).

*Random regression test-day model:* A single trait linear mixed random regression test-day model was applied to monthly test-day records of Sahiwal cattle in different lactation. To find suitable RRM, 22 model similar to Prakash *et al.* (2016) were fitted. For Random regression models with homogenous residual variance (RRM-HOM), a constant variance along days in milk (DIM) were considered. Residual effects on different DIM were uncorrelated both within and between cows. For random regression model with heterogeneous residual variance (RRM-HET), residual variance was assumed to be constant for test-day records within, but different (heterogeneous) between the followings nine lactation period classes (6 to 35, 36 to 65, 66 to 95, 96 to 125, 126 to 155, 156 to 185, 186 to 215, 216 to 245 and 246 to 276 DIM). The  $\sigma_e^2$  of

$$RRM-HET \text{ was calculated as } \sigma_e^2 = \frac{1}{n} \left( \sum_{i=1}^n \sigma_e^2 \right) \text{ where}$$

$n$  is number of residual variance class. The average lactation curve was modeled using 5<sup>th</sup> order LP. The fixed effect classes used in the model were similar to those reported by Prakash *et al.* (2015). Most suitable model was identified on the basis of residual variance and Schwarz's Bayesian information criteria (BIC) value. The analysis was performed using a single-trait random regression model by WOMBAT software (Meyer 2007).

The random regression model used in the analysis was:

$$y = Xb + Za + Wp + e$$

Where,  $y$ , vector of test-day milk yields of Sahiwal cows;  $b$ , vector of fixed effects (season of test-day recording month, year of calving, age at calving and days open class);  $X$ , incidence matrix relating test-day milk yields to fixed effects;  $p$ , vector of permanent environmental random regression coefficients (RRC);  $a$ , vector of additive genetic RRC;  $Z$  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,  $G$ , variance-covariance matrix of additive genetic RRC;  $A$ , additive genetic relationship matrix among the animals,  $\otimes$ , Kronecker product function;  $P$ , variance-covariance matrix of permanent environmental RRC;  $I$ , Identity matrix and  $R$  is diagonal matrix of residual variances.

*Estimation of 305-day milk yield variance from random regression test-day estimates:* As the test-day records were taken up to 276<sup>th</sup> day of lactation, so in order to calculate the 305-day milk yield variance, the covariance function were transformed to actual days in milk (six to 305 day) as suggested by Kirkpatrick *et al.* (1990).

Additive genetic variance of 305-day milk yield =

$$\sigma_{\alpha 305MY}^2 = Z'_{c305MY} G Z'_{c305MY}$$

Permanent environmental variance of 305-day milk yield

$$= \sigma_{pe305MY}^2 = Z_{c305MY} P Z'_{c305MY}$$

Where,  $Z_{c305MY}$ , vector of the summations of LP corresponding to 6<sup>th</sup> to 305<sup>th</sup> day lactation milk production; P, permanent environmental RRC matrix and G is additive genetic RRC matrix.

*Error variance ( $\sigma_e^2$ ) of 305-day milk yield:* Random regression model with homogenous residual variance

$$= \sigma_e^2 = \sigma_e^2 * (300)^2$$

Random regression model with heterogeneous residual variance =  $\sigma_e^2 = \frac{1}{n} \left( \sum_{i=1}^n \sigma_{e_i}^2 \right) * (300)^2$  where n is residual variance class.

*Estimation of daily breeding value:* EBV for an animal (j) in any given days in milk (t) was calculated as

$$EBV_{jt} = C_i \hat{a}_j$$

Where,  $\hat{a}_j$ , solutions for additive genetic RRC of animal j and  $C_i$  is vector of covariates of legendre polynomials on day t of lactation.

*Prediction of 305-day milk yield breeding value from daily breeding value:*

$$EBV_{305} = \sum_{i=6}^{305} EBV_{jt}$$

where,  $EBV_{jt}$ , expected breeding value of j<sup>th</sup> animal in any day in milk t.

*Spearman's rank correlation:* The Spearman's rank correlations between breeding values of sires and cows derived from various methods was used to judge the effectiveness of different methods. The rank correlation was estimated as per Spearman (1904). The significance of rank correlation was tested by t-test with (n-2) degrees of freedom

$$\text{as given } t = r \sqrt{\frac{n-2}{1-r^2}}$$

## RESULTS AND DISCUSSION

*Model selection:* Model with lower residual variance and lower BIC value was preferred. The model with a third order LP to model additive genetic effect and sixth order LP to

model permanent environmental effect was found best for all lactation from RRM-HOM. However, from RRM-HET, for first lactation, model with fourth order LP to model additive genetic effect and fifth order LP to model permanent environmental effect were found best. For second and third lactation, model with third order polynomial to model additive genetic effect and fifth order polynomial to model permanent effect were found best (Table 2).

*Variance component and heritability of 305-day milk yield:* Details of various estimates from different methods, and its comparison are presented in Table 3. From RRM, it was possible to account for permanent environmental variance arising due to individual milk yield variations during lactation. However, analysis using standard univariate animal model, the environmental variances were not split in permanent and temporary. Therefore, to compare estimates obtained from the three methods, the environmental temporary and permanent variances obtained from RRM were summed. In general, genetic variances and variances of the permanent environmental effect went up substantially between the first and second lactation. The differences between the permanent environmental variances in the second and third lactations were small.

Perusal of Table 3 shows that 305-day milk yield additive genetic variance estimated from RRM was higher compared to univariate animal model for all lactation. Among the random regression model, RRM-HET gave higher additive genetic variance estimates. Contrary to results of present study, others have found higher estimates of additive genetic variance for first lactation 305-day milk yield (Sawalha *et al.* 2005, Kaygisiz 2013).

The heritability estimates were low especially for first and third lactation 305-day milk yield from all three methods. For second lactation, low heritability estimate from univariate animal model and moderate estimate from different random regression models were obtained. However, higher heritability estimates were obtained by Zavadilova *et al.* (2005), Kumar *et al.* (2009), Hossein-Zadeh (2012), Gupta (2013) for first lactation 305-day milk yield. The estimate obtained in the present study for second lactation 305-day milk yield was similar to estimate of Hossein-Zadeh (2012) but lower than the estimates reported by Zavadilova *et al.* (2005), Gupta (2013). For third lactation, 305-day milk yield, higher estimates were reported by Zavadilova *et al.* (2005), Hossein-Zadeh (2012),

Table 2. Description of best PZ random regression test-day models fitted for different lactation in Sahiwal cattle

Order of fit (k)	Number of parameters	First lactation	Second lactation	Third lactation				
$k_a$	$(k_p)$	$(N_p)$	BIC value	$\sigma^2_e$	BIC value	$\sigma^2_e$	BIC value	$\sigma^2_e$
<i>Random regression model with homogeneous residual variance (RRM-HOM)</i>								
3	6	28	10657.92	0.89	8648.77	1.31	6292.54	1.20
<i>Random regression model with heterogeneous residual variance (RRM-HET)</i>								
3	5	30	-	-	8597.90	1.67	6255.88	1.67
4	5	34	10635.24	1.04	-	-	-	-

$\sigma_e^2$ , error variance;  $k_a$ , order of fit for additive genetic effect;  $k_p$ , order of fit for permanent environmental effect.

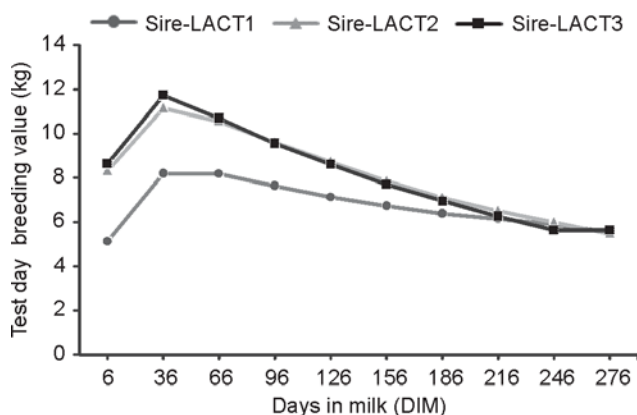


Fig. 1. Trend of sire test-day breeding value estimated from best RRM-HOM for different lactation in Sahiwal cattle. LACT1, first lactation; LACT2, second lactation; LACT3, third lactation; BV, breeding value; RRM-HOM, random regression model with homogeneous residual variance.

Gupta (2013). Comparatively lower estimate of heritability may be attributed to data structure and limited variability present in the Sahiwal herd at NDRI, Karnal. Over the years, very limited introduction of variability has been made in the herd by purchase of Sahiwal cows from breeding tract. Differences between the estimates from the present study and other studies may also be due to including different functions related to the changes in lactation curve in the model, sire model vs. animal model and different assumptions used (heterogeneous residual variance vs. constant residual variance across lactations).

**Breeding value and rank correlations:** The trend of sire breeding value of test-day milk yields along different period of lactation is given in Fig. 1. Similar trends were seen for cow breeding value. The average test-day breeding value of sires ranged from 5.12 kg to 8.20 kg for first lactation, from 5.46 kg to 11.15 kg for second lactation, and from 5.63 kg to 11.72 kg for third lactation by RRM-HOM model. Similar estimates were obtained from RRM-HET model. The average test-day milk yield breeding value for cow ranged from 5.10 kg to 8.16 kg for first lactation, from 5.44 kg to 11.10 kg for second lactation, and from 5.63 kg to 11.72 kg for third lactation by RRM-HOM model. Similar estimates were obtained from RRM-HET model. Similar test-day breeding value pattern was obtained by Singh (2014) in Karan-Fries cattle.

The rank correlation among the cows as well as sire 305-day milk yields breeding values, from RRM and lactation model were moderate to high. For all lactations, rank correlation was more than 0.78 (Table 3). It was found that among the top ten sires, six sire for first lactation, seven sire for second lactation and eight sire for third lactation were common in all three lists (Table 4). From all three methods, same bull was ranked on the top. Sawalha *et al.* (2005) also found that predicted breeding values with 305-day lactation records were moderately correlated with those with test-day models, and the correlation ranged from 0.71 to 0.87 for sires and 0.80 to 0.87 for cows. Zavadilova *et*

Table 3. Comparison of 305-day milk yield variance, heritability and rank correlation of breeding values estimated with different models in Sahiwal cattle

Component	Lactation model	RRM-HOM	RRM-HET
<i>First lactation 305-day milk yield</i>			
Additive genetic variance (kg <sup>2</sup> )	25635.80	28447.80	31752.80
PE variance (kg <sup>2</sup> )	-	278797.70	274126.60
Residual variance (kg <sup>2</sup> )	330732.00	79790.40	93371.00
Phenotypic variance (kg <sup>2</sup> )	356367.80	387035.90	399250.40
Heritability (V <sub>a</sub> /V <sub>p</sub> )	0.072	0.073	0.079
Spearman's rank Correlation			
Sire breeding value	1	0.82	0.85
Cow breeding value	1	0.80	0.81
<i>Second lactation 305-day milk yield</i>			
Additive genetic variance (kg <sup>2</sup> )	64420.60	128538.80	128911.50
PE variance (kg <sup>2</sup> )	-	340865.20	341357.50
Residual variance (kg <sup>2</sup> )	383987.00	118089.00	155070.00
Phenotypic variance (kg <sup>2</sup> )	448407.60	587493.00	625339.10
Heritability (V <sub>a</sub> /V <sub>p</sub> )	0.144	0.219	0.206
Spearman's rank Correlation			
Sire breeding value	1	0.78	0.79
Cow breeding value	1	0.79	0.79
<i>Third lactation 305-day milk yield</i>			
Additive genetic variance (kg <sup>2</sup> )	51216.6	52783.19	51358.09
PE variance (kg <sup>2</sup> )	-	383089.30	387344.40
Residual variance (kg <sup>2</sup> )	383142.00	108027.00	150200.00
Phenotypic variance (kg <sup>2</sup> )	434358.60	543899.50	588902.50
Heritability (V <sub>a</sub> /V <sub>p</sub> )	0.112	0.097	0.087
Spearman's rank Correlation			
Sire breeding value	1	0.90	0.89
Cow breeding value	1	0.89	0.89

RRM-HOM, random regression model with homogeneous residual variance; RRM-HET, random regression model with heterogeneous residual variance; V<sub>a</sub>, additive genetic variance; V<sub>p</sub>, phenotypic variance; PE, permanent environmental variance

*al.* (2005) observed high rank correlation of 0.98 between test-day model and lactation model in Holstein cows. Cilek *et al.* (2008) found correlation between breeding values estimated using test-day milk yields and breeding values of 305-day milk yield as 0.876 in Simmental cattle.

It can be concluded that 305-day milk yield estimates from random regression model gave higher additive genetic variance, and moderate to high rank correlation for expected breeding value. It was also observed that random regression can account for lactation curve variation of the animals, and major part of the residual variance is partitioned into permanent environmental effects. Thus, random regression model can replace the conventional lactation model for genetic evaluation of Sahiwal cattle.

Table 4. Estimated breeding value (EBV) of 305-day milk yield (kg) as deviation from population mean and rank of top ten Sahiwal sires using different models

Method	RRM-HET	RRM-HOM	Lactation Model			
<i>First lactation</i>						
Rank	Sire ID	EBV	Sire ID	EBV	Sire ID	EBV
1	526*	336.659	526*	301.124	526*	238.261
2	1404*	276.514	1404*	240.770	1404*	165.073
3	1090*	251.442	1090*	212.914	1090*	159.877
4	1421*	222.792	1558	187.577	310*	152.392
5	1572*	192.082	1421*	177.661	1376	139.472
6	310*	189.790	310*	176.921	1424	127.039
7	1558	188.724	1572*	167.011	1421*	124.384
8	1563	171.986	1563	166.327	1069	124.368
9	1337	160.079	446	149.674	1564	122.579
10	1564	156.580	453	144.199	1572*	115.31
<i>Second lactation</i>						
1	1424*	798.614	1424*	798.131	1424*	325.654
2	1404*	613.568	1404*	573.492	530*	286.097
3	1376*	561.973	1376*	553.186	1220*	221.512
4	1510*	533.665	1069*	536.332	1376*	214.85
5	1069*	498.711	1510*	527.573	1069*	191.341
6	530*	493.952	530*	495.200	1510*	158.286
7	1502	443.460	1502	422.403	327	136.293
8	1220*	382.077	1220*	386.142	619	135.723
9	1301	376.690	1090	344.607	1404*	134.757
10	1572	364.926	1301	343.848	1060	133.593
<i>Third lactation</i>						
1	252*	214.651	252*	219.804	252*	140.163
2	1572*	172.451	1572*	179.843	1572*	107.168
3	1424	163.212	1424	147.230	327*	98.601
4	1565*	120.425	327*	130.902	494	81.133
5	327*	118.605	1565*	128.319	MISTRI	76.076
6	1220*	108.392	1220*	108.959	1220*	68.230
7	1337*	106.346	1033*	104.982	1565*	62.184
8	530*	92.096	1337*	99.879	530*	54.549
9	1033*	86.052	530*	93.965	1337*	46.010
10	506	83.851	MISTRI	88.349	1033*	40.988

\*Denotes common sire; RRM-HOM, random regression model with homogeneous residual variance; RRM-HET, random regression model with heterogeneous residual variance

#### ACKNOWLEDGEMENT

The authors are thankful to Director, ICAR-NDRI, Karnal for providing necessary facilities for research work and to ICAR for granting financial assistance in form of ICAR-SRF (PGS) for PhD studies. First authors also acknowledges Director, ICAR-CSWRI, Avikanagar for granting study leave for PhD Study.

#### REFERENCES

- Cilek S, Orhan H, Kaygisiz A and Sahin E H. 2008. Estimation of breeding values of Anatolian population of Simmental cows using monthly test-day milk yields. *Archive Zootechnica* **11**: 79–85.
- Gupta A. 2013. 'Genetic evaluation of Sahiwal cattle using lactation curve models'. Ph.D Thesis, National Dairy Research Institute, Karnal, Haryana, India.
- Hossein-Zadeh N G. 2012. Estimation of genetic parameters and trends for energy-corrected 305-d milk yield in Iranian Holsteins. *ArchivTierzucht* **55**: 420–26.
- Jaffrézic F, White I M S, Thompon R and Hill W G. 2000. A link function approach to model heterogeneity of residual variances over time in lactation curve analyses. *Journal of Dairy Science* **83**: 1089–93.
- Kaygisiz A. 2013. Estimation of genetic parameters and breeding values for dairy cattle using test-day milk yield records. *Journal of Animal and Plant Science* **23**: 345–49.
- Kirkpatrick M, Lofsvold D and Bulmer M. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* **124**: 979–93.
- Kumar A, Gandhi R S and Paul A K. 2009. Genetic evaluation of lifetime traits in Sahiwal cattle. *Indian Journal of Dairy Science* **62**:374–76.
- Meyer K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University Science B* **8**: 815–21.
- Prakash V, Gupta A K, Gupta A, Gandhi R S and Kumar A. 2015. Milk yield variation in first three lactations and factors affecting milk yield in Sahiwal cattle. *Indian Journal of Animal Sciences* **85**: 1267–69.
- Prakash V, Gupta A K, Gupta A, Gandhi R S, Singh A and Chakravarty A K. 2016. Random regression model with heterogeneous residual variance reduces polynomial order fitted for permanent environmental effect but does not affect genetic parameters for milk production in Sahiwal cattle. *Animal Production Science* <http://dx.doi.org/10.1071/AN15347>.
- Sawalha R M, Keown J F, Kachman S D and Van Vleck L D. 2005. Genetic evaluation of dairy cattle with test-day models with autoregressive covariance structures and with a 305-d model. *Journal of Dairy Science* **88**: 3346–53.
- Schaeffer L R. 2004. Application of random regression models in animal breeding. *Livestock Science* **86**: 35–45.
- Singh A. 2014. 'Genetic evaluation of Karan-Fries cattle using test-day milk yields by random regression models'. M.V.Sc. Thesis, National Dairy Research Institute, Karnal, Haryana, India.
- Spearman C. 1904. The proof and measurement of association between two things. *American Journal of Psychology* **15**: 72–101.
- Zavdilova L, Jamrozik J and Schaeffer L R. 2005. Genetic parameters for test-day model with random regressions for production traits of Czech Holstein cattle. *Czech Journal of Animal Science* **50**: 142–54.