Estimation of Breeding Value of Najdi Cattle
Breed Using Random Regression Model

M. Nazari, M.T. Beigi Nassiri, J. Fayazi and S. Tabatabaei
Department of Animal Science, Faculty of Animal and Food Science,
University of Ramin Agricultural and Natural Resources, Khozestan, Mollasani, Iran

Abstract: In order to estimate variance components 8843 test day milk and fat yields records from first lactation Najdi cattle were used. Data obtained between years of 1989-2005 by animal breeding station of Najdi cattle located in Shooshtar city. Random regression models were studied in respect of different orders of fitting for fixed and random regression. Also, different methods of residual variance in the statistical model assume that included assumption of constant residual variance and different assumption about variable residual variance during lactation. According to the obtained results, the assumption of heterogeneous residual variance during lactation improved characteristics of statistical model. A suitable model with (3, 3) orders of fit for additive genetic and permanent environmental covariance functions to analyze TD fat and milk yield records were recognized. Minimum of heritability for fat and milk yield was estimated at the beginning of lactation (0.1, 0.15), respectively. The amount of this parameter increased to mid lactation and almost in the 5th month of lactation reached to maximum level (0.34 for milk yield and 0.44 for fat yield), then decreased to the end of lactation.

Key words: Random regression, milk and fat yields, Najdi cattle, residual variance, Iran

INTRODUCTION

Traditionally, aggregated 305 days yields have been used as the breeding goal for production in dairy cattle. The 305 days yields have been estimated from individual test-day records according to relatively simple rules given by Sargent *et al.* (1968).

That the 305 days yields are predicted from few observations may give rise to bias (Anderson *et al.*, 1989). Furthermore, short lactations on culled cows or records in progress must be extended, which also may lead to bias. Extension procedures used in the Danish evaluation system have been described already and later updated by Aamanted *et al.* (1999). Biases in the extended records may be due to individual variation in the shape of the lactation curve (Shanks *et al.*, 1981; Pander and Hill, 1993).

Furthermore, accumulated yields may be influenced by different temporal environmental factors that are difficult to account for in the model for genetic evaluation. Use of a test-day model may alleviate some of the problems seen in models for 305 days lactation yields. In a test-day model, extension procedures are not needed and temporal environmental effects of individual test days can be taken into account (Meyer *et al.*, 1990; Ptak and Schaeffer, 1993; Van Raden, 1997). Several models have

been used for estimation of variance components for test-day yields. A repeatability model has been applied. This model assumes that the sequence of measurements of an individual is repeated measurements of the same trait.

A complete multitrait model has been applied as well (Meyer *et al.*, 1990). In this model, every test day is considered as a separate trait. However, a limiting factor for this kind of analysis is the number of traits allowed simultaneously and the large number of parameters to be estimated. In random regression models, curves are estimated for random effects.

In the earliest random regression test-day models, only the additive genetic effect was modeled with random regressions and the permanent environmental effect was modeled as in a repeatability model. However, this model gave unrealistic genetic variances especially in the ends of the lactation.

The purpose of the research reported in this study was to estimate (co) variance components for production traits in first parity Najdi cattle, using random regression model. This study aims at evaluating alternative RRMs using a broad range of model comparison criteria, including also models that allow different orders of fit for the Additive Genetic (AG) and Permanent Environmental (PE) components for first lactation data.

Corresponding Author: M. Nazari, Department of Animal Science, Faculty of Animal and Food Science,
University of Ramin Agricultural and Natural Resources, Khozestan, Mollasani, Iran

MATERIALS AND METHODS

The performance of different RRMs was evaluated using independent data sets composed of first lactation Test Day milk record (TD) from Najdi cows calving from 1989-2005 in Iran. Data editing for the intervals between TD records in successive controls was based on the Iranian official milk recording system. For this study, lactation period was divided according to DIM at test to give 10 traits. Traits indicate TD milk production at a certain stage of lactation and are therefore referred to as Interval Milk Production (IMP). The (co) variance components for all models were estimated with the restricted maximum likelihood method, employing an accelerated expectation maximization algorithm, using the REML3.1 program (Meyer, 1995). For this study, Ten alternative RRMs were analyzed. All models included in their equation the same systematic effects and temporary measurements errors were taken to be independently distributed and with constant variance along Days in Milk (DIM). The equation for all the models analyzed can be written in scalar notation as:

$$y_{ijklmpq} = M_i + N_j + TD_k + S_l + Y_m + \sum_{n=1}^{5} b_n (age_{ijklmpq})^n$$
$$+ \sum_{n=0}^{k} \beta_n \varphi_n (dim^*_{ijklmpq}) + \sum_{n=0}^{k_a-1} \alpha_{pn} \varphi_n (dim^*_{ijklmpq})$$
$$+ \sum_{n=0}^{k_p-1} \gamma_{pn} \varphi_n (dim^*_{ijklmpq}) + e_{ijklmpq}$$

Where:

$Y_{ijklmpq}$ = The q-th TD of animal p, M_i is the fixed effect of calving month (I = 1, 2,..., 12)

N_j = Fixed effect of kind of milking (j = 1, 2, 3)

Td_k = Fixed effect of Test-day (k = 1, 2,..., 60)

S_l = Fixed effect of calving season (l = 1, 2, 3, 4)

Y_m = Fixed effect of calving year (m = 1, 2,..., 16)

b_n = The n-th fixed regression coefficient, nested within the j-th class of the age-season of calving month ($Xt_m = 1$, $X_{2m} = DIM_m/c$, $X_{3m} = (DIM_m/c)^2$, $X_{4m} = In(c/DIM_m)$, $X_{5m} = In(c/DIM_m)^2$, c = 305)

$X_m(t)$ = The m-th covariate from the Ali-Schaeffer function evaluated at DIM t

β_n = The n-th fixed regression coefficient

The terms α_{pn} and γ_{pn} are the n-th Random Regression Coefficient (RRC) associated with the AG and PE effects of animal p, respectively and φ_n are the n-th legendra polynomials depending on DIM t, based on the functions fitted to the AG and PE effects, respectively.

Finally, e_{ijkl} is the temporary measurement error. Defining a as the vector that contains the r RRC for the AG effects for all animals in the pedigree, p as the vector, which includes the s RRC fitted to the PE effects for all animals with records and e as the vector containing the temporary measurement errors, the following (co) variance structure was assumed:

$$var \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & R \end{bmatrix}$$

Where:

$$G = \sum_a \otimes / A \text{ and } P = \sum_p \otimes I, \sum_a \text{ and } \sum_p$$

being matrices containing the (co) variances among the r and s RRC fitted over the AG and PE effects, respectively. R is a diagonal matrix containing the constant variance of the temporary measurements errors or Residual Variance (RV). Several alternative sub models were considered to define the competing RRM analyzed in this study. Firstly, lactation models fitting to both AG and PE effect were studied. In addition, 10 RRMs fitting orthogonal Legendre polynomials of different orders (Kirkpatrick *et al.*, 1990) (RRL) to both AG effects were considered. We will refer to order of fit either for the AG or for the PE effects, as the number of RRC used to describe each effect. The notation that will subsequently be used to refer to RRM based on Legendre polynomials is L (i, j), where i (i = 3, 3) indicates the order of fit for the sub model fitted to the AG effects and j = (j = 3, 3) is the order of fit for the sub model fitted to the PE effects. The (co) variance components for all models were estimated with the restricted maximum likelihood method, employing an accelerated expectation maximization algorithm, using the REML3/1 program (Meyer, 1995). Estimates of (Σ_a) the (co) variance matrix among the RRC for the AG effects in all models were used to obtain the daily AG (co) variances along the lactation from the covariance function

$$g(t_i, t_j) = z(t_i)' \sum_a z(t_j)$$

Where:

$Z(t_i), Z(t_j)$ = The vectors of covariates evaluated at times

t_i and t_j = The sub model fitted to the AG effect

The same procedure was applied to the PE effect. Several different functions have been applied to fixed regression models as well as random regression models. To be a suitable function in random regressions, it must be linear in the parameters and have as few parameters as possible.

RESULTS AND DISCUSSION

Estimated daily additive variance for data sets, in Table 1 for lactation models, corresponding RRL models,

Table 1: Shows order of fit for fix regression and random regression for milk and fat yield

Traits	Model (M)	Order of fit for fix regression		Order of fit for random regression	Order of error variance (n)	No. of parameter	Log of likelihood function	Error variance
		AG	PE					
	1	3	3	3	1	13	-2445	4.238
	2	4	3	3	1	13	-2285	3.03
Milk	3	5	3	3	1	13	-2234	2.95
	4	5	3	3	10	22	-1448	-
	5	5	3	4	1	17	-2225	2.05
	6	5	4	4	1	21	-2230	2.01
Fat	1	5	3	3	10	22	-15500	-

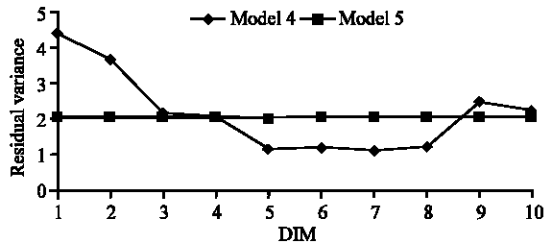PE = Permanent Environment; AG = Additive Genetic

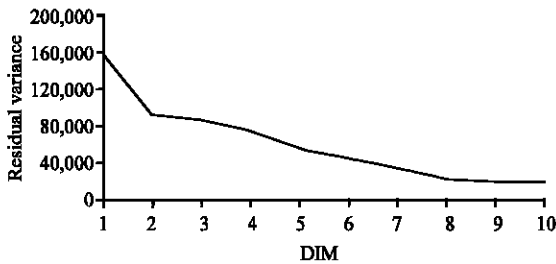Fig. 1: Residual variance of model 4 and 5 for milk traits

Fig. 2: Residual variance of L (3, 3) for fat traits

L (3, 3) and L (3, 4), plus model L (4, 4). Random regression models were studied in the respect of different orders of fitting for fixed and random regression and also different methods of assuming residual variance in the analysis include assumption of constant residual variance and different assumption about variable residual variance during lactation. According to the obtained results, the assumption of heterogeneous residual variance during lactation improved characteristics of statistical model (M4). Maximum residual variance for milk and fat yield traits was estimated in the early lactation (Fig. 1 and 2). With assumption of constant residual variance during lactation cause to estimation lower than actual value for residual variance at the beginning and the end of lactation and cause to estimation higher than actual value for residual variance at mid lactation (Fig. 1). Observed residual variance for fat yield were similar to milk yield. When the order of fit fixed regression was high, this caused to increasing log-likelihood (M1, M2 and M3). Additionally, when the order of fit PE effect was high, increasing the additive order of fit hardly decrease

Table 2: Shows estimates of variance components for daily milk yields on selected test-days with random regression model for model 4

No.	σ_A^2		σ_C^2		σ_P^2	
	M4	M5	M4	M5	M4	M5
1	4.89	2.35	22.50	21.05	31.8	25.45
2	7.60	2.82	21.46	4.30	32.74	9.17
3	9.93	3.74	20.48	4.41	32.58	10.20
4	10.06	4.10	20.49	6.15	32.6	12.30
5	10.90	4.20	20.49	4.75	32.52	11.00
6	10.20	4.50	20.52	8.35	32.14	14.90
7	9.81	5.90	20.67	6.56	31.62	14.51
8	10.27	5.36	20.06	8.39	31.56	15.80
9	8.32	8.20	21.88	15.70	32.69	23.90
10	8.76	8.70	23.36	14.95	34.35	25.70

Phenotype variance, i.e., sum of σ_A^2, σ_C^2 and σ_e^2

residues variance (Table 1) (Lopez-Romero and Carabano, 2003). A suitable model 4 with (L (3, 3) orders of fit for additive genetic and permanent environmental covariance functions to analyze TD fat and milk yield records were recognized (Lopez-Romero and Carabano, 2003). In this research, mode 4 and model 5 were compared. The values of log likelihood on convergence were -1448, -2225 and -2230 for models M4, M5 and M6, respectively. The likelihood ratio test of M4 over M5 and M6 gives 9 degrees of freedom. The model 4 fits clearly better to the data than the simpler models of 5 and 6. Estimation of variance components with the RR model predicts variance covariance structure of RR coefficients. Minimum and maximum of additive genetic variance for both traits were estimated at beginning and 5th month of lactation (mid lactation), respectively. Phenotypic variance was not constant during lactation and its value is higher at the beginning and the end of lactation (Table 2).

For milk models, values from M5 were higher (0.09-0.33) than those from M4 (0.15-0.26). Heritability was estimated to be lowest at the beginning and end of lactation. Jamrozik *et al.* (1997) also reported that estimates of heritability decreased, when RR function for PE effects was included in the model. Minimum of heritability for fat and milk yield was estimated at the beginning (0.1, 0.15) of lactation, respectively. The amount of this parameter increased to mid lactation and almost in the 5th month of lactation reached to maximum level (0.34 for milk yield and 0.44 for fat yield), then decreased to the end of lactation (Table 3).

Table 3: Estimates of heritability for daily milk yields on selected test-days with orthogonal polynomial random regression model for M4 and M5

Days	Milk		Fat
	M4	M5	
1	0.15	0.09	0.10
2	0.23	0.30	0.13
3	0.30	0.36	0.21
4	0.31	0.33	0.29
5	0.34	0.38	0.36
6	0.32	0.30	0.33
7	0.31	0.40	0.41
8	0.33	0.33	0.44
9	0.25	0.34	0.37
10	0.26	0.33	0.22

RRM assumes heterogeneous additive genetic effects throughout the course of lactation. This is done by regression of the additive genetic effect on individual DIM via a lactation curve function. Thus, a 305 days lactation estimate of a cow's breeding value (u_{RRM}) corresponds to the area under this curve:

$$u_{RRM} = \sum_{i=1}^{305} \left(\sum_{j=1}^{np} \alpha_j x_{ij} \right)$$

RRM allows for differences between cows in the shape and level of the distribution of the additive genetic effect throughout lactation. RRM requires more test day records to give an accurate estimate of the breeding value as it fits two more parameters in lactation curve function.

CONCLUSION

Use of RRM to analyze longitudinal traits makes it possible to study changes over time and in a dairy cattle context, a better understanding of the genetics of lactation. A suitable model with (3, 3) orders of fit for additive genetic and permanent environmental covariance functions to analyze TD fat and milk yield records were recognized. The assumption of heterogeneous residual variance during lactation improved characteristics of statistical model. With assumption of constant residual variance during lactation cause to estimation lower than actual value for residual variance at the beginning and the end of lactation and cause to estimation higher than actual value for residual variance at mid lactation. The highest value of heritability, additive genetic and permanent environmental variance for to use records of mid lactation for individual evaluation. These records have more heritability and lower variance, therefore cause better estimate for breeding fat and milk yield were estimated at the mid lactation. According to these results, it's better value and its accuracy. Then, when computationally feasible, RRM is recommended for the routine genetic evaluation as more flexible.

REFERENCES

Aamanted, G.P., J. Pedersen, U.S. Nielsen, J. Jensen, P. Madsen, J.R. Thomasen and L.G. Christensen, 1999. Animal Model for Ydelse. Danish Agricultural Advisory Centre, Skejby, Denmark, pp: 71.

Anderson, S.M., I.L. Mao and J.L. Gill, 1989. Effect of frequency and spacing of sampling on accuracy and precision of estimating total lactation milk yield and characteristics of the lactation curve. J. Dairy Sci., 72: 2387-2394.

Jamrozik, J., L.R. Schaeffer, Z. Liu and G. Jansen, 1997. Multiple trait random regression test day model for production traits. Proceeding of the 16th Interbull Meeting, Aug. 28-29, Vienna, Austria, pp: 43-47.

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

Lopez-Romero, P. and M.J. Carabano, 2003. Comparing alternative random regression models to analyses 1st-lactation daily milk yield data in Holstein-friesian cattle. Livest. Prod. Sci., 82: 81-96.

Meyer, K., 1995. Estimation of genetic parameter for mature weight of Australian beef cattle and its relationship to early growth and skeletal measures. Livest. Prod. Sci., 44: 125-125.

Meyer, K., H.U. Graser and K. Hammond, 1990. Estimates of genetic parameters for first lactation test day production of Australien black and white cows. Livest. Prod. Sci., 21: 177-199.

Pander, B.L. and W.G. Hill, 1993. Genetic evaluation of lactation yield from test day records on incomplete lactation. Livest. Prod. Sci., 37: 23-36.

Ptak, E. and L.R. Schaeffer, 1993. Use of test day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci., 34: 23-34.

Sargent, F.D., V.H. Lytton and O.G. Wall, 1968. Test interval method of calculating dairy herd improvement association records. J. Dairy Sci., 51: 170-179.

Shanks, R.D., P.J. Berger, A.E. Freeman and F.N. Dickinson, 1981. Genetic aspects of lactation curves. J. Dairy Sci., 64: 1852-1860.

Van Raden, P.M., 1997. Lactation yields and accuracies computed from test day yields and (co)variances by best prediction. J. Dairy Sci., 80: 30-30.