Estimation of covariance functions for growth trait from birth to 180 days of age in Iranian Baluchi sheep

Mohammad Safaei [1], Ali-Reza Fallah-Khair [1], Reza Seyed-Sharifi [2]*, Hadi Haghbin Nazarpak [3]
Abdollah Sobhabi [2] and Taher Yalchi [2]

[1] *University of Zanjan, Zanjan, Iran.* [2] *Feculty of Agriculture, University of Mohaghegh Ardabili, 56199-11367Ardabil, Iran.*
[3] *Feculty of Veterinary Medicine, Islamic Azad university, Garmsar Branch, Garmsar, Iran.*
*e-mail: reza_seyedsharifi@yahoo.com

Abstract

Weights of Iranian Baluchi sheep, from birth to 180 d of age, recorded every month, were analyzed using Random Regression Models. Independent variables were Legendre polynomials of age at recording. Up to four sets of random regression coefficients were fitted for animals' direct and maternal additive genetic and permanent environmental effects. Changes in measurements of error variances by age were modeled through a variance function. Orders of polynomial fit from three to five were considered and resulted in up to 36 parameters to be estimated. Direct heritability estimates increased after birth and tended to be the highest at ages at which maternal effect estimations tended to be the lowest. Maternal heritability estimations decreased fast after birth to about 30 of age and decreased slowly after that. Estimations of direct and maternal additive genetic correlation between weights of birth with other ages were moderate and between other ages were high.

Key words: Baluchi sheep, covariance functions, growth, heritability.

Introduction

Recently, covariance functions and random regression models (RRM) have been proposed as an alternative in modeling traits that are recorded repeatedly during the animal's life (i.e., longitudinal data [5, 14]. In particular, random regression models accommodate 'repeated' records for traits with changes, gradually and continually, over time, and do not require stringent assumptions about constancy of variances and correlations [2-4]. Meyer and Hill [11] and Meyer [7] showed that random regression models are a special case of covariance functions, and covariance functions coefficients can be estimated directly from random regression models by restricted maximum likelihood [7, 8, 10].

R RM [a] has been used for some times in the analysis of human growth curve data. Laird and Ware [19] outlined a random regression mixed model for longitudinal data which comprised both growth curve models and repeatability models as special cases. Andersen and Pedersen [18] showed how RR can be employed in modeling growth curves of pigs on a phenotypic level. RRM are expected to give more accurate estimations of genetic parameters and predictions of breeding values than the conventional multi trait model considering a limited number of points, such as birth, weaning, yearling and final weights, used nowadays [9, 11, 15].

Most countries are now in the process of introducing RRM for evaluating dairy cattle. RR models have also been used to describe food intake and weight gain in pigs [18] and growth and mature weight of beef cows [8, 9]; but the benefit of the methodology in the analysis of weight data in sheep are as yet unknown. This paper describes the application of RR techniques to live weight recorded on a population of Baluchi sheep. Also the objective of this study was to estimate genetic direct and maternal genetic covariance functions using a RR model for weights from birth to 180 d of age in Baluchi sheep.

Material and Methods

Data: Data consisted of weights of sheep collected as part of a selection experiment at the Baluchi Sheep Research Station (BSHRS), located in Iran. Since 1980, BSHRS has been running a genetic improvement program considering growth, fertility and maternal ability traits. After basic edits, there were 19,821 sheep weights recorded between 1997 and 2007 (Table 1).

Table 1. Summary of data structure.

Category	Value
No. of records	19809
No. of animal	3952
No. of dams with progeny	1182
No. of sires with progeny	142
Minimum no. of records per animal	2159
Maximum no. of records per animal	3240
Mean live weight overall (kg)	19.35 (3.98)
Mean live weight at weaning (60 days; kg)	14.72 (3.72)
Mean live weight at end of test (180 days; kg)	30.04 (5.53)

Standard deviation showed in parentheses.

---

[a] Random Regression Model.

Random regression analysis: Covariance between random regression coefficients was estimated by REML using DFREML algorithm and program of DXMRR [7]. Legendre polynomials of age at recording were used as independent variables. Analyses yielded estimations of covariance among random regression coefficients and estimations of variances due to temporary environmental variances, so-called measurement error variances. The general model can be represented as follows:

$$Y_{ijk\ln p} = \mu + YR_i + SX_j + BT_k + AG_l + SA_n + \sum_{m=0}^{3} \beta_{pm} \varphi_m(\alpha^*_{ijk\ln p}) + \sum_{m=0}^{k_A-1} \alpha_{pm} \varphi_m(\alpha^*_{ijk\ln p}) +$$

$$\sum_{m=0}^{k_M-1} \gamma_{pm} \varphi_m(\alpha^*_{ijk\ln p}) + \sum_{m=0}^{k_C-1} \delta_{pm} \varphi_m(\alpha^*_{ijk\ln p}) + \sum_{m=0}^{k_Q-1} \rho_{pm} \varphi_m(\alpha^*_{ijk\ln p}) + \varepsilon_{ijk\ln p}$$

where $Y_{ijk\ln p}$ = any of weight records; μ = total average; YR_i = fixed effects of i th birth year (i = 1, 2, 3, ..., 11); SX_j = jth fixed effects of birth sex (j = 1 and 2); BT_k = fixed effects of kth birth procedure (k = 1 and 2); AG_l = fixed effects of lth ewe age (l = 1, 2, ..., 8); SA_n = fixed effects of nth birth season (n = 1, 2, ..., 4); $\alpha^*_{ijk\ln p}$ = standardized age in range of +1 to -1; $\varphi_m(\alpha^*_{ijk\ln p})$ = the mth Legendre polynomial of age; β_m = m^{th} fixed regression coefficients; α_{pm}, γ_{pm}, δ_{pm} and ρ_{pm} are m^{th} random regression coefficients for direct additive genetic, maternal additive genetic, animal and maternal permanent environmental effects, respectively, related to p^{th} animal. k_A, k_M, k_C and k_Q denote the corresponding orders of fit; and $\varepsilon_{ijk\ln p}$ = the temporary environmental effect related to $Y_{ijk\ln p}$.

Above statistical model matrix form was as follows:

$$Y = XB + Z_1\alpha + Z_2\gamma + W_1\delta + W_2\rho + \varepsilon$$

$$v\begin{bmatrix} \alpha \\ \gamma \\ \delta \\ \rho \\ \varepsilon \end{bmatrix} = \begin{bmatrix} k_A \otimes A & 0 & 0 & 0 & 0 \\ 0 & k_M \otimes A & 0 & 0 & 0 \\ 0 & 0 & k_C \otimes I & 0 & 0 \\ 0 & 0 & 0 & k_Q \otimes I & 0 \\ 0 & 0 & 0 & 0 & R \end{bmatrix}$$

Y= observations vector; B = the vector of fixed effects and β_m; and α = the vector of additive genetic direct random coefficients; γ = the vector of maternal additive genetic random coefficients; δ = the vector of animal permanent environmental coefficients; ρ = the vector of maternal permanent environmental coefficients; X, Z_1, Z_2, W_1, and W_2 = incidence matrices; and ε = the vector of residuals; k_A, k_M, k_C, and k_Q are the matrices of coefficients of the covariance function for additive direct and maternal effects and animal and permanent environmental effects respectively. A is the additive numerator relationship matrix and I is identity matrix.

Residual variance is considered in several procedures in models. In some models, residual variance in different ages supposed fixed and stable in which obtained an estimation for σ_e^2. In some other ones, residual variance in different ages supposed variable and non-uniform. In this study by supposing independency of temporary environmental effects from each other, residual variance by using a variance function (VF) degree 2 and 3 was considered as follows in computation:

$$\sigma_j^2 = \sigma_0^2 \exp\left\{1 + \sum_{r=1}^{v-1} b_r(a^*_{ij})\right\}$$

where σ_j^2 = residual variance in jth age, σ_0^2 = intercept or error variance at the mean age, b_r and v = coefficients and numbers of variance functions parameters, respectively.

Model comparison: Models with different orders of fitting and random effects were compared by log-likelihood ratio test (LRT). The log-likelihood ratio test only allows comparisons between nested models and tends to favor models with higher number of parameters [9, 12]. Also selection of models was based on Akaike's Information Criterion [20]. Akaike [20] proposed a simple and useful criterion called Akaike's Information Criterion (AIC) for selecting the best-fit model among alternative models.

AIC= -2 log (maximum likelihood) +2 (No. of model parameters). Differences between AIC values are important, not the absolute size of AIC values. The model with the lowest AIC is considered as the best one. Various experiences verify the applicability of AIC in model selection [16, 21].

Results and Discussion

Means and numbers of records for the various ages are shown in Fig.1. Body weights have an increasing pattern by ages from birth to 180 days of age, with a decrease in growth rate after weaning.
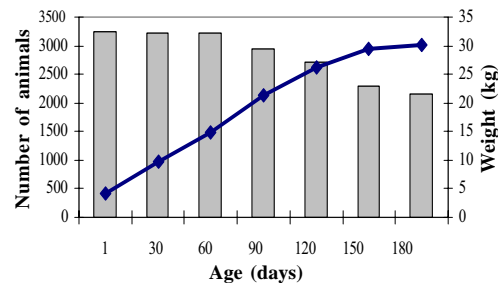
Figure 1. Number of records and weights average in different ages.

Totally, eight models were fitted to the data and a description of each with the corresponding Log L and AIC values are given in Table 2. Increasing the order of fit for direct and maternal permanent environmental effects from 3 to 4 (Model 3) improved the fit to the data as indicated by an increase in Log L (p<0.01) over the other models. Meyer [10] reported similar findings. Modeling measurement error variances (σ_e^2) using a variance function, assuming a log linear model, produced a better fit than assuming homogeneous. Both LRT and AIC suggested a model with k = 3, 3, 4, 4 and v = 4, with a total of 36 parameters.

Estimations of (co)variances and correlations between RR coefficients for models with k = 3, 3, 4, 4 and v = 4 are presented in Table 3. In all cases, intercept of the polynomial regression explained the highest proportion of variation and there were positive correlation between the linear and intercept coefficients for all effects with the exception of maternal permanent environmental effect. Variance functions for residual were as follows:

$$\sigma_j^2 = 2/204 \exp[1 - 0/3624(a^*_{ij}) - 1.6418(a^*_{ij})^2 + 1/8835(a^*_{ij})^3]$$

Estimation of variance component: Variance component estimations for weights from birth to 180 d of age, obtained with the model chosen by LRT and AIC (k = 3,3,4,4) are presented in Fig. 2. Decreasing the order of fit for genetic direct effects slightly

Table 2. Order of fit for animal direct (k_a) and maternal (k_m) genetic effects, animal (k_q) and maternal permanent (k_c) environmental effects, and residual (v-1), number of parameters (np), log likelihood function (log L), Akaike's Information Criterion (AIC), number of time stages (or metamer) considered for estimating residual variance component(n), and order of fit of fixed regression (O.F.F).

Model	k_A	k_M	k_Q	k_C	$v-1$	np	n	O.F.F	LogL	AIC
1	5	3	5	3	3	46	..	3	-29423.0	58938.0
2	3	4	4	3	3	36	..	3	-26155.7	52383.4
3	3	3	4	4	3	36	..	3	-25298.4	50668.8
4	3	3	3	3	3	28	..	3	-25467.6	50991.2
5	3	3	3	3	..	25	1	3	-27750.0	55550.0
6	3	3	3	3	..	31	7	3	-25206.4	50474.8
7	3	3	3	3	2	27	..	3	-25765.7	51585.4
8	3	3	3	3	..	28	1	4	-27692.0	55440.0

Table 3. Estimations of variances (diagonal), covariance (below diagonal), and correlations (above diagonal) between random regression coefficients for model with order of fit of 3,3,4,4 for additive genetic direct and maternal effects and animal and maternal permanent environmental effects respectively and fitting a cubic variance function for measurement errors.

| 0	1	2	3
		K=3,3,4,4	
Additive direct effect			
7.006	0.939	-0. 357	
3.295	1.757	-0.0200	
-0.374	-0.0103	0.156	
Additive maternal effect			
1.233	0.930	-0.931	
0.338	0.107	-0.809	
-0.189	-0.0487	0.0336	
Animal permanent environmental effect			
1.381	0.971	0.223	-0.863
1.617	2.009	0.273	-0.922
0.052	0.077	0.0401	0.066
-0.479	-0.618	0.006	0.223
Maternal permanent environmental effect			
11.121	0.944	0.997	0.386
3.119	0.981	-0.918	0.061
-1.870	-0.511	0.316	-0.452
0.231	-0.018	-0.046	0.032

changed the partitioning of animal effect variances. The differences between the two models were largest after 60 d (not shown). Direct additive genetic variance increased steadily over the trajectory. A similar pattern of variation for direct additive genetic variance was reported [2, 6, 17]. Maternal additive genetic variance increased after birth to a peak around 120 d and decreased thereafter. Maternal permanent environmental variance increased after birth to a peak around 120 d and decreased after that. Generally, maternal variances tended to be higher at younger ages and declined with age, particularly beyond post-weaning ages (>120 d).

Estimation of direct and maternal heritability: Direct and maternal heritability estimates for Model 3 are presented in Fig. 3. Direct heritability estimations increased after birth (0.1) until animals were about 150 d old (0.35) and increased faster after that. Maternal heritability estimations decreased faster from birth (0.14) to about 30 d of age and decreased slowly with age thereafter. At weaning (90 d), maternal heritability was estimated by 0.55. As

Baluchi sheep were weaned around 90 d of age, results show that maternal genetic effects started to decrease before weaning. These trends are similar to that described by Lewis and Brotherstone [6]. However, the peak in maternal heritability estimations occurred closer to weaning, around 60 to 90 d of age. For data set, direct heritability estimations tended to be higher when maternal effect estimations were lower.

Maternal permanent environmental effects: Maternal permanent environmental variances as a proportion of the phenotypic variance (C) for using a model with k = 3, 3, 4, 4 (Fig. 3) increased rapidly from birth (0.1) to about 45 d of age (0.54) remained almost the same with a slight trend to decrease until 90 d of age and thereafter decreased markedly with age. Most studies reported the decrease of maternal permanent environmental effect as time lapses post weaning. Albuquerque and Meyer [1] and Meyer [10] described a similar pattern for weights from birth to weaning in the Zebu and Wokalup breeders.

Estimation of correlations: Estimations of direct genetic and animal permanent environmental correlation between weights from birth to 180 d of age are presented in Table 4. In general, correlations tended to decrease by increasing number of days between records. Direct genetic correlation estimations decreased by increasing age between weights. In addition, direct genetic correlations among body weights for all ages were positive; indicating that selection for weaning weight would also increase birth weight and other weights. The results we have obtained with RRM are, overall, similar to those reported by other authors [7, 8, 13]. Maximum of correlations of direct additive genetic, maternal genetic, animal permanent environmental and maternal permanent environmental effects were estimated between adjacent ages (0.99) and this parameter ratio decreased parallel to increase of distance between age days in most cases. In a way that mentioned correlations ratio between far ages were low and allocated minimum correlations as birth weight and weight of 60 days of age (0.63). In general, genetic and permanent environmental correlations decreased by increasing number of days between records. These trends were reported by Lewis and Brotherstone [6] in weight records analysis. Comparison between correlations in two models indicated that estimation of correlations were similar for any two models except maternal additive genetic correlations between birth weight and body weight in other ages that were more in Model 4 relating to Model 3.
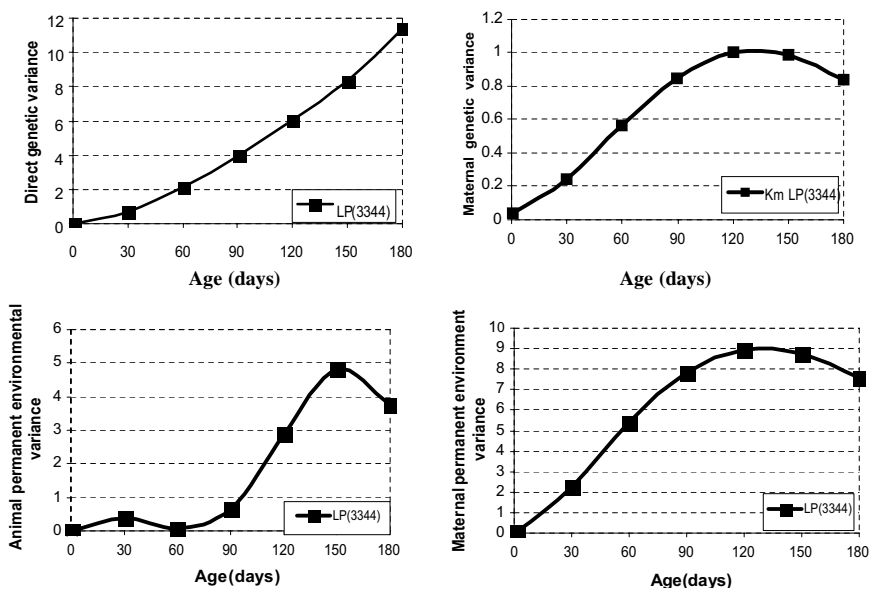
Figure 2. Direct additive genetic variance (top left), maternal additive genetic variance (top right), maternal permanent environmental variance (bottom right) and animal permanent environmental variance (bottom left) estimated for body weight based on Models 3 and 4.
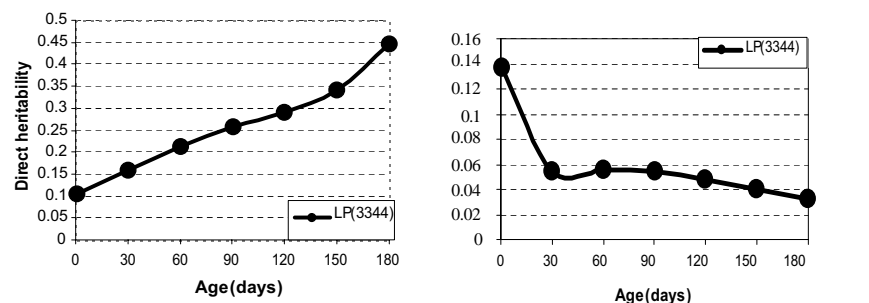
Figure 3. Direct (top left) and maternal (top right) heritability estimates and maternal (bottom) variance component estimates as proportions of phenotypic variance for data from random regression analysis with k = 3, 3, 4, 4.

Table 4. Direct genetic (above diagonal) and animal permanent environmental (below diagonal) correlation between weights at selected ages (in days).

Age	1	30	60	90	120	150	180
1	1	0.67	0.66	0.68	0.70	0.72	0.72
30	0.68	1	0.99	0.98	0.95	0.89	0.78
60	0.70	0.77	1	0.99	0.97	0.91	0.82
90	-0.39	-079	-0.23	1	0.98	0.95	0.88
120	-0.54	-0.88	-0.41	0.98	1	0.98	0.94
150	-0.62	-0.88	-0.42	0.96	0.99	1	0.98
180	-0.72	-0.73	-0.30	0.86	0.89	0.94	1

Conclusions

Direct heritability estimations increased after birth and tended to be highest at ages which maternal effect estimations tended to be lowest. Maternal heritability estimations decreased after birth to 30 d of age and decreased slowly thereafter. Similar results were found by Lewis and Brotherstone [6]. A higher response to selection for maternal ability would be expected if selection was based on pre weaning weights.

A model with RR modeled the pattern of (co)variances in the data adequately, by estimations similar to those obtained with univariate analysis. Increasing the order of polynomial fit increased the flexibility of the curve. However, applying polynomials with a high order of fit increased computational requirements, made convergence difficult to reach and increased sampling problems [1,5].

Covariance functions that give the covariances between any two ages within the range of ages in the data can be estimated by applying a RRM and REML [7], and impose a structure on the covariance matrices. Instead of predicting breeding value (BV) for each recorded weight, BV is predicted for additive direct and maternal regression coefficients.

Using the genetic regression coefficients, BV can be predicted for the growth curve as a whole, for any age (within the range in the data), even those for which the animal had no records and for functions of the growth curve. Animals with only one or even without records can be evaluated using information from relatives. Random regression models are expected to give more accurate estimations of genetic parameters and predictions of breeding values than the conventional multi trait model.

Applying RRM instead of multi trait models for genetic evaluations of growth traits will increase the number of mixed-model equations, the coefficient matrix of random effects will be denser, and, consequently, computing requirements will increase accordingly. Alternatives to decrease the order of polynomial fit must be investigated.

References

[1] Albuquerqe, L. G. and Meyer, K. 2001. Estimates of covariance function for growth from birth to 630 days of age in Nelore cattle. J. Anim. Sci. **79**:2776-2789.

[2] Ghafouri Kesbi, F., Eskandarinasab, M. P. and Shahir, M. H. 2008. Estimation of direct and maternal effects on body weight in Mehraban sheep using random regression models. Arch. Tierz., Dummerstorf **51**(3):235-246.

[3] Huisman, A. E., Verkamp, R. F. and Van Arendonk, J. A. M. 2002. Genetic parameters for different random regression models to describe the weight data of pigs. J. Animal Sci. **8**:575-582.

[4] Jamrozic, J. and Schaeffer, L. R. 1997. Estimates of genetic parameters for test day model with random regressions for production of first lactation Holsteins. J. Dairy Sci. **80**:762-770.

[5] Kirkpatrick, M., Hill, W. G. and Thompson, R. 1994.Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. Genet. Res. **64**:57-69.

[6] Lewis, R. M. and Brotherstone, S. 2002. A genetic evaluation of growth in sheep using random regression techniques. J. Animal Sci. **74**: 63-70.

[7] Meyer, K. 1998. "DXMRR" is a program to estimate covariance functions for longitudinal data by restricted maximum likelihood. Proc. 6th World Cong. Genet. Appl. Livest. Prod., Armidale, NSW, Australia, **27**:465-466.

[8] Meyer, K. 1999. Estimates of genetic and phenotypic covariance functions for post weaning growth and mature weight of beef cows. J. Anim. Breed. Genet. **116**:181-205.

[9] Meyer, K. 2000. Random regression to model phenotypic variation in monthly weights of Australian beef cows. Livest. Prod. Sci. **65**:19-38.

[10] Meyer, K. 2001. Estimates of direct and maternal covariance function for growth of Australian beef calves from birth to weaning. Genet. Sel. Evol. **33**:1-28.

[11] Meyer, K. and Hill, W. G. 1997. Estimation of genetic and phenotypic covariance function for longitudinal or repeated records by restricted maximum likelihood .Livest. Prod. Sci. **47**:185-200.

[12] Olori, V. E., Hill, W. G., McGuirk, B. J. and Brotherston, S. 1999. Estimating variance components for test day milk record by restricted maximum likelihood with random regression animal model. Livest. Prod. Sci. **61**:53-63.

[13] Pool, M. H. and Meuwissen, T. H. E. 2000. Reduction of the number of parameters needed for a polynominal random regression test day model. Livest. Prod. Sci. **64**:133-145.

[14] Schaeffer, L. R. and Dekkers, J. C. M. 1994. Random regressions in animal models for test-day production in dairy cattle. Proc. 5th Word Congr. Genet. Appl. Livest. Prod., Guelph, Canada, **18**:443-446.

[15] Swalve, H. H. 2000. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. **83**:1115-1124.

[16] Burnham, K. P. and Anderson, D. R.1998. Model Selection and Inference: A Practical Information-theoretic Approach. Springer-Verlag, New York.

[17] Fischer, T. M., Gilmour, A. R. and Van der Werf, J. H. J. 2004.Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. Genetics Selection Evolution **36**:363–369.

[18] Andersen, S. and Pedersen, B.1996. Growth and food intake curves for group-housed gilts and castrated male pigs. Animal Science **63**:457-464.

[19] Laird, N. M. and Ware, J. H. 1982. Random-effects model for longitudinal data. Biometrics **38**:963–974.

[20] Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In Petrov, B. N. and Csaki, F. (eds). Proceedings of the Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267-281.

[21] Wada, Y., and Kashiwagi, N. 1990. Selecting statistical models with information statistics. J. Dairy Sci. **73**:3575–3582.