

THE PREDICTION OF BREEDING VALUE IN A DAIRY SHEEP POPULATION USING DIFERENT TEST DAY ANIMAL MODELS

H. GROSU, M. GRAS, Cristina LAZAR, Rodica PELMUS
National Research & Development Institute
for Biology and Animal Nutrition
Balotesti, Ilfov,
Romania

ABSTRACT

The objective of this study was to predict the breeding values in a dairy sheep population in order to find the best individuals for the next generation. The breeding values were predicted using two methods: a) *random regression test day animal model & fixed regression test day animal model* and b) *auto-regressive test day animal model*.

The data set consisted of 833 TD records from 174 ewes in the first lactation. The whole population had 362 individuals, which the following structure: 48 sires, 140 dams and 174 offspring's (ewes with own performances). Totally, 174 ewes had records. The average number of TD per lactation was about 5.3.

Data was edited and TD records were deleted if ewe' ID was unknown, if lactation number was not specified and if days in lactation for the TD record was < 60 or > 175 days. Also, a TD class must have at least 4 observations.

The methods were compared using the accuracy of prediction of breeding values, Spearman correlation coefficient and the percentage of squared bias. For the data set accounted the best model was the random regression test days with the three order Legendre polynomials.

INTRODUCTION

Best linear unbiased prediction (BLUP) applied to an animal model (AM) is the standard procedure for genetic evaluation. It has the advantage that all known information is optimally taken into account and selection or special mating has a small or no effect on the evaluation. This procedure is even more valuable in dairy sheep because, with natural mating, the number of progeny per ram is relatively small, which makes information from other relatives more important. Consequently, an increasing number of national evaluation systems is based on AM in dairy sheep: since 1990-1992, AM evaluation systems have been implemented in France for Lacaune, Basco-Bearnaise, Manech and Corsica breeds (Gabina et Barillet, 1991; Barrilet et al, 1992; Sana et al.,1993; Pagnacco et al., 1991; Pinelli et al.,2000; cited by 4).

In recent years, attention has been drawn to use of test day (TD) records directly instead of using cumulative lactation yield calculated from them. There are several potential benefits from using this technology. First, conventional recording systems are costly and ways for a simplification of production recording schemes are desirable. Second, generation intervals can be reduced as genetic evaluations can be performed sooner using all test day records available at a given time instead of waiting for complete lactation records. Third, the traditional approach of using complete lactation records has been criticized as inconsistent since record taken at defined locations and time are aggregated in a rather arbitrary way and are subsequently subjected to quite sophisticated statistical analyses targeted toward an optimum differentiation of all genetic and environmental effects.

The objective of this study was to predict the breeding values in a dairy sheep population in order to find the best individuals for the next generation. The breeding values were predicted using three methods: a) random regression test day animal model (RRM), b) auto-regressive test day animal model (ARM) and c) fixed regression animal model (FRM).

MATERIALS AND METHODS

The data set consisted of 833 TD records from 174 ewes in the first lactation. The whole population had 362 individuals, which the following structure: 48 sires, 140 dams and 174 offspring (ewes with own performances). Totally, 174 ewes had records.

Data was edited and TD records were deleted if ewe' ID was unknown, if lactation number was not specified and if days in lactation for the TD record was < 60 or > 175 days. Also, a TD class must have at least 4 observations.

Table 1

Numbers of test day (TD) records, fixed effects, animals, and mean daily milk yield in first lactation for a local dairy sheep line

No crt	Specification	
1	Total test day records, no	833
2	TD fixed effects	23
3	Animal, no	362
4	Ewes with records, no	174
5	Mean TD record per ewe, no	5.3
6	Mean daily yield, gram	453.55
7	Standard deviation daily, gram yield	341.75

Preselection of Models

Model comparison criteria were the percentage of squared bias (PSB; Ali and Schaeffer, 1987, cited by 5) and the accuracy of evaluations. The PSB was computed as:

$$PSB = \frac{\sum (y_{ijkl} - \hat{y}_{ijkl})^2}{\sum (y_{ijkl})^2} \times 100$$

where n is the number of records, y_{ijkl} is the observed record, and \hat{y}_{ijkl} is the record predicted by particular model in question.

The methods were compared using the accuracy of breeding values prediction. For the data set accounted the best model was the random regression test days with the three order Legendre polynomials.

The accuracy of evaluation was calculated from inverse elements of the mixed model equations for the diagonal block corresponding to animal genetic effects (Jamrozik et al., 2000).

Rank correlation (Spearman) was used for calculate the correlation between the ranks occupied by the same animal on different models. The ordinary correlation was computed between the breeding values using the two models. (Co)variance components were estimated using REML method.

Models

The first type of model used for genetic evaluation is a **random regression model**:

$$y_{ijkl} = TD_i + \sum_{m=1}^3 b_{mj} X_m + \sum_{m=1}^3 \alpha_{mk} X_m + p_k + e_{ijkl}$$

where:

yijkl = record l on ewe k made on DIM jl of first lactation for a ewe belonging to class j of age at lambing; **TDi** is the fixed effect due to ewe tested in the same test-date i ;

pk is random permanent environmental effect associated with all TD yields of ewe k

within lactation; **eijkl** is random residual effect; **bmj** and **amk** are fixed and random regression coefficients, respectively.

Third-order Legendre polynomials (X_m) were used for both fixed and random regressions on the scale from 60 to 175 DIM

In matrix notation, the model can be written as:

$$Y = X_1 b_1 + X_2 b_2 + Z_1 \alpha + Z_2 p + e$$

The Mixed Model Equations for this model are:

$$\begin{bmatrix} X_1' \cdot X_1 & X_1' \cdot X_2 & X_1' \cdot Z_1 & X_1' \cdot Z_2 \\ X_2' \cdot X_1 & X_2' \cdot X_2 & X_2' \cdot Z_1 & X_2' \cdot Z_2 \\ Z_1' \cdot X_1 & Z_1' \cdot X_2 & Z_1' \cdot Z_1 + G^{-1} \otimes A^{-1} & Z_1' \cdot Z_2 \\ Z_2' \cdot X_1 & Z_2' \cdot X_2 & Z_2' \cdot Z_1 & Z_2' \cdot Z_2 + I \cdot k \end{bmatrix} \cdot \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \hat{\alpha} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X_1' \cdot y \\ X_2' \cdot y \\ Z_1' \cdot y \\ Z_2' \cdot y \end{bmatrix}$$

The **auto-regressive test day animal model** (TDAM) developed by J. Carvalheira (Carvalheira et.al., 1998) has been used to estimate breeding values for all animals in the data. The computer software is based on a series of programs that build the incidence matrices according to the structure of the data, and compute the inverse of the genetic additive relationship matrices to be incorporated into the coefficient matrix of the BLUP mixed model equations. The model portions the random environmental component

into short-term environmental (STE) effects (following a first-order auto-regressive process across lactation), long-term environmental (LTE) effects (following a first-order auto-regressive process across lactations), and independent residuals. The model used to describe the data was as follows:

$$y_{ijkmn} = TD_i + Age_j + DIM_k + a_l + t_{m(l)} + e_{ijkmn}$$

where: y_{ijkmn} is milk yield of an individual TD sample from the n^{th} ewe; TD_i is the fixed effect due to i^{th} test day control; age_j is the fixed effect of the j^{th} age class at calving; DIM_k is the fixed effect of the k^{th} DIM class in the lactation; a_l is the random effect of n^{th} animal; $t_{m(l)}$ is the random effect of STE nested within ewe and assumed to follow a first-order auto-regressive process, and e_{ijkmno} is the random residual term.

In matrix notation, the model can be written as:

$$Y = Xb + Za + Qt + e$$

The Mixed Model Equations for this model are:

$$\begin{bmatrix} X' \cdot X & X' \cdot Z & X' \cdot Q \\ Z' \cdot X & Z' \cdot Z + A^{-1} \cdot k & Z' \cdot Q \\ Q' \cdot X & Q' \cdot Z & Q' \cdot Q + S^{-1} \cdot k \end{bmatrix} \cdot \begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{t} \end{bmatrix} = \begin{bmatrix} X' \cdot y \\ Z' \cdot y \\ Q' \cdot y \end{bmatrix}$$

Similar to repeatability model, at the genetic level, the **fixed regression model** assumes that test day records within a lactation are repeated measurements of the same trait, that is, a genetic correlation of unity among test day observations. Usually the permanent environmental is included in the model to account for environmental with permanent effects on all test day yields within lactation.

The fixed regression model has the form:

$$y_{ijkl} = TD_i + \sum_{m=1}^3 b_{mj} X_m + a_k + p_k + e_{ijkl}$$

where y_{ijkl} is the record of ewe k made on DIM jl of first lactation for a ewe belonging class j of age at lambing within TD subclass i ; TD_i is the fixed effect due to ewe tested in the same test-date I ; a_k is vector of animal additive genetic effects of animal k ; p_k is random permanent environmental effect associated with all TD yields of ewe k within lactation; e_{ijk} is random residual effect; b_{mj} is fixed regression coefficient; X_m third-order Legendre polynomials were used for fixed regression on the scale from 42-274 DIM.

In matrix notation, the model can be written as:

$$Y = Xb + Qu + Zpe + e$$

The Mixed Model Equations for this model are:

$$\begin{bmatrix} X' \cdot X & X' \cdot Q & X' \cdot Z \\ Q' \cdot X & Q' \cdot Q + A^{-1} \cdot \alpha_1 & Q' \cdot Z \\ Z' \cdot X & Z' \cdot Q & Z' \cdot Z + \alpha_2 \end{bmatrix} \cdot \begin{bmatrix} \hat{b} \\ \hat{u} \\ p\hat{e} \end{bmatrix} = \begin{bmatrix} X' \cdot y \\ Q' \cdot y \\ Z' \cdot y \end{bmatrix}$$

RESULTS AND DISCUSSION

Covariance components. The parameters for the genetic covariance matrix of the random regression coefficients (G), genetic variance, short term environmental variance and error variance are shown in table 2.

Table 2

Estimates of genetic (co)variance for random regression coefficients, permanent environmental (pe) and residual variances for milk yield

Parameters		Milk
α_0	α_0	1.4895
α_0	α_1	0.2683
α_0	α_2	-0.6239
α_1	α_1	0.4160
α_1	α_2	-0.1306
α_2	α_2	0.4540
Genetic variance		29198
Short term environmental variance		5840
Error variance		81754

Predicted Breeding Values (PBV). The breeding values and accuracy of EBV were predicted for all animals in the data set. Using BLUP methodology, these estimates are adjusted for all other effects included in the model. All effects included in the model are simultaneously estimated and predicted each others.

For the ewe evaluation with yields and pedigree information, all three sources of information could be available. Progeny performance is adjusted for merit of the mate by subtracting half the mate's breeding value. For ram evaluation, only pedigree and progeny portions of the equation are used.

A sample of results is presented in table 3, representing the best 15 animals based on their breeding value and their accuracy, for all methods.

Table 3

The best 15 sheep ordered after the models

ID	Type	No TD	BV RRM	Rank of BV RRM	BV ARM	Rank of BV ARM	ACC AR M	BV FRM	Rank of BV FRM	ACC FRM
54178	P	5	16.58	1	273.88	2	0.79	255.84	1	0.76
74107	P	6	15.45	2	274.78	1	0.81	244.54	2	0.73
84163	P	5	14.04	3	231.27	5	0.77	237.22	3	0.75
24025	S	0	12.31	4	186.85	8	0.64	161.83	9	0.37
84177	P	5	12.15	5	151.27	12	0.77	168.81	8	0.71
84055	P	3	10.01	6	183.03	9	0.69	178.10	6	0.76
44165	P	6	9.89	7	136.97	15	0.80	126.35	14	0.63
74131	P	6	9.21	8	150.78	13	0.81	140.19	12	0.76
64026	S	0	9.06	9	91.46	38	0.62	111.53	20	0.56
54077	P	6	8.96	10	116.34	25	0.81	89.50	32	0.76
64018	P	4	8.46	11	202.60	6	0.75	175.58	7	0.69
74208	P	5	8.31	12	157.00	11	0.79	136.68	13	0.75
74006	P	6	8.21	13	109.81	27	0.82	105.94	21	0.73
54023	S	0	7.97	14	55.63	74	0.78	54.24	74	0.56
54178	P	5	7.83	15	120.23	24	0.79	87.14	34	0.75

From the data shown in Table 3 it can be noticed that both types of models yield close results. The criterion of ranking was the breeding value calculated using the random regression model.

It can be observed that 10 of the 15 animals (about 67%) are among those ranked by their breeding value calculated using the auto-regressive test day. In the same way we can observe a different rank for breeding value calculated using fixed regression model similarly with previous model (67% of animals ranked in with RRM are also in the group ranked with FRM).

The individuals subtracted from the best 15 individuals (ranked by RRM) are the same in the both compared models (ARM and FRM), means individuals from the places 9, 10, 13, 14, 15 from EBV with RRM.

The accuracy of selection, P.S.B. and rank correlation. For whole population the average accuracy of selection was about 0.607 for auto-regressive model and 0.584 for fixed regression model. The gain obtained is about 4%. The rank correlation coefficient between methods was very high ($r_{AR,FR}=0.98$, $r_{AR,RR}=0.93$ and $r_{FR,RR}=0.95$).

P.S.B. was 8.2331% for the random regression model, 8.503% for the auto-regressive test day and 9.2747% for fixed regression model. Considering, however, the very high correlation (0.905 between EBV_{RRM} and EBV_{ARM} ; 0.930 between EBV_{RRM} and EBV_{FRM} ; 0.984 between EBV_{ARM} and EBV_{FRM}) existing between the breeding values calculated with the models as well as the high rank correlation ($r_{AR,FR}=0.98$, $r_{AR,RR}=0.93$ and $r_{FR,RR}=0.95$), it results that any of the models can be used in the practice of dairy sheep genetic evaluation.

CONCLUSIONS

1. The TD animal models offer the opportunity to improve the genetic evaluation of dairy sheep.
2. TD models can account more precisely for the environmental factors that could affect the ewes differently during lactation.
3. TD model allows a ewe to be evaluated on the basis of any number of TD records during lactation.
4. The selection done using the models has an accuracy of 58.4-60.7%.
5. P.S.B. was 8.2331% for the random regression model, 8.503% for the auto-regressive test day and 9.2747% for fixed regression model which means that there are almost the same differences between the measured and estimated phenotypic values for studied models.
6. The correlation between the breeding values calculated using the models is very high (0.905 between EBV_{RRM} and EBV_{ARM} ; 0.930 between EBV_{RRM} and EBV_{FRM} ; 0.984 between EBV_{ARM} and EBV_{FRM}). The rank correlation displayed the same trend ($r_{AR,FR}=0.98$, $r_{AR,RR}=0.93$ and $r_{FR,RR}=0.95$)
7. Any of the two variants of models investigated during this study can be used in the practice of dairy sheep genetic evaluation.

Acknowledgement:

This work was financially supported from PNCDI2 –
UEFISCDI project no. 52163 OPECGEN, National
Authority for Scientific Researches

REFERENCES

1. **Barillet, F.** 1992. Use of an animal model for genetic evaluation of the Lacaune dairy sheep. *Livest. Prod. Sci.*, 31:287-299
2. **Barillet, F.** 1997. Genetics of milk production. In "The Genetics of Sheep" edited by L. Piper and A. Ruvinsky, CAB International.
3. **Carvalho J., R. W. Blake, E. J. Pollak, R. L. Quaas, and C. V. Duran-Castro.** 1998. Application of an Autoregressive Process to Estimate Genetic Parameters and Breeding Values for Daily Milk Yield in a Tropical Herd of Lucerna Cattle and in US Holstein Herds. *J. Dairy Sci.*, 81:2738-2751.
4. **Grosu H., P.A. Oltenacu, J. Carvalho.** 2001. Efficiency of the Test Day Animal Model in Dairy Ewes. The Third Joint Meeting of the Universities and Research Institutions for Animal Science of the Balkan Countries, Tekirdag, Turkey.
5. **Grosu H., 2007** – The prediction of breeding value in a dairy sheep population using a test day animal model. BALNIMALCON, Thessaloniki, Greece.
6. **Harville, D.A.** 1979. Recursive estimation using mixed linear model with autoregressive random effects. Pages 157-159 in *Variance Components and Animal Breeding. Proc. Conf. in Honor of C.R. Henderson*, Cornell Univ. , Ithaca, NY.
7. **Jamrozik J., L.R. Schaeffer and J.C.M. Dekkers.** 1997. Genetic Evaluation of Dairy Cattle Using Test Day Yields and Random Regression Model.
8. **Jamrozik J., L.R. Schaeffer and G.B. Jansen.** 2000. Approximate accuracies of prediction from random regression models. *Livest. Prod.*: 66: 85-92.
9. **Kachman, S.D., and R.W. Everett.** 1989. Test day model with individual herd correction factors. *J.Dairy Sci.* 72.
10. **Pineli, F., P.A. Oltenacu, G. Iannolino, H. Grosu, A.D. D'Amico, M. Scimonelli, G. Genna, G Calagna and V. Ferrantelli.** 2000. Design and Implementation of Genetic Improvement Program for Comisana Dairy Sheep in Sicily. *Sheep Breeding*. Guelph University.
11. **Wade, K.M., and R.L. Quaas.** 1993. Solutions to a system of equations involving a first-order autoregressive process. *J.Dairy Sci.* 76: 3026-3031