

Random Regression Model for the Genetic Evaluation of Production Traits of Dairy Cattle in the UK

R. Mrode¹, S. Brotherstone², I. White², G. Swanson³, M. Coffey¹, H. Jones¹ and R. Thompson⁴

¹SAC, Sir Stephen Watson Building, Penicuik, EH26 0PH, UK; ²Edinburgh University, Edinburgh; ³MDC, Cirencester; ⁴Rothamstead Research, Harpenden & Roslin Institute, Roslin

Introduction

A test day model (TDM) for the genetic evaluation of production traits in the UK was presented by Mrode *et al.* (2003). The methodology presented allowed for the joint evaluation of test day (TD) and 305 day records. In this first analysis there was evidence that the contemporary group effect (CG) for test day records varied over the course of lactation depending on the days in milk (DIM). For instance plots of heritability (h^2) by days in milk indicated that h^2 was increasing towards the end of the lactation. Secondly, in the animal model used for the evaluation of lactation records in the UK, a sire-herd interaction effect is included. This has the effect of avoiding over-prediction of farmer-owned bulls used in one or few herds. The absence of this effect in the initial TDM resulted in a number of farmer-owned bulls

being over-evaluated and placed in several top bull lists. The TDM was therefore up-dated to account for these effects. This paper presents the new TDM implemented in the UK and some results for the Holstein and Jersey breeds.

Material and Methods

Parameter estimation

The genetic parameters used for the evaluation were estimated fitting a sire model consisting of 823 bulls with 22343 daughters using ASReml (Gilmour *et al.*, 2002).

Milk, fat and protein in the first three parities were analysed as 9 traits with the model for the trait k in the j^{th} parity as:

$$y_{ijkltc} = htd_{kl}(DIM_t) + hd_{kc}(DIM_t) + FIX_{ijkt} + \sum_{r=1}^5 v_{tr} \beta_{jk} + \sum_{r=1}^3 v_{tr} s_{mjk} + \sum_{r=1}^3 v_{tr} w_{ijk} + e_{ijkltc} \quad [1]$$

where y_{ijkltc} is the TD yield of trait k in the j^{th} lactation of cow i , daughter of sire m , made on day t in milk, within herd-test-day (htd) subclass l in the c^{th} herd (hd); Htd (random) and hd (fixed) effects depended linearly on DIM, with an intercept and slope for each trait, independent of lactation number; FIX are other fixed effects in the model and details have been given by Mrode *et al.* (2003); β_{jk} are the fixed regressions coefficients, s_{mjk} and w_{ijk} are vectors of random regression (RR) coefficients for sires and cow (permanent environmental) effects respectively; v_t is the vector of Legendre polynomials for day t in milk and e is the random error.

The traits were analysed in sets of three with milk in the first lactation always included to allow for selection. There were 28 runs required to analyze all 9 traits, each run estimating a 9x9 submatrix of the complete 27x 27 between sire and within-sire matrices. Complete matrices were reconstituted from overlapping submatrices so that they were positive definite. This was achieved by using Cholesky factors of submatrices as 'pseudo-data' for the full set of traits (Thompson *et al.*, 2005).

Genetic evaluation

In [1], 9 RR coefficients were used to define additive effects for each sire and cow (within sire) effects for each trait across the three lactations. Given that the correlations between these RR coefficients ranged from medium to high, it can be expected that the same effects could be described with fewer coefficients.

$$htd_{kl}(DIM_t) + \sum_{r=1}^6 q_{tkr} \mathbf{a}_{ikr} + \sum_{r=1}^6 d_{tkr} \mathbf{pe}_{ikr} + hs_{kj} + \mathbf{e}_{iklt} \quad [2]$$

where $\mathbf{a}_{ikr}$ and $\mathbf{pe}_{ikr}$ are vectors of random regression coefficients for animal and permanent environmental effects respectively for animal i and trait k ; q and d are covariables associated with $\mathbf{a}$ and $\mathbf{pe}$ respectively and are obtained from eigenfunctions corresponding to the 6 largest eigenvalues used to describe animal and permanent environmental effects; hs_{kj} is the random sire-herd interaction effect for trait k in parity j and $\mathbf{e}$ is the random error. Based on other analyses, the variance for sire-herd interaction for each trait was assumed to be 5% of the total phenotypic variance.

The major difference between this model and the one presented by Mrode *et al.* (2003), is that CG effects now depend linearly on DIM, with an intercept and slope for each trait. The model for the 305d records for trait k in the three parities was similar to that described for TD records but with the first two terms in [1] replaced by herd-year-season.

Solving the mixed model equations (MME)

The system of equations was solved using the Mix99 software (Lidauer *et al.*, 1999) on a workstation with 8 processors and 32 GB RAM. The convergence criterion was the relative difference between the right-hand and left-hand sides of the equations for animal effects. Equations were considered to have converged when this was lower than 10^{-4} . The structure of data analysed for the Holstein and Jersey breeds are in Table 1. After convergence, predicted

Therefore a multi-lactation reduced rank random regression TDM was fitted to analyse milk or fat or protein in the first three lactations. The detail of the method for the rank reduction has been described by Mrode *et al.* (2003). The fixed part of the model for the genetic evaluation is the same as given in [1] but the random part of the model for TD records was:

transmitting abilities (PTAs) were computed from the RR coefficients

Results and Discussion

Genetic Parameters

Estimates of heritabilities for 305d yield for milk, fat and protein yields computed from the covariances for the regression coefficients were medium to high (Table 2) varying from 0.35 to 0.61. These figures are slightly higher than the h^2 used in Canada but similar to estimates used in Netherlands (De Roos *et al.*, 2001). It should be noted that there are variations in the method used to compute h^2 for 305d yield and this could affect the value obtained. Assuming 10 regular monthly samplings over the length of the lactation, Mäntysaari (2002), used the formula: $10 \times 30^2 \times \sigma_e^2$ to compute residual variance for 305 day yield. An alternative approach involves summing daily estimates of residual variances over 305 days and this will result in a different value. The method used by Mäntysaari (2002) has been used to obtain the estimates shown in Table 2. Using the alternative approach, for instance, gave h^2 estimates of 0.59 and 0.55 for milk yield in first and second parities. It would be useful if Interbull were to recommend a standard procedure to ensure that consistent estimates of 305d yields are published.

A plot of daily heritability for the old TDM model compared with the current (Figure 1) shows the problem of the rising heritabilities at the end of the lactation and indicates that this has

been resolved by the modification to the structure of the CG.

The estimates of genetic correlations for 305d yield for milk, fat and protein are given in Table 3. These varied from 0.39 to 0.93 and are consistent with previous estimates.

Genetic evaluation

Each trait for the Holstein breed converged after about 240 iterations and real time per iteration was about 3.5 minutes. The Jersey breed converged after about 290 iterations with real time of about 0.08 minute per iteration.

The correlations of PTAs for bulls between the TDM and IAM in both the Holstein and Jersey breeds were high at about 0.97 for milk and fat and 0.98 for protein for bulls with reliability of 50% or higher in the IAM. These correlations are similar to those reported by Van Doormaal and Kistemaker, 1999 and Lidauer *et al.*, 2000.

Correlations for cow PTAs in the TDM and IAM were 0.92, 0.93 and 0.95 respectively for milk, fat and protein in Holstein cows with at least 30% reliability in the IAM. Corresponding estimates for the Jersey breed were 0.89, 0.94 and 0.93. However if cows with at least 65% reliability were considered, then these correlations dropped slightly to 0.89, 0.92 and 0.93 for milk, fat and protein respectively. Equivalent estimates for Jersey were 0.87, 0.94 and 0.92. Cows with 65% reliability in the IAM are likely to have up to 5 lactations in their evaluations but the TDM currently includes only the first 3 parities. This might explain the slight drop in correlations for this group of cows. The correlations for cows is higher than those reported by Lidauer *et al.* (2000) but similar to those obtained in Canada (VanDoormaal and Kistemaker, 1999).

Conclusion

The inclusions of CG that depend linearly on DIM improved the fit of the model to the data by accounting for additional management or environmental effects. The less than unity correlations resulted in changes in bull ranking especially for younger bulls.

Some of these differences may be due to differences in parameters, better correction for environmental effects and other effects included in the TDM.

References

- De Roos, A.P.W., Harbers, A.G.F. & De Jong, G. 2001. Random regression test-day in the Netherlands. *Interbull Bulletin* 27, 155-158.
- Gilmour, A.R., Goet, B.J., Cullis, B.R., Welham, S.J. & Thompson, R. 2002 *ASREML User Guide, Release 1.0*. VSN International: Hemel Hemstead. 290 pp.
- Lidauer, M.I., Strandén, I. & Mäntysaari, E. 1999. *Mixed Model Equations Solver. MIX99 Manual*. Animal Production Research, Jokioinen, Finland.
- Lidauer, M., Mäntysaari, E.A. Strandén, I. & Pösö, J. 2000. Multiple-trait random Regression Test-Day Model for all lactations. *Interbull Bulletin* 25, 81-86.
- Mäntysaari, E.A. 2002. Combining Test day and full lactation records in Prediction of Breeding values. *7WCGALP*, Montpellier. Communication No 01-01.
- Mrode R.A., Swanson, G.J.T & Paget, M.F. 2003. Implementation of a test day model for production traits in the UK. *Interbull Bulletin* 31, 193-196.
- Thompson, R., Brotherstone, S. & White, I.M.S. 2005. Estimation of quantitative genetic parameters. *Phil. Trans. of the Royal Soci.* |In press.
- Van Doormaal, B.J. & Kistemaker, G.J. 1999. Application of the Canadian test day model. *Interbull Bulletin* 22, 100-103.

Table 1. The Data structure for Jersey and Holstein breeds.

Parity	Jersey			Holstein		
	Cows	TD	305d (Cows)	Cows	TD	305d(Cows)
1	58591	511242	71084	2871890	25977994	2937909
2	48478	415430	58716	2381926	21268615	2307356
3	38426	326300	47857	1865672	16475955	1745486

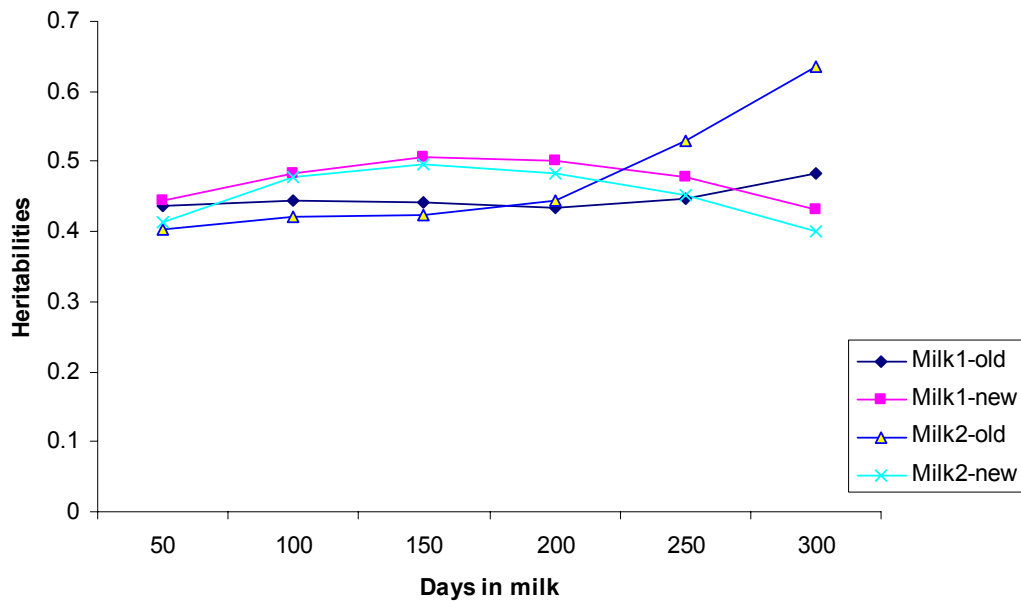
Table 2. Heritabilities for 305 day for Milk, Fat and Protein yields by lactation.

Parity	Milk	Fat	Protein
1	0.61	0.50	0.56
2	0.56	0.41	0.49
3	0.60	0.35	0.47

Table 3. Genetic correlations of 305 day Milk, Fat and Protein yields in the first three lactations

	M1	F1	P1	M2	F2	P2	M3	F3	P3
M1	1.00	0.72	0.93	0.75	0.50	0.77	0.68	0.51	0.74
F1		1.00	0.81	0.57	0.87	0.72	0.39	0.64	0.57
P1			1.00	0.69	0.61	0.80	0.60	0.68	0.73
M2				1.00	0.64	0.90	0.95	0.66	0.91
F2					1.00	0.75	0.49	0.80	0.61
P2						1.00	0.81	0.75	0.88
M3							1.00	0.63	0.88
F3								1.00	0.71
P3									1.00

where M_i , F_i and P_i refer to milk, fat and protein in the i^{th} parity.

**Figure 1.** Daily heritabilities by days in milk.