

Attempts to Reconcile Calibration Formulae

R. Thompson and R. Mrode

IACR-Rothamsted, Harpenden, Herts, UK

and

Animal Data Centre, Chippenham, Wilts, UK

Summary

Analyses are carried out to attempt to reconcile different calibration equations for prediction performance of US bulls in the UK. One estimate has a very wide confidence interval and the other is biased mainly because of correlation between US proofs with some group correction terms.

Introduction

Initial analysis (1995) showed differences between conversion formulae for USA bulls in the UK from the accepted and INTERBULL recommended international conversion procedure, the Goddard method (Goddard, 1985) with intercept and slope values of 490 and 0.32 to convert US to UK, and regression formulae developed from interbull Multiple Across Country Evaluations (MACE, Schaeffer, 1994) with intercept and slope values of 407 and 0.25. In an attempt to reconcile these different formulae, we have carried out several analyses.

Material and methods

Analyses were carried out on various subsets of data sent to INTERBULL in January 1996:

- (A) Data sent from USA (23298 bulls) and UK (4784 bulls). Data of this type is used by INTERBULL to calculate genetic variances.
- (B) Data on 170 bulls with proofs in both countries and on 4258 close relatives (3/4 sibs). Data of this type is used by INTERBULL to calculate genetic correlations.

- (C) Data on 170 bulls with proofs in both countries.

- (D) Data on 40 bulls with reliabilities greater than 0.85 in both countries. This data is used to mimic the data used in the original calibration method.

Data set C was used as an intermediate between the MACE-based method and the original calibration method. Data sets A, B, C were analysed using deregressed proofs using pedigree information.

Data sets D were deregressed ignoring pedigree information. To investigate the effect of using pedigree information in deregressing proofs, we used a data set (C2) based on C but deregressed ignoring pedigree information. Analyses were carried using Residual Maximum Likelihood to give estimates of genetic variances and covariances. We used programs written by Gilmour *et al.* (1997) that used approximate second differentials (Gilmour *et al.*, 1995) which converged, in most cases, with a small number of iterations. For data sets A, B and C we used pedigree information in the analysis. For data set C2 and D we ignored pedigree information. In order to compare the effect of pedigree information in the analysis, data set C was also analysed ignoring pedigree information (denoted C1). Previous MACE analyses have been carried out using heritabilities from the appropriate populations. Values for heritabilities of 0.25 for US and 0.35 for UK were used. Analyses were carried out for:

- (1) UK variances
- (2) US variances
- (3) UK and US covariance using estimates for variances based on (1) and (2)
- (4) combined UK and US estimating variances and covariances.

In one sense, the original calibration method is implicitly estimating heritabilities and so analyses for (1), (2) and (4) were repeated also estimating heritabilities ((5), (6), (7)).

Results

Estimates for genetic variance and covariance for milk yield (measured in lbs in the US and kg in the UK) are given in Table 1 for the six data sets and 4 analyses assuming heritabilities are known. In Table 2, corresponding estimates are given when heritabilities are also estimated.

The results show that as the size of the data set reduces, the estimate of genetic variance in the UK population reduces, but the US estimate is less variable. Ignoring pedigree information in the analyses leads to an approximate doubling of the estimate of genetic variance, presumably due to ignoring genetic group effects. There are corresponding changes in the genetic covariance (σ_{UKUS}) as the data set changes, leading to smaller changes in genetic regression. The genetic correlation (r) is remarkably constant. The size of a likelihood-based confidence interval increases as the data set reduces and is very wide for set D.

When heritabilities are estimated (Table 2), genetic variances, covariances, regressions and correlations are similar to those in Table 1, with the major changes being in the estimates of residual variances (σ_{eUK}^2 and σ_{eUS}^2). The relatively large changes in residual variances and small

changes in likelihoods ($L(a) - L(b)$) from the two different models show that it is difficult to estimate both environmental and genetic variances from this type of data. The genetic regression estimate (b) for set D is very similar to the estimate from regression of UK daughter deviation on US proof.

These analyses concentrate on estimating the genetic regression, b , in different populations. The constructed calibration equation from regression of UK proof on US proof from a MACE analysis for bulls with daughters only in the US, b_c , will depend on this b value but also because

$$\begin{aligned} \text{UK proof} &= b_{\text{US proof}} + \& (\text{sire UK} - b \text{ sire US proof}) \\ &+ ' (\text{MGS UK proof} - b \text{ MGS US proof}) \\ &+ \text{group correction} \end{aligned}$$

on the correlation between US proof with the other 3 terms in the prediction.

Using 1996 data, this regression, b_c , is 12% less than the genetic regression, mainly because of the correlation of the US proofs with group correction terms. This difference is approximately a quarter of the difference in calibration slope.

Overall, these analyses suggest the MACE analysis is a more complete and appropriate analysis, being based on much more data and a more comprehensive model.

References

- Gilmour, A.R., Thompson, R., Cullis, B.R. and Welham, S. 1997. *ASREML Manual*, N.S.W. Department of Agriculture.
- Gilmour, A.R., Thompson, R. and Cullis, B.R. 1995. *Biometrics* 51, 1440-1450.
- Goddard, M. 1985. *Livestock Prod. Sci.* 13, 321-331.
- Schaeffer, C.R. (1994). *J. Dairy Sci.* 77, 2671-2678.

TABLE 1

Estimates of genetic parameters for milk yield using various data sets and models of 0.35 (UK) and assuming heritabilities 0.25 (US).

Set Analysis		A	B	C	C1	C2	D
1	σ^2_{UK}	39145	57915	48359	81669	86652	68059
	σ^2_{eUK}	408233	603906	504269	851606	903562	709685
2	σ^2_{US}	531040	551761	553465	1012250	1270770	565819
	σ^2_{eUS}	7965200	8276000	8760358	15185300	19063400	8488130
3	σ_{UKUS}	125260	162023	148750	272680	315440	178860
	b	0.236	0.293	0.268	0.273	0.248	0.316
	r	0.87	0.91	0.91	0.95	0.95	0.91
4	σ^2_{UK}	39150	58240	59470	87610	99730	67780
	σ^2_{US}	531190	551550	604080	955550	1167140	566590
	σ_{UKUS}	126340	161570	173000	275470	327750	178590
	σ^2_{eUK}	408260	607320	620120	832610	1042180	706820
	σ^2_{eUS}	7967330	8273260	9060800	12905300	17507700	8498770
	b	0.238	0.292	0.286	0.288	0.280	0.315
	Clb	0.222-0.245	0.274-0.300	0.246-0.318	0.266-0.310	0.260-0.301	0.265-0.365
	r	0.88	0.90	0.91	0.95	0.96	0.91
	L(4)-L(3)	0.10	0.34	0.24	0.80	0.74	6.13

TABLE 2

Estimates of genetic parameters for milk yield using various datasets and models

Set Analysis		A	B	C	C1	C2	D
5	σ^2_{UK}	35986	52270	50666	8492	75624	18914
	σ^2_{eUK}	624044	1043160	152000	255647	3198810	44518700
	L(5) - L(1)	2.51	1.43	0.12	0.11	0.36	2.86
6	σ^2_{US}	600767	634760	403458	665221	616703	570000
	σ^2_{eUS}	5170098	4771060	29181600	71728900	125465000	1710000
	L(6) - L(1)	36.00	9.70	4.08	0.89	5.72	0.24
7	σ^2_{UK}	36690	53280	46210	81960	85650	67480
	σ^2_{US}	589500	611060	429622	876510	889700	534520
	σ_{UKUS}	126530	169820	131275	255680	26486	176560
	σ^2_{eUK}	596900	1015640	1214450	11609500	1047340	697510
	σ^2_{eUS}	5486990	5695460	18193600	18181010	34235150	46621810
	b	0.214	0.270	0.310	0.291	0.298	0.330
	r	0.86	0.91	0.93	0.95	0.96	0.93
	L(7) - L(4)	30.06	7.66	5.50	0.79	5.70	0.30