Comparison of Alternative Strategies for an International Genetic Evaluation of Beef Cattle Breeds

F. Phocas[1,2], K. Donoghue[1] and H. Graser[1]
[1] Animal Genetics and Breeding Unit, University of New England, NSW 2350 Armidale, Australia
[2] on leave from SGQA, Institut de la Recherche Agronomique, 78352 Jouy-en-Josas Cedex, France

Introduction

Although the Limousin breed has been developing in several countries for the last four decades, the genetic evaluation of seed stock is usually performed within country. However some foreign cattle are simultaneously evaluated with the French animals (Italy, Luxembourg, Hungary) or with the Australian animals (New Zealand). Breeders may seek to compare their domestic to overseas seed stock in order to enlarge the choice of animals that better fit their own selection objectives. In collaboration with the Irish Cattle Breeding Federation (Ireland), the Meat and Livestock Commission (UK), the Institut de l'Elevage (France) and INTERBULL sub-committee of ICAR, a research program was conducted at INRA (SGQA, France) with the purpose of developing an European genetic evaluation in British, French and Irish Charolais and Limousin breeds (Renand *et al.,* 2003) and at UNE (AGBU, Australia) with the aim of comparing the different strategies for such an international genetic evaluation of beef cattle breeds. It is this second part of the research program which is reported here and which relates to the benefit of using a complex model (which treats raw data in each country as correlated trait) versus a simpler model (that includes only the heterogeneity of the environment and genetic variances) on raw data or on de-regressed proofs.

Population data

A total of 1,213,163 adjusted weaning weights, recorded in France between 1971 and 2002, and in Oceania between 1976 and 2002, were initially provided for the study (Table 1).

Table 1. Weaning weight statistics.

	n	Mean (kg)	s.d. (kg)
France	1,175,152	247.2	43.0
Oceania	38,011	232.3	40.8

Weaning weight (WW) was adjusted to 200 days in Oceania (Oc) and 210 days in France (Fr). The numbers of sires registered were 31,399 in France and 2,395 in Oceania. Consequently, there were about 37 French progeny weaning weights records per sire while there were only about 16 progeny weaning weight records per sire in Oceania. The number of herds was 6,242 in France and 347 in Oceania. In Oceania, 263 out of the 347 herds had recorded weaning weights for calves sired by either a French sire or a French maternal grandsire. From the 71 French bulls identified as sires of calves recorded in BREEDPLAN (Australian genetic evaluation), 66 were also used as maternal grandsires (MGS) in BREEDPLAN. Out of these 71 French bulls, only the more recent 36 bulls (born after 1978) were known in the French pedigree file used in IBOVAL (French genetic evaluation) and 27 bulls had their own weaning weight recorded in France. The direct connections between the two countries (that are established through bulls siring calves in both countries) were not numerous: they consist of 16 sires born between 1988 and 1994 (Table 2).

Table 2. Number of (grand)progeny of connection sires across France and Oceania.

French ID	Progeny		Grand-progeny	
	Oc	Fr	Oc	Fr
FR1988004715	114	21,061	17	9,202
FR8791002491	97	4,803	30	1,452
FR8791051813	61	104	13	31
FR8792004759	56	122	2	2
FR8793007496	47	108	15	82
FR8791002604	36	374	7	70
FR1992000066	29	196	9	98
FR3693000206	25	1,227	0	12
FR1692111209	23	13,058	1	611
FR1991000217	20	1,338	5	742
FR8790014838	15	3,806	0	515
FR2392021153	13	8	6	4
FR1992014267	11	887	6	6
FR2394011796	7	2,824	0	7
FR8790011897	4	9	3	5
FR1992002007	3	367	0	7

These bulls accounted for only 4% of the recorded calves, although some sired a large number of calves: 2/3 of the 561 Australian connection progeny records came from 5 of the 16 connection sires while 2/3 of the 50,325 French connection progeny records came from only 2 sires. These 16 connection sires were used as MGS in France and 12 were MGS in both countries. However, indirect connections could be found in the back pedigree of the 20 males only used as sires and MGS in Oceania, but identified in the French pedigree. Another source of genetic links could be the trade of replacement heifers, but original female identification numbers were not available for the current study.

Estimation of genetic parameters

Edits

Table 3 gives a summary of the dataset under consideration.

Table 3. Summary of the sampled dataset.

Number	France	Oceania
Weaning weights	65,543	17,846
Herds	203	165
Contemporary groups	8,707	3,743
Dams	22,470	7,821
Sires	1,571	844
Maternal Grand Sires	1,627	800

Records before 1993 were discarded, because very little information was present before that time to connect the French and the Australian datasets. The 10-year period 1993-2002 corresponded to the period of birth for 99% of the Australian calves sired by the connection bulls. A minimum of 5 and 10 calves per sire or MGS was required in Oceania and France, respectively. Finally, contemporary groups with a single record were eliminated from the analysis of Australian data and those with less than 3 records were eliminated from the analysis of French data. Moreover a sampling of French herds was considered due to their large number in comparison to the number of Australian herds. Indeed, there were 2,275 French herds with progeny records of the connection bulls. Because of the great difference in the progeny numbers of the connection bulls, it was decided to select the 203 herds with at least 200 records over the last 20 years and where number of French progeny of the connection bulls was between 10 and 500. In the Australian dataset, herds without any weaning weights for calves sired by either a French sire or a French MGS or with less than 20 records over the period 1983-2002 were removed from the analysis. Only 49 out of the 165 remaining herds had progeny of the connection sires.

Thus, after the edits, there were only 13 bulls that directly connected the 2 countries in the sampled datasets (the 3 connection sires with less than 10 records in each countries were removed). The remaining connection progeny were 468 in herds of Oceania and 7,261 in French herds, accounting for 9% of the calves recorded in the sampled dataset. Only 9 bulls had at least 20 progeny in each country and only 4 bulls at least 40 progeny in each country. Furthermore, there were only 7 of these 13 bulls who were common MGS across countries in the sampled datasets. Only 3 bulls had at least 10 grandprogeny in each country, and only one bull had at least 20 grandprogeny in each country.

Estimation within country

In a first step, an animal model with maternal effects was run separately on the French sub-dataset (performance vector $\mathbf{y_1}$) and the Oceanic sub-dataset (performance vector $\mathbf{y_2}$).

The model included fixed effects (b: contemporary group, birth month and age of dam), direct (u_d) and maternal (u_m) genetic effects and permanent environment provided by the dam (e_m):

$$[1] \quad \mathbf{y_i = Xb + Z_d u_d + Z_m u_m + W_m e_m + e}$$

The second moments of the random effects, $\mathbf{var}\begin{bmatrix} \mathbf{u_d} \\ \mathbf{u_m} \end{bmatrix} = \mathbf{G_0} \otimes \mathbf{A}$, with $\mathbf{G_0} = \begin{bmatrix} \sigma^2_d & \sigma_{dm} \\ \sigma_{dm} & \sigma^2_m \end{bmatrix}$, and var[$e_m$] = $I \cdot \sigma^2_{em}$ and var[e] = $I \cdot \sigma^2_e$ were REML estimates with ASREML (Gilmour *et al.,* 2000).

In both countries, the contemporary group (CG) was defined as a “herd-year-season-management group-sex of the calf” combination, where the management group relates to calf-dam couples which are all managed in a homogeneous way and in the same paddock. It should be noted that the CG definition used in real French national evaluations does not account for the sex nor the birth season of the calf. But, it was shown on the current dataset that considering these two effects only as average effects across years and herds gave biased EBVS due to sire by sex and sire by season significant interactions. In France, the season effect defined within CG was a 2 month-period starting by grouping January and February together and so on. An overall birth month effect was also fitted for the French data. In Oceania, the season effect was a 45 day-period starting at the first calving date within herd-year. Although Oceanic weaning weights were already pre-corrected for age of dam, this effect was fitted for both countries because some extra-variation due to the age of the dam remained in the Oceanic data. The age of the dam was classified into 11 classes according to significantly different LS means estimates in the French data: < 28 months; 28-30 months; 31-33 months; 34-36 months; 37-42 months; 43-48 months; 49-60 months; 61-108 months; 109-144 months; 145-156 months; > 156 months.

The estimates of genetic parameters within country are reported in Table 4. In the French herds sampled in this study, the estimate of direct genetic variance is about 60% over the Oceanic estimate, whereas the estimates of the permanent environment and maternal genetic variance are close. The French estimate of the correlation between direct and maternal genetic effects is more negative than the Oceanic estimate.

Table 4. Estimates of genetic parameters within countries (standard errors).

	H^2d	H^2m	Rdm	C^2	σ_p^2
Fr.	.31 (.02)	.10 (.01)	-.21 (.06)	.10 (.01)	582
Oc.	.22 (.03)	.13 (.02)	-.10 (.12)	.13 (.02)	509

Estimation across countries

The estimation could have been done by a bi-trait animal model by considering each country’s records as a different trait with a diagonal residual variance matrix across traits. However, the memory requirements, the low convergence speed and the CPU time needed did not allow us to run this bi-trait model with the animal and dam genetic effects. Consequently, we used an equivalent single trait animal model accounting for across country interactions (AMACI model) with fixed or genetic effects.

$$[2] \quad \mathbf{y = Xb + Z_d u_d + Z_m u_m + W_m e_m + e}$$

where **y** is the vector of performance ordered by country $\mathbf{y' = (y_1', y_2')}$, **b** contains the CG and month effects within country and the interaction terms between country and age of the dam. This model took into account the relationship matrix across animals and dams (2 generations pedigree), and the across countries heterogeneity of residual and permanent environment variances. In order to test the heterogeneity of the genetic and residual variances across countries and to estimate the genetic correlations, a heteroskedastic model was used where $\mathbf{Var(e)} = \bigoplus_{i=1}^{2} \mathbf{I}_{ni}\sigma^2_{ei}$, being n_i the number of records in the i^{th} country and $\mathbf{Var(u) = G_{12} \otimes A}$, where **u’** is the transposed vector of direct and maternal genetic effects in each of the two countries $\mathbf{(u_{d1}, u_{m1}, u_{d2}, u_{m2})'}$, **A** is the relationship matrix between animals and dams and $\mathbf{G_{12}}$ is the genetic variance matrix across countries. As shown on a Charolais study (Quintanilla *et al.,* 2002), the heteroskedastic model was highly significant as compared to a model with homogeneous (co)variances.

Because of the lack of information to estimate the correlation between maternal effects across countries, it was necessary to ask for a positive definite genetic matrix in the estimation procedure, leading to the impossibility to get standard errors for the estimates. The estimated genetic parameters under an AMACI model are reported in Table 5. As already observed in the within country analyses, the residual and direct genetic variances were higher in French data. The estimate of the correlation between direct genetic effects across countries was 0.86, while the estimate of the correlation between maternal genetic effects across countries was 0.80. Applying Meyer and Hill's (1992) method for approximating the standard errors of direct and maternal genetic correlation estimates gave standard errors of 0.25 and 0.50, respectively.

Table 5. Estimates of genetic parameters under an AMACI model.

	France	Oceania
σ^2p	583	510
σ^2e	306	279
h^2d	0.31	0.22
h^2m	0.10	0.13
c^2	0.10	0.13
Rdm	-0.21	-0.12
$Rd_{Fr}d_{Oc}$	0.86	0.86
$Rm_{Fr}m_{Oc}$	0.80	0.80
$Rd_{Fr}m_{Oc}$	0.19	0.19
$Rm_{Fr}d_{Oc}$	-0.25	-0.25

Alternative international evaluations

An international evaluation under an animal model across country that assumes genetic correlations between countries are not unity (AMACI model) was compared to two simpler strategies for international evaluation of beef cattle. The first alternative was an evaluation based on a single-trait animal model across countries (SAMAC) accounting for heterogeneity of variances, but assuming direct or maternal genetic correlations between countries are equal to 1, while correlations between direct and maternal effects are set to 0. A common ranking of animals across countries was performed with this strategy, whereas the use of the AMACI model generated country specific estimated breeding value (EBV) distributions. The second alternative was to perform separate analyses of direct and maternal de-regressed EBVs under a Multiple-trait Across Country sire Evaluation (MACE). MACE was the model initially proposed by Schaeffer (1993) for international evaluation of dairy sires, where traits corresponded to de-regressed national EBV of bulls considered as correlated traits in the different countries. The current procedure used to de-regress national sire proofs makes the observations independent of genetic group effects and relationship among animals (to avoid their double counting), as well as accounting for accuracy of national proofs by considering weights based on the information used to derive national EBVs. Two MACE evaluations were run separately for de-regressed direct proofs and maternal proofs. Two different sets of weights had been considered to de-regress proofs: 1) the number of sire progeny for direct proofs and the number of daughters' progeny (maternal grand progeny) for maternal proofs; 2) the "MTEDC" weights proposed by Liu *et al.* (2003). Another method for weighting the proofs, derived from Tier and Meyer's (2004) reliability approximation method, was used to clarify the differences observed between the two previous weighting methods. Results for this last method are not presented here, because they were intermediate results, however closer to method 1, especially for maternal proofs. Its maternal weights were indeed highly correlated to the total number of daughters' progeny, while maternal weights derived by method 2 are more correlated to the number of daughters with at least a calf recorded.

Softwares and parameters

ASREML was used for most of the evaluations. Several different programs were necessary to run the MACE evaluation. The de-regression program used is the one that runs INTERBULL evaluations in dairy cattle and was sent by Thomas Mark for the purpose of this research. Zengting Liu also sent his program to derive weights by the MTEDC method. The MACE parameters have been derived using an AI-REML algorithm initially written by Ignacy Misztal and modified by Tom Druet (INRA-SGQA, personal communication) to weight each observation. Only a selected sample of bulls with enough information could be used from the dataset described in Table 3 for the variance component estimation by MACE on de-regressed proofs. This sample included bulls born between 1986 and 1995 with at least 5 progeny and 5 maternal grandprogeny recorded for weaning weight. It contained 642 bulls in France and 266 in Oceania, of which 7 bulls are connection sires.

The previous estimates (Table 5) of across country genetic correlations were considered for the AMACI evaluation, while, under the SAMAC evaluation, across country genetic correlations were set to 1 and direct-maternal correlations were set to 0. Phenotypic and residual variances remained constant over the two evaluations, because the AMACI direct-maternal covariances were equally carried over direct and maternal variances: the French direct and maternal variances became 166 and 49 kg^2 in the SAMAC model instead of 177 and 60 kg^2 in the AMACI evaluation; the variances in Oceania were only slightly modified because of the low value of the covariance (-10 kg^2).

Comparison of national and international evaluations (AMACI model)

When running separate national evaluations from the studied dataset, there were 28,180 animals (844 sires) evaluated within Oceania and 92,130 animals (1,571 sires) evaluated within France. In the across country evaluation, there were 121,277 animals (2,402 sires). When looking at the overall population, correlations between EBVs (either direct or maternal effects) predicted for each country by a national evaluation and by an AMACI international evaluation were over 0.99 (Table 6). Hence, in a given country, animals were ranking in the same manner under both evaluations. This was not the case when considering sires (r~0.93-0.95), especially connection sires. Their ranking in Oceania was strongly changed when including French information: the correlation between Oceanic AMACI and national direct (maternal) EBVs was only 0.86 (0.62).

Table 6. Correlations between EBV predicted within (W) and across (A) country (AMACI).

	All animals	All Sires	Connection sires
Dir. EBVs			
$Wd_{Fr}Ad_{Fr}$	0.998	0.953	0.934 (a)
$Wd_{Oc}Ad_{Oc}$	0.995	0.950	0.861 (b)
$Wd_{Fr}Ad_{Oc}$	0.994	0.950	0.945 (a)
$Ad_{Fr}Ad_{Oc}$	0.989	0.985	0.975 (b)
$Ad_{Fr}Wd_{Oc}$	0.968	0.917	0.764 (b)
Mat. EBVs			
$Wm_{Fr}Am_{Fr}$	0.997	0.942	0.956 (a)
$Wm_{Oc}Am_{Oc}$	0.988	0.930	0.619 (b)
$Wm_{Fr}Am_{Oc}$	0.853	0.727	0.613 (a)
$Am_{Fr}Am_{Oc}$	0.878	0.850	0.816 (b)
$Am_{Fr}Wm_{Oc}$	0.960	0.891	0.651 (b)

(a) 13 connection sires; (b) 28 Oc. sires identified.

Comparison of AMACI and SAMAC models for international animal evaluation

Comparing AMACI and SAMAC international evaluations revealed a slight re-ranking of animals for maternal EBVs, and almost the same ranking of animals for direct EBVs (Table 7). The slight re-ranking on maternal EBVs was due to the nul correlations required between direct and maternal genetic effects under the SAMAC model. Unpublished results showed indeed that this re-ranking increased as the direct-maternal correlation within country became more negative: if the AMACI evaluation was run assuming Rdm=-0.35 instead of –0.21 within France, the correlation between AMACI and SAMAC maternal French EBVs reduced to 0.85. As a consequence, a SAMAC model is not a satisfying strategy for international evaluation when direct-maternal correlations are not close to zero within country, even if genetic correlations across countries are close to unity.

Table 7. Correlations of AMACI & SAMAC EBVs.

	121,277 animals	2,402 sires
Fr. Direct EBV	0.994	0.993
Oc. Direct EBV	0.990	0.993
Fr. Mat. EBV	0.962	0.949
Oc. Mat. EBV	0.952	0.949

Comparison of AMACI and MACE models for international sire evaluation

Table 8 presents the parameters estimated by a MACE model depending on the weighting method used to de-regress the proofs. The direct and maternal genetic correlation estimates between de-regressed proofs were much smaller (about 0.5) than the estimates obtained on raw data. It can be most likely explained by the fact that all information on the dam proofs and pedigree were not accounted for in MACE.

Table 8. MACE parameters (*) according to the weighting method.

	Progeny numbers	MTEDC weights
Vd_{Fr} (Vr)	35.0 (241.2)	35.0 (555.2)
Vd_{Oc} (Vr)	20.8 (118.3)	18.0 (327.5)
Vm_{Fr} (Vr)	7.8 (250.5)	5.9 (481.4)
Vm_{Oc} (Vr)	10.2 (237.7)	7.8 (452.7)
$Rd_{Fr}d_{Oc}$ (s.e.)	0.45 (.23)	0. 64 (.29)
$Rm_{Fr}m_{Oc}$ (s.e.)	0.49 (.62)	0.48 (1.40)

(*) Vd and Vm are the sire variances for direct and maternal de-regressed proofs. Vr are the corresponding residual variances.

The Interbull de-regression program gave an estimate of sire variance for a known heritability of the trait. In this study, these within country estimates were quite different from the MACE estimates, especially for the MTEDC method. Deriving sire variances from within country analyses by Druet's program or ASREML proved that the MACE estimates were correct. In consequence, the heritabilities that can be estimated from the bulls' de-regressed proofs are very different from the heritabilities assumed in the de-regression step. In a first run, these heritabilities were the estimates obtained on raw data (Table 4). Due to the inconsistency observed between the estimates of sire variance from the de-regression and the MACE programs, the idea of using "corrected" heritability accounting for the fact that direct proofs should not be de-regressed for the random permanent environment and maternal effects (and vice-versa for maternal proofs), because direct and maternal de-regressed proofs are evaluated separately by MACE. This idea was proposed by Liu *et al.* (2003) for deriving the direct and maternal weights. Results became more consistent especially for maternal proofs when using these corrected heritabilities. The estimate of the genetic correlation based on de-regressed proofs was far more sensitive to the sampling of bulls (number of cohorts, minimum information requirement) for the MTEDC method than for the progeny numbers. When comparing MACE EBVs derived from the two weighting methods, results were very similar with correlations either for direct or maternal EBVs about 0.98 within country. Correlations between the national and MACE direct EBVs were over 0.99 for the two countries and the two weighting methods. For each of the two countries, correlations between the national and MACE maternal EBVs were 0.998 when considering the grandprogeny numbers in the de-regression and only about 0.98 when considering the MTEDC weighting method. In the subsequent comparisons, only MACE results based on progeny numbers were therefore presented.

Table 9 presents the correlations between national and international EBVs of the 901 sires born between 1986 and 1995 that were evaluated with all international models. The two international EBVs (AMACI & MACE) were highly correlated to their corresponding national EBVs, either for direct (over 0.99) or maternal (over 0.97) effects.

Table 9. Correlations between within (W) & across country sire EBVs by AMACI (A) or MACE (M).

	Direct EBV	Maternal EBV
$W_{Fr}A_{Fr}$	1.00	1.00
$W_{Oc}A_{Oc}$	0.99	0.97
$A_{Fr}A_{Oc}$	0.98	0.90
$W_{Fr}M_{Fr}$	1.00	1.00
$W_{Oc}M_{Oc}$	0.99	1.00
$M_{Fr}M_{Oc}$	0.78	0.81

However, consequent re-ranking of sires between countries was observed for some top bulls. The choice of the top 100 of sires across countries would be really different depending on the implemented model (Tables 10 & 11) and on the selection criterion (direct or maternal EBVs). The MACE model favours within country sire selection, because correlations between French and Oceanic MACE EBVs are only about 0.8, while correlations between French and Oceanic AMACI EBVs are over 0.9 (Table 9).

Table 10. Sire's origin in the top 100 of sires ranked on EBV predicted by AMACI.

Sire's Origin	Top Ad_{Fr}	Top Ad_{Oc}	Top Am_{Fr}	Top Am_{Oc}
Fr.(connection)	77 (4)	70 (4)	75 (1)	63 (1)
Oc.	23	30	25	37

Table 11. Sire's origin in the top 100 of sires ranked on EBV predicted by MACE.

Sire's Origin	Top Md_{Fr}	Top Md_{Oc}	Top Mm_{Fr}	Top Mm_{Oc}
Fr (connection)	99 (4)	43 (3)	94 (2)	41 (1)
Oc.	1	57	6	59

Conclusion

The estimation of genetic correlations and also the comparability of EBVs rely on the existing genetic links across countries. These links are very tenuous when considering only Limousin male pedigree across France and Oceania. Emphasis should be given initially to the generation of a valid cross reference file and tracing back the pedigree of all imported seed stock. With environment and genetic (co)variances shown to differ markedly across countries and the fact that the hypothesis of G*E interaction could not be rejected, an animal model with maternal genetic effects, heterogeneous variances and different genetic correlations between countries was the model of choice (AMACI). Consequently, breeders will have to deal with sires that have different breeding values for the same genetic effects. The AMACI model allows the prediction of international EBV for all animals: sires, dams and calves. However, if this model cannot be implemented either for political reasons (unpublished raw data) or computational reasons (too many countries or animals), the MACE approach may be a sufficient alternative to compare sires across countries.

Acknowledgments

The authors would like to thank Z. Liu, T. Mark, S. Minery and B. Tier for providing part of the MACE programs and for helping in the understanding of the results.

References

Gilmour, A.R., Cullis, B.R., Welham, S.J. & Thompson, R. 2000. *ASREML Reference Manual.*

Liu, Z., Pasman, E., Reinhardt, F. & Reents, R. 2003. Application of the Multiple Trait Effective Daughter Contribution to a Linear Model with Direct and Maternal Genetic Effects for Calving Traits. *Interbull Bulletin 30,* 75-81.

Meyer, K. & Hill, W.G. 1992. Approximation of sampling variances and confidence intervals for maximum likelihood estimates of variance components. *J. Anim. Breed. Genet. 109,* 264-280.

Quintanilla, R., Laloë, D. & Renand, G. 2002. 7th WCGALP Montpellier, France, CD-ROM communication n°18-10.

Renand, G, Laloë, D., Quintanilla, R. & Fouilloux, M.N. 2003. A first attempt of an international genetic evaluation of beef breeds in Europe. *Interbull Bulletin 31,* Roma, Italy.

Schaeffer, L.R. 1993. Multiple-trait, Across Country Evaluation of dairy sires. *Interbull Bulletin 8.* 21p.

Tier, B. & Meyer, K. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. *J. Anim. Breed. Genet. 121,* 77-89.