

Genotype by Environment Interaction for Milk Production Traits in Guernsey Cattle

Freddy Fikse¹, Romdhane Rekaya² and Kent Weigel²

¹ Interbull Centre, Sweden

² University of Wisconsin, USA

Introduction

Current methods for routine international evaluations are based on pre-processed information. This puts several restrictions on the method of analysis that can be used. For example, comparisons need to be made on genetic evaluation unit bases. Neighboring countries with similar production environments cannot be treated as genetically identical traits, and environmental variation within country cannot be utilized. Also, the genetic correlation between countries does not only depend on genotype by interaction, but on differences between national evaluation systems as well.

Access to individual performance records allows for international evaluations that can bypass some of the shortcomings of pre-processed input. Two procedures that consider variation in production environments between and within country, and in addition allow for more uniform processing of data, have been suggested for international genetic evaluation using performance records. The multiple-trait herd cluster model (Weigel & Rekaya, 2000) stratifies herds according to production environment based on information on environmental descriptors. In the genetic evaluation model production in each herd cluster is treated as genetically separate trait. Reaction norms (Kolmodin *et al.*, 2001) models the phenotype as a function of environmental descriptor.

The objective of this study was to compare the results from the multiple-trait herd cluster model and the reaction norm model with results from models that treat country as production environment.

Material and Methods

Data

First lactation test day observations on milk yield for Guernsey cows with a first lactation between January 1990 and December 1999 was available from four countries: Australia (AUS), Canada (CAN), United States (USA) and South Africa (ZAF). Data files contained information on test day yield, days in milk at test day, calving date, herd and region, date of birth and sire, dam, maternal grandsire and granddam. Table 1 presents the number of cows, test day records and herds in each data set prior to edits.

Table 1. Descriptive statistics of data before edits.

	# cows	# test day records	# herds
AUS	10,343	147, 197	606
CAN	7,097	97,641	154
USA	87,608	922,931	1,863
ZAF	3,030	32,216	56

After data edits, incomplete lactation records were extended to 305 day lactation length, and analyses were based on 305 day lactation yields. For a description of the data edits and procedure for extension see Fikse *et al.* (2001).

Lactation yields were only included in subsequent analyses for cows sired by bulls with daughters in at least 10 herds. This was done to assure that pedigree information was available for bulls. Pedigree information was completed with records for bulls taken from the February 2000 routine international evaluation for the Guernsey breed.

Statistical Models

Model ST:

$$y = \text{fixed} + \text{sire} + e$$

The simplest model in this study was a single trait sire model, with one sire and one residual variance common to all countries. Fixed effects included were herd×year×season of calving, age at calving and heterosis class.

Model SThet:

$$y_i = \text{fixed}_i + \text{sire} + e_i, \quad i=1 \dots 4$$

In this case, country specific residual variance and fixed effect solutions were estimated, relaxing some of the assumptions of the previous model. That is, for each country a separate set of solutions for age at calving and heterosis class and residual variance was computed. A sire variance common to all countries was assumed.

Model MT:

$$y_i = \text{fixed}_i + \text{sire}_i + e_i, \quad i=1 \dots 4$$

The performance records equivalent to Mace allowed for a country specific sire variance in addition to residual variance and fixed effect solutions. This model also yielded genetic correlations between countries.

Model HC:

$$y_i = \text{fixed}_i + \text{sire}_i + e_i, \quad i=1 \dots C$$

Herds were grouped into clusters with the K-means algorithm that minimized the within-cluster sums of squares (NAG, 19xx). The distance was computed as the weighted squared difference for nine environmental descriptors. These environmental descriptors were found to be able to distinguish between production environments in another study (Fikse *et al.*, 2001). Weight given to each variable was based on both the genetic correlation between milk production in extreme environments (5% and 95% percentile values for herd averages of environmental descriptor) and the phenotypic correlation between herd averages for all environmental descriptors, taken from the study by Fikse *et al.* (2001).

Lactation yields were subsequently analyzed with a multiple trait sire model, in which each cluster was treated as a separate trait. The same three fixed effects were included as for the previous models. Fixed effect solutions, sire and residual variance were specific for each cluster.

Model RR:

$$y = \text{fixed} + \text{sire}(\text{env}) + e$$

Reaction norms for milk production were estimated with a random regression model. As environmental gradient the herd average for milk peak yield was chosen, which was the environmental descriptor with the lowest genetic correlation between milk production in extreme environments (Fikse *et al.*, 2001). The regression model included both an intercept and linear term.

Implementation

Inferences about dispersion parameters were made from posterior distributions obtained with Gibbs sampling. Bounded uniform priors were used for the fixed effects. For the residual and sire variance weakly informative prior distributions were used.

A single chain with length 150,000 was run for all analyses. Burn-in and thinning parameters were determined with the Gibanal program (VanKaam, 1998).

Model comparison

Support for each model was determined with the Bayesian Deviance Information Criterion (DIC). DIC combines goodness of fit and model complexity, and was computed as:

$$DIC = D(\bar{\theta}) + 2p_D, \quad ,$$

where p_D is the effective number of parameters, and $D(\theta)$ the expectation of Bayesian deviance. For computational details the reader is referred to Spiegelhalter *et al.* (1998).

Results and Discussion

The lowest average milk production is realized in the countries practicing grazing systems (Table 2). Standard deviation of 305d milk production was very similar for all four countries (~ 1,200 kg).

Table 2. Number, mean and standard deviation of lactation records per country.

	#	μ	σ
AUS	8,432	4,210	1,175
CAN	6,417	5,385	1,220
USA	51,493	5,750	1,276
ZAF	2,279	4,450	1,215

ST, SThet & MT

Estimated heritability ranged between .30 and .33 for CAN and USA for the ST, SThet and MT analyses (Table 3). Estimated sire and residual variances for both CAN and USA were also consistent between all three analyses (Table 4 & 5). A common sire variance for all four countries was estimated for the ST and SThet model, but the residual variance was left to vary in the SThet model, and the lower residual variance in that case resulted in a high heritability for AUS and ZAF. Residual variance for the SThet and MT model were consistent for all models, yielding lower heritability for AUS and ZAF compared to the SThet analysis. These results indicate that both sire and residual variance are heterogeneous between countries.

The genetic correlation between countries obtained with the MT model were between .86 and .90 except for CAN-ZAF (Table 6).

HC

Three clusters were identified. One cluster could be characterized as large herds with much variation of production within herd. The other two clusters contained smaller herds, but differed in intensity determined from the within-herd standard deviation and days in milk when the peak yield was realized.

Table 3. Estimated heritability for the models ST, SThet and MT.

	ST	SThet	MT
AUS		.53	.41
CAN	.32	.33	.31
USA		.30	.32
ZAF		.47	.15

Table 4. Estimated sire variance for the models ST, SThet and MT.

	ST	SThet	MT
AUS			47.6
CAN			66.7
USA	68.5	68.3	71.7
ZAF			22.8

Table 5. Estimated residual variance for the models ST, SThet and MT.

	ST	SThet	MT
AUS		435.9	436.7
CAN		741.4	735.8
USA	754.9	809.4	809.9
ZAF		501.0	506.3

Table 6. Estimated genetic correlations for the MT model.

	CAN	USA	ZAF
AUS	.90	.87	.87
CAN		.87	.78
USA			.86

Estimated heritability in cluster 1 and 3 was .36 and .34, respectively, but was smaller for the second cluster (Table 7). The genetic correlations between the second and the other clusters were .91 and .92 (Table 7), indicating that the production environment was slightly different.

Table 7. Estimated genetic parameters for the HC model.

	Cluster 1	Cluster 2	Cluster 3
σ_s^2	64.6	63.8	75.2
σ_e^2	700.1	787.8	790.5
h^2/r_g	.36	.92	.95
		.29	.91
			.34

RR

Sire variance for the intercept and the slope of the reaction norms was 66.4k and 4.8k, respectively. The genetic correlation between both terms was .57. The residual variance was with a value of 755.7k very similar to that of the ST model (Table 5). Heritability in an average environment equaled .36, and the estimated correlation between milk production in extreme environments was .62.

Model comparison

The ST model had poorest goodness of fit judged from the residual sums of squares (Table 8). However, improved fit for some of the other models was at the expense of an increased effective number of parameters. The SThet model performed best as it had the lowest value for DIC (-749.).

Conclusions

Results indicated that variances were heterogeneous between production environments, whether classified by country or through environmental descriptors. There was little justification for treating production environments as separate traits, and we hypothesize that a model that allows for heterogeneous sire and residual variances will have a better fit than any of the models presented in this study.

Table 8. Results from the comparison between models.

	Deviance	Eff. no. parameters	DIC
ST	0.	6372.	0.
SThet	-788.	6411.	-749
MT	-796.	6433.	-734.
HC	-181.	6464.	-89.
RR	-64.	6491.	77.

References

- Fikse, W.F., Rekaya, R. & Weigel, K.A. 2001. Assessment of environmental descriptors of studying genotype by environment interaction. *52th Meeting of the EAAP*, Budapest, Hungary, 25-29 August 2001.
- Kolmodin, R., Strandberg, E., Madsen, P., Jensen, J. & Jorjani, H. 2001. Genotype by environment interaction in Nordic dairy cattle studied by use of reaction norms.
- Spiegelhalter, D.J., Best, N.G. & Carlin, B.P. 1998. Bayesian deviance, the effective number of parameters and the complexity of arbitrary complex models. *Technical report, Medical Research Council, Biostatistics Unit*, Cambridge, UK.
- VanKaam, J.B.C.H.M. 1998. Gibanal - Analyzing program for Markov Chain Monte Carlo sequences. Department of Animal Sciences, Wageningen Agricultural University.
- Weigel, K.A. & Rekaya, R. 2000. A multiple-trait herd cluster model for international dairy sire evaluation. *J. Dairy Sci.* 83, 815-821.