Rothamsted Repository Download

A - Papers appearing in refereed journals

Calculation of multiple-trait sire reliability for traits included in a dairy cattle fertility index

G. Banos[1]†, S. Brotherstone[2, 3], R. Thompson[4, 5], J. A. Woolliams[5], E. Wall[2] and M. P. Coffey[2]

[1]*Department of Animal Production, School of Veterinary Medicine, Aristotle University of Thessaloniki, Box 393, GR-54124 Thessaloniki, Greece*
[2]*Sustainable Livestock Systems, Scottish Agricultural College, Bush Estate, Penicuik, Midlothian EH26 0PH, UK*
[3]*Institute of Cell, Animal and Population Biology, University of Edinburgh, Ashworth Laboratories, King's Buildings, Edinburgh EH9 3JT, UK*
[4]*Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK*
[5]*Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK*

† Corresponding author; e-mail : banos@vet.auth.gr

Abstract

The advent of genetic evaluations for fertility traits in the UK offers valuable information to farmers that can be used to control fertility problems and safeguard against involuntary culling. In addition to estimated genetic merit, proof reliabilities are required to make correct use of this genetic information. Exact reliabilities, based on the inverse of the coefficient matrix, cannot be estimated for large data sets because of computational restrictions. A method to calculate approximate reliabilities was implemented based on a six-trait sire model. Traits considered were interval between first and second calving, interval between first calving and first service, non-return rate 56 days post first service, number of inseminations per conception, daily milk yield at test nearest day 110 and body condition score. Sire reliabilities were calculated in four steps. Firstly, the number of effective daughters was calculated for each bull, separately for each trait, based on total number of daughters and daughter distribution across herd-year-seasons. Secondly, multiple-trait reliabilities were calculated, based on bull daughter contribution, applying selection index theory on independent daughter groups. Thirdly, (great-) grand-daughter contribution was added to the reliability of each bull, using daughter-based reliability of sons and maternal grandsons. An adjustment was made to account for the probability of bull and son or grandson having daughters in the same herd-year-season. Without the adjustment, reliabilities were inflated by proportionately 0·15 to 0·25. Finally, parent (sire and maternal grandsire) contribution was added to the reliability of each bull. The procedure was first tested on a data subset of 28 061 cow records from 285 bulls. Approximate reliabilities were compared with exact estimates based on the inverse of the coefficient matrix. Mean absolute differences ranged from 0·014 to 0·020 for the six traits and correlation between exact and approximate estimates neared unity. In a full-scale application, sire reliability for the fertility traits increased by proportionately 0·47 to 0·79 over single-trait estimates and the number of bulls with a reliability of 0·60 or more increased by 42 to 115%.

Keywords: *dairy cattle, fertility index, reliability, sire evaluation.*

Introduction

Long-term selection on milk yield combined with the unfavourable genetic correlation between production and fertility has led to declining dairy cattle fertility in the UK (Royal *et al.*, 2000). Infertile or sub-fertile cows are involuntarily culled, thereby inflating milk production costs, decreasing the profitability of dairy cow farming and increasing the environmental impact of dairying through cow loss.

Genetic evaluation for fertility traits may provide useful selection tools to help a) farmers to monitor the fertility of their cows and b) assess and enhance the genetic trend of the population as a whole. The recently developed UK fertility index (Wall *et al.*, 2003) will make such information available to the domestic and global industry. This index is based on a six-trait analysis of four fertility traits, milk yield and body condition score.

Besides estimated breeding values (or predicted transmitting abilities), however, reliabilities are required to make best use of the genetic evaluation information and manage the usage of bulls with favourable genetic merit but low accuracy evaluation. Reliability of the genetic evaluation is related to the correlation between the estimated and the true (unknown) genetic merit of the animal.

Exact asymptotic reliability estimation requires the inverse of the coefficient matrix of the mixed model equations. With large datasets, however, computational demands become prohibitive with this course of action. Several approximations have been proposed in the literature based on predictive functions of the diagonal of the inverse (Weller *et al.*, 1985; Greenhalgh *et al.*, 1986; Meyer, 1987; Boichard and Lee, 1992) and cumulative contribution from various information sources (Misztal and Wiggans, 1988; Harris and Johnson, 1998a).

The objective of this study was to develop and implement a computationally feasible procedure for the calculation of sire reliabilities for traits included in the UK dairy cattle fertility index.

Material and methods

Data
Insemination and calving records of primiparous Holstein cows collected from 1992 to 2002 were obtained from the Cattle Information Service, National Milk Records and Holstein UK databases. Fertility traits were defined as in Wall *et al.* (2003). Briefly, traits included were the number of days between first and second calving (CI), the number of days between first calving and first service (DFS), the number of inseminations per conception (INS) and an all-or-none trait measuring non-return to service 56 days after the first insemination (NR56). The latter was scored 0 if the cow had returned and 1 if the cow had not returned to service. Milk yield (MILK), defined as daily yield at test nearest day 110, and body condition score (BCS) records were also obtained. Body condition score was adjusted for the recording officer by scaling records so that individual field officer standard deviations were equal to the mean standard deviation of all field officers

(Brotherstone, 1994). Both MILK and BCS are genetically correlated with fertility (Pryce *et al.*, 2000; Brotherstone *et al.*, 2002) and were included in the analysis to improve the accuracy of the genetic evaluation and, partially, to account for selection based on milk production. Following a series of edits and validations described by Wall *et al.* (2003), the final data set comprised 1 793 460 records of first lactation Holstein cows that were daughters of 27 765 bulls. In addition, 3801 sires and maternal grandsires of these bulls, having no daughter records of their own in the database, were included in the analysis. Cow records were distributed in 282 162 herd-year-season classes. All cows had MILK record but some were missing the other traits, as shown in Table 1.

Reliability calculation
Calculation of approximate sire reliability was completed in four steps.

Step 1. Number of effective daughters was calculated for each bull, based on total number of daughters and daughter distribution across herd-year-seasons. The result was conceptually equivalent to absorbing herd-year-season into sire equations in a mixed model analysis. The number of effective daughters was calculated separately for each trait.

Step 2. Multiple-trait reliabilities were calculated based upon bull daughter contribution, applying selection index theory on independent daughter groups. Sire and residual parameters assumed were those estimated by Wall *et al.* (2003) and are shown in Table 2. All bulls were required to have daughters with milk yield. Number of effective daughters per bull and trait (step 1) was processed in descending order and was expressed as a deviation from the number of effective daughters for the previous trait. This created independent daughter groups for the calculation of prediction error variance. Different

Table 1 *Data description for daily milk yield at test nearest day 110* (MILK), *number of days between first and second calving* (CI), *number of days between first calving and first service* (DFS), *non-return rate 56 days after first service* (NR56), *number of inseminations per conception* (INS) *and body condition score* (BCS)

Trait (units of measurement)	No. of records	Mean	s.d.	Value range Min.	Value range Max.	No. of bulls
MILK (kg)	1 793 460	22·56	5·79	5	60	27 765
CI (days)	1 218 620	393·41	55·06	300	600	18 540
DFS (days)	1 594 232	83·29	31·31	20	200	25 359
NR56 (0/1)	1 594 232	0·68	0·47	0	1	25 359
INS (count)	1 159 097	1·66	1·00	1	10	17 609
BCS (score)†	214 882	4·46	1·66	−1·50	10·50	7 898

† Adjusted for recording officer as per Brotherstone (1994).

Table 2 *Estimates of heritability* (*diagonal*), *genetic correlations* (*above diagonal*) *and residual correlations* (*below diagonal*), *for the six traits* (*source Wall* et al., *2003*)

	MILK	CI	DFS	NR56	INS	BCS
MILK	0·33	0·27	0·49	−0·25	0·06	−0·44
CI	0·05	0·03	0·67	−0·45	0·61	−0·14
DFS	0·04	0·48	0·04	0·24	−0·12	−0·63
NR56	−0·03	−0·34	0·02	0·02	−0·94	−0·30
INS	0·04	0·67	−0·06	−0·68	0·02	0·29
BCS	−0·16	−0·06	−0·08	0·01	−0·01	0·24

MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.

trait patterns were allowed for different bulls. A detailed description of this procedure is given in **Appendix 1**.

Step 3. (Great-)grand daughter contribution was added to the reliability of each bull, based on daughter reliability (step 2) of sons and maternal grandsons. The Information Source method of Harris and Johnson (1998a) was followed, but an adjustment procedure was added to account for the probability of bull and son or grandson having daughters in the same herd-year-season. This difference between our approach and that of Harris and Johnson (1998a) was introduced to decrease bias in the reliability calculation of bulls that had daughters in the same herd-year-season as did their male progeny. Animals were processed from youngest parent to oldest. A detailed description of this procedure is given in **Appendix 2**.

Step 4. Parent (sire and maternal grandsire) contribution was added to the reliability of each bull, following the Information Source method of Harris and Johnson (1998a). Animals were processed from oldest to youngest. A detailed description of this procedure is included in **Appendix 2**.

Test application

The approximate reliability procedure was first tested on a randomly chosen subset of 28 061 records and 285 bulls. The same data set was also analysed with the Prediction and ESTimation (PEST) software package (Groeneveld, 1990), which computes standard errors of bull solutions, from which exact reliabilities may be derived. A multivariate analysis of the six traits using the models for national genetic evaluation for fertility (Wall *et al.*, 2003) was performed. Briefly, models included the effects of herd-year-season of calving (or visit for BCS), linear and quadratic regression on age at calving, linear and quadratic regression on days in milk (for MILK and BCS) and sire of the cow. Genetic parameters given in Table 2 were assumed. Comparisons between approximate and exact reliabilities included mean absolute difference, standard deviation of difference, product moment correlation and linear regression of the approximate reliability on the exact.

Results

Comparisons between approximate and exact reliabilities are shown in Tables 3 and 4. Results in Table 3 consider only daughter information, therefore approximate reliabilities were intermediate estimates from steps 1 and 2 only, and the exact reliabilities assumed all sires were unrelated. Thus,

Table 3 *Comparison between intermediate estimates of approximate* (*steps 1 and 2*) *and exact* (*excluding genetic relationships among sires*) *multiple-trait reliabilities, by trait; intercept and slope refer to the regression of the former on the latter*

Trait†	Mean absolute difference	Standard deviation of difference	Range of difference		Product moment correlation	Intercept‡	Slope§
			Min.	Max.			
MILK	0·005	0·004	−0·015	0·015	0·999	−0·007	1·021
CI	0·002	0·003	−0·008	0·011	0·999	0·008	1·010
DFS	0·003	0·003	−0·010	0·009	0·999	0·000	1·014
NR56	0·009	0·007	−0·003	0·025	0·999	0·006	1·020
INS	0·008	0·006	−0·003	0·028	0·999	0·005	1·018
BCS	0·003	0·003	−0·010	0·009	0·999	0·002	1·004

† MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.
‡ Standard error = 0·002 to 0·004.
§ Standard error = 0·003 to 0·009.

Table 4 *Comparison between final estimates of approximate and exact multiple-trait reliabilities, by trait; intercept and slope refer to the regression of the former on the latter*

Trait†	Mean absolute difference	Standard deviation of difference	Range of difference		Product moment correlation	Intercept‡	Slope§
			Min.	Max.			
MILK	0·014	0·006	–0·012	0·028	0·999	–0·009	1·030
CI	0·016	0·013	–0·036	0·045	0·995	0·008	1·018
DFS	0·014	0·010	–0·028	0·064	0·997	0·001	1·025
NR56	0·020	0·013	–0·028	0·074	0·995	0·007	1·033
INS	0·020	0·014	–0·027	0·059	0·994	0·008	1·034
BCS	0·014	0·010	–0·038	0·039	0·998	0·007	1·013

† MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.
‡ Standard error = 0·002 to 0·005.
§ Standard error = 0·004 to 0·0010.

Table 3 exact reliabilities were computed for comparison purposes only and do not represent final values. Results in Table 4 consider all information available, therefore approximate reliabilities were the final estimates derived from steps 1 to 4 and exact reliabilities were computed from the inverse of the full coefficient matrix including genetic relationships among sires.

Table 3 results readily validate steps 1 and 2 of the method, which include the computation of the number of effective daughters and the application of selection index theory to multiple-trait reliability calculation based on daughter contribution. Mean absolute difference and standard deviation of difference were near zero suggesting very little variation in the results. Maximum bias was 0·025 and 0·028 for NR56 and INS, respectively, which were the two traits with the lowest heritability (0·02). Product moment correlation between approximate and exact reliability was practically unity, whereas the regression intercept and slope of the former on the latter were near expectation (0 and 1, respectively).

Table 4 results suggest that final approximate reliabilities were very close to the exact estimates. Compared with intermediate results of Table 3, however, approximate reliabilities appeared to overestimate exact reliabilities in some cases, especially for the low heritability traits. This bias was traced to step 3, where the reliability of a sire was updated to include information from his (great-)grand-daughters, and affected bulls with many sons and/or maternal grandsons. Slight upwards biases associated with the Information Source method, applied here, were also reported by Harris and Johnson (1998a) for milk production traits. The bias shown in Table 4 was after adjustment, in step 3, for the probability of sires and sons or grandsons having daughters in the same herd-year-season. Harris and Johnson (1998a) did not make this adjustment. Without adjustment, mean absolute differences were higher by 0·002 to 0·005 and maximum biases were larger by 0·002 to 0·026, compared with results in Table 4.

The method was applied to a large national database of 1·8 million first lactation records. Software was written in FORTRAN 90 and run in a UNIX environment. The entire process required 75 min of computer processing time, including all editing and preparatory steps, of which the actual reliability calculation took less than 1 min. The process also required 24 Mb of memory.

Table 5 shows the distribution of the calculated sire reliabilities for six traits. Although some bulls did not have daughters with CI, DFS, NR56, INS or BCS records (Table 1), they too had genetic evaluation and reliability estimated from the multiple-trait analysis. Reliability estimates of all 27 765 bulls, which had daughters with MILK records, are considered in Table 5. Because of this and the higher heritability (0·33), as many as 6389 sires had MILK reliability larger than 0·60, compared with 847, 1597, 840, 721 and 1683 for CI, DFS, NR56, INS and BCS, respectively. The heritability of BCS was the second highest (0·24) but only 28% of the bulls had daughters with BCS recorded, hence the relatively low reliability estimates for this trait. Amongst the four fertility traits, DFS had the highest heritability (0·04) and, more importantly, the largest number of validated records. In fact, this trait was available in 89% of the cows and 91% of the bulls had daughters with DFS recorded. For this reason, mainly, about twice as many bulls had reliability greater than 0·60

Table 5 *Number of bulls (N) and percentage (%) of total of 27 765 bulls, in parenthesis, by reliability range and trait*

Reliability range	Trait†					
	MILK	CI	DFS	NR56	INS	BCS
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
< 0·10	1026 (3·7)	3279 (11·8)	2317 (8·4)	3707 (13·4)	3964 (14·3)	2835 (10·2)
0·10–0·19	1600 (5·8)	4918 (17·7)	3549 (12·8)	4988 (18·0)	5262 (18·9)	4650 (16·7)
0·20–0·29	4230 (15·2)	8578 (30·9)	7058 (25·4)	8531 (30·7)	9887 (35·6)	7026 (25·3)
0·30–0·39	7286 (26·2)	8085 (29·1)	8983 (32·4)	7584 (27·3)	6431 (23·2)	8351 (30·1)
0·40–0·49	4167 (15·0)	1488 (5·4)	2949 (10·6)	1479 (5·3)	1141 (4·1)	2169 (7·8)
0·50–0·59	3067 (11·0)	570 (2·1)	1312 (4·7)	636 (2·3)	359 (1·3)	1051 (3·8)
0·60–0·69	2161 (7·8)	238 (0·9)	720 (2·6)	236 (0·9)	200 (0·7)	678 (2·4)
0·70–0·79	1571 (5·7)	202 (0·7)	282 (1·0)	221 (0·8)	194 (0·7)	520 (1·9)
0·80–0·89	1447 (5·2)	207 (0·7)	283 (1·0)	205 (0·7)	185 (0·7)	271 (1·0)
0·90–0·99	1210 (4·4)	200 (0·7)	312 (1·1)	178 (0·6)	142 (0·5)	214 (0·8)

† MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.

for DFS compared with CI. For the latter trait, although the heritability (0·03) was close to the heritability of DFS, only 67% of bulls had daughters with validated records. It should also be noted that DFS had a stronger genetic correlation with MILK and BCS than CI, which further enhanced its reliability estimates. The amount of NR56 data was the same as DFS but because of the lower heritability (0·02), generally weaker genetic correlations with MILK and BCS and, possibly, its binary nature, reliability estimates were lower for this trait. Finally, INS had the smallest amount of data, lowest heritability and weakest correlation with MILK and BCS, amongst all fertility traits, hence the lowest reliability estimates were for this trait.

In Table 6, average reliabilities estimated after each step are shown, reflecting the contribution of various sources of information to the final estimate. Average reliability estimates are low because many of the 27 765 bulls had very few effective daughters (even less than 1), whereas a certain proportion of them had daughters with fertility and BCS records missing (Table 1). Multiple-trait analysis (step 2) increased the average reliability for fertility traits by proportionately 0·47 to 0·79 but had minimal effect on MILK. As expected, multiple-trait analysis will be of value mainly to low heritability traits with missing observations. On average, (great-)grand-daughters (step 3) made little contribution to bull reliability because only a few bulls had a considerable number of sons and maternal grandsons. However, in specific cases, the effect of this contribution was sizeable, e.g. a certain bull had multiple-trait daughter-based reliability of 0·807 for MILK, which increased to 0·816 with the contribution of his great-grand-daughters,

via his five maternal grandsons, and to 0·910 with his grand-daughter contribution, via his 22 sons. In this particular example, average multiple-trait daughter-based reliability of grandsons and sons was 0·733 and 0·827, respectively. Bull parents (step 4) contributed considerably to the average reliability, as a result of the low mean value and its prevalence amongst the different sources of information. The theoretical maximum reliability of a pedigree index based on sire and maternal grandsire is 0·312 and occurs when both have a completely known breeding value. Its relevance, however, will decrease as information from other sources accumulates. Thus, the average pedigree contribution here ranged from 0·131 to 0·173 for the six traits. Moreover, in the previous example, the bull's parent contribution only

Table 6 *Average reliability of 27 765 bulls for six traits, estimated after each step of the procedure*

Trait†	Daughter contribution		(Great-) grand-daughter contribution (Step 3)	Parent contribution (Step 4, final estimate)
	Single-trait (Step 1)	Multiple-trait (Step 2)		
MILK	0·307	0·311	0·317	0·448
CI	0·062	0·091	0·095	0·268
DFS	0·080	0·143	0·149	0·316
NR56	0·054	0·087	0·092	0·263
INS	0·046	0·072	0·077	0·250
BCS	0·074	0·129	0·136	0·298

† MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.

Trait†	Daughter contribution		(Great-) grand-daughter contribution (Step 3)	Parent contribution (Step 4, final estimate)
	Single-trait (Step 1)	Multiple-trait (Step 2)		
MILK	5093	5093 (0%)	5102 (<1%)	6389 (25%)
CI	596	688 (15%)	739 (24%)	847 (42%)
DFS	743	1069 (44%)	1137 (53%)	1597 (115%)
NR56	529	684 (29%)	733 (39%)	840 (59%)
INS	455	579 (27%)	626 (38%)	721 (58%)
BCS	1051	1282 (22%)	1351 (29%)	1683 (60%)

† MILK = daily milk yield at test nearest day 110; CI = number of days between first and second calving; DFS = number of days between first calving and first service; NR56 = non-return rate 56 days after first service; INS = number of inseminations per conception; BCS = body condition score.

increased the final MILK reliability estimate from 0·910 to 0·912.

Table 7 shows the number of bulls that reach a minimum reliability of 0·60 after each step of the procedure, for each trait. Such arbitrary thresholds often determine whether a sire evaluation is considered official and his semen is marketed as semen of a proven bull. The number of sires with fertility trait reliability greater or equal to 0·60 increased by 42 to 115% at the end of the procedure compared with single-trait reliability. For MILK this increase was 25% and occurred mainly in step 4, whereas for BCS it was 60%.

Discussion

Previously, Weller *et al.* (1985) tested various functions of the reciprocal of the diagonal elements to estimate prediction error variance and sire reliability. Although some functions yielded accurate results for single-trait models, no function adequately estimated sire reliability in multiple-trait model analysis (Weller *et al.*, 1985). Other multiple-trait approximations (Greenhalgh *et al.*, 1986), based mainly on the inverse of individual and block diagonals, underestimated the actual prediction error variance, although some of them were highly correlated with the actual sire reliability. More recently, the Information Source method has been used in the estimation of approximate reliabilities (Harris and Johnson, 1998a). This method sequentially combines reliabilities from an animal's parent average, own records and progeny records. Harris and Johnson (1998a) showed that the

Information Source method is more accurate than previous methods based on the approximation of the diagonal element of the coefficient matrix inverse only. The method is now used in the New Zealand cattle genetic evaluation. A modification of the method (Harris and Johnson, 1998b) is used in the across-country genetic evaluations of Interbull.

The procedure outlined here combines aspects of inverse diagonal prediction and Information Source. The method seems to provide a good approximation of sire reliabilities for traits included in the fertility index of dairy cattle in the UK. Reliability calculations were completed in four steps. The first two steps featured estimation of number of effective daughters, calculation of daughter contribution and multiple-trait analysis based on selection index theory. Approximate reliabilities resulting from the first two steps were very close to the exact estimates. In the last two steps, sire reliabilities were updated to include (great-)grand-daughter and parent contribution, based on the Information Source method. The last two steps depended largely on the work of Harris and Johnson (1998a), with the exception of the adjustment for sires and male progeny having daughters in the same herd-year-season. This adjustment was introduced to remove the upwards biases in the reliability calculation that were also observed by Harris and Johnson (1998a). Approximate reliabilities after adjustment were close to exact estimates, although a slight overestimation remained for low heritability traits, notably NR56 and INS. The bias is due to the way the Information Source method updates a bull's reliability to include (great-)grand-daughter information using the reliability estimate of his sons and grandsons, and affects mainly the most popular bull-sires. In a sire mixed model coefficient matrix, when herd-year-season equations are absorbed into sire equations, the resulting off-diagonal elements of sire-son or sire-maternal grandson depend on whether the two animals have daughters in the same herd-year-season. The Information Source method accounts correctly for these off-diagonals when observations are in different herd-year-seasons but not otherwise. Thus, this method would yield no bias if bulls and their male progeny have all their daughters in different management groups. The adjustment applied here for the probability of a bull and son or maternal grandson having daughters in the same herd-year-season proportionately decreased the bias by an average 0·15 to 0·25 but for sires with more than 20 male progeny in the system bias in calculated reliability decreased by a proportion as great as 0·40.

The approximate reliability method was successfully applied to the multivariate analysis of six traits in a

database of 1·8 million first lactation cows, with minimal processing time and computing memory requirements. Compared with single-trait analysis, average reliability for the four fertility traits increased by proportionately 0·47 to 0·79. Furthermore, the number of bulls with reliability 0·60 or higher increased by 42 to 115%. Results indicate that it is possible to accurately estimate the genetic merit of bulls for the low heritability fertility traits. Increasing the number of usable, validated fertility records, as observed in the case of calving to first service interval, would increase the number of proven sires with acceptable reliability.

A possibly attractive feature of the method is its straightforward extension to include fertility data from later lactations as separate traits. Treating fertility in first and later lactations as different traits seems reasonable given the less than unity genetic correlation of the trait in early and late life of the cow (Jansen *et al.,* 1987; Bagnato and Oltenacu, 1993; Roxström *et al.,* 2001).

Acknowledgements

References

Bagnato, A. and Oltenacu, P. A. 1993. Genetic study of fertility and production traits in different parities in Italian Friesian cattle. *Journal of Animal Breeding and Genetics* **110:** 126-134.

Boichard, D. and Lee, A. J. 1992. Approximate accuracy of genetic evaluation under single-trait animal model. *Journal of Dairy Science* **75:** 868-877.

Brotherstone, S. 1994. Genetic and phenotypic correlations between linear type traits and production traits in Holstein-Friesian dairy cattle. *Animal Production* **59:** 183-187.

Brotherstone, S., Banos, G. and Coffey, M. P. 2002. Evaluation of yield traits for the development of a UK fertility index for dairy cattle. *Proceedings of the seventh world congress on genetics applied to livestock production, Montpellier, France,* CD-ROM communication no. 01-28.

Greenhalgh, S. A., Quaas, R. L. and Van Vleck, L. D. 1986. Approximate prediction error variances for multiple trait sire selection. *Journal of Dairy Science* **69:** 2877-2883.

Groeneveld, E. 1990. *PEST user's manual.* Federal Agricultural Research Centre, Germany.

Harris, B. and Johnson, D. 1998a. Approximate reliability of genetic evaluations under an animal model. *Journal of Dairy Science* **81:** 2723-2728.

Harris, B. and Johnson, D. 1998b. Information source reliability method applied to MACE. *Proceedings of the 1998 Interbull meeting, Rotorua, New Zealand* **17:** 31-36.

Jansen, J., Werf, J. van der and Boer, W. de. 1987. Genetic relationships between fertility traits for dairy cows in different parities. *Livestock Production Science* **17:** 337-349.

Meyer, K. 1987. Approximate accuracy of genetic evaluations under an animal model. *Livestock Production Science* **21:** 87-110.

Misztal, I. and Wiggans, G. R. 1988. Approximation of prediction error variance in large-scale animal models. *Journal of Dairy Science* **71:** (suppl. 2) 27-32.

Pryce, J. E., Coffey, M. P. and Brotherstone, S. 2000. The genetic relationship between calving interval, body condition score and linear type and management traits in registered Holsteins *Journal of Dairy Science* **83:** 2664-2671.

Roxström, A., Strandberg, E., Berglund, G., Emanuelson, U. and Philipsson, J. 2001. Genetic and environmental correlations among female fertility traits and milk production in different parities of Swedish Red and White dairy cattle. *Acta Agriculturæ Scandinavica, Section A (Animal Science)* **51:** 7-14.

Royal, M. D., Darwash, A. O., Flint, A. P. F., Webb, R., Woolliams, J. A. and Lamming, G. E. 2000. Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Animal Science* **70:** 487-501.

Wall, E., Brotherstone, S., Woolliams, J. A., Banos, G. and Coffey, M. P. 2003. Genetic evaluation of fertility using direct and correlated traits. *Journal of Dairy Science* **86:** 4093-4102.

Weller, J. I., Norman, H. D. and Wiggans, G. R. 1985. Estimation of variances of prediction error for best linear unbiased prediction models with relationships included. *Journal of Dairy Science* **68:** 930-938.

Appendix 1

Multiple-trait reliability estimation based on daughter contribution

The objective of this Appendix is to formulate the calculation of multiple-trait sire reliability based on daughter contribution and correlation estimates among traits.

Let T be the number of traits, G the genetic (co)variance matrix and R the residual (co)variance matrix between the traits. For T traits there are $C = 2^T-1$ possible combinations of traits measured on an individual. Let R_c be the residual matrix for the cth combination of traits and $\mathbf{n}_{ic}$ a vector of number of daughters of sire i for trait combination c. The prediction error variance (PEV) for this sire can be written as:

$$PEV_i^{-1} = (\sum_{c=1}^{C} n_{ic} R_c^{-} + G^{-1})$$ where R_c^{-} is a T $\times$ T matrix with the (r,s) element containing the relevant elements of R_c^{-1} if the traits r and s are measured in the cth combination of traits. For a sire model, G is equivalent to the sire (co)variance and R includes $\frac{3}{4}$ of the genetic variance.

If $PEV_h = \sum_{c=1}^{C} \mathbf{n}_{hc} R_c^{-}$ and $PEVHS_{hi} = \sum_{c=1}^{C} \mathbf{n}_{ihc} R_c^{-}$ with n_{hc} and n_{ihc} being the numbers of daughters with trait combination c in the hth herd and ith sire-hth herd combination, respectively, then a corrected predicted error variance (PEVC) for the ith sire is given by:

$$PEVC_i^{-1} = PEV_i^{-1} - (\sum_{h=1}^{H} PEVHS_{hi} (PEVH_h)^{-1} PEVHS_{hi}$$

where H is the total number of herds. This requires the formation and inversion of several T $\times$ T matrices. As an approximation the number of effective daughters (ne) per trait may be calculated and traits may then be ordered in descending number of effective daughters, assuming that if trait t is measured then so also are traits t-1, t-2, …, 1. Then an approximate PEV is given by:

$$PEVC_i^{-1} = \sum_{t=1}^{T} ne_t^* R_{1-t}^{-1} + G_i^{-1})$$

where $ne_t^* = ne_t - ne_{t+t} \dots$, R_{1-t}^{-1} the inverse of the subset of the R matrix for traits 1 to t with the extra elements associated with traits $t + 1,\dots,T$, equal to zero and G_i the G matrix reordered according to the order of traits by descending number of daughters for sire i.

In our case, all sires had maximum number of effective daughters with milk yield record. For the other traits, the descending order of number of effective daughters was not the same for all sires, therefore, traits and matrices had to be re-ordered accordingly by the software developed.

Example

Let three traits for which a sire has the following number of effective daughters = ne_t = (10 8 3)

$$R = \begin{pmatrix} 1\cdot11 & -0\cdot43 & -0\cdot33 \\ -0\cdot43 & 1\cdot12 & 1\cdot84 \\ -1\cdot33 & 1\cdot84 & 25\cdot60 \end{pmatrix} \quad \text{and } G = \begin{pmatrix} 0\cdot28 & -0\cdot17 & -0\cdot68 \\ -0\cdot17 & 0\cdot18 & 0\cdot50 \\ -0\cdot68 & 0\cdot50 & 4\cdot64 \end{pmatrix}$$

Then

$$R_{1-1} = \begin{pmatrix} 1\cdot11 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad R_{1-2} = \begin{pmatrix} 1\cdot11 & -0\cdot43 & 0 \\ -0\cdot43 & 1\cdot12 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad R_{1-3} = R$$

$$ne_t^* = (2\ 5\ 3) \qquad PEV_i^{-1} = 2R^{-1}_{1-1} + 5R^{-1}_{1-2} + 3R^{-1}_{1-3} + G^{-1}$$

$$PEV_i = \begin{pmatrix} 0\cdot077 & -0\cdot0405 & -0\cdot1601 \\ & 0\cdot0678 & 0\cdot1400 \\ & & 2\cdot478 \end{pmatrix}$$

From the diagonal elements of PEV and G, sire reliability is 0·73 for trait 1, 0·62 for trait 2 and 0·47 for trait 3.

Appendix 2

Information source method to add (great-)grand-daughter and parent information to bull reliability

Bull (great-)grand-daughter contribution (Harris and Johnson, 1998a). Let Rb, Rs and Rms be the multiple-trait daughter-based reliability of a bull, his son and his maternal grandson. Bull reliability is then updated to consider son progeny contribution, as follows (Harris and Johnson, 1998a):

$$\text{Rb} = \frac{\text{Rb} + \frac{1}{4} \cdot \text{Rs} - \frac{1}{2} \cdot \text{Rb} \cdot \text{Rs}}{1 - \frac{1}{4} \cdot \text{Rb} \cdot \text{Rs}}$$

Similarly, bull reliability is updated to consider maternal grandson progeny contribution, as follows (Harris and Johnson, 1998a):

$$\text{Rb} = \frac{\text{Rb} + \frac{1}{16} \cdot \text{Rms} - \frac{1}{8} \cdot \text{Rb} \cdot \text{Rms}}{1 - \frac{1}{16} \cdot \text{Rb} \cdot \text{Rms}}$$

The process works from the youngest parent to the oldest, to allow all generations to be incorporated, and repeats for each trait (Harris and Johnson, 1998a).

Adjustment for probability of bull and son or grandson having daughters in same herd-year-season.
Prior to applying the above formulae, Rb, Rs and Rms are changed as follows:
Let ndb, nhb, nds and nhs be number of actual daughters and number of herd-year-seasons for bull and son, respectively. Let k be the residual to sire variance ratio and hysz the average herd-year-season size. Compute:

$$\text{neb} = \frac{k \cdot \text{Rb}}{1 - \text{Rb}} \qquad \text{nes} = \frac{k \cdot \text{Rs}}{1 - \text{Rs}}$$

Then compute: $\text{neb} = \text{neb} - \dfrac{(w \cdot ((^{ndb}/_{nhb} \cdot {^{nds}/_{nhs}})/_{hysz}))^2}{\text{nes}} \qquad \text{nes} = \text{nes} - \dfrac{(w \cdot ((^{ndb}/_{nhb} \cdot {^{nds}/_{nhs}})/_{hysz}))^2}{\text{neb}}$

where $w = (\text{minimum}(nhb, nhs))/(1 + h^2)$, h^2 = trait heritability. In these formulae the factors subtracted from neb and nes are functions of the off-diagonal elements pertaining to the sire and son, after absorbing the herd-year-season sub-matrix into the sire equations. This approximation assumes average herd-year-season size (hysz) in all cases and introduces a probability that sire and son have daughters in common herd-year-seasons. The maximum possible number of common herd-year-seasons is the smallest of nhb and nhs.

Finally compute:

$$\text{Rb} = \frac{\text{neb}}{k + \text{neb}} \qquad \text{Rs} = \frac{\text{nes}}{k + \text{nes}}$$

If Rb or Rs above become smaller than effective daughter-based reliability (for very small neb or nes), the latter replaces it. Similar adjustments are made for every bull-maternal grandson pair.

Bull parent contribution (Harris and Johnson, 1998a).
Let Rb, Rs and Rms be the reliability of a bull, his sire and his maternal grandsire. Bull reliability is updated as follows (Harris and Johnson, 1998a):

$$\text{Compute:} \quad \text{Rs} = \frac{\text{Rs} - \frac{1}{4} \cdot \text{Rb}}{\text{Rs} \cdot \frac{1}{4} \cdot \text{Rb} + 1 - \frac{1}{2} \cdot \text{Rb}} \qquad \text{Rms} = \frac{\text{Rms} - \frac{1}{16} \cdot \text{Rb}}{\text{Rms} \cdot \frac{1}{16} \cdot \text{Rb} + 1 - \frac{1}{8} \cdot \text{Rb}}$$

Then compute parent average reliability: $\text{Rpa} = 1/4 \cdot (\text{Rs} + \frac{1}{4} \cdot \text{Rms})$

Finally compute: $\text{Rb} = \dfrac{\text{Rb} + \text{Rpa} - 2 \cdot \text{Rb} \cdot \text{Rpa}}{1 - \text{Rb} \cdot \text{Rpa}}$

The process works from the oldest to the youngest, to allow all generations to be incorporated, and repeats for each trait (Harris and Johnson, 1998a).