Application of the Multiple Trait Effective Daughter Contribution Method to a Linear Model with Direct and Maternal Genetic Effects for Calving Traits

Z. Liu, E. Pasman, F. Reinhardt, and R. Reents
VIT, Heideweg 1, D-27283 Verden, Germany

Introduction

Secondary traits such as calving ease are gaining importance in international bull genetic evaluations. A previous study (Pasman and Reinhardt, 2002) showed that it is feasible to conduct international bull comparisons for the calving traits, despite the fact that diverse statistical models are utilised in national genetic evaluations, including a linear model with direct and maternal genetic effects (Clément et al. 2001; Carnier et al., 1997; Niskanen and Juga, 1997; Steinbock et al., 2000), threshold model (Emanuelson et al., 1993; Luo et al., 2002; Wiggans et al., 2002), and heteroskedastic threshold model (Foulley and Gianola, 1996; Ducrocq, 2000). Most of the participating countries apply a linear model with direct and maternal genetic effects to analyse calving ease scores or stillbirth data. Because of the correlation between direct and maternal genetic effects, the original Interbull's effective daughter contribution (EDC) method developed for single trait models with one genetic effect is no longer applicable. Therefore, the objective of this study was to extend the multiple trait effective daughter contribution (MTEDC, Liu et al., 2001) method to a linear animal model with correlated direct and maternal genetic effects for the purpose of approximating EDC of bulls for international genetic evaluations.

Methods

A linear animal model is assumed for analysing stillbirth or calving ease data transformed with Snell score:

$$y_{ijkl} = f_i + d_j + m_k + p_k + e_{ijkl} \qquad [1]$$

where y_{ijkl} is the trait value of calf *j* recorded for calving *l* of dam *k*, belonging to fixed effect *i*; f_i represents fixed effect *i*; d_j is additive direct genetic effect of calf *j*; m_k is additive maternal genetic effect of dam *k*; p_k is permanent environmental effect of dam *k*; and e_{ijkl} is residual effect. In EDC calculation only one main fixed effect, e.g. herd-year classes, may be chosen, if the other fixed effects have neglectable effects on the accuracy of EDC. Let σ_d^2, σ_m^2, σ_p^2, and σ_e^2 represent variance of direct genetic effect, maternal genetic effect, maternal permanent environmental effect and error effect, respectively. Define variance ratio $\lambda = \sigma_e^2 / \sigma_p^2$ for permanent environmental effect. Let r_{dm} represent correlation between direct and maternal genetic effects. Genetic (co)variance matrix can be written

as: $\mathbf{G}_0 = \begin{bmatrix} \sigma_d^2 & r_{dm}\sigma_d\sigma_m \\ r_{dm}\sigma_d\sigma_m & \sigma_m^2 \end{bmatrix}$.

Following MTEDC (Liu et al. 2001), data contribution is computed for every calf and every cow with calving records followed by summing up progeny contribution for bulls.

Computing own data contribution for calves and cows with calving records: Fixed effects are absorbed in order to compute effective number of records for calves. If only one main fixed effect is considered, effective number of records for calf *j* for direct genetic effect is: $n_j = 1 - \frac{1}{h}$ with *h* being the number of calves in the contemporary group, e.g. herd-year class. Unknown calf with identified parents is treated as a known animal in the calculation of its own data contribution. When a calf is unknown, male or female without calving records in genetic evaluation, own data contribution of the calf is characterised as:

$$\boldsymbol{\Psi}_{\mathrm{O}} = \begin{bmatrix} n_j & 0 \\ 0 & 0 \end{bmatrix} \sigma_e^{-2} . \qquad [2]$$

If a female calf became dam, then both direct calving information of her as a calf and maternal calving information of her as a dam must be absorbed into her genetic effects. Let assume that the cow has q calving records, effective number of records for this cow after absorbing fixed effects is: $n_q = \sum_{l=1}^{q} (1 - \frac{1}{h_l})$ with h_l being number of calves in corresponding contemporary group for calving record l of the cow. After fixed and permanent environmental effects have been absorbed, data contribution for this cow is computed as:

$$\boldsymbol{\Psi}_{\mathrm{O}} = \begin{bmatrix} n_j & 0 \\ 0 & \dfrac{n_q \lambda}{n_q + \lambda} \end{bmatrix} \sigma_e^{-2} . \qquad [3]$$

Her reliability matrix due to own data is

$$\mathfrak{R}_{\mathrm{Y}} = \mathbf{I} - (\boldsymbol{\Psi}_{\mathrm{O}} \mathbf{G}_0 + \mathbf{I})^{-1} \qquad [4]$$

and corresponding EDC matrix is

$$\boldsymbol{\Psi}_{\mathrm{Y}} = 4 \boldsymbol{\Psi}_{\mathrm{O}} . \qquad [5]$$

If a cow has calving records but is not identified as a calf in genetic evaluation, then $n_j = 0$ in formula 3. Compared to other traits in dairy cattle, service sires have own performance records as calves for direct genetic effect. To avoid double counting, the contribution of own record to own estimated breeding values (EBV) is ignored for service sires, however, the records of service sires as calves contribute to the direct genetic effects of their sires. If daughter yield deviation, instead of deregressed proof, is used in international genetic evaluation for calving traits, then the contribution of own record to own EBV must be ignored as well, which corresponds to the calculation of EDC. Should related calves, such as progeny of the same bull, belong to the same contemporary group, the reduction of effective data information must be accounted for using a similar procedure as the one in Interbull EDC method (Interbull, 2000).

Computing progeny contribution for both direct and maternal effects of bulls: After data contribution has been calculated for both calves and cows, EDC matrices from all progeny of a bull can be summarised. Bull's reliability contributed by a progeny adjusted for his mate is (Liu et al., 2001):

$$\mathfrak{R}_{\mathrm{P-M}} = \tfrac{1}{4}[\mathbf{E} - \mathbf{E}(\mathbf{E} + (\mathbf{I} - \tfrac{1}{4}\mathfrak{R}_{\mathrm{M}}^{*})^{-1})^{-1}\mathbf{E}] , \qquad [6]$$

where $\mathfrak{R}_{\mathrm{M}}^{*}$ is mate's reliability calculated by excluding EDC of this progeny and $\mathbf{E} = (\mathbf{I} - \mathfrak{R}_{\mathrm{Y}})^{-1} - \mathbf{I}$. The reliability contributed by this progeny is converted to EDC with $\boldsymbol{\Psi}_{\mathrm{P\text{-}M}} = 4[(\mathbf{I} - \mathfrak{R}_{\mathrm{P\text{-}M}})^{-1} - \mathbf{I}]\mathbf{G}_0^{-1}$ for the bull. Final EDC matrix for the bull is the sum of the EDC matrices of all his progeny: $\boldsymbol{\Psi}_{\mathrm{T}} = \sum \boldsymbol{\Psi}_{\mathrm{P\text{-}M}}$, and corresponding reliability matrix for this bull is calculated with: $\mathfrak{R}_{\mathrm{T}} = \mathbf{I} - (\tfrac{1}{4}\boldsymbol{\Psi}_{\mathrm{T}} \mathbf{G}_0 + \mathbf{I})^{-1}$.

In case that direct and maternal calving traits will be evaluated jointly using a multiple trait multiple country model at the international level, the EDC matrix $\boldsymbol{\Psi}_{\mathrm{T}}$ for individual bulls is needed. However, when direct or maternal calving trait will be analysed separately with a single trait multiple country model, the contribution by correlated genetic effect will not be able to be considered in separate MACE evaluations. Therefore, in order to calculate proper EDC for single trait multiple across country evaluation (MACE), reliability value for each of both genetic effects of a bull must be summarised with:

$$R_d^2 = ([1 \quad 0]\mathbf{G}_0 \mathfrak{R}_{\mathrm{T}} [1 \quad 0]') / ([1 \quad 0]\mathbf{G}_0 [1 \quad 0]') = ([1 \quad 0]\mathbf{G}_0 \mathfrak{R}_{\mathrm{T}} [1 \quad 0]') / \sigma_d^2 \qquad [7]$$

and $$R_m^2 = ([0 \quad 1]\mathbf{G}_0 \mathfrak{R}_{\mathrm{T}} [0 \quad 1]') / ([0 \quad 1]\mathbf{G}_0 [0 \quad 1]') = ([0 \quad 1]\mathbf{G}_0 \mathfrak{R}_{\mathrm{T}} [0 \quad 1]') / \sigma_m^2 . \qquad [8]$$

Note that the weighting vectors in formulae 7 and 8 can be modified to obtain reliability value of any linear function of direct and maternal EBV. Based on the reliability values for single genetic effects, EDC for direct and maternal effects for the bull can be computed and used in single trait MACE evaluations:

$$\psi_d = k_d\, R_d^2/(1-R_d^2)\,, \qquad [9]$$

and $$\psi_m = k_m\, R_m^2/(1-R_m^2)\,, \qquad [10]$$

where $$k_d = (\sigma_e^2 + \tfrac{3}{4}\sigma_d^2)/(\tfrac{1}{4}\sigma_d^2) = [(4-h_d^2) - (4h_m^2 + (8\sigma_{dm} + \sigma_p^2)/\sigma^2)]/h_d^2\,, \qquad [11]$$

and $$k_m = (\sigma_e^2 + \tfrac{3}{4}\sigma_m^2)/(\tfrac{1}{4}\sigma_m^2) = [(4-h_m^2) - (4h_d^2 + (8\sigma_{dm} + \sigma_p^2)/\sigma^2)]/h_m^2 \qquad [12]$$

where $\sigma^2 = \sigma_d^2 + 2\sigma_{dm} + \sigma_m^2 + \sigma_p^2 + \sigma_e^2$.

For comparison, EDC of direct and maternal genetic effects were calculated using the Interbull EDC method for single trait models with single genetic effect (Interbull, 2000) as well, in which the correlation between both genetic effects is ignored. Due to the correlated effects, variance ratio $k = (4-h^2)/h^2$ defined in the Interbull EDC method is no longer valid. Instead, formulae 11 and 12 are appropriate for the variance ratios.

The definition of one EDC depends on the models used in international genetic evaluations. If both direct and maternal effects are evaluated jointly in a multiple trait MACE for calving traits, then one EDC means a cow identified both as dam of a calf and as a calf in genetic evaluation, having dam missing, having one record as calf and one calving record as dam in contemporary groups with infinite number of calves. In case the two genetic effects are evaluated separately at the international level, one EDC for direct genetic effect is characterised by a calf having one record with an infinite number of contemporaries and having dam missing. Analogously, one EDC for maternal genetic effect for single trait MACE evaluation is represented by a cow with one calving record in a contemporary group with an infinite number of calves and dam of the cow is assumed to be missing. Since genetic correlation is considered in national genetic evaluations, formulae 8 and 9 based on MTEDC should lead to higher EDC than single trait EDC method (Interbull, 2000).

Results and Discussion

The presented MTEDC method for a linear animal model with correlated genetic effects was applied to calving ease data used in August 2002 national genetic evaluation for German dairy cattle breeds, including Holstein, Red, Jersey and cross-breds. Genetic parameters assumed in the calving ease national genetic evaluation are: $\sigma_d^2 = 1, \sigma_m^2 = 1, r_{dm} = -0.1, \sigma_p^2 = 3$ and $\sigma_e^2 = 15$, resulting in a heritability of 0.05 for both genetic effects and a repeatability of 0.24 for maternal effect. Table 1 gives descriptive statistics of the pedigree and data sets for the calving trait. Even for the large data set, the calculation with the MTEDC method took only limited computing resources in terms of CPU and RAM.

Table 1. Descriptive statistics of the pedigree and data sets for calving ease from August 2002 national genetic evaluation for German dairy cattle breeds.

Item	Calving records	Herd-year classes	Cows with calving records	Unidentified calves with records	Identified animals in total	Service sires and maternal grandsires	Total equations
Size	17,971,652	681,197	8,462,574	5,803,816	24,679,213	138,964	58,504,337

Table 2 shows number of AI bulls born from 1985 onwards and percentage of bulls with data information on direct or maternal genetic effects by birth year. Quite high proportions of old bulls born between 1985 and 1988 had fewer calves than daughters with calving records, because only calving records from 1990 were used in the national genetic evaluation. On the contrary, young bulls born after 1998 had no data information for maternal genetic effect but only for direct genetic effect. The majority of bulls born between 1989 and 1997 were used both as service sires and maternal grandsires (MGS) of calves.

Table 2. Number and percentage of AI bulls used as service sires and MGS of calves by birth year.

Birth year of bulls	Number of bulls	Bulls as service sires (%)	Bulls as MGS (%)	Bulls as service sires and MGS (%)	Bulls as service sires only (%)	Bulls as MGS only (%)
1985	1421	41	92	33	8	58
1986	1644	49	93	41	7	51
1987	1608	56	90	46	10	43
1988	1639	64	92	56	7	36
1989	1381	80	87	67	13	20
1990	1437	81	90	72	9	18
1991	1505	79	91	70	9	21
1992	1577	80	91	71	9	20
1993	1676	83	89	72	11	17
1994	1669	86	86	72	13	14
1995	1634	89	87	77	13	10
1996	1577	95	90	85	9	5
1997	1512	98	78	76	22	2
1998	1290	100	14	14	86	0
1999	1124	100	0	0	100	0
2000	246	100	0	0	100	0

The influence of genetic correlation on EDC was investigated by setting the correlation to zero ($r_{dm} = 0$), nearly one $r_{dm} = -0.99$ and an intermediate value $r_{dm} = -0.5$. The three scenarios were compared to the routine evaluation with the original correlation $r_{dm} = -0.1$. In two other validation studies the size of all herd-year classes was set to infinite and all mates of bulls were assumed to be unknown. Results of the validations suggest that larger contemporary group and higher reliability value of mates result in higher EDC for both effects, and the influence of contemporary group size is more evident than reliability of mates.

It can be seen in Table 3 that single trait EDC and MTEDC with $r_{dm} = 0$ lead to identical EDC values, confirming the correctness of the MTEDC method. For single trait EDC and MTEDC with zero correlation, EDC for direct effect is smaller than but close to the actual number of calves, which can be explained by the size of herd-year classes and reliabilities of mates of the bulls. Comparing direct effect EDC obtained from four different genetic correlations, we can conclude that EDC increases along with the correlation between both genetic effects. When bulls were used as both service sires and MGS (denoted as BOTH in Table 3), where the proportion of daughters with both types of records is higher than all bulls (ALL), the increase in EDC is slightly higher. The old bulls, born between 1985 and 1988 and having relatively many daughters with calving records and relatively few calves of their own as a result of discarding calving data before 1990, have the highest ratio of direct effect EDC to actual number of calves among all bull groups. Young bulls born after 1998 have no daughters with calving records yet, therefore equal EDC for direct effect was obtained in all scenarios, regardless of the different genetic correlations. Note that in Table 3 bulls without own calves were excluded from the calculation of the ratio of EDC to number of calves. Should such bulls have daughters with calving records, their direct effect EDC is greater than zero.

Similar results were found for maternal genetic effect in Table 4 as for direct genetic effect. In spite of two different methods identical EDC was resulted from single trait EDC and MTEDC with $r_{dm} = 0$. Maternal effect EDC is, in most cases, greater than number of daughters, which is caused by multiple calving records per daughter. Younger bulls tend to have smaller ratio than older ones, because their daughters have fewer calving records than the daughters of older bulls. As of the genetic correlation, bulls identified as service sires and MGS (BOTH) have higher ratio than all bulls (ALL), and the difference between the two ratios becomes greater as the genetic correlation increases. As observed for direct effect EDC, maternal effect EDC increases with the genetic correlation as well, and the increment is substantial when the correlation approaches to one. Young bulls benefit the most in terms of EDC increase. The ratio is highest, when young bulls were intensively used as service sires but had only few daughters with calving records so far. Bulls without daughters, in particular young bulls born after 1998, were not included in the calculation of the ratio, though their EDC for maternal genetic effect is greater than zero.

Summary

The multiple trait effective daughter contribution method was extended to linear models with correlated direct and maternal genetic effects. For two possible statistical models used in international bull comparison, single trait and multiple trait MACE models, EDC formulae were developed for direct and maternal genetic effects and the definition of one EDC was derived. As an application of the new MTEDC method, calving ease data from August 2002 German national genetic evaluation were analysed. Four different values of the genetic correlation were assumed to study its effect on EDC. In order to verify the new MTEDC method, Interbull EDC method for single trait model with one genetic effect was applied to the same data set as well. When correlation between both genetic effects was set to zero, MTEDC gave identical EDC as the single trait EDC method, proving the correctness of the MTEDC method. As expected, higher correlation between both genetic effects results in higher EDC for both effects. The increase in maternal effect EDC due to higher correlation was most evident for young bulls that have many more calves than daughters with calving records. Since only calving records from 1990 onwards are used in the national calving ease genetic evaluation, old bulls born before 1988 had fewer calves than daughters and consequently the highest ratio of direct effect EDC to number of calves among all bull groups. Ignoring the genetic correlation clearly leads to lower EDC for both effects. The presented MTEDC for linear models with correlated genetic effects is recommended for calculating EDC for international bull evaluations of calving traits. Extension of the presented MTEDC method to threshold or other non-linear models needs to be accomplished.

Literature cited

Clément, V., Bibé, B., Verrier, É., Elsen, J.-M., Manfredi, E. & Hanocq, J.B. 2001. Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. *Genet. Sel. Evol. 33,* 369-395.

Ducrocq, V. 2000. Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects. *Interbull Bulletin 25,* 123-130.

Emanuelson, U., Oltenacu, P.A. & Gröhn, Y.T. 1993. Nonlinear mixed model analyses of five production disorders of dairy cattle. *J. Dairy Sci*. 76, 2765-2772.

Foulley, J.L. & Gianola, D. 1996. Statistical analysis of ordered categorical data with threshold model. *Genet. Sel. Evol. 15,* 201-224.

Interbull 2000. New weighting factors for the international genetic evaluation, revised July 2000.

Liu, Z., Reinhardt, F. & Reents, R. 2001. The effective daughter contribution concept applied to multiple trait models for approximating reliability of estimated breeding values. *Interbull Bulletin 27,* 41-47.

Luo, M.F., Boettcher, P.J., Schaeffer, L.R. & Dekkers, J.C.M. 2002. Estimation of genetic parameters of calving ease in first and second parities of Canadian Holsteins using Bayesian methods. *Livest. Prod. Sci. 74,* 175-184.

Niskanen, S. & Juga, J. 1997. Calving difficulties and calf mortality in Finnish dairy cattle population. *Interbull Bulletin 18,* 17-20.

Pasman, E. & Reinhardt, F. 2002. Feasibility of international genetic evaluation for calving ease and stillbirth in Holstein. *Interbull Bulletin 29,* 49-54.

Steinbock, L., Johansson, K., Näsholm, A., Berglund, B. & Philipsson, J. 2000. Stillbirths and dystocia at first and second parities in Swedish holstein cattle. *Interbull Bulletin 25,* 131-134.

Wiggans, G.R., Van Tassell, C.P., Philpot, J.C. & Misztal, I. 2002. Comparison of dystocia evaluations from sire and sire-maternal grandsire threshold models. *Proc. 7WCGALP,* Communication no. 20-17.

Table 3. Ratio of direct effect EDC to number of calves for all bulls (ALL) as well as bulls used as both service sires and MGS (BOTH).

Birth year	Single trait EDC or MTEDC with $r_{dm} = 0$	MTEDC with original $r_{dm} = -0.1$		MTEDC with $r_{dm} = -0.5$		MTEDC with $r_{dm} = -0.99$	
	ALL	ALL	BOTH	ALL	BOTH	ALL	BOTH
1985	.93	1.00	1.01	2.76	3.17	17.37	21.06
1986	.94	.99	1.00	2.41	2.67	13.35	15.56
1987	.92	.96	.96	2.04	2.27	10.04	11.94
1988	.90	.94	.95	2.32	2.51	11.52	12.89
1989	.93	.95	.95	1.44	1.53	5.61	6.52
1990	.95	.95	.95	1.04	1.06	1.78	1.89
1991	.96	.96	.96	1.05	1.07	1.73	1.82
1992	.96	.96	.96	1.02	1.03	1.49	1.55
1993	.96	.96	.97	1.01	1.02	1.34	1.40
1994	.96	.96	.97	1.00	1.02	1.29	1.36
1995	.96	.96	.97	1.00	1.01	1.23	1.28
1996	.97	.97	.97	1.00	1.01	1.19	1.22
1997	.97	.97	.97	.99	1.00	1.07	1.11
1998	.97	.97	.98	.97	.99	.98	1.02
1999	.96	.96		.96		.96	
2000	.94	.94		.94		.94	

Table 4. Ratio of maternal effect EDC to number of daughters for all bulls (ALL) as well as bulls used as both service sires and MGS (BOTH).

Birth year	Single trait EDC or MTEDC with $r_{dm} = 0$	MTEDC with original $r_{dm} = -0.1$		MTEDC with $r_{dm} = -0.5$		MTEDC with $r_{dm} = -0.99$	
	ALL	ALL	BOTH	ALL	BOTH	ALL	BOTH
1985	1.34	1.35	1.37	1.50	1.78	6.98	16.72
1986	1.38	1.38	1.39	1.60	1.87	9.23	18.94
1987	1.41	1.42	1.40	1.67	1.88	10.89	19.65
1988	1.43	1.44	1.41	1.55	1.59	3.47	4.73
1989	1.38	1.39	1.38	1.57	1.62	5.51	6.74
1990	1.35	1.36	1.36	1.55	1.60	6.16	7.38
1991	1.32	1.32	1.32	1.52	1.57	5.91	7.25
1992	1.29	1.29	1.29	1.48	1.54	3.93	4.67
1993	1.26	1.27	1.28	1.53	1.60	5.76	6.82
1994	1.20	1.21	1.23	1.58	1.67	10.12	11.85
1995	1.11	1.12	1.13	1.38	1.42	5.02	5.53
1996	.96	.97	.97	1.24	1.25	4.70	4.92
1997	.82	.86	.86	1.92	1.95	16.41	16.80
1998	.82	1.01	1.01	6.62	6.75	90.63	92.64