

Validation of Genetic Trend Estimates from Test-Day Model Using Individual Daughter Deviations

M. Lidauer, E. A. Mäntysaari and I. Strandén

Agrifood Research Finland MTT, Animal Production Research, FIN-31600 Jokioinen, Finland

1. Introduction

International genetic evaluation of dairy bulls via Multiple Across Country Evaluation method relies on correct national breeding value estimation models. Therefore, unbiased estimation of national genetic trends is one main demand by Interbull to accept national bull proofs. Boichard et al. (1995) proposed three methods for validation of genetic trend estimates of national evaluations. All three methods are currently used by Interbull.

Method 1 compares evaluations from data of all lactations with those of first lactation only; Method 2 tests within-bull variation of daughter yield deviations (DYD); and Method 3 analyzes the variation in bull evaluations with time. The first two methods are referred to be more reliable (Boichard *et al.*, 1995) and, thus, Method 2 can be recommended for any linear multiple trait model. Implementation of random regression test-day models (RRM) into national evaluations abandoned use of Method 2 because the calculation of DYD under a RRM was not obvious. Mrode and Swanson (2004) generalized the calculation of DYD (VanRaden and Wiggans, 1991). Under RRM, sires will obtain DYD functions on days in milk (DIM) from which daily DYD can be calculated. Liu *et al.* (2003) have addressed the calculation of reliabilities associated with DYD functions.

There are, however, some shortcomings which may hamper use of DYD functions for Method 2: a DYD does not include the non-genetic animal effects; a DYD function is only defined for the lactation stage in which the daughters produce; and software for calculation of reliabilities associated with DYD functions are seldom available during model development work. Aim of this study was to investigate the usefulness of individual daughter deviations (*idd*) for Method 2.

2. Material and Methods

2.1. Individual daughter deviations (*idd*)

For simplicity a single trait RRM is considered:

$$y_{ijk} = \mathbf{x}'_k \mathbf{b}_i + \boldsymbol{\psi}'_d \mathbf{p}_j + \boldsymbol{\phi}'_d \mathbf{a}_j + e_{ijk}, \quad [1]$$

where, y_{ijk} is the test-day (TD) yield k of cow j , made on DIM d , in an environment that is described by the fixed effects in \mathbf{b}_i , \mathbf{x}_k is a design vector; $\boldsymbol{\psi}_d$ and $\boldsymbol{\phi}_d$ are covariable vectors for DIM d associated with regression coefficients \mathbf{p}_j and \mathbf{a}_j for the non-genetic animal effects and additive genetic animal effects, respectively; and e_{ijk} is the measurement error. The variances for \mathbf{p} , \mathbf{a} and \mathbf{e} are $\mathbf{I} \otimes \mathbf{K}_p$, $\mathbf{A} \otimes \mathbf{K}_a$, and $\mathbf{I} \sigma_e^2$, respectively, where \mathbf{K}_p and \mathbf{K}_a are (co)variance matrices and \mathbf{A} is the additive genetic relationship matrix.

The additive genetic animal effects \mathbf{a}_j can be expressed as an average of the sire's and dam's additive genetic effects plus a Mendelian sampling (MS) term:

$$\mathbf{a}_j = 1/2 \mathbf{a}_{sire} + 1/2 \mathbf{a}_{dam} + \mathbf{a}_{MS} \quad [2]$$

From [1] and [2] an *idd* was formed as:

$$\begin{aligned} idd_{ijk} &= y_{ijk} - \mathbf{x}'_k \hat{\mathbf{b}}_i - 0.5 \boldsymbol{\phi}'_d \hat{\mathbf{a}}_{jdam} \\ &= \boldsymbol{\psi}'_d \hat{\mathbf{p}}_j + 0.5 \boldsymbol{\phi}'_d \hat{\mathbf{a}}_{jsire} + \boldsymbol{\phi}'_d \hat{\mathbf{a}}_{jMS} + \hat{e}_{ijk} \end{aligned} \quad [3]$$

The average of *idds* of an animal is defined as daughter deviation (DD). This is different as in Mrode and Swanson (2004) where TD yields, adjusted for environmental and non-genetic animal effects, are used to form weighted cow wise regressions, which are then corrected for half of the dam's additive genetic effect. On the contrary, the non-genetic animal effects were included into the *idd* to ensure correct results from Method 2. A weight was

associated with each DD to account for the number of *idds* per cow:

$$w_j = \frac{\text{var}(DD_{\text{Standard}})}{\text{var}(DD_j)}, \quad [4]$$

where w_j is the ratio of the variance of a DD from a cow with a standard lactation, i.e. having ten TD yields on DIM 15, 45, 75, ..., and 285, and the variance of the DD of cow j . Applying selection index theory to [3], the variance of the DD of cow j with n_j TD yields becomes:

$$\begin{aligned} \text{var}(DD_j) = & \frac{1}{n_j^2} \mathbf{I}' \Psi_j \mathbf{K}_p \Psi_j' \mathbf{I} \\ & + \frac{3}{4n_j^2} \mathbf{I}' \Phi_j \mathbf{K}_a \Phi_j' \mathbf{I} + \frac{\sigma_e^2}{n_j}, \end{aligned} \quad [5]$$

where Ψ_j and Φ_j comprise of n_j rows of ψ' and ϕ' specific to each *idd* of cow j .

2.2. Adaptation of Method 2 for RRM

Instead of DYD as derived for RRM (Mrode and Swanson, 2004) average daughter deviations (ADD) were used. The ADD_{lm} of bull l was calculated as a weighted average of DDs of daughters that calved the first time in the bull's production year m :

$$ADD_{lm} = \frac{305}{\sum_{j=l}^{n_{lm}} w_j} \sum_{j=l}^{n_{lm}} DD_j w_j \quad [6]$$

As proposed by Boichard *et al.* (1995), a weighted analysis of within-bull variation of ADDs was carried out:

$$ADD_{lm} = \text{bull}_l + m\beta + \varepsilon_{lm}, \quad [7]$$

where, bull_l is the fixed effect of bull l and β is the regression coefficient on m . The bull production year is $m = f - f_0 + 1$, where f is the first calving year of a daughter and f_0 is the first calving year with at least ten daughters of

bull l . The weight associated with ADD_{lm} should account for the reliability of the ADD

and was approximated as: $w_{lm} = \frac{k_{lm}}{k_{lm} + \alpha}$,

where $k_{lm} = \sum_{j=l}^{n_{lm}} w_j$ with w_j being the weight

for the DD of daughter j and n_{lm} the number of daughters in bull l 's production year m ; and

$\alpha = \frac{4 - 2h^2}{h^2}$ with h^2 being the heritability of a 305-day yield.

2.3. Comparison of Method 2 with Method 3

Method 2 and 3 were applied to the Finnish Ayrshire evaluation. The Finnish multiple-trait RRM (Lidauer *et al.*, 2000) models first lactation milk, protein, and fat, and later lactation milk, protein, and fat as six different traits. Effects in the model are: age at calving; days carried calf; days dry; year \times month of production; fixed regressions on DIM nested within calving year \times calving season \times parity; herd \times year; random herd-test-day; random regressions for additive genetic animal effects and for non-genetic animal effects within lactation and across traits as well as within later lactations. Data from the April 2004 evaluation were used and comprised of 2.7 million animals and 36.7 million TD records.

Methods were tested with three models: the official Finnish model (RRMofficial), a model that should overestimate genetic trend (RRMover), and another that should underestimate genetic trend (RRMunder). For RRMover, the effects herd \times year, year \times month of production, and calving year \times calving season interaction were removed. For RRMunder, the effect of calving age was removed. For Method 3, a second set of evaluations was obtained by removing the most recent four years of data. Method 3 analysis included 337 bulls born in the years 1990 to 1992, whereas all 2481 bulls with daughters in at least 10 herds were included in the analysis of Method 2.

3. Results and Discussion

Compared to RRMofficial, estimated genetic progress in cows was on average of all traits 106% larger or 26% smaller when applying RRMover or RRMunder, respectively (Figure 1).

Changes in genetic trends (Table 1) were well detected by Method 3. However, regression coefficients of Method 3 (Table 2) indicated much higher increases in genetic trends (Table 4) than those observed from use of RRMover. For RRMunder, regression coefficients were in better agreement with the observed changes in genetic trends.

Method 2 indicated lower increases in genetic trends for RRMover than the observed increases in estimated genetic trends between RRMover and RRMofficial. With RRMunder, Method 2 failed to detect the underestimation of genetic trends (Table 2). Regression coefficients suggested increases in genetic trends, although the trends were significantly lower than those from RRMofficial (Table 1).

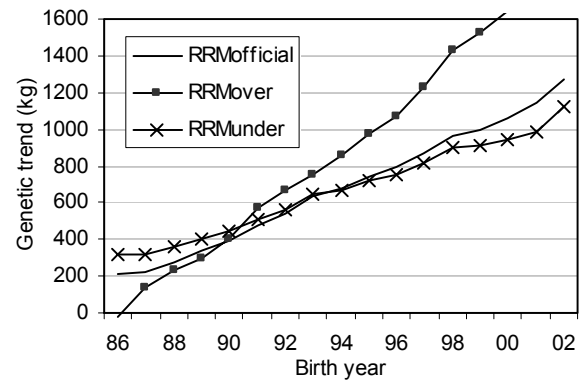


Figure 1. Estimated genetic trend for first lactation milk yield for Finnish Ayrshire using the official model (RRMofficial) and two tampered models (RRMover and RRMunder).

Table 1. Change in genetic trend (in kg/year) compared to trend estimates from official model when evaluation was carried out with two tampered models: RRMover and RRMunder.

Trait	RRMover	RRMunder
Milk, 1. lactation	+60.09	-17.48
Protein, 1. lactation	+2.95	-0.58
Fat, 1. lactation	+2.66	-0.92
Milk, later lactations	+71.39	-15.21
Protein, later lactations	+3.23	-0.52
Fat, later lactations	+3.22	-0.94

Table 2. Regression coefficients (in kg) from Method 2 and Method 3 for the three different models. RRMofficial is the official model; RRMover is modified to overestimate genetic trend; RRMunder is modified to underestimate genetic trend; and σ_{a305} is the genetic standard deviation of 305-day yield.

Trait	$0.02\sigma_{a305}$	RRMofficial		RRMover		RRMunder	
		Method 2	Method 3	Method 2	Method 3	Method 2	Method 3
Milk, 1. lactation	10.64	3.76	5.22	21.37	118.05	8.44	-8.53
Protein, 1. lactation	0.26	0.10	0.34	1.19	5.55	0.32	-0.05
Fat, 1. lactation	0.40	0.16	0.24	1.12	4.44	0.41	-0.65
Milk, later lactations	10.82	15.39	7.24	39.25	141.05	18.03	-14.05
Protein, later lactations	0.30	0.38	0.03	1.55	5.84	0.45	-0.58
Fat, later lactations	0.48	0.72	0.38	1.74	5.77	0.77	-1.00

Investigation of the daughters of the bull production year classes m in [7] revealed a confounding between class m and the calving age of the daughters. With increasing year class m , cows were on average calving on older age. Cows in bull production year class 2 were on average two month older than cows in bull production year 1. RRMunder increased DDs of older cows and decreased DDs of younger cows, since the calving age effect was missing in the model. This led to erroneous results from Method 2.

Bull production years by birth year

In [6] the DDs were classified by the daughter's year of first calving. However, the use of the daughter's birth year for the year classification eliminates the confounding between classes m and calving age. Further, it is reasonable to assume that classifications of the effects in the evaluation model (and also classifications of missing effects in biased models) are equally distributed over birth years. Hence, use of birth year instead of calving year was thought to make Method 2 more robust against certain types of model errors. Method 2 tests were redone using the

birth year classification. Now, Method 2 detected the underestimation of genetic trends by RRMunder correctly (Table 3). Method 2 detected a slight overestimation of official genetic trends for later lactation traits, which should be considered.

Table 4 presents for each method the difference (d) between regression coefficients in percent of the observed change Δ_i in genetic trend (Table 1): $d=100(\beta_i-\beta_o)/\Delta_i$, where β_i and β_o are the regression coefficients for the modified and the official model, respectively. Changes in genetic trends were best traced by Method 2 and using the birth year classification. For RRMover, Method 3 overestimated changes in trends on average by 76%, whereas Method 2 underestimated the changes on average by 15%. For RRMunder both methods overestimated the change in genetic trends on average by 9%.

4. Conclusions

The use of individual daughter deviations for validation of genetic trend under a RRM was found useful. Validation Method 2 had problems in detecting bias in genetic trend when daughter deviations were grouped by the year of first calving. The year of birth instead gave better results, and is therefore recommended.

Acknowledgements

We thank Nordisk Avelsvärdering (NAV) for financing this study and the Finnish Animal Breeding Association (FABA) for providing the national evaluation.

Table 3. Regression coefficients (in kg) from Method 2 when daughters were grouped by year of birth, rather than by year of first calving. RRMofficial = official model; RRMover = modified to overestimate genetic trend; RRMunder = modified to underestimate genetic trend.

Trait	RRMofficial	RRMover	RRMunder
Milk 1.	5.22	28.81	-4.01
Protein 1.	0.12	1.43	-0.15
Fat 1.	0.18	1.42	-0.29
Milk later	6.86	35.30	-2.22
Protein later	0.15	1.66	-0.14
Fat later	0.36	1.64	-0.18

Table 4. Difference between regression coefficients in % of the observed change when comparing RRMover or RRMunder to RRMofficial. M3 = Method 3; M2cy = Method 2 using calving year; M2by = Method 2 using birth year. For Method 2 values are multiplied by a factor of 2.

Trait*	RRMover			RRMunder		
	M3	M2cy	M2by	M3	M2cy	M2by
M1	188	57	79	79	-54	106
P1	177	73	88	71	-80	101
F1	158	72	93	97	-53	102
ML	187	67	80	140	-35	119
PL	180	73	93	117	-25	112
FL	168	64	80	147	-10	113
Pooled	176	64	85	109	-43	109

* M1, P1 and F1 are first lactation milk, protein and fat; ML, PL and FL are later lactation milk, protein and fat.

References

- Boichard, D., Bonaiti, B., Barbat, A. & Mattalia, S. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. *J. Dairy Sci.* 78, 431-437.
- Lidauer, M., Mäntysaari, E.A., Strandén, I. & Pösö, J. 2000. Multiple-trait random regression test-day model for all lactations. *Interbull Bulletin* 25, 81-86.
- Liu, Z., Bünger, A., Reinhardt, F. & Reents, R. 2003. Calculation and use of daughter yield deviations and associated reliabilities of bulls under multiple trait models. *Interbull Bulletin* 31, 16-25.
- Mrode, R.A. & Swanson, G.J.T. 2004. Calculating cow and daughter yield deviations and partitioning of genetic evaluations under a random regression model. *Livest. Prod. Sci.* 86, 253-260.
- VanRaden, P.M. & Wiggans, G.R. 1991. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* 74, 2737-2746.