Model validation using individual daughter deviations – statistical power

M. Lidauer[1], E. A. Mäntysaari[1], J. Pedersen[2] and I. Strandén[1]
[1]*Agrifood Research Finland MTT, Animal Production Research, FIN-31600 Jokioinen, Finland*
[2]*Danish Agricultural Advisory Service, Danish Cattle, Udkaersvej 15, Skejby, DK-8200 Aarhus N, Denmark*

1. Introduction

Testing within-sire variation of daughter yield deviations can be used for validation of genetic trend estimates (Boichard et al., 1995). This method, also know as Interbull Method 2, cannot as such be used for random regression models (RRM).

Lidauer et al. (2004) suggested an adaptation of Method 2 that accommodates RRM. Apart from RRM related modifications, they presented two methodological changes into Method 2. Firstly, daughter's year of birth rather than the year of first calving was used to classify daughters into sire production years. This would make Method 2 more robust against certain types of model errors. Secondly, the inclusion of the non-genetic animal (PE) effect into the calculation of daughter deviations. The later would increase the statistical power of Method 2 and would make results consistent with trend estimates.

Lidauer et al. (2004) tested the modified Method 2 with the Finnish RRM. Aim of this study is to repeat the tests with a different data and RRM, and to focus on the statistical power of the test.

2. Material and methods

2.1. Data and models

The currently developed RRM for Danish Holstein (Pedersen et al., 2004) was used for the model validation. The data included 19.1 million first lactation test-day (TD) records for milk, protein and fat yields and the pedigree had 3.1 million animals. As fixed effects the RRM included herd-test-day, first order polynomial on days in milk (DIM) ×

herd × time period, second order polynomial plus 2 Wilmink terms on DIM × calving year × calving season, calving age, and days carried calf. Random effects were regressions for additive genetic animal effects and for PE effects. Heterogeneous variance was accounted by the multiplicative mixed model approach (Meuwissen et al., 1996). This model was used as a control (CNTRL). The second, tampered model included the same effects as CNTRL but without the calving age effect (noCAE).

2.2. Validation methods

Method 2

In Lidauer et al. (2004) the Method 2 for RRM was based on individual daughter deviations (*idd*) rather than on daughter yield deviation functions (Mrode & Swanson, 2004). An *idd* was defined as:

$$idd_{jk} = \psi'_d \hat{p}_j + 0.5\phi'_d \hat{a}_{j_{sire}} + \phi'_d \hat{a}_{j_{MS}} + \hat{e}_{jk} \quad [1]$$

where idd_{jk} is the *idd* of cow j in TD k, which is the sum of the cow's PE effect ($\psi'_d \hat{p}_j$), half of its sire's genetic animal effect ($\phi'_d \hat{a}_{j_{sire}}$), the Mendelian sampling term ($\phi'_d \hat{a}_{j_{MS}}$) and the residual ($\hat{e}_{jk}$), pertaining to TD k on DIM d. However, in practice an *idd* was calculated by subtracting the estimates for the fixed effects and for half of the dam's genetic animal effects from the observation.

For each lactation a daughter deviation (*DD*) was formed by averaging the *idd*'s. The weight associated with each *DD,* to account for the number of *idd*'s per cow k_j, was:

$$w_j = \frac{\operatorname{var}(DD_{Standard})}{\operatorname{var}(DD_j)}, \qquad [2]$$

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, …, 285, and the variance of the DD of cow j, which was calculated as:

$$\operatorname{var}(DD_j) = \frac{1}{k_j^2} \mathbf{1}' \mathbf{\Psi}_j \mathbf{K}_p \mathbf{\Psi}'_j \mathbf{1} + \frac{3}{4k_j^2} \mathbf{1}' \mathbf{\Phi}_j \mathbf{K}_a \mathbf{\Phi}'_j \mathbf{1} + \frac{\sigma_e^2}{k_j}, \qquad [3]$$

where $\mathbf{\Psi}_j$ and $\mathbf{\Phi}_j$ comprise of k_j rows of $\boldsymbol{\psi}'$ and $\boldsymbol{\phi}'$ specific to each idd of cow j; and $\mathbf{K}_p$ and $\mathbf{K}_a$ are coefficient matrices of the covariance functions for PE and genetic animal effects, respectively.

Method 2 was calculated as a weighted analysis of within-sire variation of average daughter deviations (ADD):

$$ADD_{im} = s_i + \beta m + \varepsilon_{im} , \qquad [4]$$

where s_i is the fixed effect of the sire i; β is the effect of sire production year m; and ε_{im} is the error. The ADD_{im} of sire i was calculated as a weighted average of DD's that belonged to the sire production year m:

$$ADD_{im} = \frac{305}{\sum_{j=1}^{n_{im}} w_j} \sum_{j=1}^{n_{im}} DD_j w_j \qquad [5]$$

Lidauer et al. (2004) suggested defining the sire production year m to be:

$$m = b - b_0 + 1 \qquad [6]$$

where b is the birth year of a daughter and b_0 is the first year in which at least ten daughters of sire i were born.

The weight associated with ADD_{im} was: $\frac{k_{im}}{k_{im} + \alpha}$, where $k_{im} = \sum_{j=1}^{n_{im}} w_j$ with w_j being the weight for the DD of daughter j; n_{im} being the number of daughters of sire i in

sire production year m; and $\alpha = \frac{4 - 2h^2}{h^2}$ with h^2 being the heritability of a 305-day yield. A total of 5579 sires with at least 20 daughters and daughters in at least 10 herds were included. If the evaluation is unbiased the expectation of β in [4] is zero.

Method 3

Interbull Method 3 was carried out as reference method. In this method (Boichard et al., 1995) the breeding value C_i of sire i of the current genetic evaluation is regressed on the breeding value P_i of the sire i from the evaluation four years ago:

$$C_i = a_0 + a_1 P_i + \delta t_i + e_i , \qquad [7]$$

where t_i is a weighted calving year difference for daughters of sire i between both evaluations. If the evaluations are unbiased the expectation of δ is zero. The analysis included 1389 sires born in Denmark from 1992 to 1995.

2.3. Genetic trend analysis

Genetic trend estimates from both models were validated with Method 2 and 3. Four alternative analyses were carried out for Method 2: classifying the daughters in [6] by the birth year (BY) or by the calving year (CY); and including (PEin) or excluding (PEout) the PE effect in [1] and [3]. Analytical and empirical 95% confidence intervals (CI) were obtained for β and δ. The analytical 95% CI was calculated as parameter estimate ± 2 times the standard error. Similar to Weller et al. (2003), the empirical 95% CI was from 1000 bootstrap sets generated for each trait and analysis.

3. Results and discussion

3.1. Birth year versus calving year

Removing the calving age effect from the model reduced estimated yearly genetic

progress by 32.5, 1.19, and 1.52 kg for milk, protein, and fat yield, respectively.

Method 2 detected the observed drop in genetic progress only when the sire production year [6] was calculated from the birth years of daughters (Table 1). When calving year was used in [6], Method 2 failed. This result is consistent with results in Lidauer et al. (2004). Both studies agree that the sire production years are related with the calving ages, if the sire production year is calculated by the year of first calving. Hence, daughters in higher sire production years are older calvers than daughters in the first sire production year. If the calving age effect is underestimated by the model (noCAE), cows with high calving age will produce over their expectations and their DD's will be biased upwards.

Table 1
Observed and detected change in estimated genetic trend (kg/year) when calving age effect was removed from the model. Method 2 with four alternatives. Sire production year (SPY) calculated by birth year (BY) or by calving year (CY), and including (PEin) or excluding (PEout) non-genetic animal (PE) effect.

		Milk	Protein	Fat
Observed change		-32.5	-1.19	-1.52
Detected change by δ of Method 3		-41.4	-1.54	-1.95
Detected change by β of Method 2				
SPY by	PE			
BY	PEin	-17.1	-0.63	-0.83
BY	PEout	-3.2	-0.19	-0.22
CY	PEin	27.2	0.97	1.31
CY	PEout	4.3	0.26	0.57

Results from Method 2 with the CY alternatives are illogical, and therefore, from here onwards, only results from Method 2 with the BY alternatives are presented.

3.2. Inclusion of non-genetic animal effect

Method 2 detected the observed drop in genetic progress correctly, only when the PE effect was included (BY PEin, Table 1) into calculation of *idd's* in [1]. This was unexpected. From a theoretical point of view

PE effect solutions should not show any time trend, whether or not the evaluation is biased. PE effect solutions for 305-day milk yield were calculated from both evaluations. When PE effect solutions were grouped by the birth years of cows, yearly averages did not significantly deviate from zero.

However, when PE effect solutions were grouped by the birth years of the sires, birth year averages were not zero any more. Then, birth year averages were positive, zero, and negative for year groups that included sires with second crop daughters only, with first and second crop daughters, and with first crop daughters only, respectively. Thus, daughters of sires that returned to service had on average positive PE effects solutions.
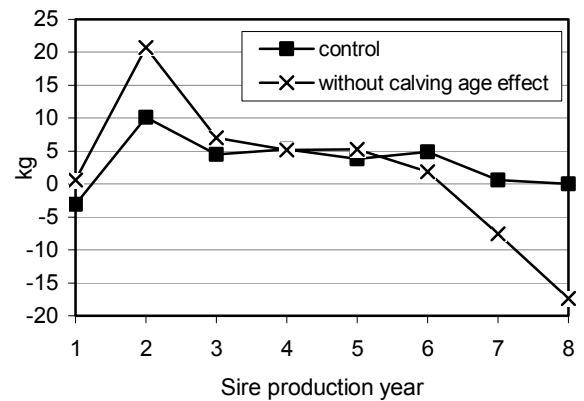

Fig. 1
Non-genetic animal (PE) effect solutions for 305-day milk yield from evaluations with the control model and the same model without calving age effect. Solutions were grouped by the daughters' sire production years, calculated using the birth year of the daughters.

In Figure 1, averages of 305-day PE effect solutions are given by sire production year. If the model is unbiased, PE effect solutions should remain unchanged over years. Solutions for the first sire production year were lower because most of the daughters had an unproven sire. A sire that returns to service can have daughters beyond sire production year 2. For a downward biased evaluation (noCAE), the sires' oldest daughters will perform better than expected and the sires' youngest daughters will perform less than expected. Results from

evaluation noCAE suggest that the deviation from the expectation go partly into the PE effect solutions of the daughters.

3.3. Statistical power

For the model CNTRL, the β estimates of Method 2 were significantly different from zero for all analysis and traits, except for fat yield (Table 2). However, estimates were within the Interbull criterion except for protein yield. The large number of sires included into Method 2 results in statistically significant deviations from zero, even if the deviation was practically insignificant.

When the PE effect was not included into Method 2, the β estimate for milk yield was still within the Interbull criterion, although the calving age effect was missing in the model (noCAE). Thus, the Interbull criterion was now too large. The currently used Interbull criterion for the β estimates of ±1% of the genetic standard deviation has been set out so that only deviations that have practical relevance are rejected. However, this requires that changes in the β estimates are consistent with changes observed in the genetic trend estimates. Based on the presented results, this would be only the case when including the PE effect into Method 2.

Table 2
Estimates of β with 95 % confidence intervals (CI) from Method 2 using the birth year classification for sire production years. Alternatively, non-genetic animal effect (PE) was included (PEin) or excluded (PEout) into calculation of individual daughter deviations.

| Model | PE | Trait | β | Analytical 95% CI | | Empirical 95% CI | | Interbull criterion |
				Lower	Upper	Lower	Upper	
CNTRL	PEin	milk	-5.89	-9.99	-1.79	-9.04	-2.83	± 6.08
	PEin	protein	-0.18	-0.30	-0.06	-0.27	-0.10	± 0.16
	PEin	fat	-0.15	-0.31	0.00	-0.29	-0.03	± 0.22
	PEout	milk	-2.25	-4.14	-0.37	-3.88	-0.65	± 6.08
	PEout	protein	-0.06	-0.10	-0.02	-0.09	-0.03	± 0.16
	PEout	fat	0.01	-0.07	0.05	-0.05	0.07	± 0.22
noCAE*	PEin	milk	-22.94	-27.16	-18.72	-26.54	-19.33	± 6.08
	PEin	protein	-0.81	-0.93	-0.69	-0.90	-0.71	± 0.16
	PEin	fat	-0.97	-1.15	-0.83	-1.11	-0.81	± 0.22
	PEout	milk	-5.43	-7.33	-3.24	-6.93	-3.96	± 6.08
	PEout	protein	-0.25	-0.29	-0.20	-0.28	-0.21	± 0.16
	PEout	fat	-0.24	-0.30	-0.17	-0.28	-0.18	± 0.22

* noCAE is the same model as CNTRL but without the calving age effect, which leads into degreased genetic trend of 32.5, 1.19, and 1.52 kg for milk, protein, and fat yield, respectively.

A practical confidence interval (Interbull criterion) is important for large populations, for which small deviations from zero are already statistically significant. In contrast, for small populations also the statistical significance is important; in case the estimate exceeds the Interbull criterion but the deviation is statistically insignificant.

The δ estimates of Method 3 were within the Interbull criterion for model CNTRL and outside the Interbull criterion for the model

noCAE (Table 3). For Method 3 the 95% CI's were consistent with the Interbull criterion but sizes of analytical and empirical 95% CI's were different, suggesting a skewed distribution of sire proofs. The empirical 95% CI was between 53% and 90% larger than the analytical 95% CI. In contrast, for Method 2 the analytical and empirical 95% CI's were consistent, which suggest that distribution of $ADD's$ was close to normal.

Table 3
Estimates of δ with 95 % confidence intervals (CI) from Method 3. Model noCAE is the same model as CNTRL but without the calving age effect.

Model	Trait	δ	Analytical 95% CI Lower	Upper	Empirical 95% CI Lower	Upper	Interbull criterion
CNTRL	milk	1.99	-4.13	8.11	-8.37	12.38	$\pm$ 12.16
	protein	0.11	-0.08	0.30	-0.19	0.45	$\pm$ 0.33
	fat	-0.10	-0.36	0.15	-0.61	0.36	$\pm$ 0.45
noCAE	milk	-39.37	-46.75	-31.99	-50.50	-27.02	$\pm$ 12.16
	protein	-1.43	-1.68	-1.19	-1.82	-1.07	$\pm$ 0.33
	fat	-2.05	-2.34	-1.76	-2.60	-1.53	$\pm$ 0.45

3.4. Weight considerations

Weights w_j used in [5] are based on selection index theory and a single trait approach. Liu et al. (2004) presented a sophisticated approximation of reliabilities for daughter yield deviations. Their method is based on a multiple trait approach and should therefore be more accurate. Such an approximation could be used to verify whether the weights we used are sufficient.

Calculation of weights could be simplified more by replacing the weight for a cow's DD [2] by a standard weight, which depends only on the number of TD's. The standard weights could be tabulated for each number of TD's using standard TD's in equation [3]. In practice, this was found to yield almost same *ADD's* in [5].

4. Conclusions

Results of this study were consistent with those from an earlier on a different data and model. Namely: classifying the daughter deviations of a sire by the birth years of the daughters makes Method 2 more robust against certain model errors, and including the PE effect into Method 2 makes Method 2 estimates consistent with changes in genetic trend estimates.

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., 2004. Validation of genetic trend estimates from test-day model using individual daughter deviations. Interbull Bulletin No. 32, 142-145.

Liu, Z., Reinhardt, F., Bünger, A., Reents, R., 2002. Derivation and calculation of approximate reliabilities and daughter yield-deviations of a random regression test-day model for genetic evaluation of dairy cattle. J. Dairy Sci. 87, 1896-1907.

Meuwissen, T.H.E., De Jong, G., Engel, B., 1996. Joint estimation of breeding values and heterogeneous variance for large data files. J. Dairy Sci. 79, 310-316.

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.

Pedersen, J., Nielsen, U.S., Aamand, G.P., Pösö, J., Mäntysaari, E., Lidauer, M., Strandén, I., Madsen, P., Johansson, K., Eriksson, J-Å., Juga. J., 2004. Joint Nordic test day model for milk production traits. Interbull Bulletin No. 32, 8-12.

Weller, J.I., Emanuelson, U., Ezar, E., 2003. Validation of genetic evaluation methodology using the nonparametric bootstrap method. Interbull Bulletin No. 31, 26-29.