
©Tiina Tahvonen

Nordic Cattle Genetic Evaluation NAV

Validation of consistency of Mendelian sampling variance in national evaluation models

A.-M. Tyrisevä[1], E.A. Mäntysaari[1] , J. Jakobsen[2], G.P. Aamand[3], J. Dürr[2], W.F. Fikse[4] and M.H. Lidauer[1]

[1]MTT Agrifood Research Finland

[2]Interbull Centre, Sweden

[3]NAV Nordic Cattle Genetic Evaluation, Denmark

[4]Dept. Animal Breeding and Genetics, SLU, Sweden

Background

- Trends in genetic variance benefit bulls coming from populations or year classes with increased genetic variance
- National evaluation centers and Interbull need a test to detect these trends
- So far two approaches to estimate within-year genetic variances proposed
  - IB4: method by Interbull (Fikse et al. 2003)
  - FMS: method by Lidauer et al. (2007)
- No test implemented yet

Aims of the research project

- Dissect behaviour of IB4 and FMS on cows and bulls by performing simulation study (Tyrisevä et al. 2011)
- Develop a test to validate consistency of Mendelian sampling variance in national evaluation models

Validation procedure

Estimation of genetic variance

- Accelerated version according to Fikse et al. 2005, RHS independent from population-wide genetic variance

$$\sigma^2_{u_i} = \frac{\sum_{k=1}^{q_i} d_k \hat{m}^2{}_k}{q_i - \sum_{k=1}^{q_i} d_k PEV(\hat{m}_k)}$$

- q_i: number of animals in year i
- d_k: inverse of proportion of genetic variance not explained by the known parents
- $\hat{m}_k$: squared estimated Mendelian sampling term of animal k
- $PEV(\hat{m}_k)$: prediction error variance of the MS term

Data

- Test can be performed either for cows or bulls
- Time period of 12 most recent years
- Number of animals with observations in the most recent year class must comprise at least 50% of average yearly size in testing period
- EBVs and reliabilities for animals and their parents → calculation of MS terms and PEV of MS

MS = EBV-PA, i.e. Mendelian Sampling term

Statistical test

- Weighted regression model is fitted on estimates of within-year genetic variances to test a possible trend

$$y_i = b0 + b1 \times year_i + e_i$$

- Empirical 95% CI by bootstrapping data with 1000 resampling within strata → 1000 $b0$ and $b1$ terms → empirical CI for a trend expressed as a percentual change in genetic variance (e.g. 100 $\times$ $b1/b0$)

Statistical test, cont.

- Possible outliers that do not fit the model identified by calculating residuals from the regression model within each bootstrap sample
- 0.025 and 0.975 quantiles to define 95% CI for residuals
- If CI does not include zero, variance estimate of that year is considered as an outlier

Tolerated level of bias

- For large populations, test has a power to detect very small deviations from zero that have no practical influence
- Tolerated level of bias is needed
- Tested population must exceed the tolerated bias before problems are reported

Validation of the method

Data

1. Data comprised Danish Hol cows from 2000 herds
   - 800 000 cows
2. Test-day observations were replaced by simulated ones
   - Alternative scenarios were generated for variance bias
3. Data were analyzed to yield 305d EBVs for cows (I) and bulls (II)
   - Sub-sample mimicing bulls in small populations (III)
   - Average number of animals in year classes:

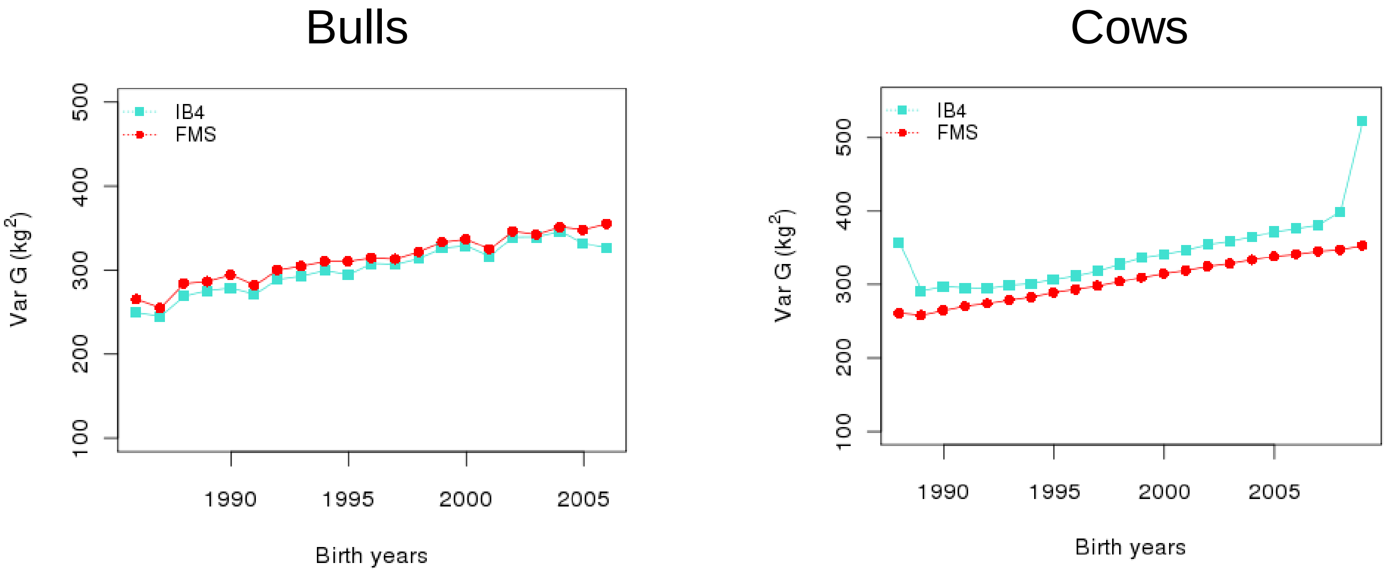
     45 000 (I)
     300 (II)
     27 (III)

Testing schemes

- Two scenarios:
  - Scenario A as control
  - For Scenario B a yearly trend of 2% in phenotypic variance
    - B1 EBVs calculated with heterogeneous variance adjustment (HV)
    - B2 EBVs calculated without adjustment (no HV)
- 20 replicates in each testing scheme for each three populations
- Genetic variances were estimated also under FMS

Results and discussion

Estimation of genetic variance



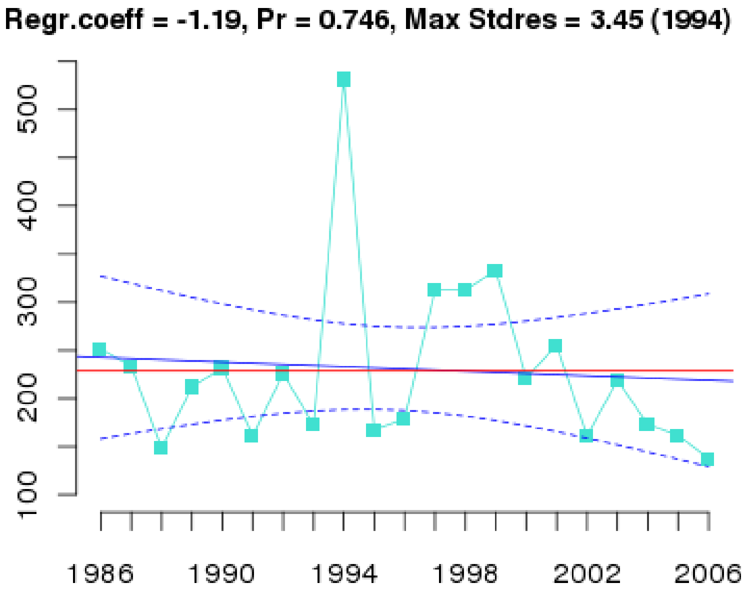
Scenario B, no HV

Estimates averaged over 20 replicates

- Reliabilities < 0.50 for majority of cows in both extremes
- By removing all cows with reliability < 0.50, peaks disappeared and all estimates were closer to FMS estimates
- Quality of estimation of genetic variance relates to value of approximated PEV, which is further associated with level of approximated reliabilities
- Carefull restriction of data, but not based on reliability causing selection bias

Sample size

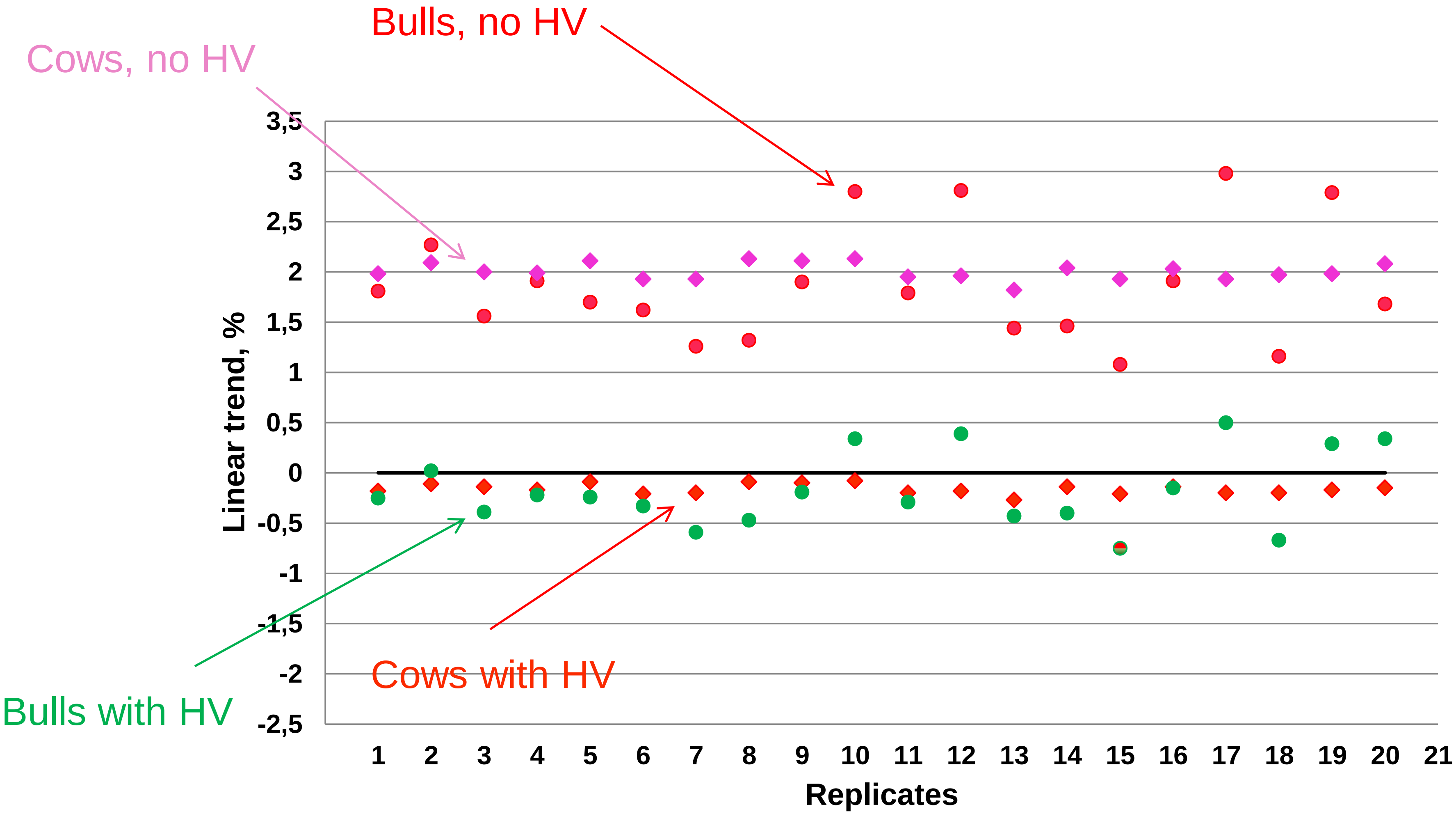
- Populations such as I and II can be used for testing
- Use of populations such as population III, might be difficult



Statistical test

- For both populations I and II, generated and not adjusted bias was detected by the test in all cases
- Applying heterogeneous variance adjustment resulted in slight decreasing trend
  - It was on average -0.16% and was detected as statistically significant deviation from zero in cow data replicates having large year classes
  - 1/20 data replicates in bulls deviated statistically significantly from zero

→ results illustrate a need to define a level of bias that has no practical influence and can be tolerated



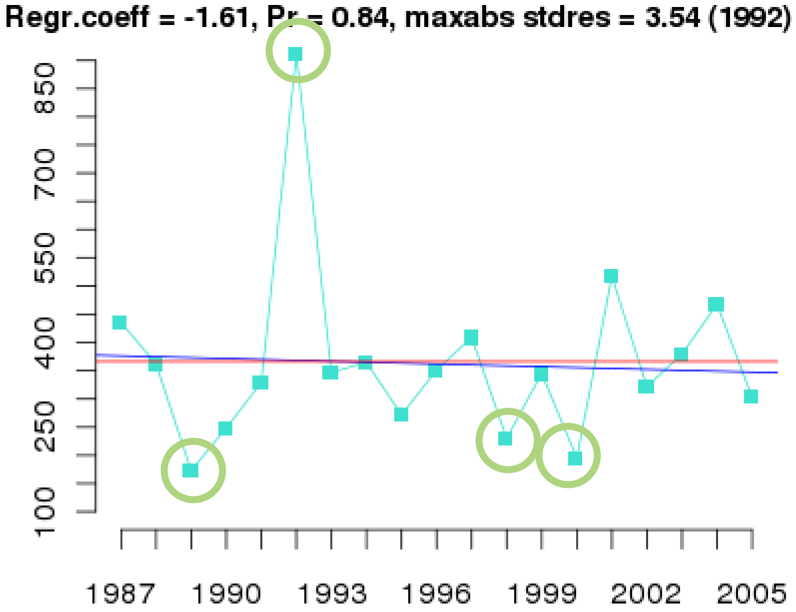
Bulls, no HV

Cows, no HV

Cows with HV

Bulls with HV

Detecting outliers



- A small bull data replicate under Scenario B without HV adjustment
- All circled cases can be considered as outliers

Conclusions

- Proposed procedure consists of following steps:
    1. Estimating within-year genetic variance by utilizing MS terms and PEV of MS
    2. Fitting regression model on variances with yearly number of animals used as weights
    3. Identifying possible outliers
    4. Defining 95% empiricial CI for a trend

- Country/evaluation center is noted for problems if trend deviates significantly from zero and exceeds tolerated level of bias

Conclusions, cont.

- Estimates of within-year genetic variances from IB4 and FMS in good agreement for bulls
- For cows, IB4 is less robust with low reliability values → data used for testing should be carefully defined
- Use of populations such as population III might be difficult