Experiences with Interbull test IV: estimation of genetic variance

M. Lidauer, K. Vuori, I. Strandén, E. A. Mäntysaari

MTT Agrifood Research Finland

Introduction

- **Time trend in genetic variance**
  - Favours top animals coming from years with higher variance
  - Causes bias in international bull comparison
- **Proposed Interbull test IV (Fikse et al., 2005)**
  - Estimates genetic variance based on EBVs and Mendelian sampling deviation PEVs
  - PEVs are approximated
  - Tests for significance of estimates' deviations from the mean
- **Unknown**
  - Quality of PEV approximation in complex model
  - Possibility of systematic bias in variance estimates

Objectives

- **Applying Interbull test IV on a complicated model**
  - Nordic test day model
- **Estimating genetic variance by a full model sampling approach**
  - Sampling of true PEV

Genetic variance estimated from EBVs of animals in stratum *i*

- **Interbull test IV (IB4), (Fikse et al., 2005)**

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

where

$$\hat{m}_k = EBV_k - \tfrac{1}{2}(EBV_s + EBV_d)$$

d_k	depends on parents known
$PEV^*(\hat{m}_k)$	prediction error variance of Mendelian sampling deviation expressed in genetic variance units
n_i	number of animals

Genetic variance estimated from EBVs of animals in stratum *i*

- **Full model sampling approach (FMS)**

$$\hat{\sigma}^2_{a_i} = \frac{\sum_{k=1}^{n_i} d_k \hat{m}_k^2}{n_i} \frac{1}{r} \sum_{j=1}^{r} \left[\frac{\sum_{k=1}^{n_i} d_k \tilde{m}_{kr}^2}{\sum_{k=1}^{n_i} d_k \hat{\tilde{m}}_{kr}^2} \right]$$

where

$$\hat{m}_k = EBV_k - \tfrac{1}{2}\left(EBV_s + EBV_d\right)$$

d_k	depends on parents known
$\tilde{m}_{kr}$	true Mendelian sampling deviation
$\hat{\tilde{m}}_{kr}$	estimated Mendelian sampling deviation
n_i	number of animals
r	number of replicates

Data and model

- **Randomly chosen 5000 Finnish Ayrshire herds**
  - 449,160 cows with observations
  - 11.2 million TD records on milk, protein and fat yield from all lactations, recorded in the years 1988 to 2006.
  - 621,775 animals in the pedigree

- **Nordic test-day model**
  - Multi-trait (9 traits for Finnish TD records)
  - Reduced rank covariance functions across traits and DIM
  - Adjustment for heterogeneous variance (Meuwissen et al., 1996)
  - 305-d EBVs weighted across first three lactations
  - Approximation of reliability (Misztal and Wiggans, 1988)

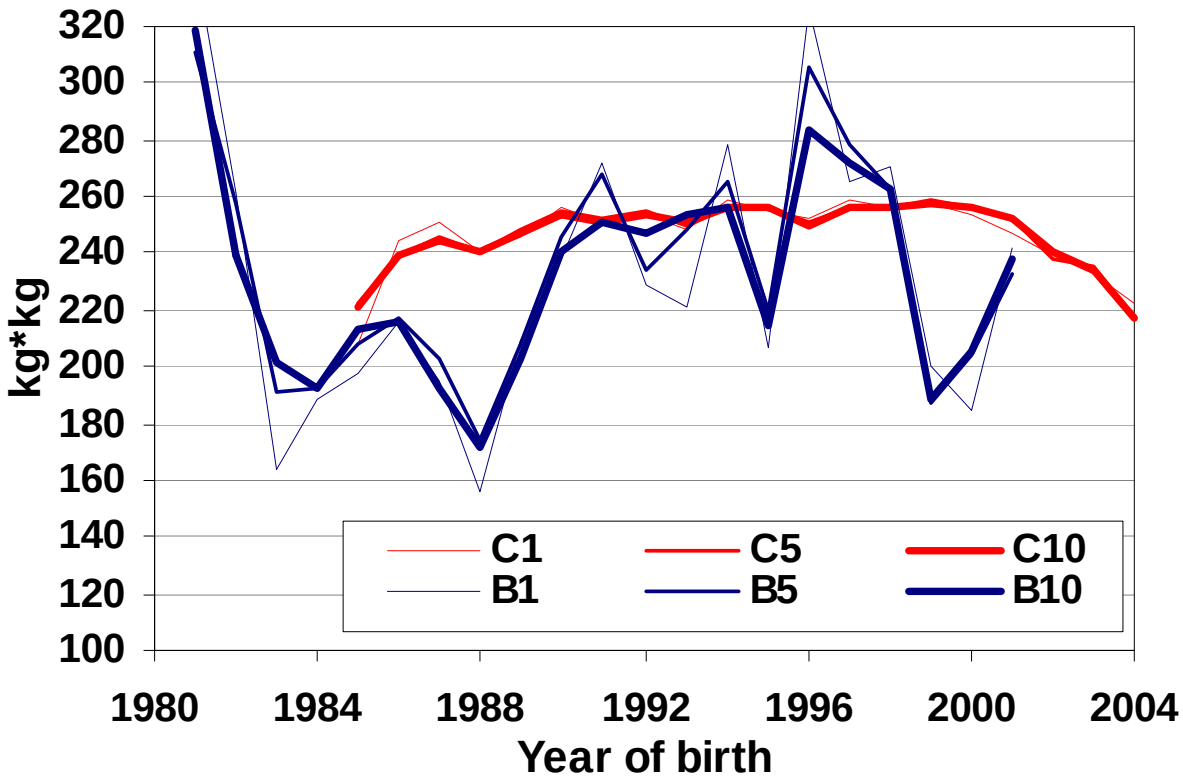
Variance estimation strata

Year	Bulls	Cows	Year	Bulls	Cows
1981	49		1993	92	25422
1982	60		1994	101	26009
1983	124		1995	119	26352
1984	150		1996	116	26430
1985	164	2504	1997	122	26037
1986	176	18177	1998	104	25679
1987	107	19611	1999	113	25663
1988	152	20301	2000	179	23959
1989	158	21741	2001	135	25232
1990	110	23723	2002	7	23459
1991	126	23336	2003		22348
1992	125	24387	2004		12048

- **Bulls**
  - At least 10 daughters with observations
  - Both parents known
- **Cows**
  - At least one observation
  - Both parents known

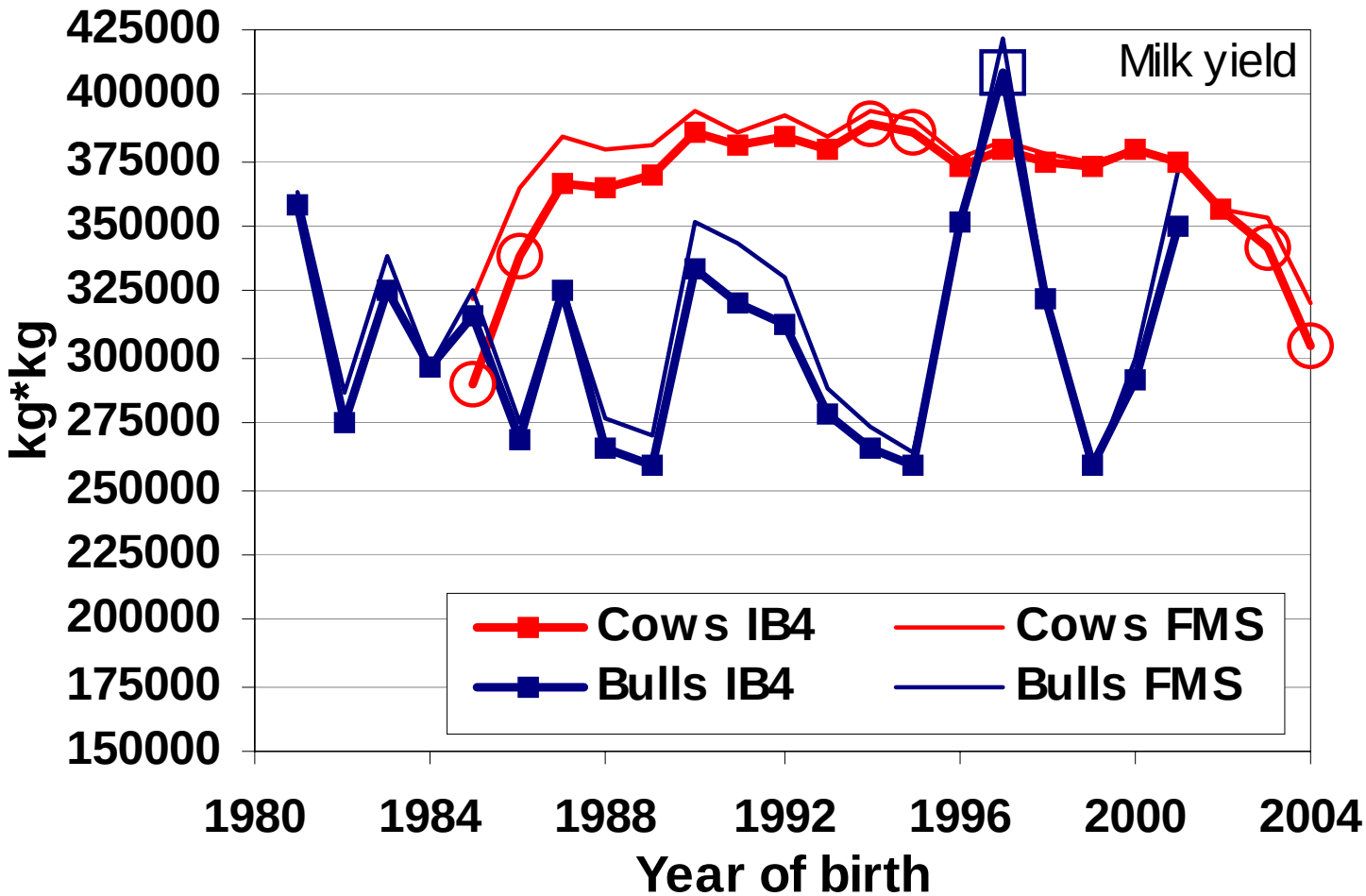
Full model sampling

Number of required replicates



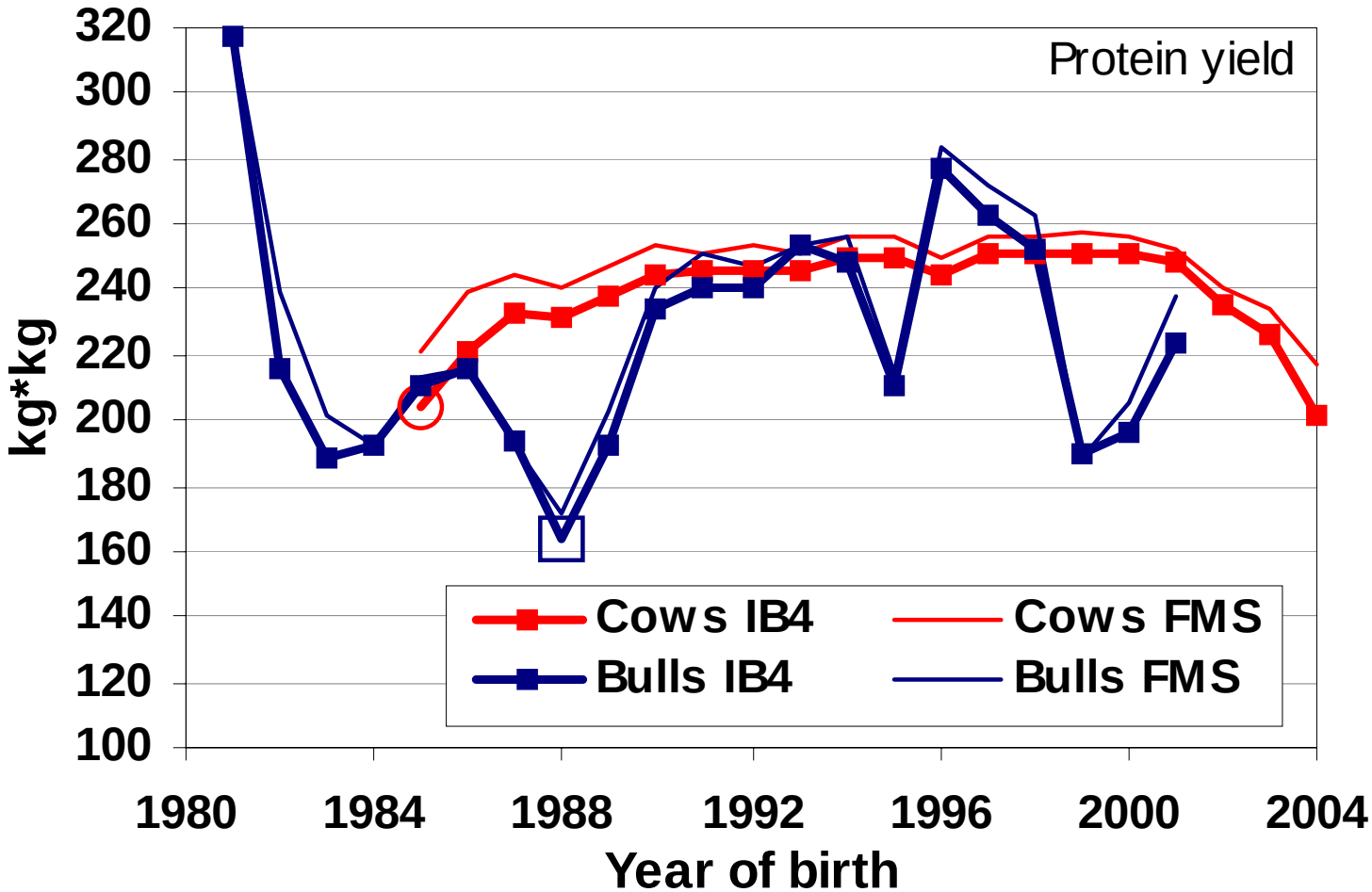
Mean of variance estimates for protein yield by cow (C) and bull (B) strata;
C1 B1, C5 B5, or C10 B10, from 1, 5, or 10 replicates.

Interbull IV *versus* full model sampling



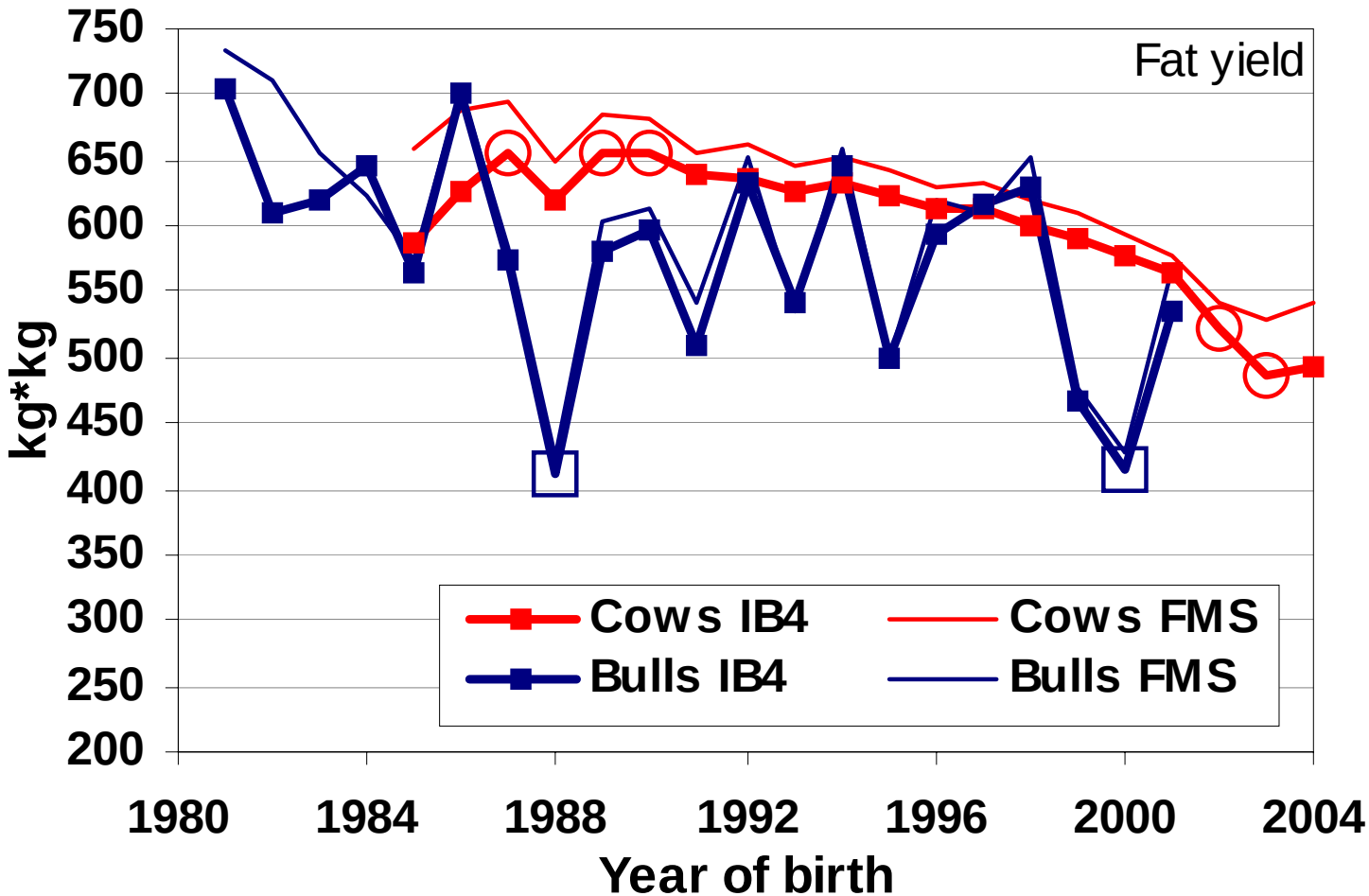
Genetic variance estimates: milk yield

Interbull IV *versus* full model sampling



Genetic variance estimates: protein yield

Interbull IV *versus* full model sampling



Genetic variance estimates: fat yield

Detection of trend in genetic variance

- **Applying IB4 test as proposed**
  - Bulls: all traits pass the test
  - Cows: only protein yield passes the test

- **Applying IB4 test only on strata without a possible systematic bias**
  - Bulls: all traits pass the test
  - Cows: milk and protein pass the test
    fat fails (which shows a clear downward trend)

Conclusions (I)

- **Overall good agreement between both methods**
  - No systematic differences between estimates based on bulls
  - Some differences between estimates based on cows
- **Upward bias in reliability estimates for cows**
  - Cows with observations at the beginning of the data
  - Cows with records in progress at the end of the data
- **Selection may affect on variance estimates**
  - Estimates for sires are lower than estimates for cows
  - Cows in the first and last year classes are not a random sample

Conclusions (II)

- **Estimates from cow strata are more consistent across years**

- **Applying IB4 to cow strata, excluding most recent and oldest year classes may increase reliability of the test**