

Simulation study on Mendelian sampling variance tests

A.-M. Tyrisevä¹, M. H. Lidauer¹, F. Fikse² and E. A. Mäntysaari¹

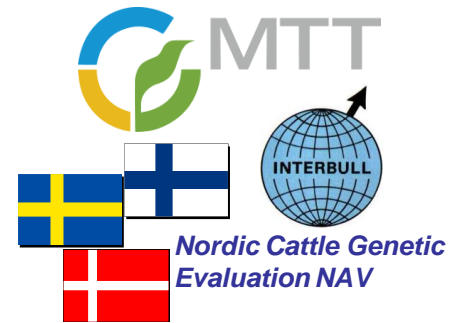
¹MTT Agrifood Research Finland

²Dept. Animal Breeding and Genetics, SLU, Uppsala, Sweden

Background

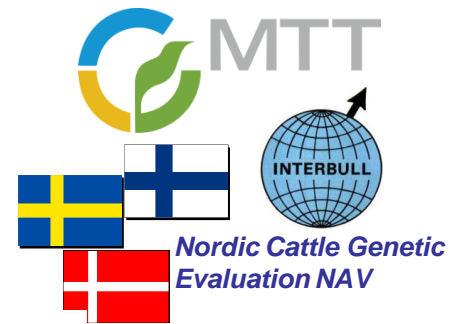
- A need for a Mendelian sampling variance test that assess quality of data sets and national evaluation models
- Two methods proposed:
 - IB4: method by Interbull (Fikse, 2003)
 - FMS: method by Lidauer et al. (2006)
- Current tolerance interval in IB4 might be too stringent
- No test statistics for FMS yet
- Both give similar estimates of genetic variance
- So far, none implemented

Aims of the study



- Dissect behaviour of IB4 and FMS on cows and bulls under two different scenarios
- Study whether MACE model residuals could be used as a quick and simple tool to check data quality

The IB4 method



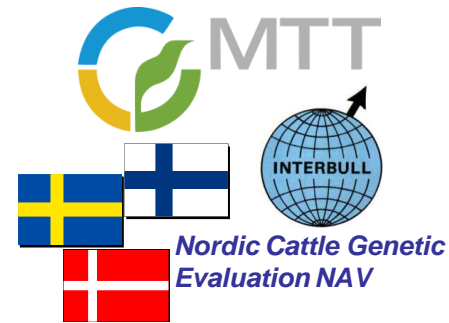
$$\sigma_{u_i}^2 = \frac{1}{q_i} \sum_{k=1}^{q_i} d_k \left[\hat{m}_k^2 + PEV(\hat{m}_k) \right],$$

where q_i is the number of animals in year i ,
 d_k is the proportion of the genetic variance not explained by the known parents (i.e. 1/2, 3/4 or 1),

\hat{m}_k^2 is the squared estimated Mendelian sampling deviation of animal k and

$PEV(\hat{m}_k)$ is the prediction error variance of the Mendelian sampling deviation.

The FMS method



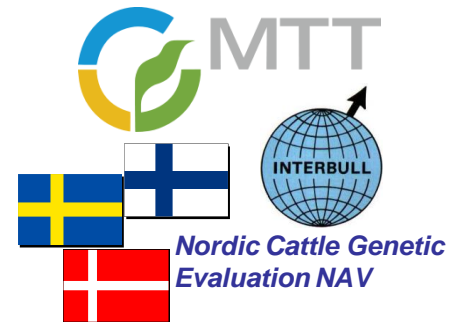
$$\hat{\sigma}_{u_i}^2 = \frac{\sum_{k=1}^{q_i} d_k \hat{m}_k^2}{q_i} \frac{1}{r} \sum_{j=1}^r \left[\frac{\sum_{k=1}^{q_i} d_k \tilde{m}_{kr}^2}{\sum_{k=1}^{q_i} d_k \hat{m}_k^2} \right],$$

where \hat{m}_k is the Mendelian sampling deviation estimated from the real data set,

\tilde{m}_{kr} is simulated true Mendelian sampling deviation of replicate r ,

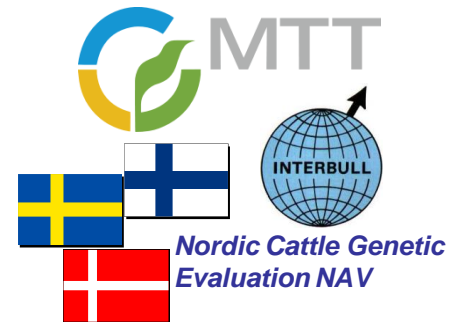
and \hat{m}_{kr} is its estimate solved from the data replicate r .

Analysis of MACE model residuals



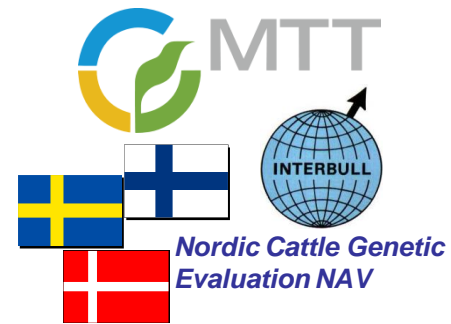
- Data from MACE analysis comes from deregression
- Model residuals should be strongly associated with Mendelian sampling terms of bulls
- Provided there is a trend in within-year estimates of $\text{Var } G$ for a country, there should be a similar trend in variances of MACE residuals

Real data set



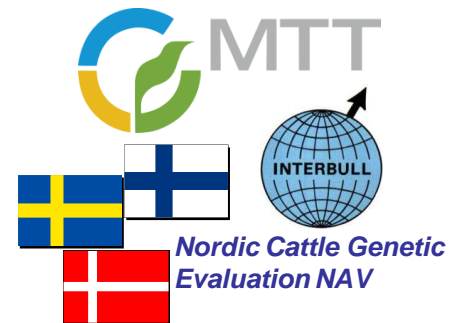
- Danish Holstein test-day data for protein yield
- 2000 herds, 760 000 cows, 13 million test-day records, 1.7 million pedigree records
- Nordic test-day model
- EBVs for first three lactations are combined to a single 305-d EBV

Simulations



- Observations of real data replaced by simulated observations
- Two scenarios, 20 replicates in each, were simulated based on variance components and systematic environmental effects from Nordic test-day model
- MiX99, valMStol softwares

Setup for simulations



EBV Prediction	Scenarios	
	A) Control	B) Yearly trend of 2% in phenotypic Var
BLUP ^a	IB4 ^c	IB4
	FMS ^c	FMS
	AMR ^c	AMR
BLUP + HV ^b	-	IB4
	-	FMS
	-	AMR

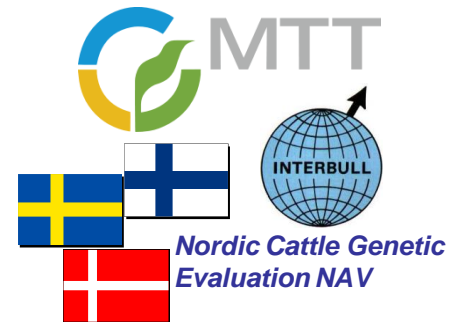
^a EBV predictions are carried out without heterogeneous variance (HV) adjustment

^b EBV predictions are carried out with the HV adjustment

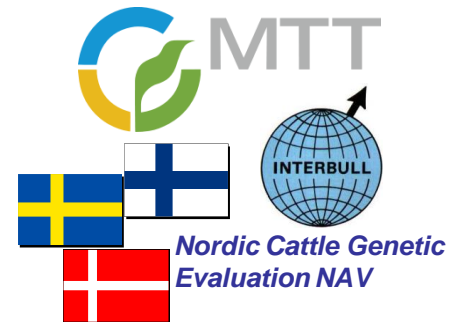
^c IB4: Interbull test IV, FMS: Full model sampling, AMR: Analysis of MACE model residuals

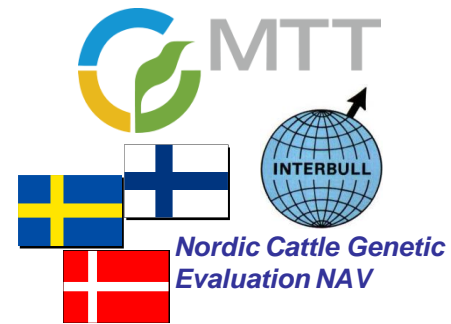
Data used for testing

- Tested bulls
 - With daughters at least in 10 herds
 - Birth years ranged from 1986 to 2006
 - Average birth year group size 291
- Tested cows
 - Cows born after 1987
 - Average birth year group size over 40 000
 - Smallest group size 2 334 in the last birth year (2009)



Results

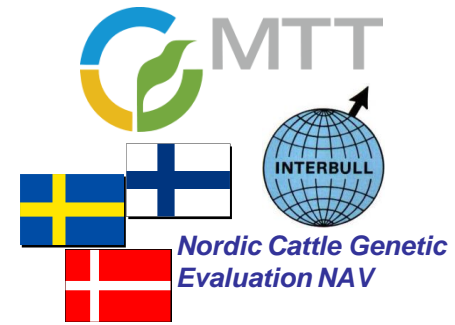
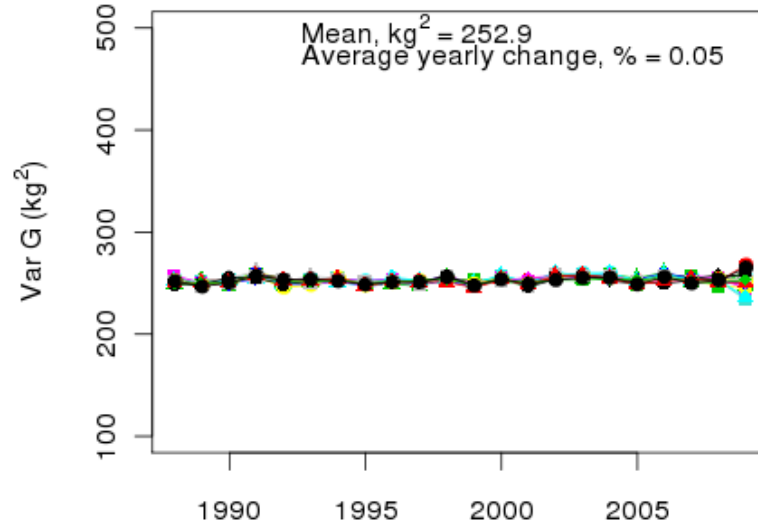




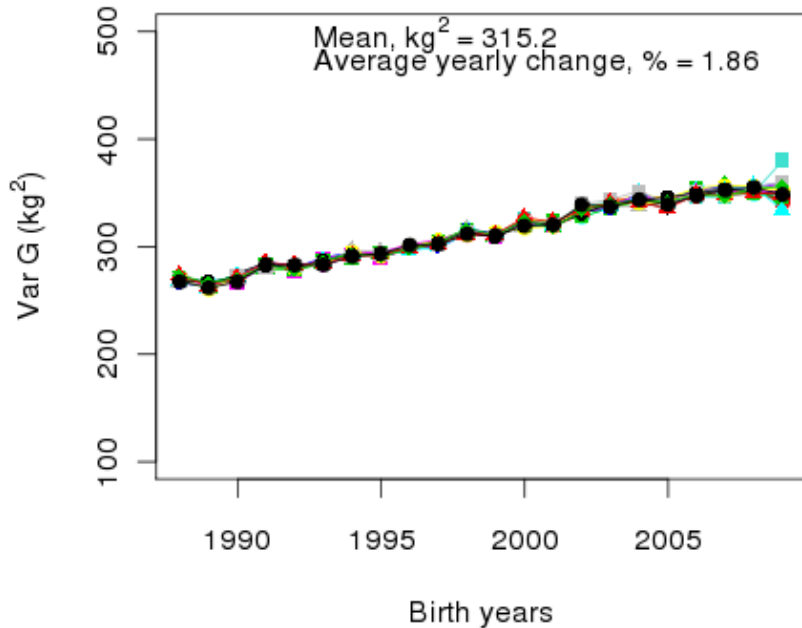
**Within-year genetic variances – provided
by the FMS method**

Cows

Scenario A

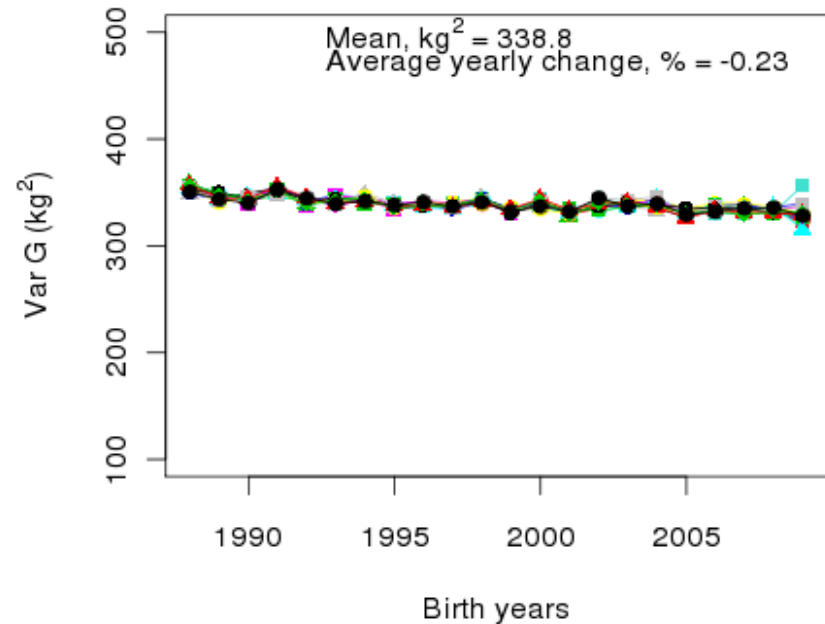


Scenario B, no HV



Birth years

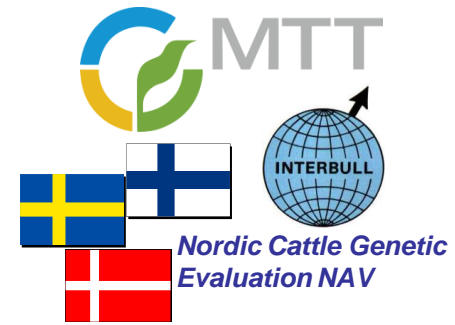
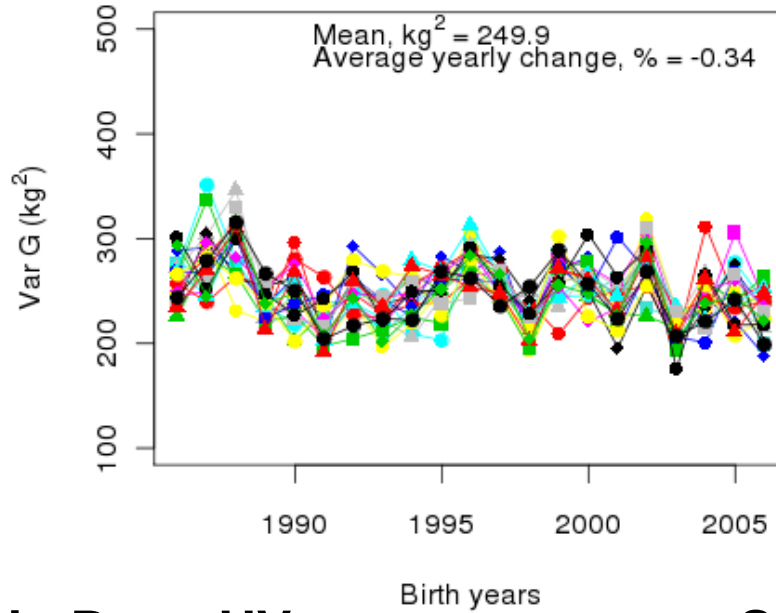
Scenario B, HV



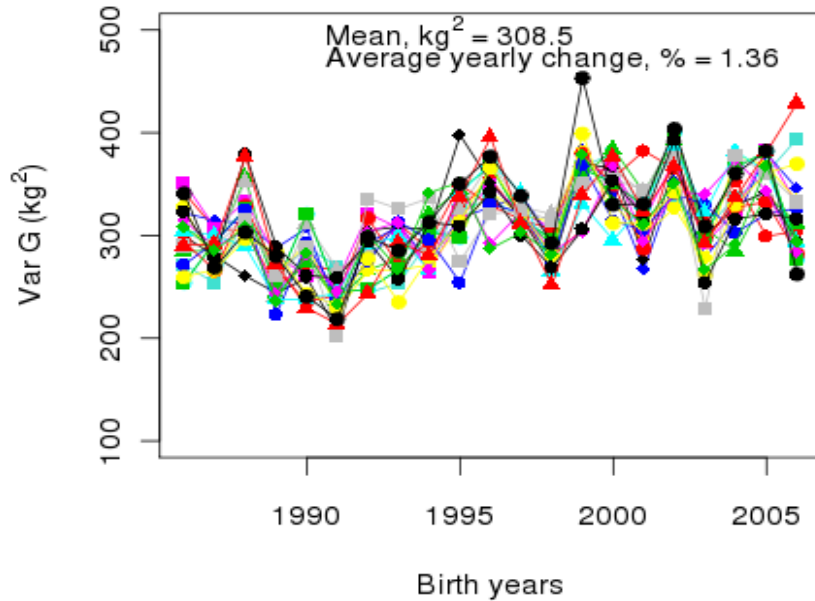
Birth years

Bulls

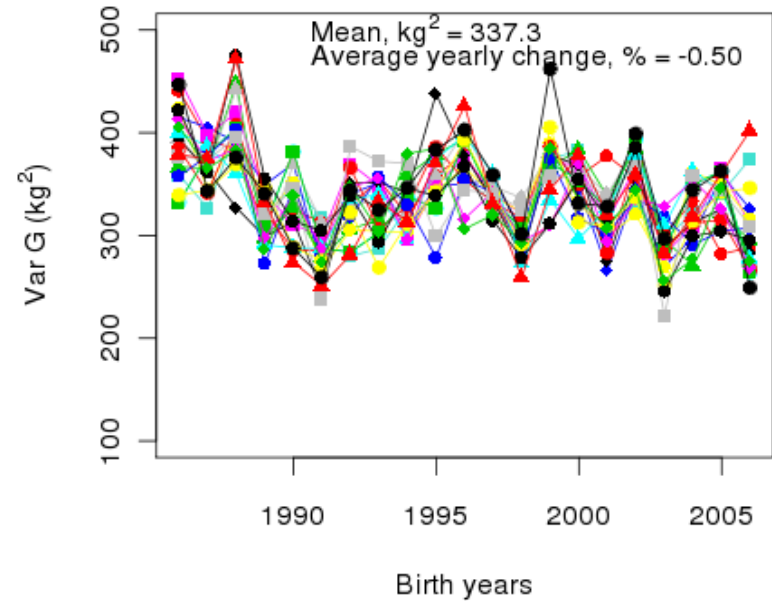
Scenario A



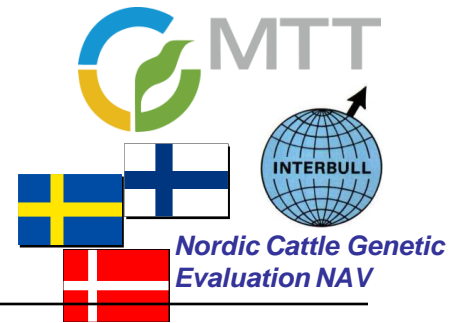
Scenario B, no HV



Scenario B, HV

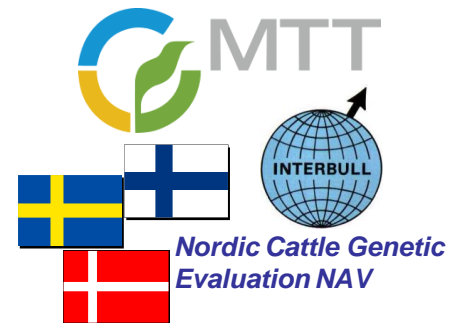


The IB4 test results



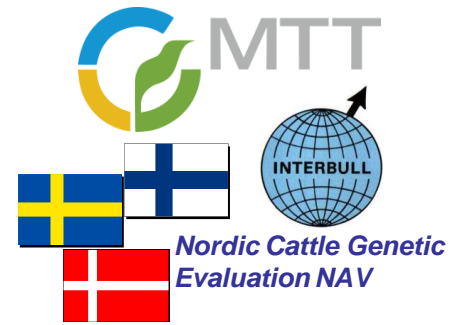
Number of failed tests (> 2 years outside tolerance interval)

A Scen		B Scen, no HV		B Scen, HV	
Cows	Bulls	Cows	Bulls	Cows	Bulls
0/20	1/20	20/20	7/20	0/20	1/20

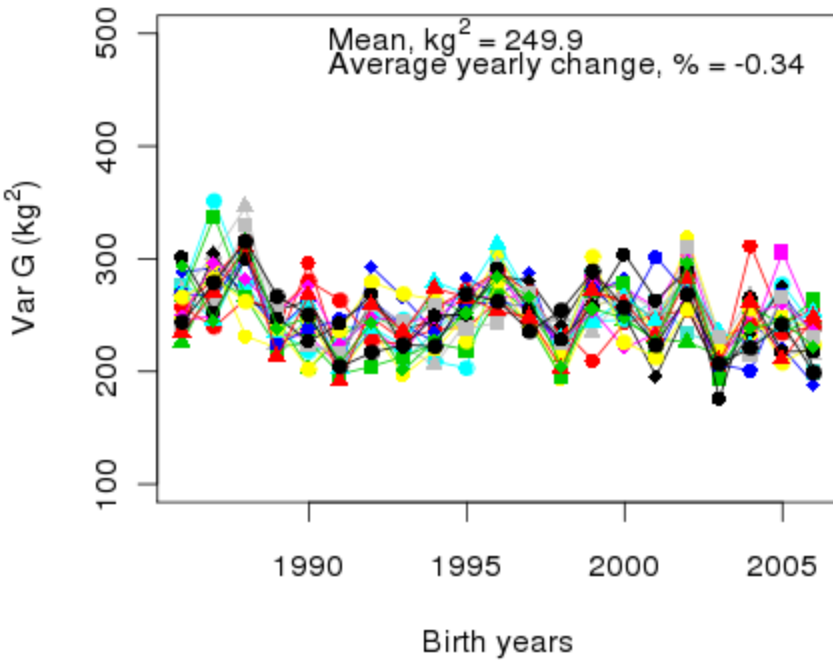


Analysis of MACE model residuals

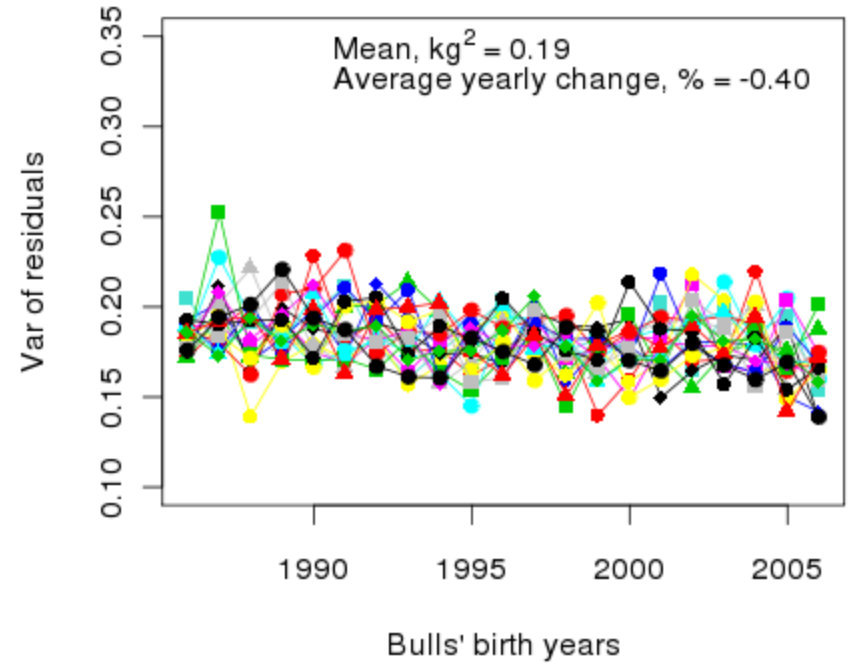
Scenario A



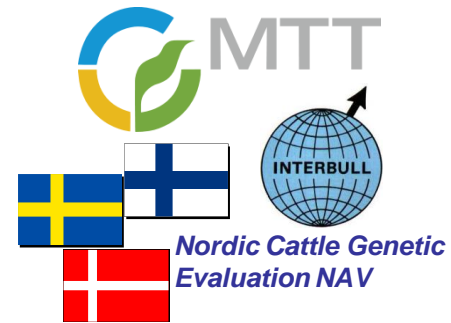
Var G



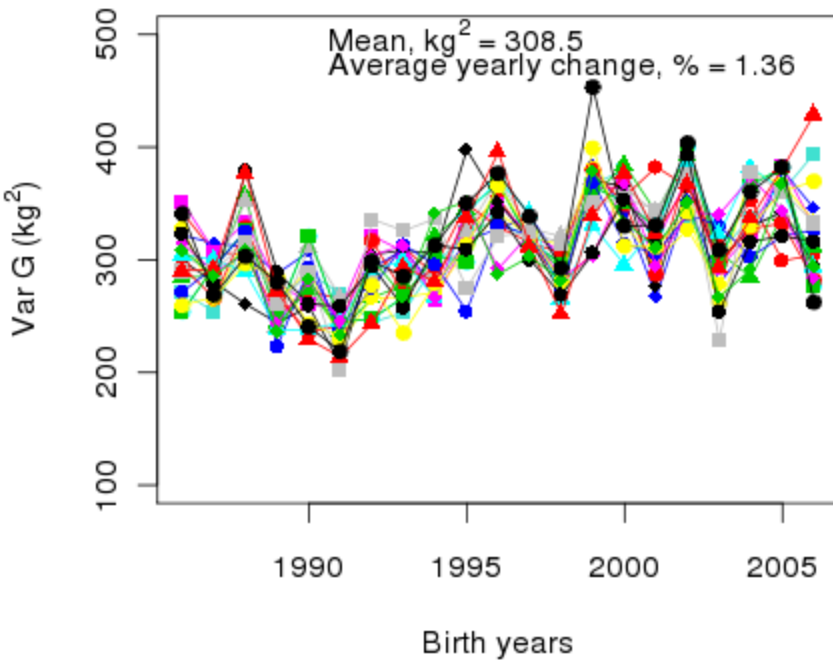
Var of MACE residuals



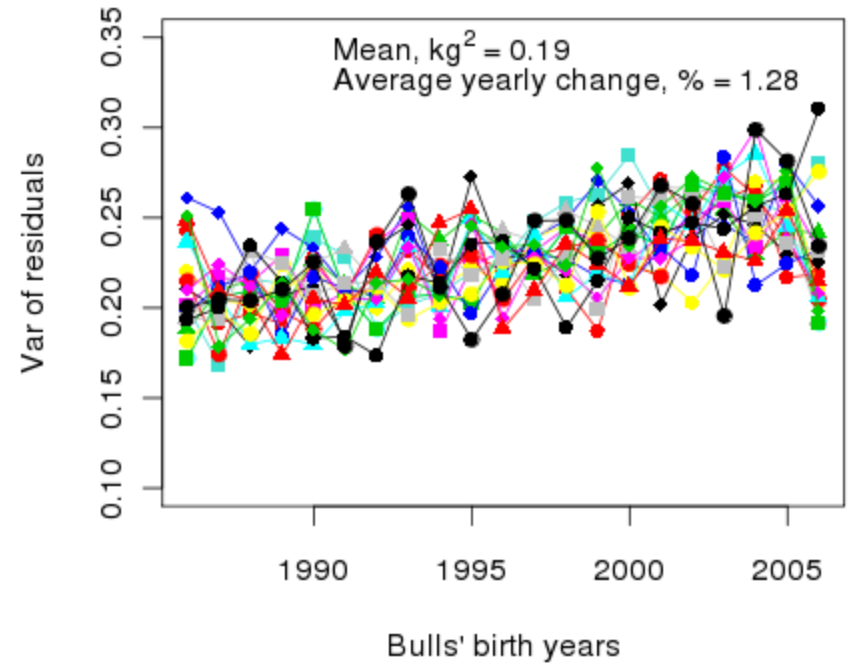
Scenario B, no HV



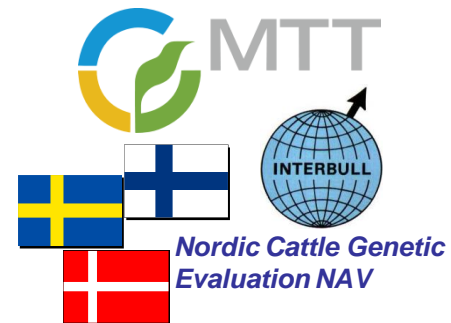
Var G



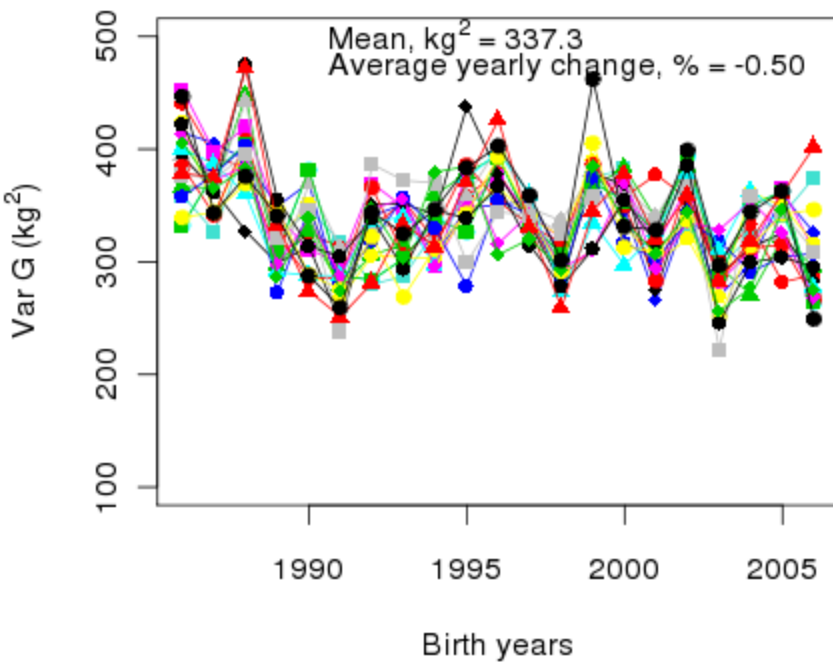
Var of MACE residuals



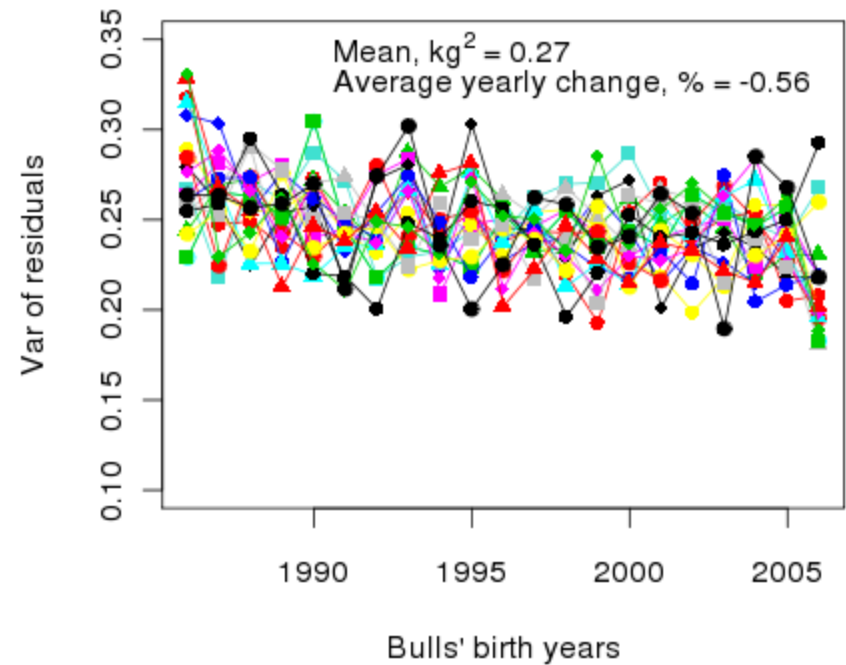
Scenario B, HV



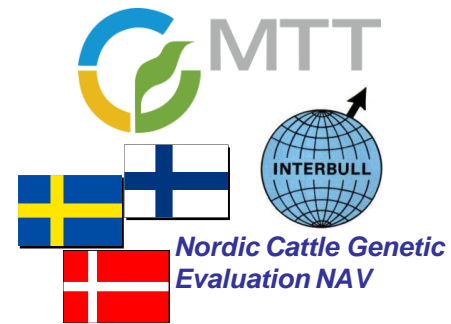
Var G



Var of MACE residuals

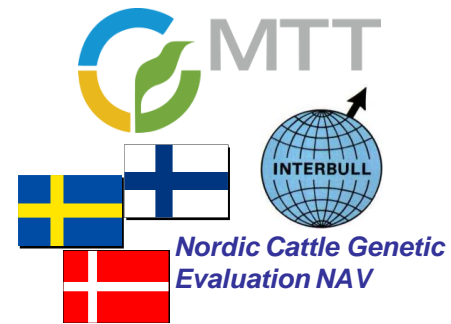


Conclusions



- IB4 is most reliable when applied on cows and it was able to detect simulated heterogeneity in Var G
- The most recent year group should be excluded when applied on cows
- Analysis of MACE residuals could be utilized as a quick and simple tool to check data quality for bulls

Work to be done



- Scenario C: a trend in genetic variance only (changing heritability)
- Test statistics
- How the tests react on small populations and pre-selection