

Estimation of GEBVs using deregressed individual cow breeding values

Esa A. Mäntysaari,

*Minna Koivula, Ismo Strandén,
Jukka Pösö and Gert P. Aamand*

Interbull Open Meeting 27.-28.8.2011

Stavanger, Norway



Background

Genomic evaluations in TWO stages

- 1) Use reference population to solve genomic model
- 2) Use genomic model to predict DGV (direct genomic value)

Combine genomic and pedigree information

$$\mathbf{GEBV} = \mathbf{b}_1 * \mathbf{EBV} + \mathbf{b}_2 * \mathbf{DGV}$$

Genomic evaluation in SINGLE step

- Use combined genomic+pedigree relationship matrix
- Solve GEBV with BLUP using original observations as depended variables

Background

Usual depended variables:

- Deregressed **sire** breeding values (DRP)
- Advantage: No need to model the environmental effects
- Can be used both in solving the genomic model and in combining/blending
- Also been used in Single step developments

Possible alternative?

Deregressed individual daughter breeding values (animal model DRP)

- Makes possible to do combining/blending in cow level
- Or, implement Single **step** approach

Objectives:

Test

1. Animal model deregression
2. Back-regression, i.e. recompute BLUP using DRPs
3. Animal model Single step
 - Interbull GEBV validation test for Bull GEBV

Using full population of Nordic Red Cattle (RDC)

Deregression Data

NAV Red Dairy Cattle (RDC) Cow EBV file:

- 4,578,942 animals in pedigree
- 3,401,346 cows with records

Model:

- MT Multi-lactation model with 27 biological TD traits
- For each cow combined EBVs for **Milk**, **Protein** and **Fat**
- For each cow each trait EBV, an EDC were estimated using absorption technique (Interbull Code of Practice April 27th 2004)
- All variance parameters in EDC calculation were as in NAV official model

EDC were computed using the MiX99 package

(Strandén, et al., 2001. INTERBULL Bulletin No. 26)

Deregression model

$$\text{DRP} = \mu + \text{EBV} + \varepsilon$$

where

- **DRP** is deregressed animal model proof for milk, protein or fat
- **EBV** is corresponding estimated breeding value
- ε are residuals with a $\text{var}(\varepsilon_i) = \sigma_e^2 / \text{EDC}_i$

Heritabilities were

- Milk 0.40
- Protein 0.28
- Fat 0.32

NOTE:

- Interbull instructs to combine dam reliability to daughter reliability
→ to be summed up for bull EDC

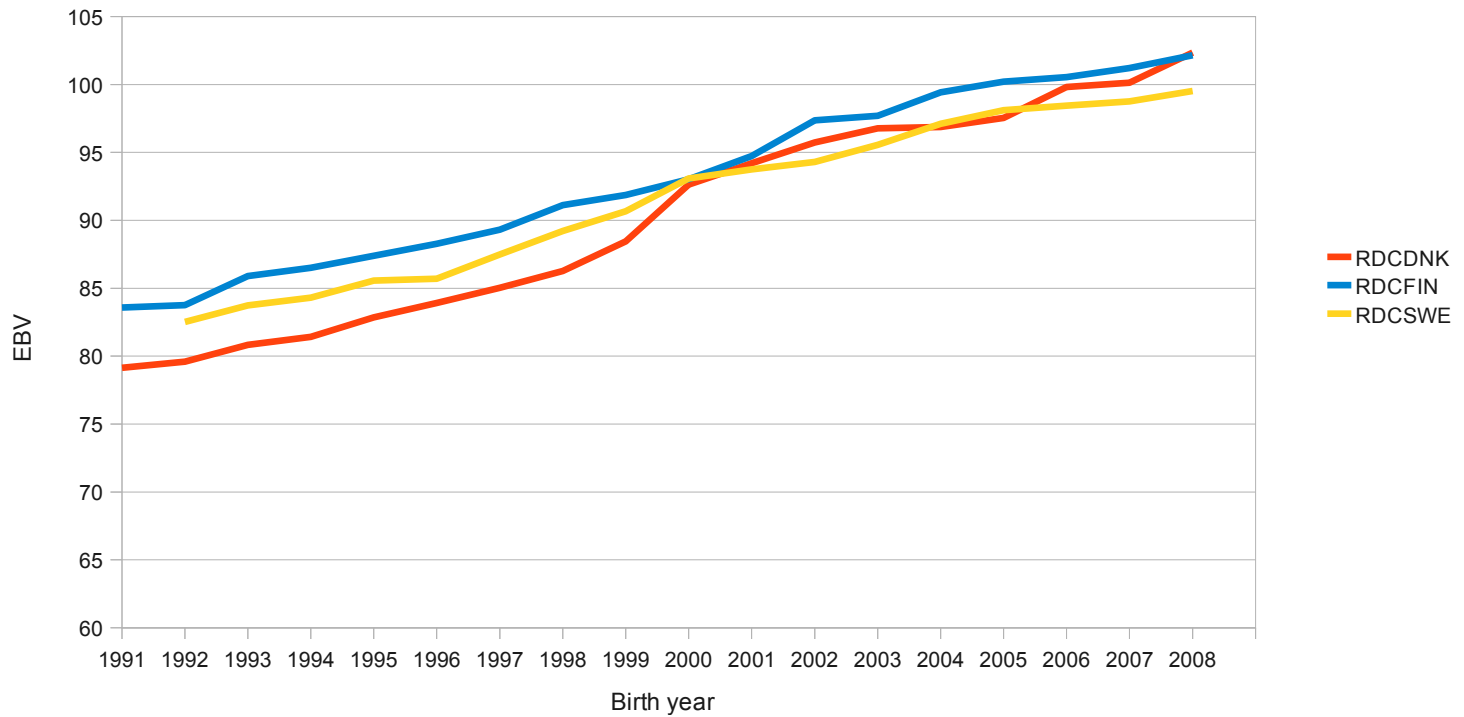
Here only individual cow EDCs were used in deregression
Needs to delete lines with no observations from analyses

Deregression was done using MiX99 package option "deregress"
(Strandén and Mäntysaari, 2010, Interbull Bulletin 42)

- MiX99s statistics
 - Secant method, 6 deregress rounds
 - Running time 13 min 38 sec (mix99s)
 - All 3 traits in same deregression run

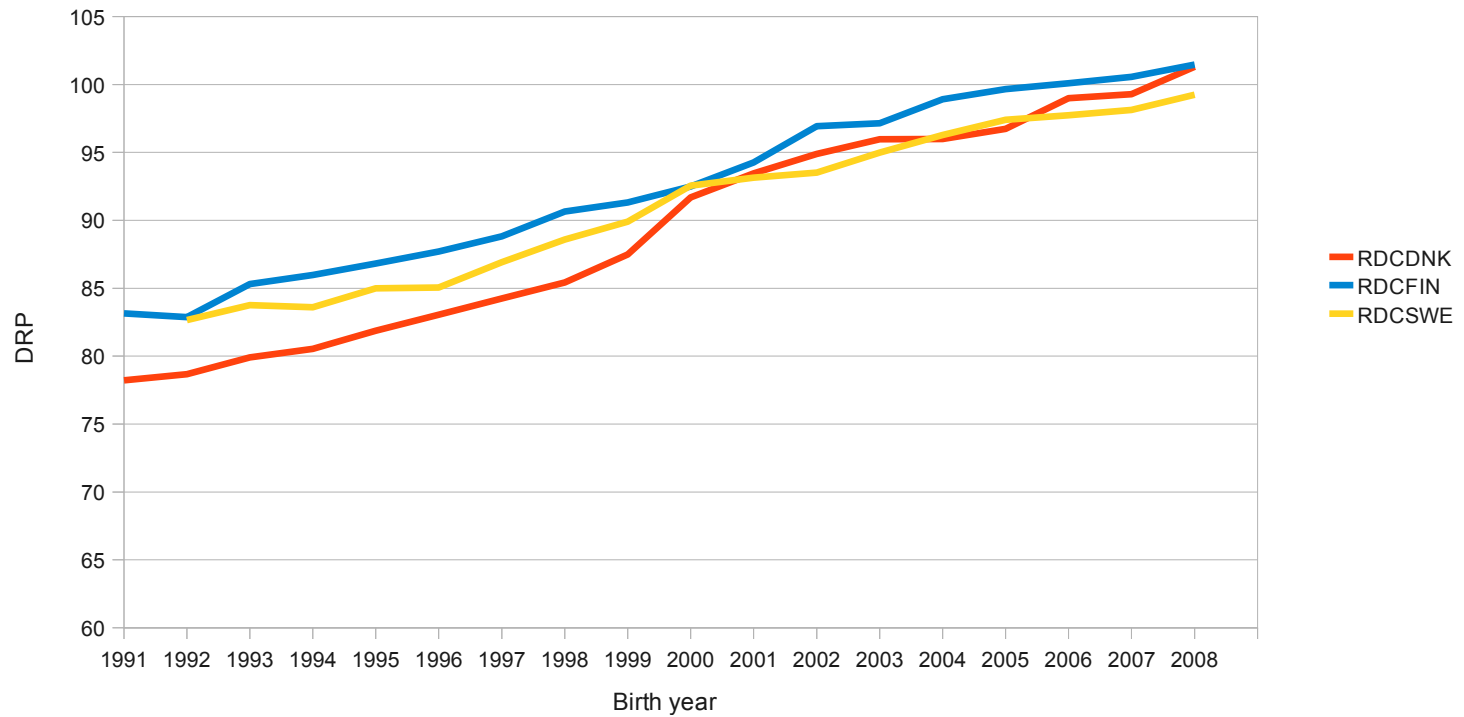
EBVs by birth year

Milk



Deregressed EBVs

Milk

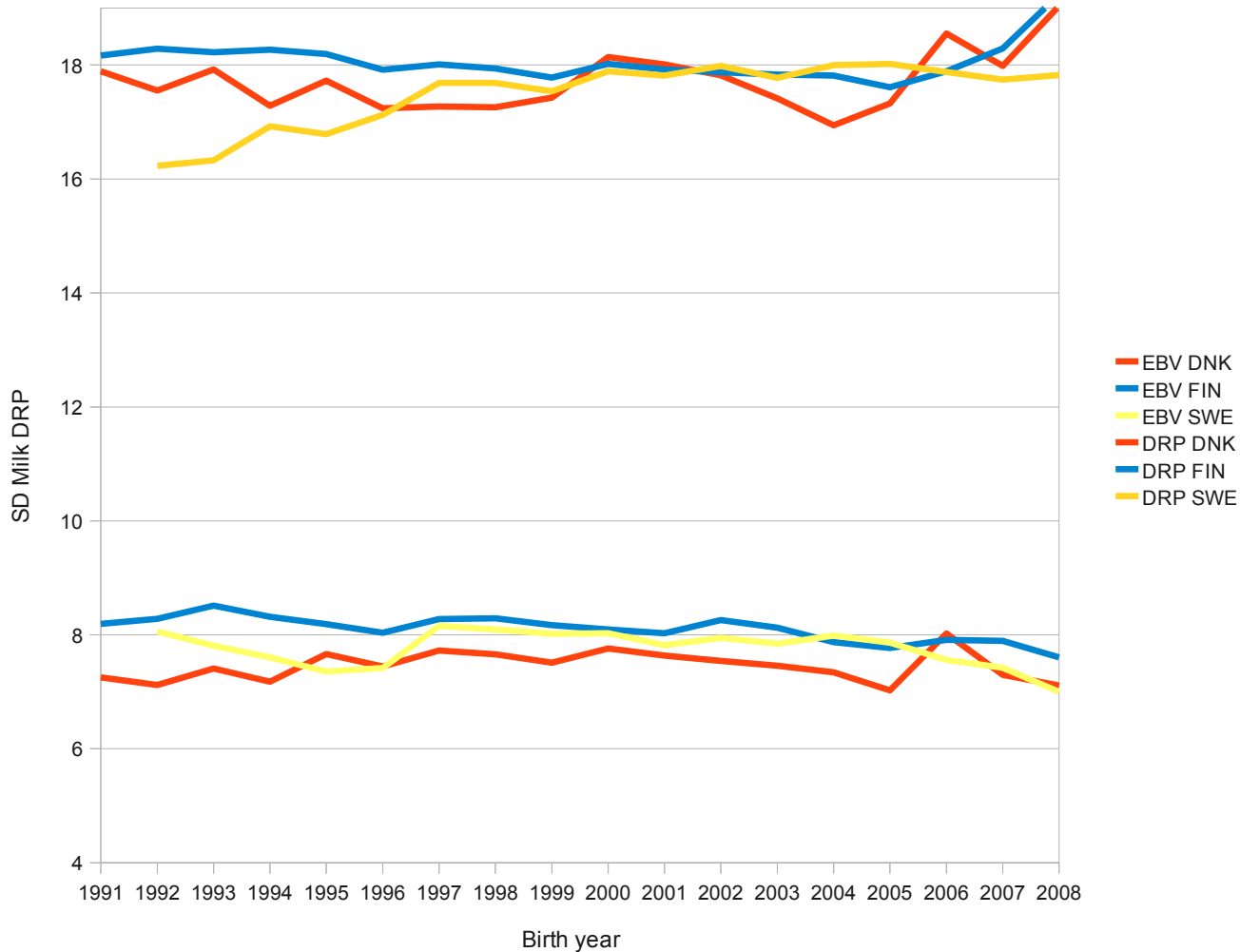


Interbull 27.-28.8.2011
Stavanger, Norway



SD of Cow DRPs

Milk



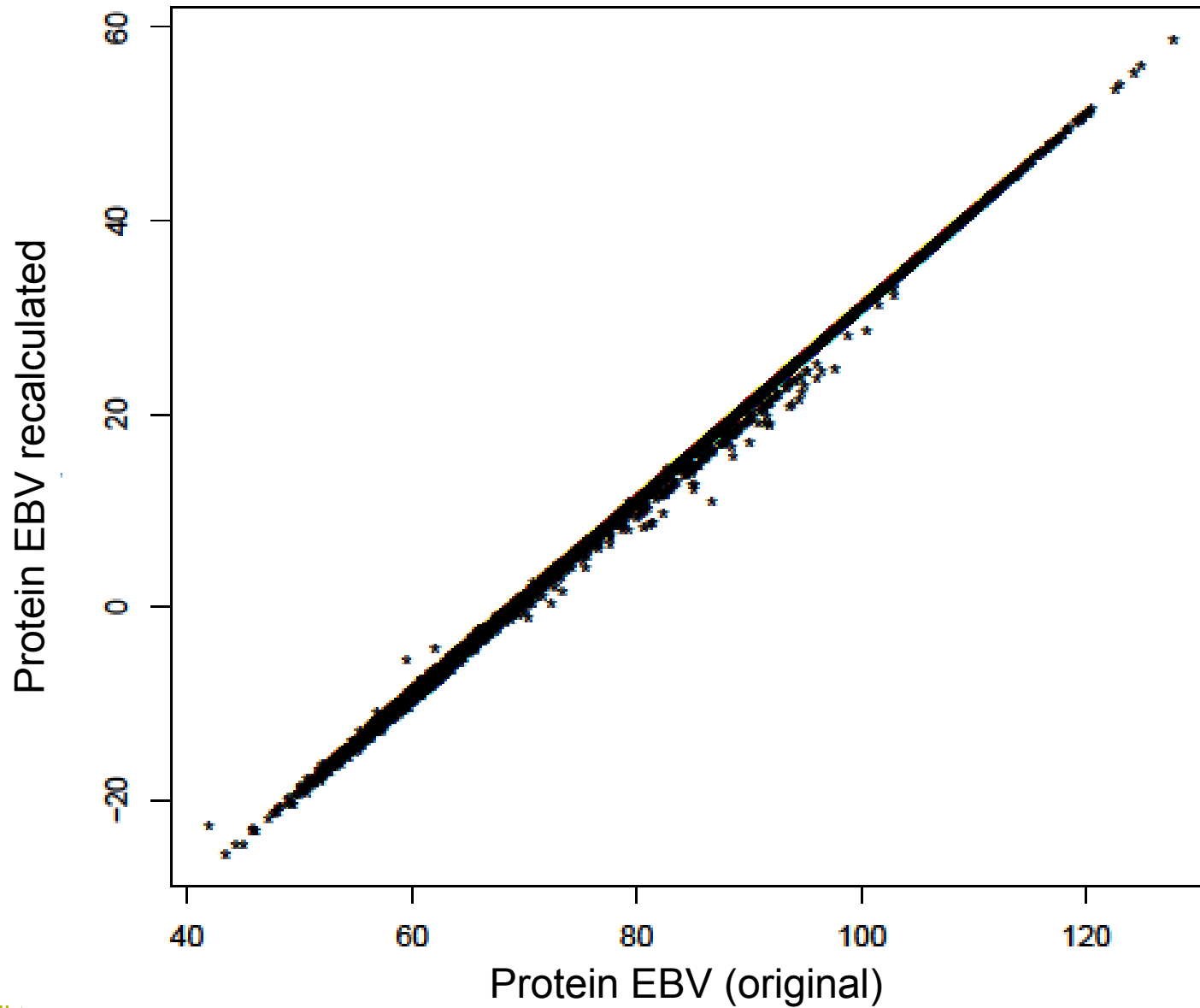
Testing the DRP in recalculation of bull EBV

Number of bulls in evaluation:	54175
RDC FIN bulls	14296
RDC DNK bulls	5619
RDC SWE bulls	17871

- Computing statistics:
 - PCG rounds 239
 - Solving time 9 min 56 sec for all 3 traits simultaneously
- Standard deviations

	Milk	Protein	Fat
EBV original	13.4	15.3	12.3
EBVresolved	13.4	15.2	12.6

Recalculated breeding values vs. Original EBV, Finnish Bulls

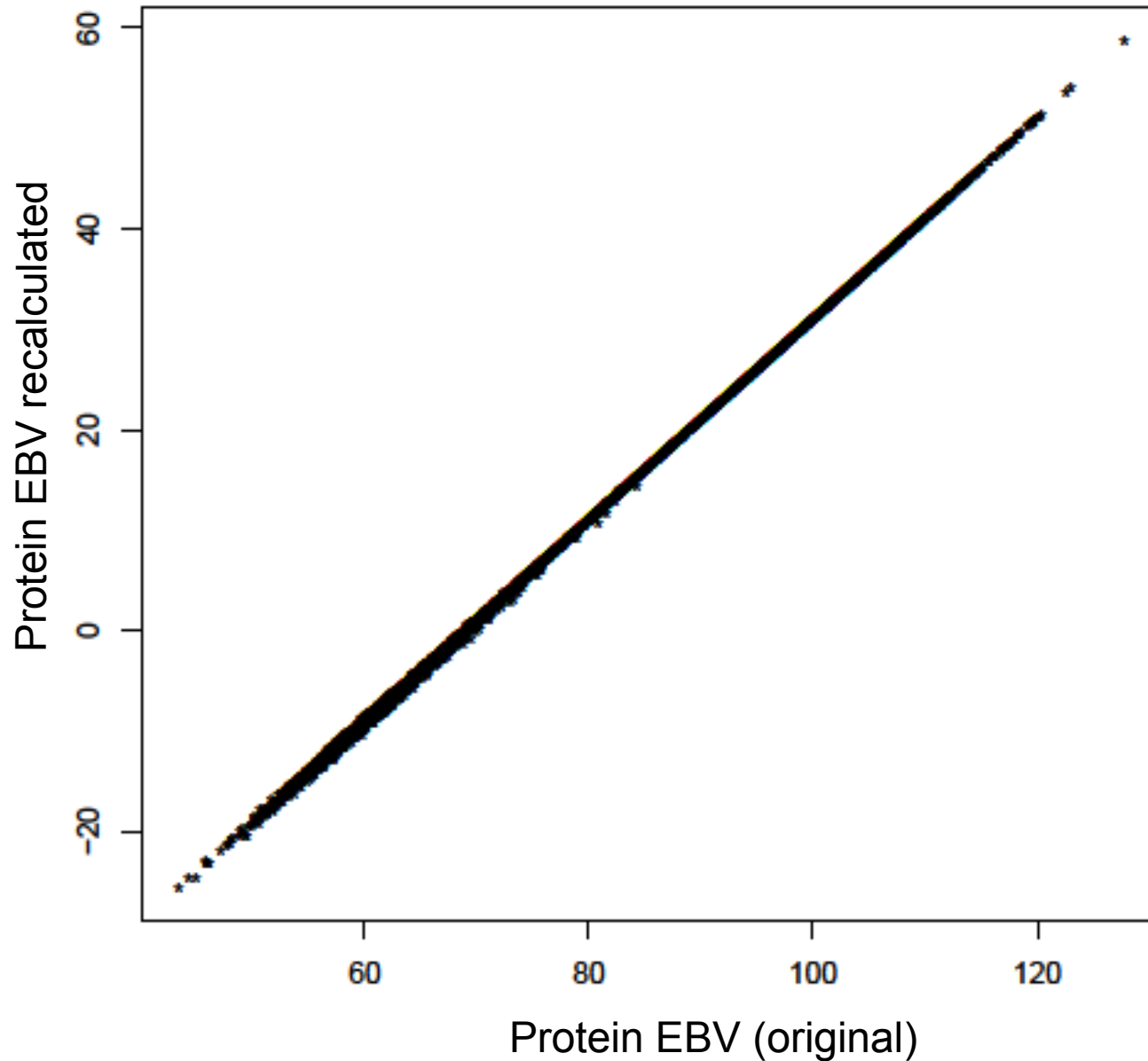


Validation of deregressed proofs

For bulls that have >20 daughters

	Correlation between original EBV and the EBV recalculated from the deregressed proofs	
	Finnish FAY	Swedish SRB
Milk	1.0000	0.9999
Protein	0.9999	0.9999
Fat	0.9999	1.0000

Recalculated breeding values vs. Original EBV, Finnish Bulls With more than 20 daughters



Solving Single-step Genomic evaluations

- H^{-1} matrix (combined A^{-1} and G^{-1})
 - Genotypes available for 4725 bulls
 - G -matrix scaled with $\sum 2pq$ and $\sum G_{ii} / \sum A_{ii}$
 - A^{-1} constructed using full pedigree file with all animals
 - 0.10 weight for polygenic variance (A^{22}) in H^{22}
- Implementation using PCG iteration
 - Iteration on data using usual pedigree file
 - Additional covariance structure (H^{22}) read in each iteration
 - Implemented in MiX99 software package

Testing single step Genomic evaluations

- Validation:
 - Bull deregressed proofs were composed from the same data
 - **809 candidate bulls** named from EBV history files:
 - Genotyped bulls with: $N_{\text{daugh}}_{2005} = 0$ AND $N_{\text{daugh}}_{2010} > 20$
Number of daughters 153,389 [on aver. 188]
 - GEBVs calculated from a data subset
(full population – daughters of candidates) “Reference pop’n”
 - Number of “reference bulls” 11,264
 - In “ the GEBV2005” evaluation 2,814,741 cow records
 - **Validation test:** $DRP_{2011} = b_0 + b_1 * GEBV_{2005} + \text{residual}$
 - Validation $R^2 = (\text{Model } R^2) / (R^2_{\text{DRP}})$

Single step analysis – implementation issues

Two alternative Single Step -model parameters were tested:

Animal_D Using the same variance component as in deregression

Animal_E Estimation of variance components from the data using sire model and animal model DRPs

- Estimated “genomic heritabilities” were higher for milk and protein, and lower for fat

- Computing time for 3 traits simultaneously: 40-50 minutes

Validation results (single step analysis)

		Protein			
		b_0	b_1	R^2	
PA		1.22	0.89	0.25	
DGV [§]		4.51	0.77	0.31	
Single step sire model [§]		4.70	0.74	0.35	
Single step Animal_D		4.23	0.81	0.38	
Single step Animal_E		5.01	0.76	0.36	

§ Koivula et al 2011. Comparison of different GEBV models

Validation results (single step analysis)

		Protein				
		b_0	b_1	R^2		
PA		1.22	0.89	0.25		
DGV [§]		4.51	0.77	0.31		
Single step sire model [§]		4.70	0.74	0.35		
Single step Animal_D		4.23	0.81	0.38		
Single step Animal_E		5.01	0.76	0.36		

§ Koivula et al 2011. Comparison of different GEBV models

Validation results (single step analysis)

		Protein			
		b_0	b_1	R^2	
PA		1.22	0.89	0.25	
DGV [§]		4.51	0.77	0.31	
Single step sire model [§]		4.70	0.74	0.35	
Single step Animal_D		4.23	0.81	0.38	
Single step Animal_E		5.01	0.76	0.36	

§ Koivula et al 2011. Comparison of different GEBV models

Validation results (single step analysis)

	Milk			Protein			Fat		
	b_0	b_1	R^2	b_0	b_1	R^2	b_0	b_1	R^2
PA	3.29	0.70	0.22	1.22	0.89	0.25	2.17	0.80	0.28
DGV [§]	3.15	0.76	0.30	4.51	0.77	0.31	2.23	0.85	0.40
Single step sire model [§]	3.67	0.69	0.32	4.70	0.74	0.35	2.69	0.80	0.44
Single step Animal_D	3.80	0.72	0.35	4.23	0.81	0.38	3.29	0.79	0.45
Single step Animal_E	3.90	0.71	0.35	5.01	0.76	0.36	2.17	0.80	0.45

§ Koivula et al 2011. Comparison of different GEBV models

Conclusions

- Animal model deregression seems to work (maybe better than sire deregression)
- Computing times for deregression and single step approach were reasonable
- Validation results for animal model single step were better than the same from the sire model
 - Information from female relatives accounted