Variance Component Re-Estimation and Model Simplification in a Single-Step Random Regression Test-Day Model for Nordic Red Dairy Cattle

T. Pitkänen, M.H Lidauer, M. Taskinen, J. Pösö, A. Kudinov, K. Byskov, H. Splittorff, M. Koivula

NAV Nordisk Avlsværdi Vurdering
Nordic Cattle Genetic Evaluation

Luke

VIKINGGENETICS®
breeding for what truly matters

Introduction

- Updates made on the Nordic RDC model
- Variance component estimation
- Results
- Conclusion

Updates made to the model

Updates made

Data and pedigree modifications

- Test day data truncated from year 2005, FIN later lactations removed, FIN HOL observations removed
- Updated UPG/metafounder grouping
- data ~34.7 million records from 1.78 million animals, pedigree ~3.1 million animals
- Genotypes from 315 448 animals (born after 2008)

Effect updates

- Calving age, heterosis calculation (EAAP 2024), herd-year-season

From country specific traits (27 traits) to 9 traits

- Country interactions included for some effects

New VC estimations and covariance functions

Variance component estimation

Milking Systems

- Test-day records are obtained from cows under conventional (CMS) and automated milking systems (AMS)
- In Finland, the protocol for AMS milk yield has changed during the years:
  - 30d average => 24h equivalent => 96h average
- In Denmark and Sweden, the protocol for AMS has been the same for all years
- The aim was to estimate residual covariance matrices for each milking system and country
  - 8 MS x Country classes

Luke

Criteria for cows and herds for VCE data sampling

For Cows:

- All records must come from a single herd
- Complete lactation sequence required:
  - Either only **1st**, **1st to 2nd** or **1st to 3rd** lactations
  - No missing lactations before the last observed one

- **For Herds**:
  - At least 8 but less than 100 suitable first-lactation cows per year (on average)
  - Number of years in data: 12 or more (starting from 2010)
  - Herd can change milking system during the years

Luke

Sampling

- The target sample size was set at 100,000 animals
- Herds were sampled independently within each Country x MS class
- The number of herds per Country X MS was allocated proportionally based on the share of the first-lactation cows in each class

Proportion of first lactation cows by country and milking system

Country	AMS	CMS
FIN	15	30
DNK	5	10
SWE	18	22

Description of the sample

			N HERD		N ani		N milk obs		N fat obs		N prot obs	
Lact	N herd	N ani	CMS	AMS	CMS	AMS	CMS	AMS	CMS	AMS	CMS	AMS
FIN												
1	223	46178	222	105	34270	13877	336330	134882	336330	134882	336330	134882
2	223	36841	223	105	26784	11642	242317	104794	242317	104794	242317	104794
3	223	24446	218	105	17356	8083	151816	70494	151816	70494	151816	70494
DNK												
1	35	16336	26	12	11103	5295	98367	47511	98367	47511	98367	47511
2	35	12769	26	11	8667	4140	69361	33757	69361	33757	69361	33757
3	35	7776	24	11	5313	2463	41095	19333	41095	19333	41095	19333
SWE												
1	138	42700	80	58	22884	19816	216344	187218	216344	187218	216344	187218
2	138	31033	80	58	16521	14512	143544	128449	143544	128449	143544	128449
3	138	18777	80	58	9811	8966	82031	76959	82031	76959	82031	76959

Note:

If herd/animal has at least 1 obs in both milking system, it will be calculated in both milking systems! Single animals in AMS herds are sometimes milked in CMS.

VCE model

Variance components were estimated using model:

$$Y = X\beta + HTD + \Phi_0 hc + \Phi_1 pe_i + \Phi_1 u_i + \epsilon,$$

Fixed effects

Heterosis, Recombination, Herd x Year x Season, YM x MS x Country,

Lactation curve, DCC, Linear herd term

VCE model, random effects

$$Y = X\beta + HTD + \Phi_0 hc + \Phi_1 pe + \Phi_1 \mathrm{u} + \epsilon,$$

Random effect	Function	Covariance matrix
Herd-test-day (HTD)	intercept	$\boldsymbol{H}_{9\times9}$
Herd-lactation curve (hc)	2nd order Legendre + Wilmink - intercept, 2nd and 3rd lactation combined	$\boldsymbol{C}_{18\times18}$
Non-genetic animal (pe)	3rd order Legendre + Wilmink	$\boldsymbol{P}_{36\times36}$
Genetic (u)	3rd order Legendre + Wilmink	$\boldsymbol{G}_{36\times36}$
Residual (e)	Lactation divided to 12 intervals	$\boldsymbol{R}$

Variance components were estimated using PBLUP model and MiX99

Parameter-Reduced Covariance Functions for the Evaluation Model

- The parameter-reduced genetic (GE) and permanent environmental (PE) effects each retain ~99% of their original variance
- Dimension of the genetic covariance matrix reduced from 36 → a CF with rank 16.

Permanent Environmental (PE) Effect

- Dimension reduced to a CF with rank 21
- The parameter-reduced PE effect CF captures variation from
  - PE, HC and residual variation within lactation

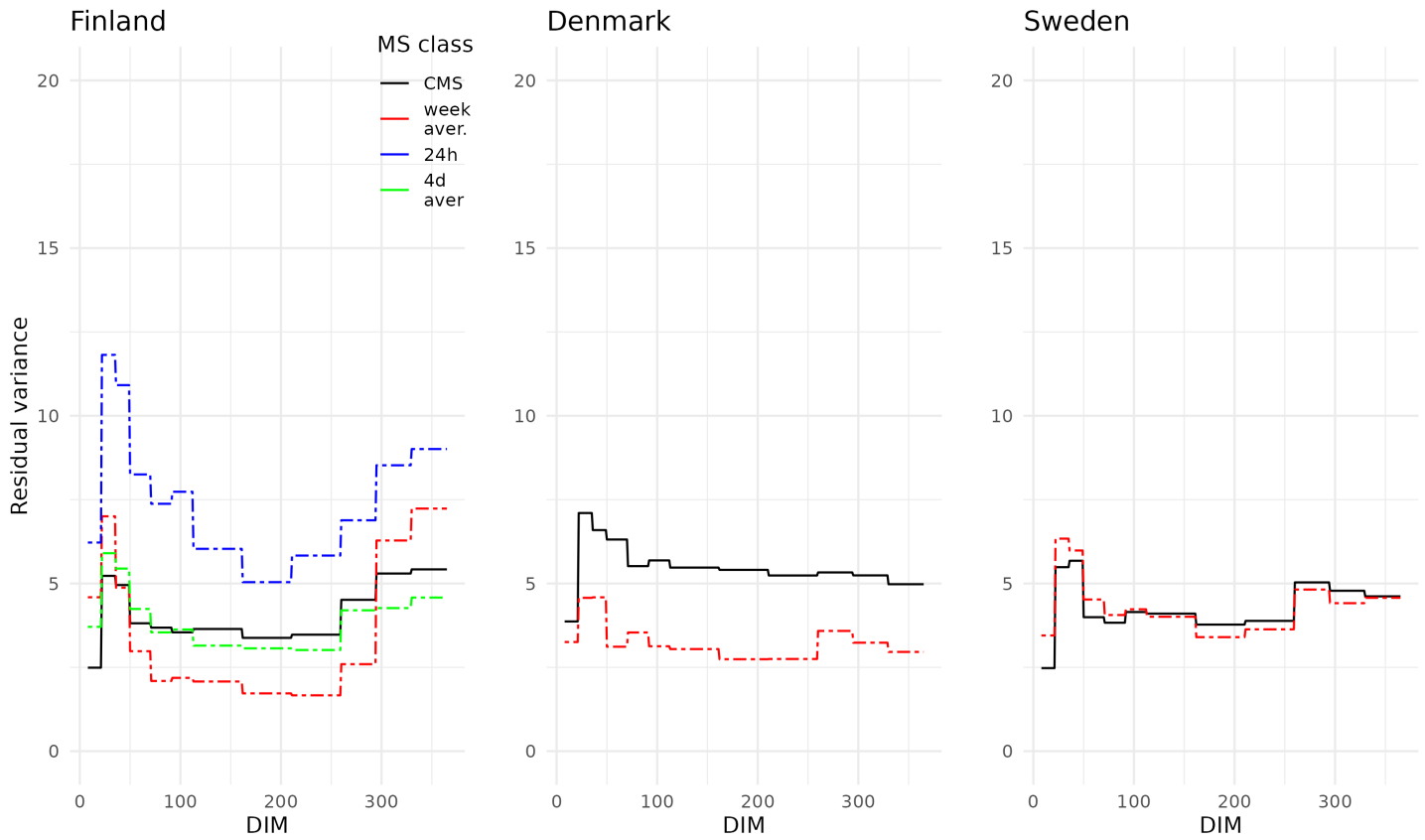
Luke

305d Heritabilities

Trait	New	Old	Change in %
Lact 1			
Milk	0.435	0.385	13
Prot	0.374	0.333	12
Fat	0.381	0.339	12
Lact 2			
Milk	0.353	0.280	26
Prot	0.320	0.268	19
Fat	0.304	0.253	20
Lact 3			
Milk	0.314	0.253	24
Prot	0.293	0.242	21
Fat	0.293	0.201	46

Luke

Residual variances for Milk first lactation



Validation

Validation results using Linear Regression § - Four years data cut

§ Regression of GEBV_FULL on GEBV_data-4

		BULLS (n=191)			COWS (n=62297)		
	Model	Bias (kg)	b_1	R^2	Bias (kg)	b_1	R^2
MILK	Old VCE	-267.21 (±377.8)	1.06	0.75	-120.82 (±279.1)	1.03	0.83
	New VCE	-296.67 (±397.7)	1.02	0.73	-147.94 (±303.6)	1.01	0.81
PROTEIN	Old VCE	-9.48 (±13.4)	0.99	0.64	-3.75 (±9.8)	0.98	0.78
	New VCE	-11.00 (±13.9)	0.96	0.63	-4.92 (±10.1)	0.96	0.77
FAT	Old VCE	-11.12 (±15.8)	0.97	0.67	-3.84 (±11.8)	0.97	0.79
	New VCE	-13.60 (±16.8)	0.97	0.68	-5.69 (±12.5)	0.97	0.80

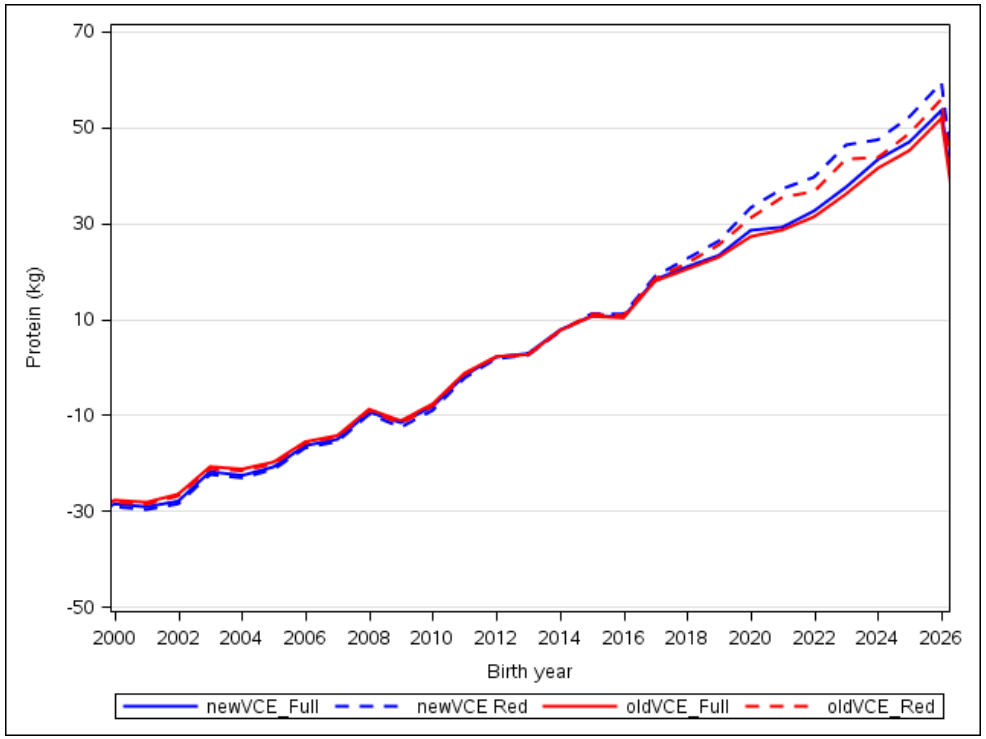
bias=mean(Full_GEBV – reduced_GEBV)

Luke

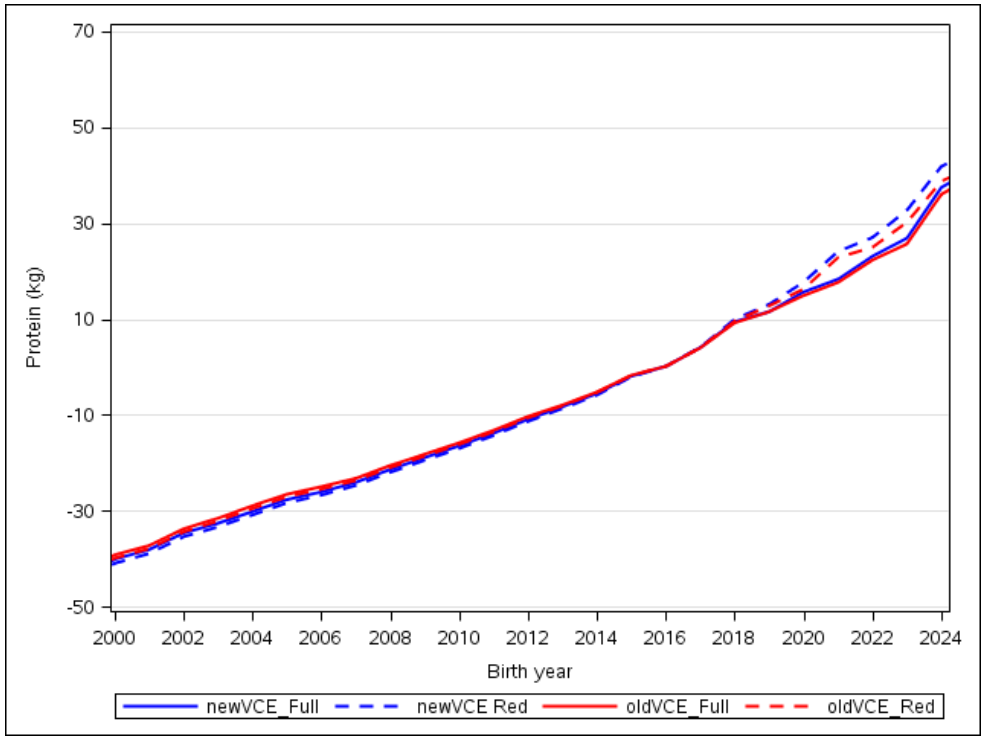
Validation bulls have at least 20 daughters in full data on no daughters in reduced data

Protein trends

Bulls



Cows



New variance components increase slightly the trend and SDs are higher

Conclusions

- The new model is simplified and models better the current population structure
- We have first Nordic variance components and residual covariances are estimated more accurately
- New variance components increase slightly the trend and SD and cause some re-ranking of animals.
- Validations give good results with the new updated model
  - Old variance components give slightly better validation results
  - However, new ones are more appropriate for the Nordic population

Luke

thank you!

luke.fi