Test-Day single-step genomic evaluation using APY algorithm

Minna Koivula, Ismo Strandén, Gert Pedersen Aamand and Esa A. Mäntysaari

NATURAL RESOURCES
INSTITUTE FINLAND

Introduction

- Number of genotyped animals has increased rapidly creating computational challenges for genomic evaluation
- The APY algorithm for core and young animals is one approach to overcome the challenges

Aims

- We compared the single-step BLUP (ssGBLUP) using APY algorithm and regular ssGBLUP for the joint Nordic Red dairy cattle (RDC) evaluations for milk, protein, and fat using test-day model

Materials and methods

- Nordic NAV RDC test-day (TD) data December 2015

  ~4 million cows with a total of 90 million records
  - 5.4 million animals in pedigree
  - 230 million equations
- Multiple trait multi-lactation models:
  - Production evaluation
  - Milk, protein and fat 305d yield (G)EBVs for all animals

Genotype data

- 56 017 genotyped RDC animals with 46 914 SNPs
  - 38 477 cows and heifers
    - » 21 170 cows with TD records
  - 17 540 bulls and bull calves
    - » 5 925 reference bulls

H-matrix in ssGBLUP

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau((1-\mathrm{w})\mathbf{G} + \mathrm{w}\mathbf{A}_{22})^{-1} - \omega\mathbf{A}_{22}^{-1} \end{bmatrix}$$

3 PARAMETERS w, τ and ω in $\tau\mathbf{G}_w^{-1} - \omega\mathbf{A}_{22}^{-1}$

where $\mathbf{G}_w = (1-w)\mathbf{G} + w\,\mathbf{A}_{22}$

w is proportion of polygenic effect not accounted by the SNPs

SS1: $\tau=1.0$ and $\omega=1.0$ and $w=0.05$

SS2: $\tau=1.0$ and $\omega=0.7$ and $w=0.05$

$\mathbf{G}^{-1}_{\mathbf{APY}}$ with Core= reference bulls in full TD data + cows born < 2012 (N=16859)

APY1: $\tau=1.0$ and $\omega=1.0$ and $w=0.00$

APY2: $\tau=1.0$ and $\omega=1.0$ and $w=0.05$

APY3: $\tau=1.0$ and $\omega=0.7$ and $w=0.05$

- Solutions via PCG iteration
  - matrix vector product $\mathbf{A}_{22}^{-1}\mathbf{d}_2$ without ever making the $\mathbf{A}_{22}$ matrix → save memory and computing time

(see Strandén et al. EAAP 2016)

Validation set up

Full TD run included all observations

1) Full data → **$\mathbf{EBV_F}$**

Reduced run – data until December 2011

(4 years of observations removed)

2) Reduced data → **$\mathbf{EBV_R}$** (PA) and **$\mathbf{GEBV_R}$**

For validation

1. Effective record contributions (ERC) calculated
2. Deregressed bull EBVs (DRP) from the EBV$_F$ for
   Milk, Protein, Fat
3. Validation bulls
   - no daughters with observations in reduced data
   - had ERC > 3 in the full TD data

   → 626 validation bulls

Correlations of protein GEBVs for the reference bulls

	EBV	SS1	SS2	APY1	APY2	APY3
EBV		0.994	0.996	0.993	0.994	0.996
SS1			0.998	0.999	0.999	0.998
SS2				0.998	0.998	0.999
APY1					0.999	0.998
APY2						0.998

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;
APY1: basic APY; APY2: w=0.05;
APY3: τ=1.0 and ω=0.7 and w=0.05

Validation results for bulls n=626

	Milk		Protein		Fat	
	b_1	R^2	b_1	R^2	b_1	R^2
PA	0.93	**0.34**	0.85	**0.28**	0.75	**0.28**
SS1	0.57	**0.37**	0.47	**0.30**	0.53	**0.36**
SS2	0.97	**0.44**	0.81	**0.35**	0.83	**0.40**
APY1	0.56	**0.36**	0.46	**0.30**	0.52	**0.36**
APY2	0.57	**0.36**	0.46	**0.30**	0.52	**0.36**
APY3	0.97	**0.44**	0.81	**0.35**	0.82	**0.40**

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;
APY1: basic APY; APY2: w=0.05;
APY3: τ=1.0 and ω=0.7 and w=0.05

Validation results for bulls n=626

	Milk		Protein		Fat	
	b_1	R^2	b_1	R^2	b_1	R^2
PA	0.93	**0.34**	0.85	**0.28**	0.75	**0.28**
SS1	0.57	**0.37**	0.47	**0.30**	0.53	**0.36**
SS2	0.97	**0.44**	0.81	**0.35**	0.83	**0.40**
APY1	0.56	**0.36**	0.46	**0.30**	0.52	**0.36**
APY2	0.57	**0.36**	0.46	**0.30**	0.52	**0.36**
APY3	0.97	**0.44**	0.81	**0.35**	0.82	**0.40**

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;

APY1: basic APY; APY2: w=0.05;

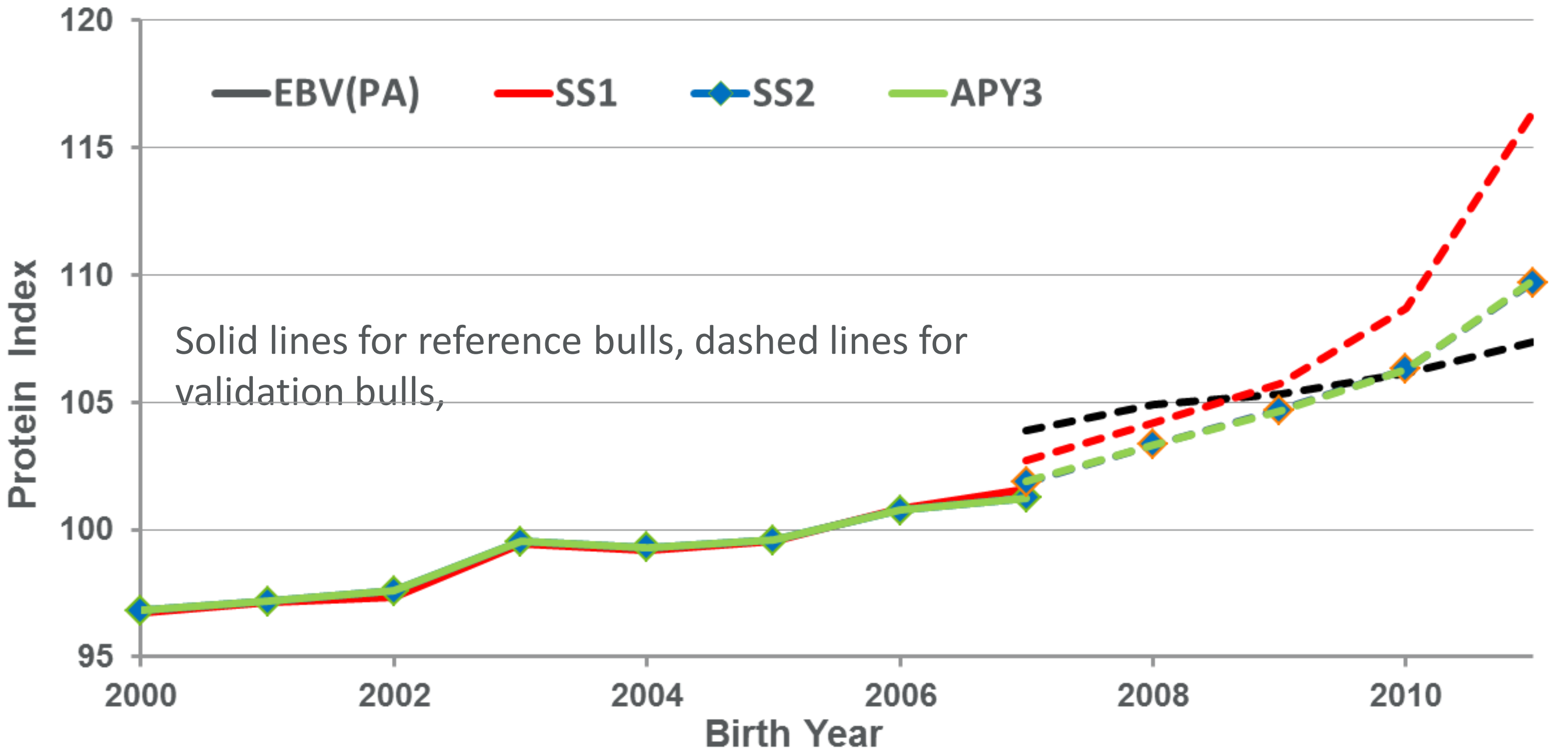
APY3: τ=1.0 and ω=0.7 and w=0.05

Validation results for bulls n=626

	Milk		Protein		Fat	
	b_1	R^2	b_1	R^2	b_1	R^2
PA	0.93	**0.34**	0.85	**0.28**	0.75	**0.28**
SS1	0.57	**0.37**	0.47	**0.30**	0.53	**0.36**
SS2	0.97	**0.44**	0.81	**0.35**	0.83	**0.40**
APY1	0.56	**0.36**	0.46	**0.30**	0.52	**0.36**
APY2	0.57	**0.36**	0.46	**0.30**	0.52	**0.36**
APY3	0.97	**0.44**	0.81	**0.35**	0.82	**0.40**

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;
APY1: basic APY; APY2: w=0.05;
APY3: τ=1.0 and ω=0.7 and w=0.05

Trends for protein (G)EBVs – reference and validation bulls

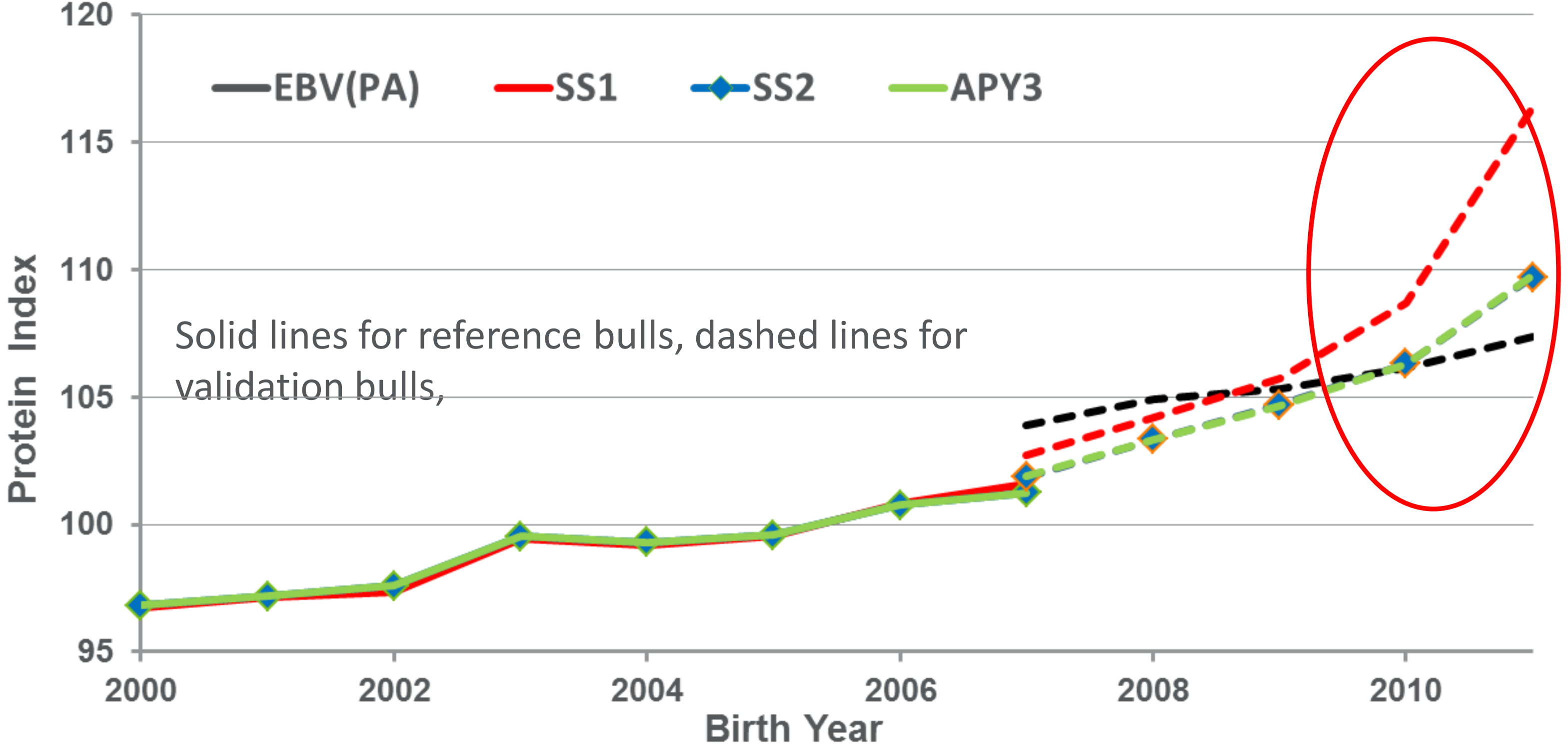


Solid lines for reference bulls, dashed lines for validation bulls,

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;

APY3: τ=1.0 and ω=0.7 and w=0.05

Trends for protein (G)EBVs – reference and validation bulls

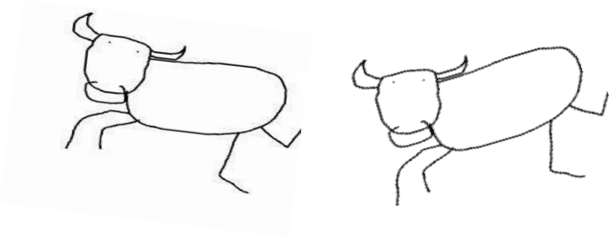


Solid lines for reference bulls, dashed lines for validation bulls,

SS1: w=0.05; SS2: τ=1.0 and ω=0.7 and w=0.05;

APY3: τ=1.0 and ω=0.7 and w=0.05

Discussion



- APY works as well as regular ssGBLUP in the TD ssGBLUP
- For reference bulls GEBVs are ~ same (correlation almost one)
- In validation APY gives similar results compared to regular ssGBLUP

Discussion


Erkki Oksanen/Luke

- However,
  - To minimize bias in validation, it is neccessary to consider the right weighting factors also in APY
  - Using only polygenic fraction w in **G** is not enough
  - ω seems to affect the bias considerably
    - corrects also the overestimation of the genetic trend

Acknowledgements


Erkki Oksanen/Luke

Phenotypic and Genomic data

NAV Nordic Cattle Genetic Evaluation,

Support:

Ministry of Agriculture and Forestry Finland,

NAV, Viking Genetics, FABA Coop and Valio Ltd

Thank you!

NATURAL RESOURCES
INSTITUTE FINLAND

Luke
NATURAL RESOURCES
INSTITUTE FINLAND