Single-step genomic predictions for yield traits in US Holsteins with UPG and phenotype-pedigree truncation

University of Georgia
1785
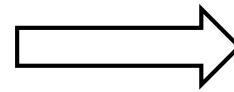College of Agricultural & Environmental Sciences

Interbull 2021 - Virtual

Daniela Lourenco

A. Cesarani, Y. Masuda, S. Tsuruta, E. Nicolazzi, P. VanRaden, and I. Misztal

Motivation

- Missing pedigree in the US Holstein data

  - ~ 10% sires

  - ~ 20% dams

  $\Longrightarrow$ Biased EBV

- QP-transformation for $\mathbf{A}^{-1}$

- (Quaas & Pollack, 1981;Westell et al., 1988)

$$\mathbf{A}^* = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{Q} \\ -\mathbf{Q}'\mathbf{A}^{-1} & \mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q} \end{bmatrix}$$

- QP-transformation for $\mathbf{H}^{-1}$ (Misztal et al., 2013)

$$\mathbf{H}^* = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} & -(\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1})\mathbf{Q}_2 \\ 0 & -\mathbf{Q}_2'(\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}) & \mathbf{Q}_2'(\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1})\mathbf{Q}_2 \end{bmatrix}$$
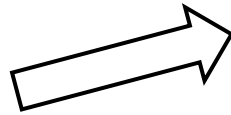
- Altered QP-transformation for $\mathbf{H}^{-1}$ (Tsuruta et al., 2019)

$$\mathbf{H}^* = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} & -(-\mathbf{A}_{22}^{-1})\mathbf{Q}_2 \\ 0 & -\mathbf{Q}_2'(-\mathbf{A}_{22}^{-1}) & \mathbf{Q}_2'(-\mathbf{A}_{22}^{-1})\mathbf{Q}_2 \end{bmatrix}$$

Motivation

- Matilainen et al. (2016): female fertility traits in Nordic Reds

- Tsuruta et al. (2019): type traits in US Holsteins

- Masuda et al. 2018 (Protein)

UPG poorly estimated
Large number of genotyped females with missing pedigree and no phenotypes

Data	UPG	R2	b1
Truncated 2011	Pedigree	0.52	0.78
	Ped. + Genomic	0.32	0.51
	No UPGs	0.50	0.78

- Lourenco et al. (2014): pedigree truncation helped to reduce bias

- due to missingness

Objectives

- Assess bias and reliability of GEBV for bulls and cows in ssGBLUP

  - Milk, Fat, and Protein

  - UPG for $\mathbf{A}^{-1}$ (SS_UPG)

  - UPG for $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}^{-1}$ (SS_UPG2)

- Six phenotype-pedigree truncation scenarios

  - 1980, 1990, or 2000

  - Pedigree depth 2 or 3

- Feasibility of ssGBLUP for dairy evaluations in the US

- US Holstein data up to December 2018

Phenotype cut-off scenario	Records		Genotypes	Animals in pedigree	
	N	Cows		Depth=3	Depth=2
Pheno1980	77.8 M	31.5 M	862 K	40.5 M	39.9 M

Analyses

- QP-transformation for $\mathbf{A}^{-1}$ in $\mathbf{H}^{-1}$

$$\mathbf{H}^*_{UPG} = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}^{-1}_{APY} - \mathbf{A}^{-1}_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Longrightarrow \qquad \text{SS\_UPG}$$

- Altered QP-transformation for $\mathbf{H}^{-1}$ (Tsuruta et al., 2019)

$$\mathbf{H}^*_{UPG2} = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}^{-1}_{APY} - \mathbf{A}^{-1}_{22} & -(-\mathbf{A}^{-1}_{22})\mathbf{Q}_2 \\ 0 & -\mathbf{Q}'_2(-\mathbf{A}^{-1}_{22}) & \mathbf{Q}'_2(-\mathbf{A}^{-1}_{22})\mathbf{Q}_2 \end{bmatrix} \qquad \Longrightarrow \qquad \text{SS\_UPG2}$$

- BLUP90IOD2OMP1 for each phenotype-pedigree truncation scenario
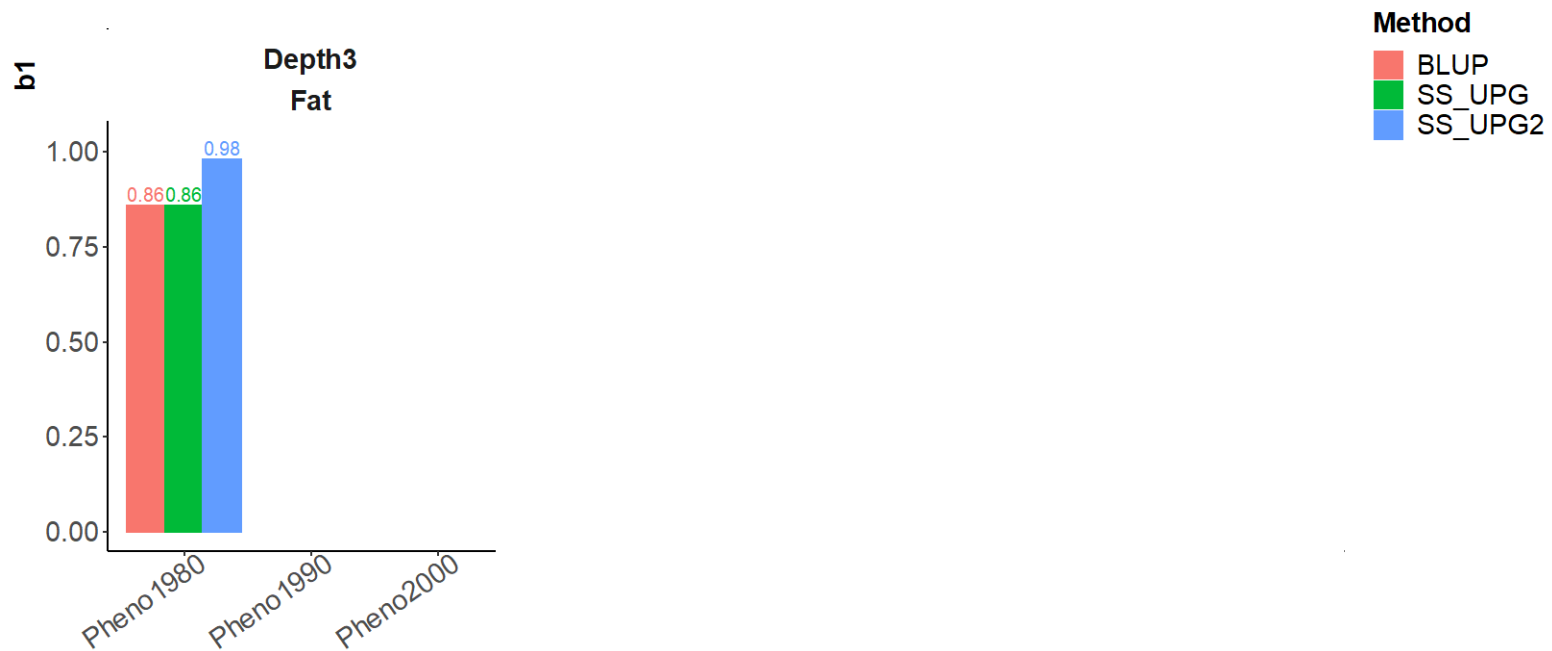 - APY with 15,000 core animals

Validation

- Complete data: 2018
- Reduced data: 2014

- 2,710 bulls
  - Reliability: $[\mathrm{CORR(DYD,(G)EBV)}]^2$
  - Dispersion: $\mathrm{DYD} = b_0 + b_1(\mathrm{G)EBV}$

- 381,779 Cows
  - Predictive ability: $\mathrm{CORR}(y_{adj}, \mathrm{(G)EBV})$
  - Dispersion: $y_{adj} = b_0 + b_1(\mathrm{G)EBV}$

Reliability for bulls

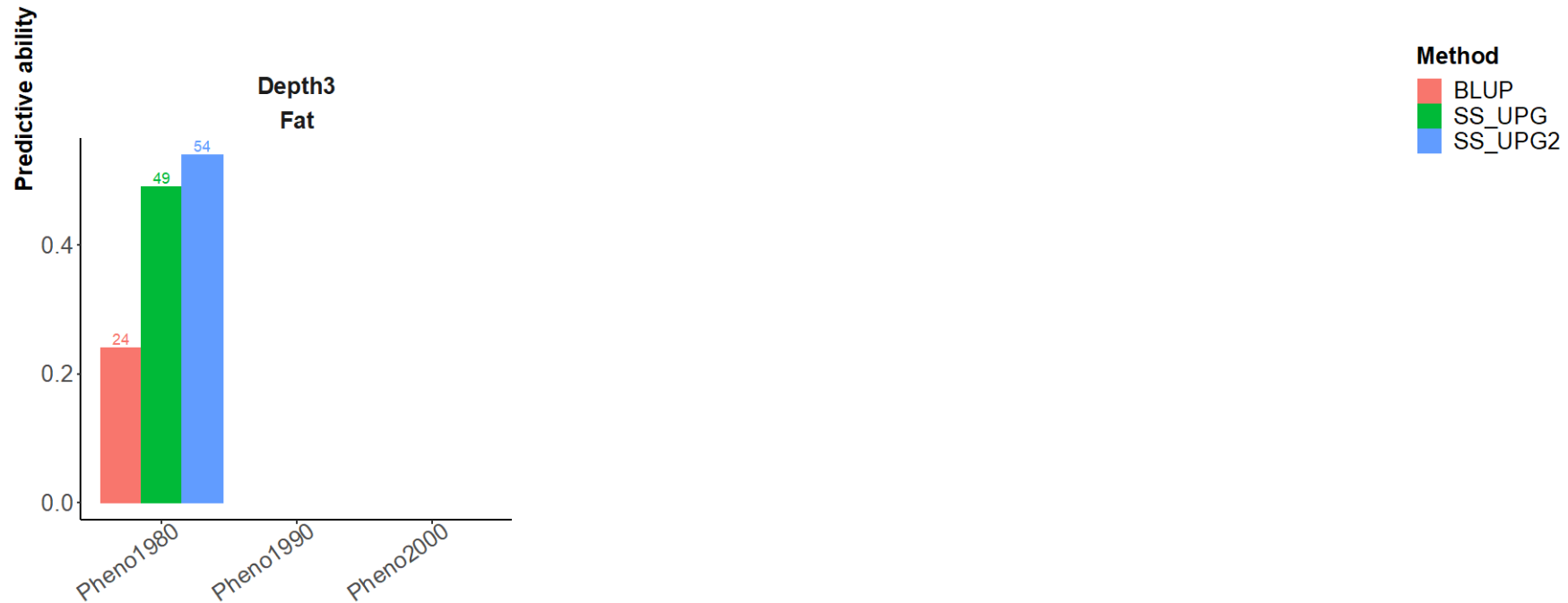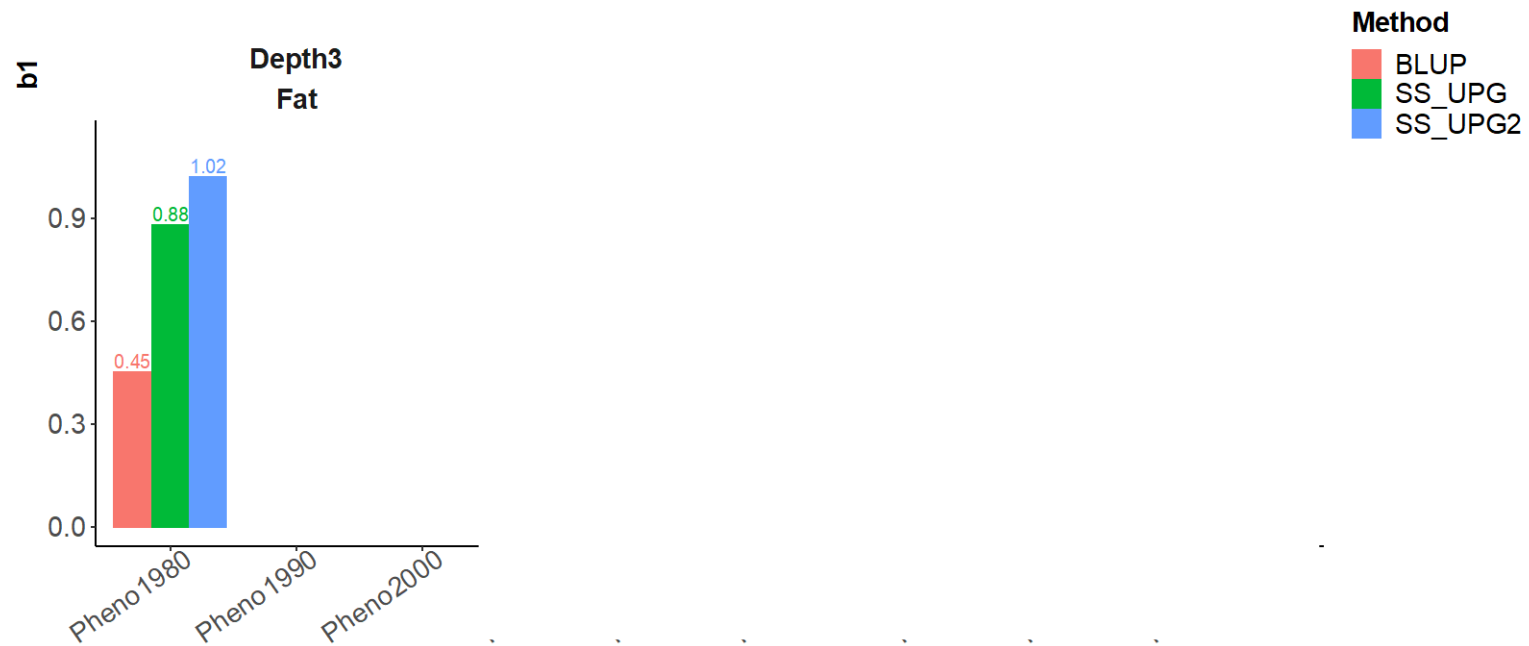# b1 for bulls

Predictive ability for cows

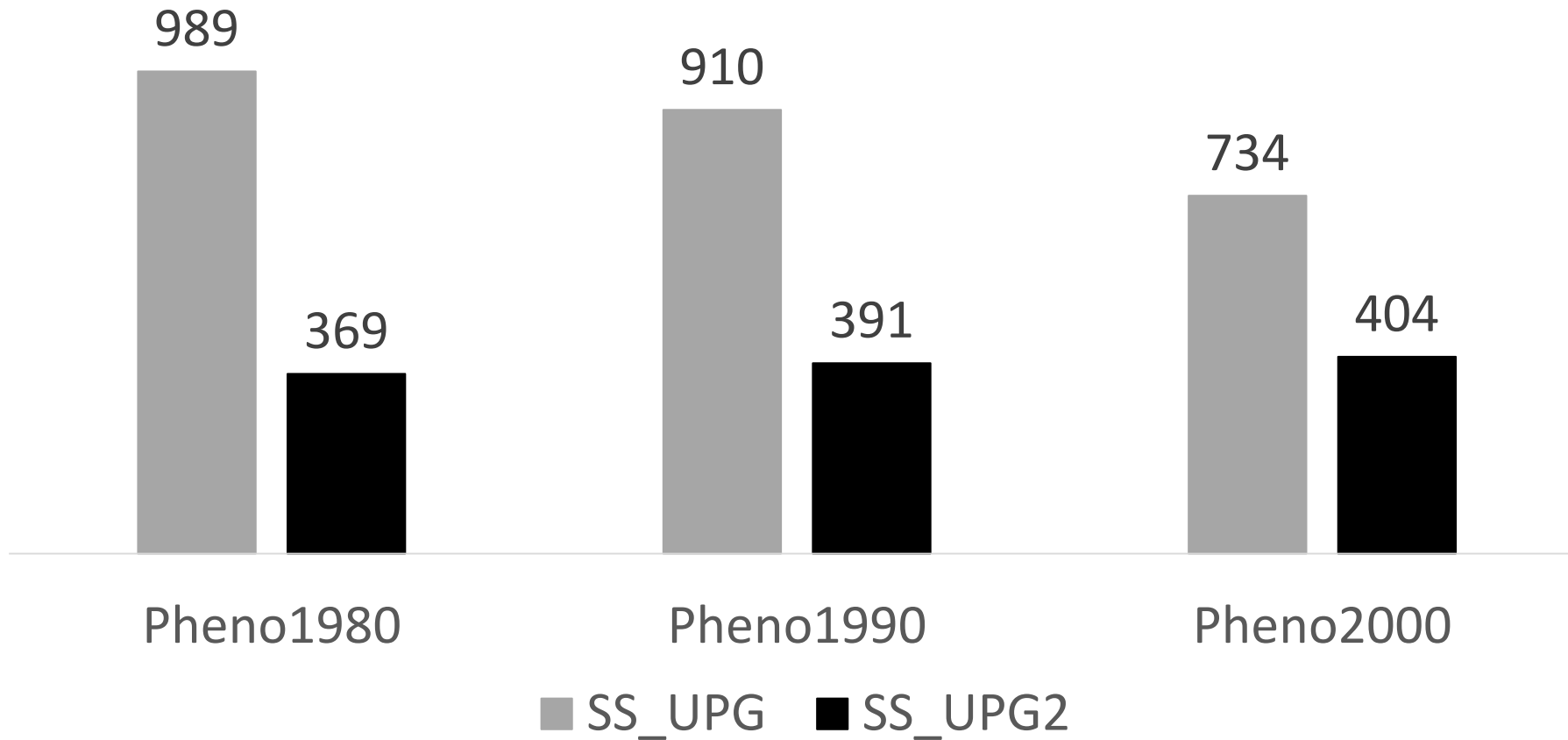# b1 for cows

Rounds to convergence

Pedigree Depth 3

	Pheno1980	Pheno1990	Pheno2000
SS_UPG	989	910	734
SS_UPG2	369	391	404

ssGBLUP with many more genotypes

	This study	A. Cesarani (Friday)
Genotypes	**862k**	**3.4M + data**
Validation bulls	1,529	1,529
Reliability		
Milk	0.72	0.81
Fat	0.69	0.80
Protein	0.67	0.77

Conclusions

- Genomic predictions for yield traits in US Holsteins using single-step

  - UPG for $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}^{-1}$

  - Reliable and unbiased

- Removing old generations of pedigree and phenotypes do not compromise predictions for young selection candidates

- Large-scale dairy genomic evaluations are feasible

  - Up to 3.4 Million genotyped animals in less than 3 days

  - Indirect predictions could further reduce computing time

Acknowledgements

Genomic predictions for yield traits in US Holsteins with unknown parent groups

A. Cesarani,[1]* Y. Masuda,[1] S. Tsuruta,[1] E. L. Nicolazzi,[2] P. M. VanRaden,[3] D. Lourenco,[1]
and I. Misztal[1]
[1]Department of Animal and Dairy Science, University of Georgia, Athens 30602
[2]Council on Dairy Cattle Breeding, Bowie, MD 20716
[3]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350