

Interim genomic prediction for young genotyped animals for different single-step evaluations

J. Vandenplas, R. Bergsma, J. ten Napel, M. Calus



Single-step approaches

- Simultaneous combination of phenotypic, pedigree and genomic information for genotyped and ungenotyped animals
- Two equivalent approaches
 - ssGBLUP: animal-based approaches
 - Genomic relationship matrix
 - ssSNPBLUP: SNP-based approaches
 - SNP effects
- Common issues
 - Solving single-step evaluations still cumbersome
 - Genotypes of young animals acquired after routine evaluations

Aim

Develop and test an approach for
interim genomic prediction
for genotyped but non-phenotyped young animals
using estimates obtained from
two single-step evaluations

Single-step approaches - ssGTBLUP

- Woodbury matrix identity applied to the inverse of the genomic relationship matrix
- SNP effects $\hat{\alpha}$ obtained after solving MME as (Liu et al., 2014):

$$\hat{\alpha} = \left(\frac{1}{w} \mathbf{M}' \mathbf{A}_{gg}^{-1} \mathbf{M} + \mathbf{I} \right)^{-1} \frac{1}{w} \mathbf{M}' \mathbf{A}_{gg}^{-1} \hat{\mathbf{u}}_g$$

$\hat{\mathbf{u}}_g$: genomic breeding values of genotyped animals

w : proportion of genetic variance explained by genomic information

\mathbf{A}_{gg}^{-1} : inverse of the pedigree relationship matrix among genotyped animals

\mathbf{M} : centred genotype matrix

Single-step approaches - ssSNPBLUP

- Liu et al. (2014)
- Equivalent to ssGTBLUP
- SNP effects $\hat{\alpha}$
 - Obtained directly from MME

GEBVs for genotyped individuals - assumptions

$$\hat{\mathbf{u}} = \mathbf{J}\hat{\mu} + \mathbf{M}\hat{\alpha} + \hat{\mathbf{a}}$$

$\hat{\mathbf{u}}$: aggregate genomic breeding values (GEBV)

$\hat{\mu}$: mean difference genotyped & non-genotyped animals

$\hat{\alpha}$: estimated SNP effects

$\hat{\mathbf{a}}$: estimated residual polygenic effects

\mathbf{J} : covariate vector (J-factor)

- -1 for genotyped animals, or imputed value for ungenotyped animals

\mathbf{M} : centred genotype matrix

Interim genomic prediction

- Based on a decomposition of \hat{u} for ssSNPBLUP and ssGTBLUP

Groups of animals			Estimates available after solving MME?			
Animals	Genotyped	Genotyped parents?	\hat{u}	$J\hat{\mu}$	$M\hat{\alpha}$	\hat{a}
Routine	Yes		Yes	Yes	Yes	Yes ($= \hat{u}_g - J_g\hat{\mu} - M\hat{\alpha}$)
	No		Yes	Yes	X	X

Interim genomic prediction

- Based on a decomposition of \hat{u} for ssSNPBLUP and ssGTBLUP

Groups of animals			Estimates available after solving MME?			
Animals	Genotyped	Genotyped parents?	\hat{u}	$J\hat{\mu}$	$M\hat{\alpha}$	\hat{a}
Routine	Yes		Yes	Yes	Yes	Yes ($= \hat{u}_g - J_g\hat{\mu} - M\hat{\alpha}$)
	No		Yes	Yes	X	X
Young	Yes	2	Yes	Yes	Yes	Yes (parent averages)
		1	X	Yes	Yes	Partly
		0	X	Yes	Yes	X

→ **Only unknown:** \hat{a} for ungenotyped animals

Approximation of $\hat{\mathbf{a}}$ for ungenotyped animals

- Theory: for genotyped animals in the routine evaluation

$$\hat{\mathbf{a}}_{old,g} = (1 - w)\mathbf{A}_{gg}\mathbf{G}^{-1}\hat{\mathbf{u}}_{old,g}$$

1. Estimate linear relationship of $\hat{\mathbf{a}}_{old,g}$ and $\hat{\mathbf{u}}_{old,g}$ for all genotyped animals in the routine evaluation

$$\hat{\mathbf{a}}_{old,g} = b_0 + b_1\hat{\mathbf{u}}_{old,g}$$

2. Use this to approximate $\hat{\mathbf{a}}_{old,n}$ for ungenotyped animals in the routine evaluation

$$\hat{\hat{\mathbf{a}}}_{old,n} = \hat{b}_0 + \hat{b}_1\hat{\mathbf{u}}_{old,n}$$

Interim genomic prediction

- Based on a decomposition of \hat{u} for **ssSNPBLUP** and **ssGTBLUP**

Groups of animals			Estimates available after solving MME?			
Animals	Genotyped	Genotyped parents?	\hat{u}	$J\hat{\mu}$	$M\hat{\alpha}$	\hat{a}
Routine	Yes		Yes	Yes	Yes	Yes ($= \hat{u}_g - J_g\hat{\mu} - M\hat{\alpha}$)
	No		Yes	Yes	X	Approximates
Young	Yes	2	Yes	Yes	Yes	Yes (parent averages)
		1	Approximates	Yes	Yes	Approximates
		0	Approximates	Yes	Yes	Approximates

Data

- Routine single-step evaluation
 - #phenotypes: 9.6 million
 - #animals in the pedigree: 18 million
 - #genotypes: 343 thousand

- Interim genomic prediction
 - 20,000 young genotyped animals
 - Both parents genotyped: 13,612
 - Only one parent genotyped: 5118
 - No parents genotyped: 1270

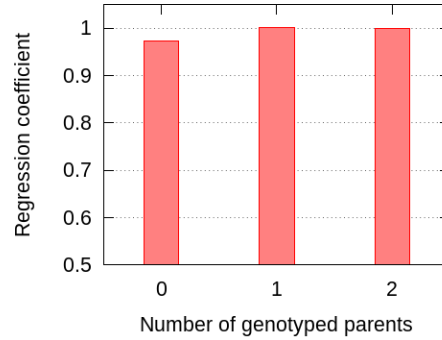
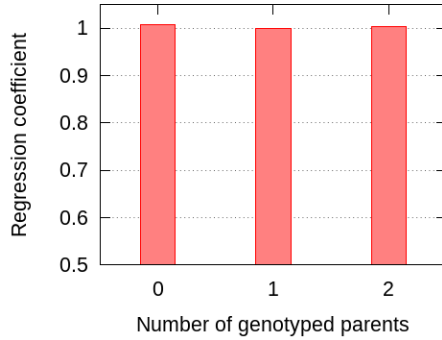
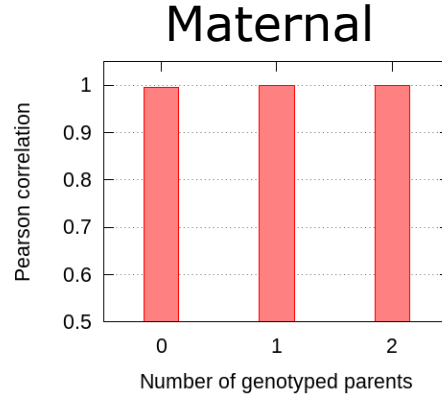
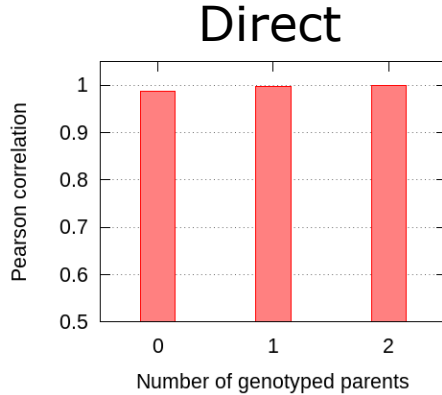
ssSNPBLUP and ssGTBLUP

- Direct and maternal genetic effects
- Residual polygenic effects: 10%
- Solved with MiXBLUP

- Reference
 - Full single-step evaluations with all available information

- (Virtually) same solutions for direct/maternal GEBVs and SNP effects obtained from ssSNPBLUP and ssGTBLUP
 - ➔ Results only shown for ssSNPBLUP

Reference vs interim GEBVs



- Correlations between reference and interim GEBV
 - All > 0.99
 - Regression coefficients of reference on interim GEBV
 - All close to 1.00
- ➔ Accurate interim GEBVs
- ➔ Interim GEBVs less accurate with increased number of ungenotyped parents

Conclusions

- Interim genomic prediction
 - Based on the decomposition of GEBVs in their different components
 - Can be easily extended to other effects (e.g., genetic groups)
 - Applicable to both ssSNPBLUP and ssG(T)BLUP
- Accurate interim GEBVs for young genotyped animals
 - Depends on the number of ungenotyped parents

Thank you for
your attention!

