Integration of foreign estimates of SNP effects into a single-step evaluation

Jeremie Vandenplas, Renzo Bonifazi

May 21, 2024

Single-step genomic evaluation

Increasingly used as routine evaluation

Output

- GEBVs
- Estimates of SNP effects
- (Approximated) Prediction Error (Co)variances (PEV, PEC)

Future: exchange of estimates of SNP effects and PEC?

- From national single-step genomic evaluations
- From SNP-MACE

How to integrate them into national single-step genomic evaluations?

WAGENINGEN
UNIVERSITY & RESEARCH

Aim

Developing and testing an approach
to integrate estimates of SNP effects and measures of precision
from a foreign genomic evaluation
into a national single-step genomic evaluation

Across-country single-step genomic evaluation - ideal

Joint evaluation

Multiple countries

All datasets (phenotypes, genotypes and pedigree) available at an international level

One trait per country

One combined pedigree-genomic relationship matrix

Example: research in beef cattle [Bonifazi et al., 2022]

Across-country single-step genomic evaluation - reality

Data (phenotypes, genotypes) **not shared** across countries

- At least for dairy cattle

Potentially shared

- Estimates of SNP effects
- (Approximated) measures of precision
  - e.g., PEV, PEC

$\Rightarrow$ To be integrated into a national single-step genomic evaluation

National single-step evaluation - without integration

Assumptions (Pop. A)

- Univariate model
- Residual polygenic effects (e.g., 30%)
- Homogeneous SNP variances

$$\mathbf{g}_A \sim MVN\left(\mathbf{0}, \mathbf{I}\sigma_{g,A}^2\right)$$

- Mixed model equations
  - single-step SNPBLUP [Liu et al., 2014]
  - GEBVs and SNP effects predicted simultaneously

WAGENINGEN
UNIVERSITY & RESEARCH

National single-step evaluation - with integration

Shared

- From Pop. B on the scale of Pop. A
- Estimates of SNP effects $\hat{\mathbf{g}}_{B,A}$
- PEC $\Delta_{B,A}$ or PEV $diag\left(\Delta_{B,A}\right)$

Integration by altering the prior means and (co)variances of $\mathbf{g}_A$:

$$\mathbf{g}_A \sim MVN\left(\mathbf{0}, \mathbf{I}\sigma_{g,A}^2\right) \Rightarrow \mathbf{g}_A \sim MVN\left(\hat{\mathbf{g}}_{B,A}, \Delta_{B,A}\right)$$

National single-step evaluation - with integration

Single-step SNPBLUP without integration

Inverted covariance matrix

$$\mathbf{H}_A^{-1} = \begin{bmatrix} \mathbf{A}^{nn}\sigma_{u,A}^{-2} & \mathbf{A}^{ng}\sigma_{u,A}^{-2} & \mathbf{0} \\ \mathbf{A}^{gn}\sigma_{u,A}^{-2} & \left(\mathbf{A}^{gg} + \left(\frac{1}{w}-1\right)\mathbf{A}_{gg}^{-1}\right)\sigma_{u,A}^{-2} & -\frac{1}{w}\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} \\ \mathbf{0} & -\frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\sigma_{u,A}^{-2} & \frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} + \mathbf{I}\sigma_{g,A}^{-2} \end{bmatrix}$$

Right-hand-side

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_n'\mathbf{y}_n \\ \mathbf{W}_g'\mathbf{y}_g \\ \mathbf{0} \end{bmatrix} + \mathbf{H}_A^{-1}\mathbf{0}$$

Single-step SNPBLUP with integration

Inverted covariance matrix

$$\mathbf{H}_{A,B}^{-1} = \begin{bmatrix} \mathbf{A}^{nn}\sigma_{u,A}^{-2} & \mathbf{A}^{ng}\sigma_{u,A}^{-2} & \mathbf{0} \\ \mathbf{A}^{gn}\sigma_{u,A}^{-2} & \left(\mathbf{A}^{gg} + \left(\frac{1}{w}-1\right)\mathbf{A}_{gg}^{-1}\right)\sigma_{u,A}^{-2} & -\frac{1}{w}\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} \\ \mathbf{0} & -\frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\sigma_{u,A}^{-2} & \frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} + \Delta_{B,A}^{-1} \end{bmatrix}$$

Right-hand-side

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_n'\mathbf{y}_n \\ \mathbf{W}_g'\mathbf{y}_g \\ \mathbf{0} \end{bmatrix} + \mathbf{H}_{A,B}^{-1}\begin{bmatrix} \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1}\mathbf{Z}\hat{\mathbf{g}}_{B,A} \\ \mathbf{Z}\hat{\mathbf{g}}_{B,A} \\ \hat{\mathbf{g}}_{B,A} \end{bmatrix}$$

National single-step evaluation - with integration

Single-step SNPBLUP without integration

Inverted covariance matrix

$$\mathbf{H}_A^{-1} = \begin{bmatrix} \mathbf{A}^{nn}\sigma_{u,A}^{-2} & \mathbf{A}^{ng}\sigma_{u,A}^{-2} & \mathbf{0} \\ \mathbf{A}^{gn}\sigma_{u,A}^{-2} & \left(\mathbf{A}^{gg} + \left(\frac{1}{w}-1\right)\mathbf{A}_{gg}^{-1}\right)\sigma_{u,A}^{-2} & -\frac{1}{w}\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} \\ \mathbf{0} & -\frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\sigma_{u,A}^{-2} & \frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} + \mathbf{I}\sigma_{g,A}^{-2} \end{bmatrix}$$

Right-hand-side

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_n'\mathbf{y}_n \\ \mathbf{W}_g'\mathbf{y}_g \\ \mathbf{0} \end{bmatrix} + \mathbf{H}_A^{-1}\mathbf{0}$$

Single-step SNPBLUP with integration

Inverted covariance matrix

$$\mathbf{H}_{A,B}^{-1} = \begin{bmatrix} \mathbf{A}^{nn}\sigma_{u,A}^{-2} & \mathbf{A}^{ng}\sigma_{u,A}^{-2} & \mathbf{0} \\ \mathbf{A}^{gn}\sigma_{u,A}^{-2} & \left(\mathbf{A}^{gg} + \left(\frac{1}{w}-1\right)\mathbf{A}_{gg}^{-1}\right)\sigma_{u,A}^{-2} & -\frac{1}{w}\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} \\ \mathbf{0} & -\frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\sigma_{u,A}^{-2} & \frac{1}{w}\mathbf{Z}'\mathbf{A}_{gg}^{-1}\mathbf{Z}\sigma_{u,A}^{-2} + \Delta_{B,A}^{-1} \end{bmatrix}$$

Right-hand-side

$$\mathbf{b} = \begin{bmatrix} \mathbf{W}_n'\mathbf{y}_n \\ \mathbf{W}_g'\mathbf{y}_g \\ \mathbf{0} \end{bmatrix} + \mathbf{H}_{A,B}^{-1}\begin{bmatrix} \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1}\mathbf{Z}\hat{\mathbf{g}}_{B,A} \\ \mathbf{Z}\hat{\mathbf{g}}_{B,A} \\ \hat{\mathbf{g}}_{B,A} \end{bmatrix}$$

Approximation of $\Delta_{B,A}^{-1}$

$\tilde{\Delta}_{B,A}$ = Approximated PEC of SNPBLUP [Gao et al., 2023]

- Genotypes of Pop. B
- Deregressed ERC of genotyped animals from Pedigree BLUP of Pop. B

$\tilde{\Delta}_{B,A}^{-1}$ = Approximation of $\Delta_{B,A}^{-1}$	
Inverse of PEV	$\left(diag\left(\tilde{\Delta}_{B,A}\right)\right)^{-1}$
Inverse of chromosome-wise PEC (version 1)	$\left(block-diag\left(\tilde{\Delta}_{B,A}\right)\right)^{-1}$
Chromosome-wise inverse of PEC (version 2)	$block-diag\left(\tilde{\Delta}_{B,A}^{-1}\right)$
Inverse of PEC	$\tilde{\Delta}_{B,A}^{-1}$

WAGENINGEN
UNIVERSITY & RESEARCH

Simulation - 2 dairy cattle populations

10 replicates

Pedigree: 420,000 animals

- 20 generations
- Generations 16-20: exchange of 8 sires per generation

Heritability: 0.30

Genetic correlation: 0.80

Residual polygenic effects: 30%

	Population A	Population B
Phenotypes	60,000	165,000
Genotypes ($\sim$ 45K SNPs)	$\sim$ 7037	$\sim$ 75,071
Selection candidates (Gen.20)	$\sim$ 1749	-

WAGENINGEN UNIVERSITY & RESEARCH   MiXBLUP

Results - Selection candidates

Reference: GEBVs of a joint (Pop. A + B) single-step evaluation
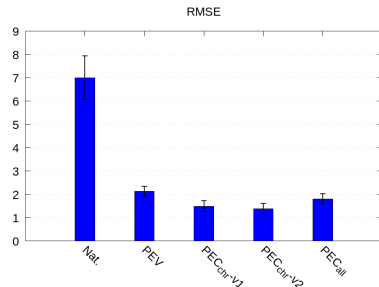Validation: GEBVs of a Pop. A single-step evaluation (+ integration)

Accurate
(even with PEV only)
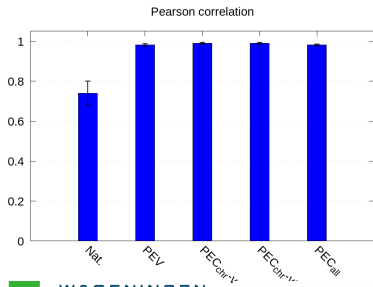
Almost no bias
(except for PEV)
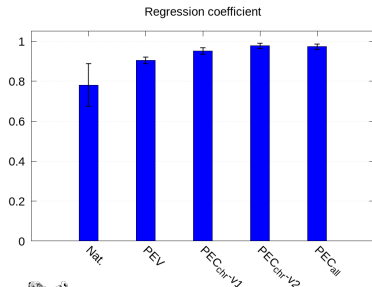
Better fit
(Chromosome-wise PEC)

Conclusions

Integration of estimated SNP effects and PEC into single-step SNPBLUP
- Accurate and (almost) unbiased

At least within-chromosome PEC should be considered

Proposed approach readily extendible
- Multi-trait single-step evaluations
- Other single-step evaluations (e.g., single-step G(T)BLUP)

Acknowledgements

Thank you for your attention!

Questions?

Conclusions

Integration of estimated SNP effects and PEC into single-step SNPBLUP

- Accurate and (almost) unbiased

At least within-chromosome PEC should be considered

Proposed approach readily extendible

- Multi-trait single-step evaluations
- Other single-step evaluations (e.g., single-step G(T)BLUP)

References I

Bonifazi, R., Calus, M. P. L., ten Napel, J., Veerkamp, R. F., Michenet, A., Savoia, S., Cromie, A., and Vandenplas, J. (2022).
International single-step SNPBLUP beef cattle evaluations for Limousin weaning weight.
Genetics Selection Evolution, 54(1):57.

Gao, H., Kudinov, A. A., Taskinen, M., Pitkänen, T. J., Lidauer, M. H., Mäntysaari, E. A., and Strandén, I. (2023).
A computationally efficient method for approximating reliabilities in large-scale single-step genomic prediction.
Genetics Selection Evolution, 55(1):1.

References II

Liu, Z., Goddard, M., Reinhardt, F., and Reents, R. (2014).
A single-step genomic model with direct estimation of marker effects.
Journal of Dairy Science, 97(9):5833--5850.