Comparing methods for approximating reliabilities in large-scale single-step genomic evaluations

Hongding Gao[1], Ismo Strandén[1], Zengting Liu[2]

Jun. 21st, 2025

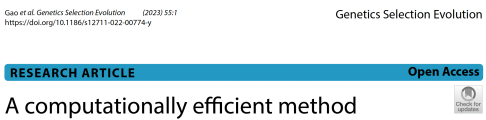
@ Interbull Annual Meeting 2025

[1]Natural Resources Institute Finland (Luke)
[2]IT-Solutions for Animal Production (vit)

Introduction

- Single-step GEBV R^2 are approximated for genotyped and non-genotyped animals separately
- Need conventional EBV R^2 for all the animals in the pedigree
- It requires propagating the genomic information to the non-genotyped animals
- Residual polygenic (RPG) effect needs to be included

To compare two methods



Gao et al. Genetics Selection Evolution (2023) 55:1
https://doi.org/10.1186/s12711-022-00774-y

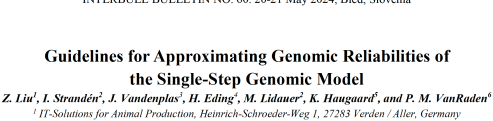
Genetics Selection Evolution

RESEARCH ARTICLE Open Access

A computationally efficient method for approximating reliabilities in large-scale single-step genomic prediction

Hongding Gao, Andrei A. Kudinov, Matti Taskinen, Timo J. Pitkänen, Martin H. Lidauer, Esa A. Mäntysaari and Ismo Strandén

Figure 1: Luke method



INTERBULL BULLETIN NO. 60. 20-21 May 2024, Bled, Slovenia

Guidelines for Approximating Genomic Reliabilities of the Single-Step Genomic Model

Z. Liu[1], I. Strandén[2], J. Vandenplas[3], H. Eding[4], M. Lidauer[2], K. Haugaard[5], and P. M. VanRaden[6]
[1] *IT-Solutions for Animal Production, Heinrich-Schroeder-Weg 1, 27283 Verden / Aller, Germany*
[2] *Resources Institute Finland (Luke), Jokioinen, Finland*
[3] *Wageningen University and Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands*
[4] *AEU CRV u.a., P.O. Box 454, 6800 AL Arnhem, The Netherlands*
[5] *Interbull Centre, Dept. of animal biosciences, Box 7023, SE-750 07, Uppsala, Sweden*
[6] *USDA-ARS Animal Genomics and Improvement Laboratory, Beltsville, MD, USA*
Zengting.liu@vit.de

Figure 2: Interbull method

Luke method

- Step 1: GEBV R^2 for the genotyped animals (via weighted-SNPBLUP)
- Step 2: Calculate genomic ERC
- Step 3: GEBV R^2 for the non-genotyped animals (via weighted-PBLUP)
- In step 1 and 3, ERC are used as weights

Step 1.1: DGV R^2 for the genotyped animals

- Simple SNPBLUP model using ERC as weights (ERC derived from EBV R^2)
- Avoid directly including the RPG effect

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{g} + \mathbf{e}$$

where

$$\mathbf{e} \sim N(\mathbf{0}, \mathbf{D}^{-1}\sigma_e^2)$$

D is a diagonal matrix with elements of ERC from genotyped animals

Step 1.2: blend with the EBV R^2

$$r^2_{GEBV} = \frac{(1-\omega)\mathbf{G}_{ii} r^2_{DGV,i} + \omega \mathbf{A22}_{ii} r^2_{EBV,i}}{(1-\omega)\mathbf{G}_{ii} + \omega \mathbf{A22}_{ii}}$$

where

ω is the proportion of the RPG

$r^{2*}_{DGV,i}$ is the DGV R^2 from SNPBLUP without RPG for animal i

$r^2_{EBV,i}$ is the conventional EBV R^2 of animal i

$\mathbf{G}_{ii}$ is the diagonal element i of the $\mathbf{G}$ matrix

$\mathbf{A22}_{ii}$ is the diagonal element i of the $\mathbf{A}_{22}$ matrix

Step 2: Calculate genomic ERC

- Calculate the ERC accounts for the genomic information

$$ERC_g = ERC_{conv} + \frac{1-h^2}{h^2} \times \left(\frac{r^2_{DGV}}{1-r^2_{DGV}} - \frac{r^2_{EBV}}{1-r^2_{EBV}}\right) \quad (1)$$

where
ERC_{conv} is conventional ERC for genotyped animals
r^2_{EBV} is EBV R^2 for genotyped animals
r^2_{DGV} is the DGV R^2 from step 1 for genotyped animals

Step 3: GEBV R^2 for the non-genotyped animals

- Use a simplified weighted-PBLUP model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{a} + \mathbf{e}$$

where

$$\mathbf{e} \sim N(\mathbf{0}, \mathbf{D}_p^{-1}\sigma_e^2)$$

$\mathbf{D}_p$ is a diagonal matrix with elements of ERC

$$\begin{bmatrix} \mathbf{ERC}_{conv} \\ \mathbf{ERC}_g \end{bmatrix}$$

Interbull method

- Step 1: calculate genomic ERC/EDC gain (φ_c)
- Step 2: propagate genomic information
- Step 3: combine the genomic R^2 gain with the conventional R^2

Step 1: calculate genomic ERC/EDC gain (φ_c)

- Require the Interbull setup
  - A list of validation bulls for Interbull GEBV test (f)
  - GEBV from full and reduced dataset
  - Theoretical R^2 from both full and reduced dataset
  - Expected EDC
  - Adjustment factor (f): ratio of the expected and theoretical EDC

Step 1: calculate genomic ERC/EDC gain (φ_c) (Cont.)

$$\varphi_i^{adj} = \frac{1-h^2}{h^2} \times \left(\frac{r_{DGV}^2}{1-r_{DGV}^2} \times f - \frac{r_{EBV}^2}{1-r_{EBV}^2}\right) \tag{2}$$

$$\varphi_c = \frac{1}{n}\sum_{1}^{n} \varphi_i^{adj} \tag{3}$$

Step 2: propagate genomic information (φ_i^{prog})

- Passing φ_c via pedigree from youngest to oldest animals
- From oldest to youngest via pedigree

Step 3: GEBV R^2

- For genotyped animals

$$\varphi_i^{total} = \varphi_i^{conv} + \varphi_c \tag{4}$$

- For non-genotyped animals

$$\varphi_i^{total} = \varphi_i^{conv} + \varphi_i^{prog} \tag{5}$$

- Final GEBV R^2

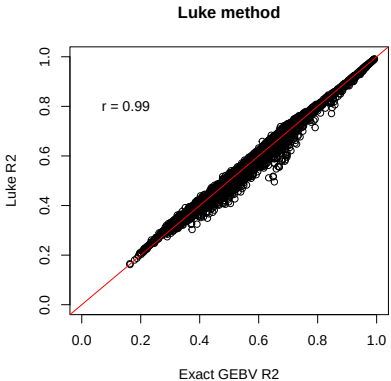
$$R_i^2 = \frac{\varphi_i^{total}}{\varphi_i^{total} + \lambda} \tag{6}$$

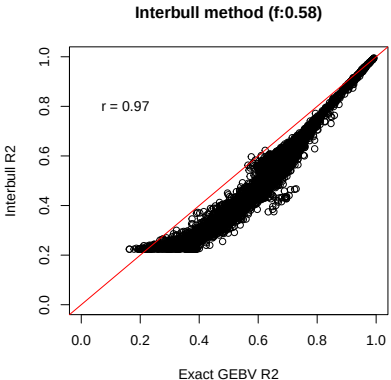
Data

- 47K Finnish Red cows with 305-day milk yield records from lactation one
- 19k genotyped animals
- 50K SNPs
- h^2: 0.44
- RPG: 0.3

Genotyped animals

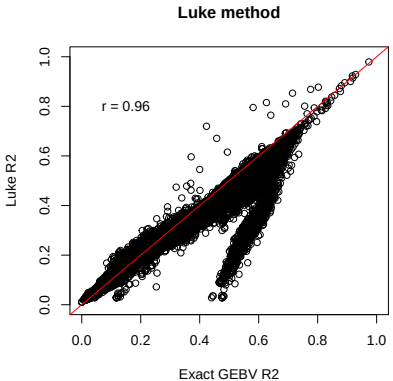
- Mean R^2: 0.66, 0.63, 0.57 for exact, Luke, and Interbull method

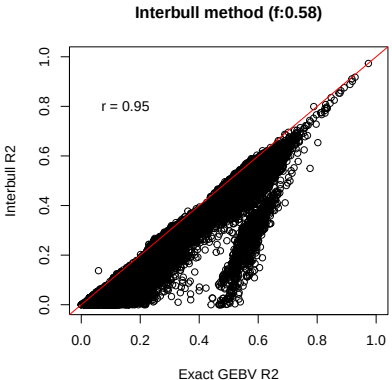




Non-genotyped animals

- Mean R^2: 0.48, 0.44, 0.43 for exact, Luke, and Interbull method





Conclusion

- Both methods provide an effective strategy for obtaining GEBV R^2 from single-step model in practice
- The approximated R^2 were in good agreement with the exact R^2