

A weekly genomic evaluation of newly genotyped selection candidates based on a single-step genomic model

H. Alkhoder and Z. Liu

IT-Solutions for Animal Production (**vit**), Germany

A weekly genomic evaluation with no new phenotypic data added

- The **current multi-step genomic evaluation** based on a SNP BLUP model
 - Difference to a full genomic evaluation: **no new phenotypic data** added, but new genotypes
 - Using SNP effect estimates from a latest genomic evaluation with the SNP BLUP model
 - Calculate DGV for new animals with genotype data
 - Parental average of conventional EBV for young candidates using a BLUP animal model
 - Compute combined GEBV with the selection index of DGV and PA_{conv} : **$GEBV = f(DGV, PA_{conv})$**
 - A **just-in-time continuous** system using DGV and a function of genomic EBV (Alkhoder et al. 2014)
 - Monthly to weekly for German Holstein since April 2019
 - 3,000 to 5,000 new genotyped animals per week
 - The weekly genomic evaluation conducted during weekends
 - Same calculation steps as the main evaluation, without the EBV deregression and SNP effect estimation



A weekly genomic evaluation based a full single-step evaluation (I)

- A **single-step SNP BLUP** model for genomic evaluation with 30% residual polygenic variance
 - No new phenotypic added, only new genotypic data
 - A full single-step evaluation for all traits INFEASIBLE to complete during a weekend
 - **GEBV = DGV + parental average of RPG** for young animals without own phenotypes (Liu et al. 2014)
 - An indirect prediction of GEBV implemented in MiX99 *Predict_GEBV* (Lidauer and Strandén, 1999)
 - **The exact method GRV** (Vandenplas et al. 2023)

- A **single trait model** for the weekly genomic evaluation
 - In contrast to more complex evaluation model in the single-step main evaluation like:
 - A multi-lactation random regression test-day model for milk yield
 - with 3 lactations x 3 coefficients = 9 operational 'traits'
 - SNP/DGV and RPG expressed on 9 random regression coefficients
 - GEBV computed only for the single trait such as:
 - A combined lactation 305-day milk yield as a linear function the 9 coefficients



A weekly genomic evaluation based a full single-step evaluation (II)

■ Calculation of direct genomic values for young selection candidates

- SNP effects from the single-step SNP BLUP model
- Allele frequencies from the main single-step evaluation
- DGV are accurate, **no approximation** involved

■ Residual polygenic effects for young selection candidates

- $\hat{\mathbf{a}}_c = \mathbf{A}_{cg} \mathbf{A}_{gg}^{-1} \mathbf{a}_g$ using RPG effects of all genotyped animals in the latest full evaluation (Liu et al. 2016)
- Equivalent to **deregress RPG** by solving the equations (Vandenplas et al. 2023)

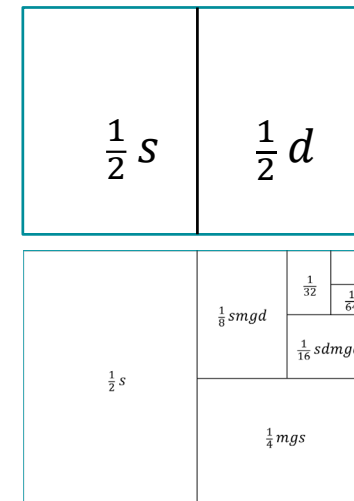
$$\begin{bmatrix} \mathbf{A}^{00} & \mathbf{A}^{0g} & \mathbf{A}^{0c} \\ \mathbf{A}^{g0} & \mathbf{A}^{gg} & \mathbf{A}^{gc} \\ \mathbf{A}^{c0} & \mathbf{A}^{cg} & \mathbf{A}^{cc} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}}_0 \\ \hat{\mathbf{a}}_g \\ \hat{\mathbf{a}}_c \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{A}_{gg}^{-1} \mathbf{a}_g \\ \mathbf{0} \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} \mathbf{A}^{00} & \mathbf{A}^{0g} & \mathbf{A}^{0c} \\ \mathbf{A}^{g0} & \mathbf{A}^{gg} & \mathbf{A}^{gc} \\ \mathbf{A}^{c0} & \mathbf{A}^{cg} & \mathbf{A}^{cc} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{0g} & \mathbf{A}_{0c} \\ \mathbf{A}_{g0} & \mathbf{A}_{gg} & \mathbf{A}_{gc} \\ \mathbf{A}_{c0} & \mathbf{A}_{cg} & \mathbf{A}_{cc} \end{bmatrix}^{-1}$$

- **Non-genotyped ancestors** of the genotyped animals and candidates receive also RPG estimates
 - $\hat{\mathbf{a}}_0 = -(\mathbf{A}^{00})^{-1} \mathbf{A}^{0g} \mathbf{a}_g$ (similar to Fernando et al. 2014 for GEBV)



A weekly genomic evaluation based a full single-step evaluation (III)

- As an alternative, a **summation** method for RPG is used (**vit**)
 - $GEBV = DGV + \mathbf{RPG}$ for new animals with genotypes
 - Case 1: $a_c = \frac{1}{2}(a_s + a_d)$ for both parents genotyped in the latest full evaluation
 - Case 2: $a_c = \frac{1}{2}(a_s + \frac{1}{2}(a_{mgs} + \frac{1}{2}(a_{smgd} + \dots)))$ for only male animals genotyped
 - Process pedigree from the candidate to oldest genotyped animal for the summation
 - Consider only genotyped animals in the latest full evaluation and new candidates
 - Equivalently $\hat{\mathbf{a}}_c \approx \mathbf{A}_{cg} \mathbf{I} \mathbf{a}_g$ versus $\hat{\mathbf{a}}_c = \mathbf{A}_{cg} \mathbf{A}_{gg}^{-1} \mathbf{a}_g$
- Application case for the summation method
 - The current *GEBV Prediction* Program (MiX99) requires that new candidates must NOT have progeny in the latest full single-step evaluation, as their RPG effects may be incorrectly estimated
 - In practice, animals with new genotypic data may be older ones through country exchange programmes
 - Genotype data October 2023: 1,318,720 in full run 2304, 110,015 young candidates + **5135 older animals**



Data materials for a comparison of the GEBV prediction methods

- A full single-step evaluation with data from April 2023
 - 1,318,720 genotyped Holstein animals divided in two groups
 - Born from 2005 to March 2023: 1,318,720
 - **Old 'reference animals'**: born till March 2022: 1,169,502
 - **Young 'genotyped candidates'**: born from April 2022: 149,218
 - Pedigree for the genotyped animals: 3,427,852 (incl. 2,109,132 non-genotyped ancestors)
 - 2,024,081 ancestors of reference animals
 - 85,051 ancestors of genotyped candidates

- For 69 traits or indices of all trait groups
 - The exact method GRV and the approximate summation method
 - Compared to the candidates GEBV from the full evaluation

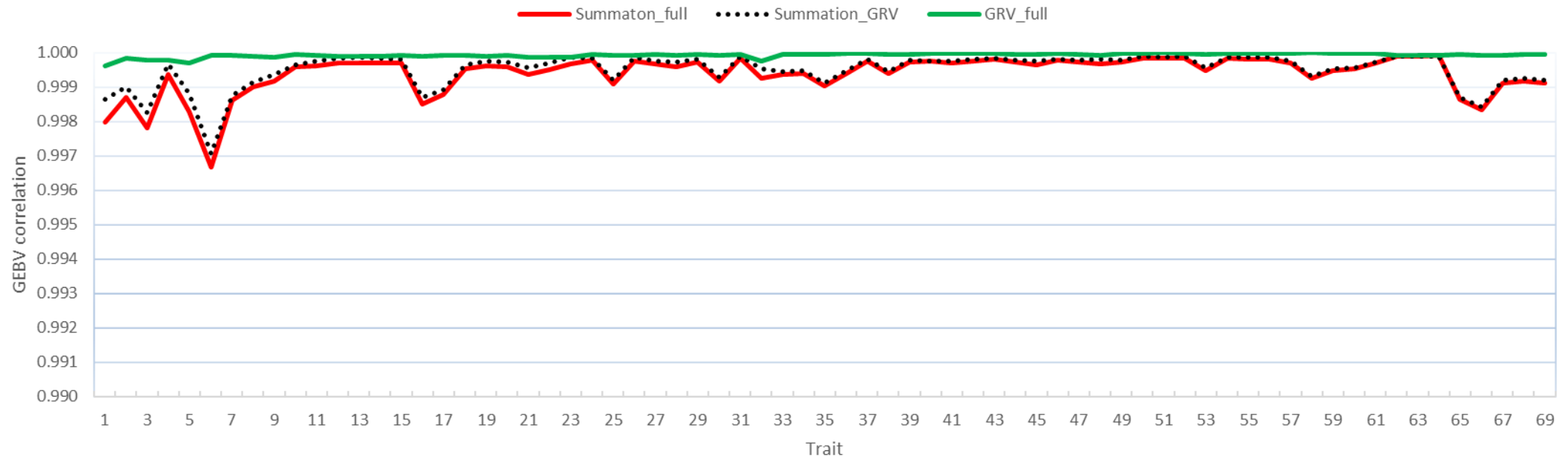


Results: the exact method GRV for the weekly evaluation

- SNP effect estimates of 69 traits/indices from obtained the full single-step evaluation Apr 2023
- CPU and RAM usage using MiX99 *GEBV_Prediction*
 - 15.5 GB RAM
 - 38 minutes on 46 cores
- **Reference animals** GEBV correlations with DGV and RPG components for all 69 traits
 - Defined as genotyped animals in the latest main single-step evaluation
 - $\text{Corr}(\text{GEBV}, \text{DGV})$: average 0.969, minimum 0.935, maximum 0.986, $b_1(\text{GEBV}|\text{DGV}) = [0.990 \ 1.138]$
 - $\text{Corr}(\text{GEBV}, \text{RPG})$: average 0.406, minimum 0.248, maximum 0.650, $b_1(\text{GEBV}|\text{RPG}) = [0.886 \ 3.169]$
- **Candidate animals** GEBV correlations with DGV and RPG components for all 69 traits
 - Defined as newly genotyped animals that were not included in the latest single-step evaluation
 - $\text{Corr}(\text{GEBV}, \text{DGV})$: average 0.976, minimum 0.944, maximum 0.991, $b_1(\text{GEBV}|\text{DGV}) = [0.974 \ 1.118]$
 - $\text{Corr}(\text{GEBV}, \text{RPG})$: average 0.304, minimum 0.064, maximum 0.589, $b_1(\text{GEBV}|\text{RPG}) = [0.324 \ 3.102]$



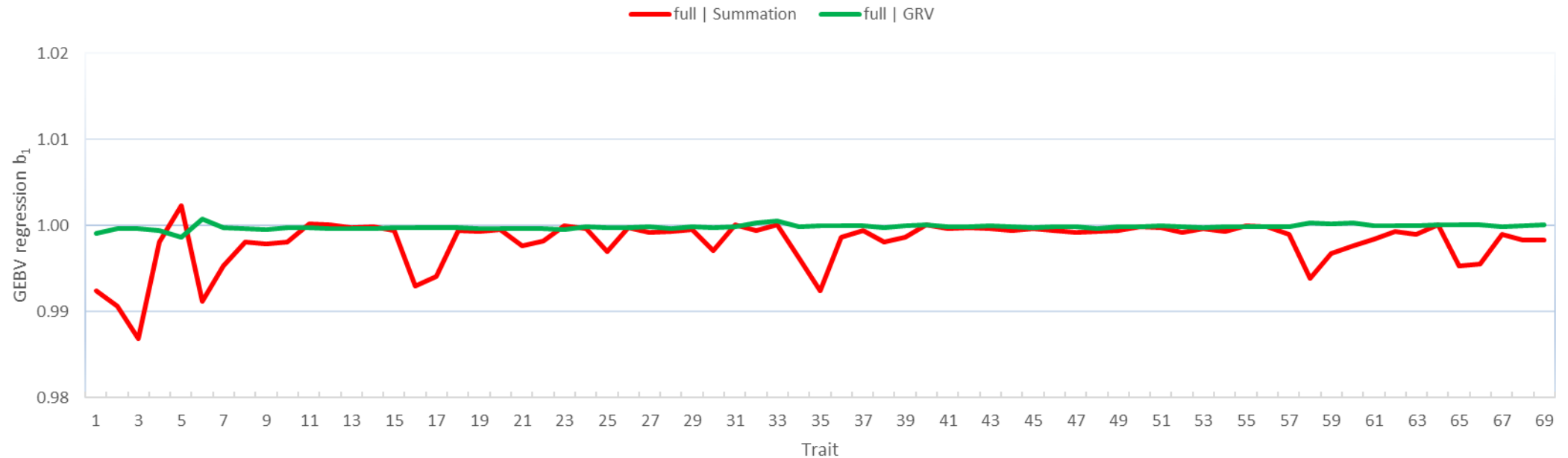
Results: GEBV correlations of the new candidates with their GEBV from the latest single-step evaluation



Correlation of single-step GEBV weekly with single-step GEBV full evaluation



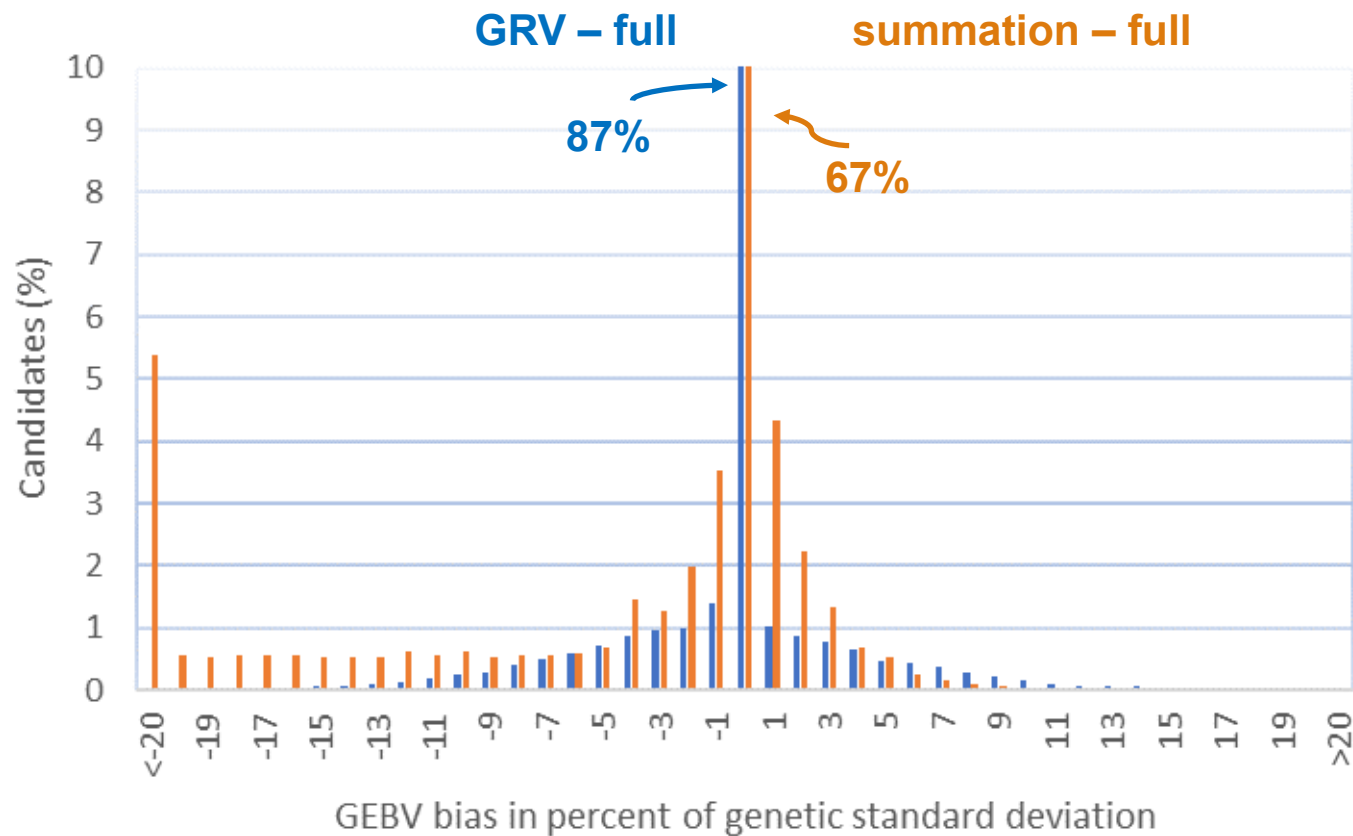
Results: Regression of GEBV of new candidates from the latest single-step evaluation on their weekly evaluation



Regression slope of single-step GEBV full evaluation on single-step GEBV weekly



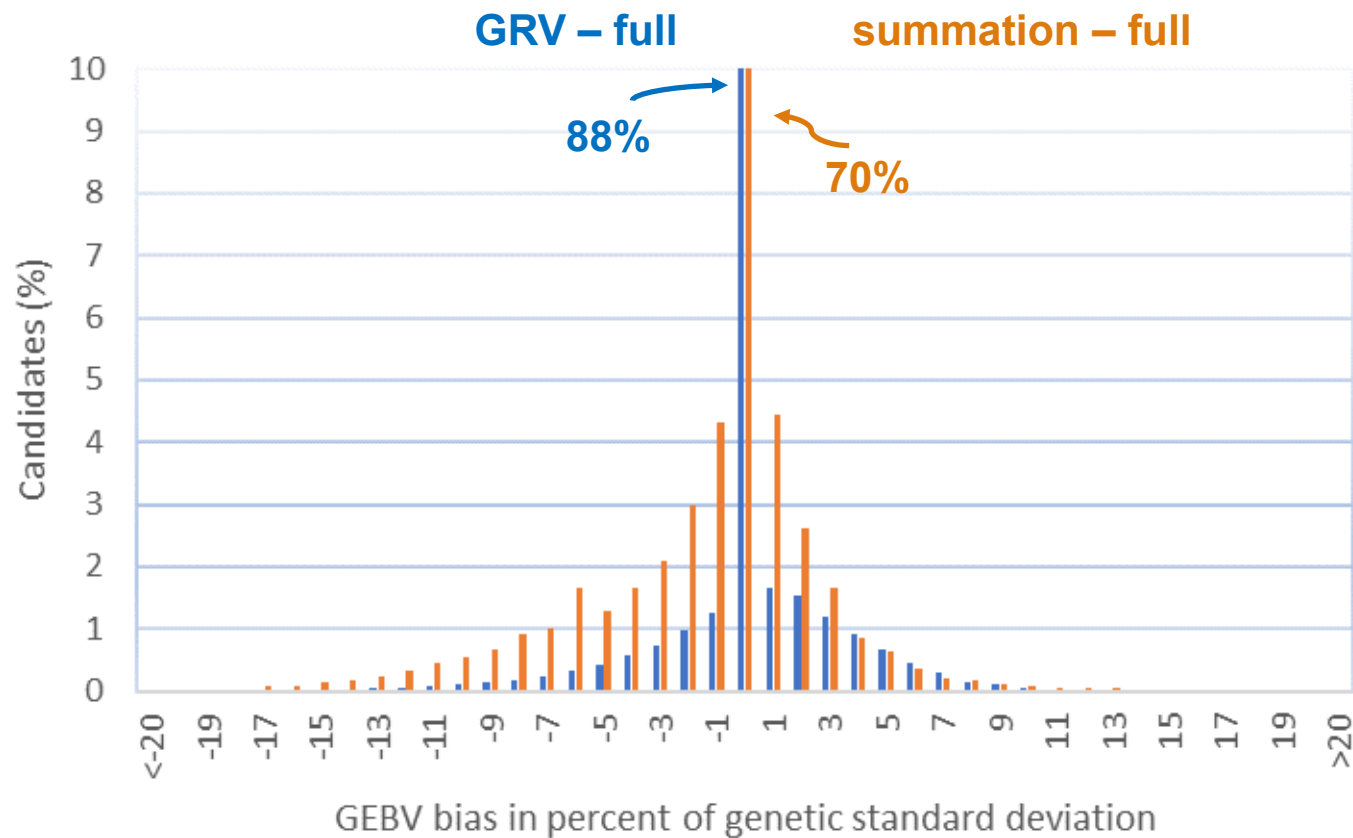
Distribution of GEBV bias for the candidates: Trait #3



Protein yield



Distribution of GEBV bias for the candidates: Trait #4



Summary and Conclusions (I)

- Genomic prediction for newly genotyped candidates at a weekly basis
 - Essential for routine genomic selection
 - Different methods for calculating GEBV:
 - Multiple-step model: selection index of DGV and conventional parental average
 - Single-step model: sum of DGV and parental average of residual polygenic effects

- Calculating GEBV components: direct genomic breeding values DGV
 - SNP effect estimates from the latest main evaluation under the ssSNPBLUP model
 - Alternatively, SNP effect estimates back-solved from single-step GEBV of genotyped animals
 - Single trait model DGV values are accurate for complex models like the random regression test-day model
 - Little need for DGV of 9 coefficients (3 lactations x 3 RRCs) for test-day milk yield



Summary and Conclusions (II)

- Two methods for computing residual polygenic effects of newly genotyped animals
 - Single trait RPG values instead of original complex multi-trait models
 - The exact GRV method proven to be nearly unbiased
 - Current limitation: new candidates must not have genotyped progeny in latest single-step evaluation
 - The summation method also accurate but slightly biased downward with incomplete genotyped ancestry in latest single-step evaluation
 - Both methods computationally efficient and feasible for a genomic evaluation at weekend
 - Use the exact GRV method whenever possible and apply the summation method for rejected animals
- Genomic reliability approximation following Interbull GREL Guidelines
 - Identical calculation steps as in main evaluation, except
 - Conventional reliabilities for candidates calculated starting with EDC and ERC instead of raw data



Thank you for your attention!



IT-Solutions for Animal Production