

Computing strategies for national beef cattle evaluations

M. Bermann, D. Lourenco, I. Misztal

06/28/2022

2022 ASAS National Meeting



UNIVERSITY OF
GEORGIA

College of Agricultural &
Environmental Sciences

*Animal Breeding and
Genetics Group*

Beef cattle datasets in the US

- > 12 millions animals
- > 30 million records for growth traits
- ~ 1.2 million genotypes
- 22 traits with different models
- One week from raw data to results

Genetic evaluations and genetic progress

- New information every week
- Increased accuracy for selection candidates
- Selection decisions at earlier age
- Higher genetic progress

$$\Delta G = \frac{i h \sigma_A}{L}$$

Single-step evaluations

- Combines phenotypic, pedigree, and genomic information
- Less biased than multi-step methods
- Flexible for various models:
 - Repeatability
 - Multiple-trait
 - Maternal
 - Threshold
 - Random regression
- Software availability

Single-step genetic evaluations

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \sigma_g^{-2}\mathbf{H}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{pmatrix}$$

Pedigree

Genomics

Single-step with APY

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{pmatrix} \quad \longrightarrow \quad \mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{\text{APY}}^{-1} - \mathbf{A}_{22}^{-1} \end{pmatrix}$$



\mathbf{G}^{-1}



$\mathbf{G}_{\text{APY}}^{-1}$

Inexpensive Computation of the Inverse of the Genomic Relationship Matrix in Populations with Small Effective Population Size

Ignacy Misztal¹

Animal and Dairy Science, University of Georgia, Athens, Georgia 30602

ORCID ID: 0000-0002-0382-1897 (I.M.)

From raw data to EPDs

- Read genotypes
- Quality control
- Setup single-step matrices



Genomic setup

- Read data and pedigree
- Solve the MME
- Approximate accuracies

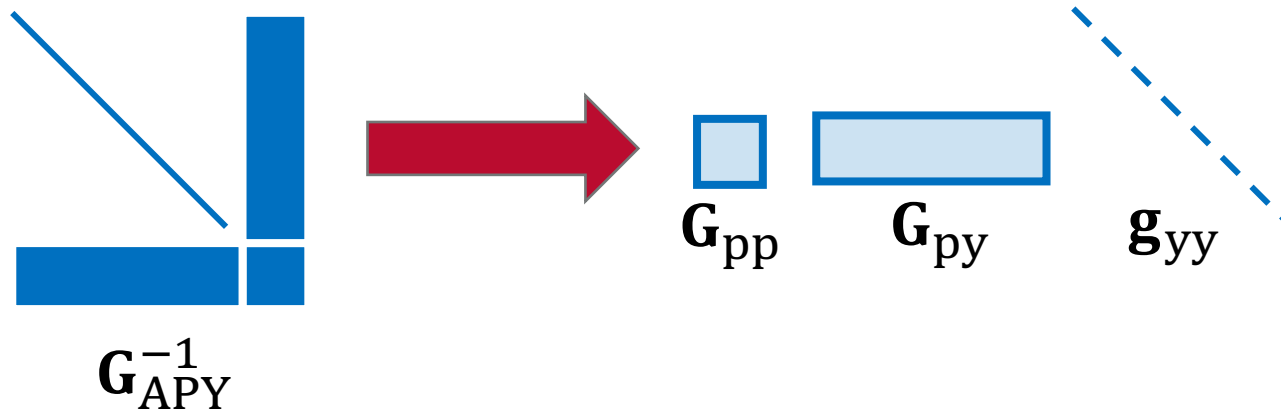


Breeding value estimation

Challenges with more genotyped animals

1. Set up APY blocks
2. Blending the GRM and compatibility with pedigree information
3. Calculating accuracy

Genomic setup: APY blocks



- Dense matrices
- Optimized libraries
- Parallel computing

Intel[®] oneAPI Math Kernel Library

OpenMP[®]

Genomic setup: blending

- Make the GRM invertible
- Add a residual polygenic effect
- Main bottleneck

Implementation of genomic recursions in single-step genomic best linear unbiased predictor for US Holsteins with a large number of genotyped animals

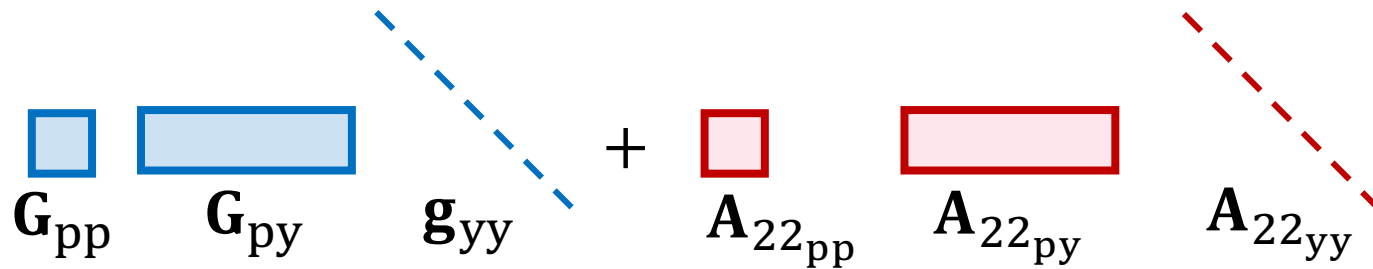
Y. Masuda,^{*1} I. Misztal,^{*} S. Tsuruta,^{*} A. Legarra,[†] I. Aguilar,[‡] D. A. L. Lourenco,^{*} B. O. Fragomeni,^{*} and T. J. Lawlor[§]

^{*}Department of Animal and Dairy Science, University of Georgia, Athens 30602

[†]Institut National de la Recherche Agronomique, UMR1388 GenPhySE, 31326 Castanet Tolosan, France

[‡]Instituto Nacional de Investigación Agropecuaria, Canelones, Uruguay 90200

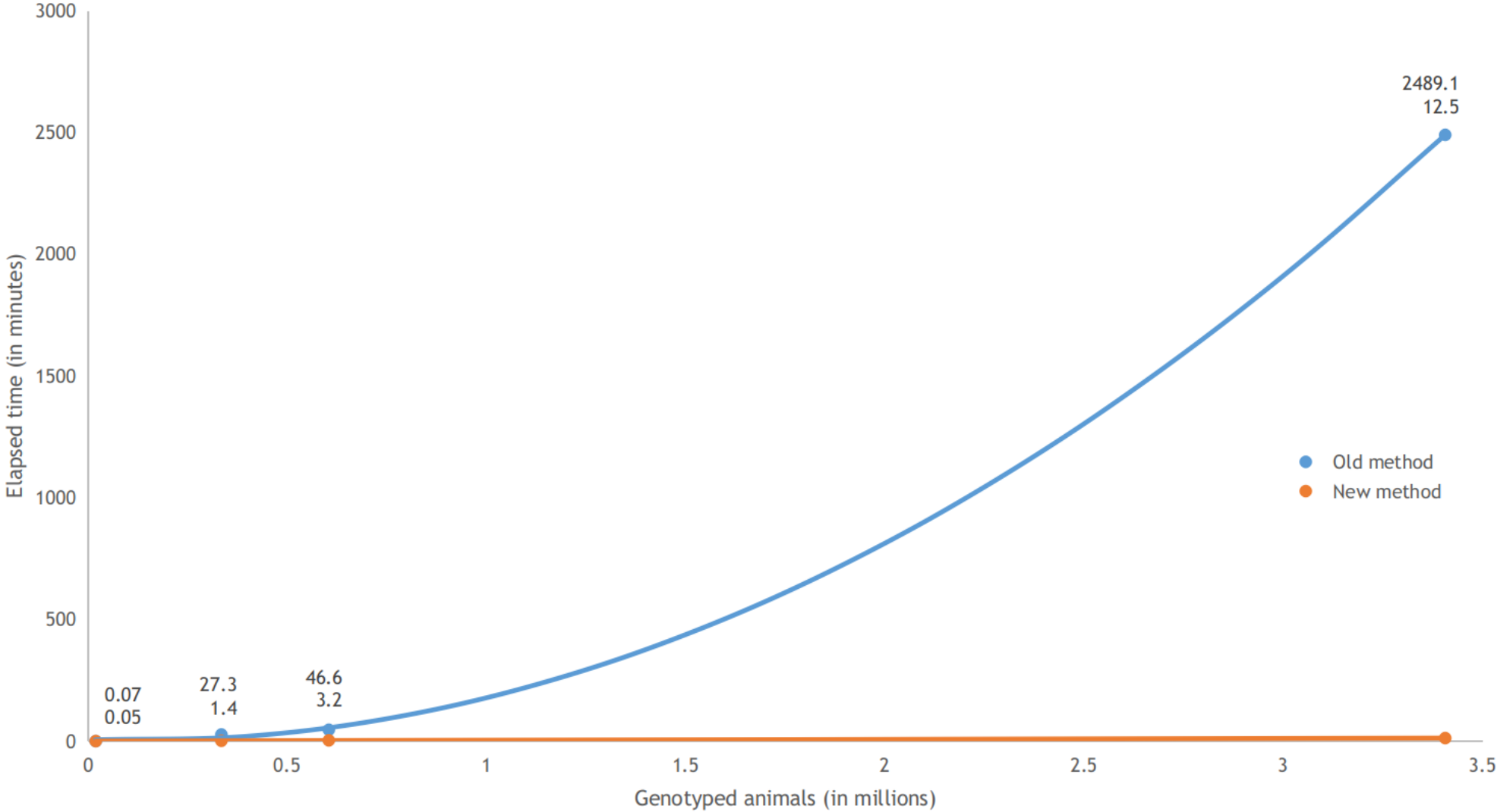
[§]Holstein Association USA Inc., Brattleboro, VT 05301



Impact of blending the genomic relationship matrix with different levels of pedigree relationships or the identity matrix on genetic evaluations

Journal:	<i>JDS Communications</i>
Manuscript ID	JDSC.2022-0229.R2
Article Type:	Short Communication
Date Submitted by the Author:	06-Jun-2022
Complete List of Authors:	Hollifield, Mary; University of Georgia, Animal and Dairy Science Bermann, Matias; University of Georgia, Department of Animal and Dairy Science Lourenco, Daniela; University of Georgia, Animal and Dairy Science Misztal, Ignacy; University of Georgia
Key Words:	large-scale genetic evaluations, prediction accuracy, residual polygenic effect, ssGBLUP

Elapsed time in minutes for old and new method



BV estimation: compatibility between G_{APY}^{-1} and A_{22}^{-1}

- Automatic scaling
- No pre-processing
- One more equation in the MME

$$\mathbf{H}_{\text{auto}}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & 0 \\ \mathbf{A}^{21} & \mathbf{A}^{22} + \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} & -\mathbf{G}^{-1}\mathbf{1} \\ 0' & -\mathbf{1}'\mathbf{G}^{-1} & \mathbf{1}'\mathbf{G}^{-1}\mathbf{1} \end{bmatrix}$$



J. Dairy Sci. 104:2027–2031

<https://doi.org/10.3168/jds.2020-18969>

© 2021 American Dairy Science Association®. Published by Elsevier Inc. and FASS Inc. All rights reserved.

Technical note: Automatic scaling in single-step genomic BLUP

M. Bermann,*  D. Lourenco,  and I. Misztal 

¹Department of Animal and Dairy Science, University of Georgia, Athens, 30602

BV estimation: accuracies approximation

- Combine different sources of information
 - Records
 - Pedigree
 - Genomics

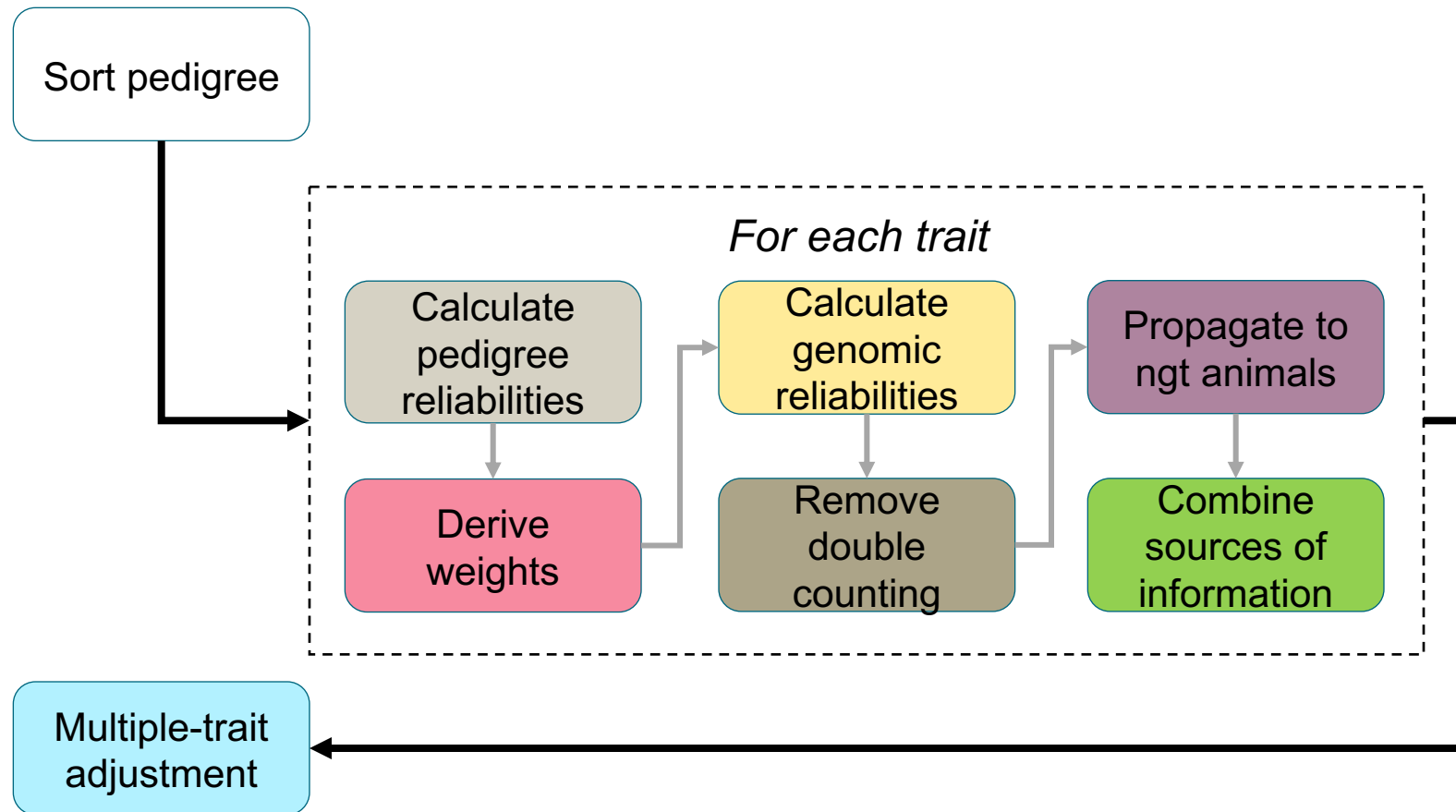
Efficient approximation of reliabilities for single-step genomic best linear unbiased predictor models with the Algorithm for Proven and Young

Matias Bermann,¹  Daniela Lourenco, and Ignacy Misztal 

Department of Animal and Dairy Science, University of Georgia, Athens, GA, USA

¹Corresponding author: mbermann@uga.edu

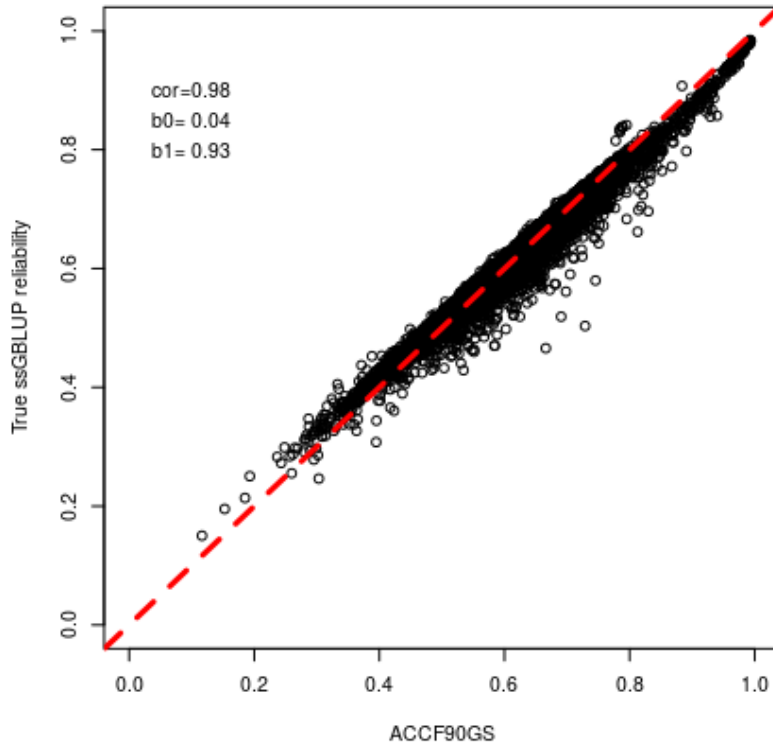
ACCF90GS2 flowchart



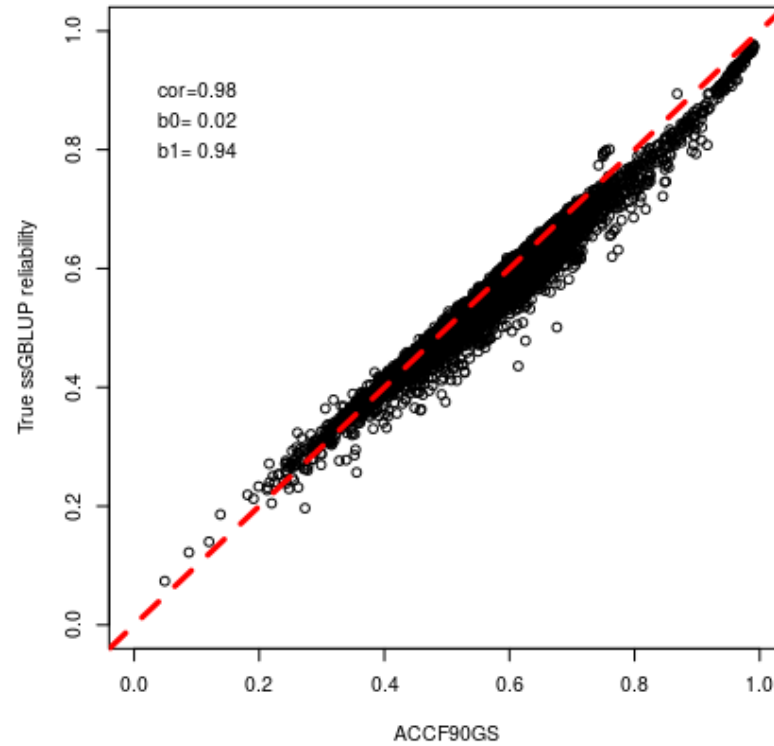
Accuracies approximation

Multiple trait model

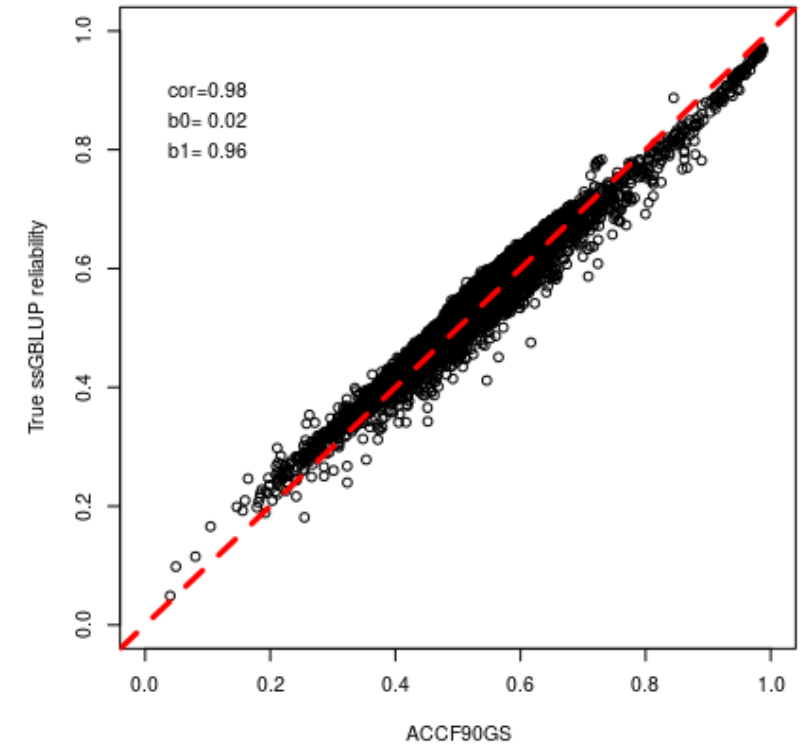
BW



WW



PWG



Future challenges: more evaluations and information

- Convergence
- New information
- Complex models

Challenges: convergence

- Seconds per round
- Number of rounds:
 - Preconditioning
 - Removing old data
 - Genetic groups

ANIMAL GENETICS AND GENOMICS

Changes in genomic predictions when new information is added

Jorge Hidalgo,^{†,1} Daniela Lourenco,[†] Shogo Tsuruta,[†] Yutaka Masuda,[†] Stephen Miller,[‡] Matias Bermann,[†] Andre L. S. Garcia,[†] and Ignacy Misztal[†]

[†]Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, [‡]Angus Genetics Inc., St. Joseph, MO 64506

ANIMAL GENETICS AND GENOMICS

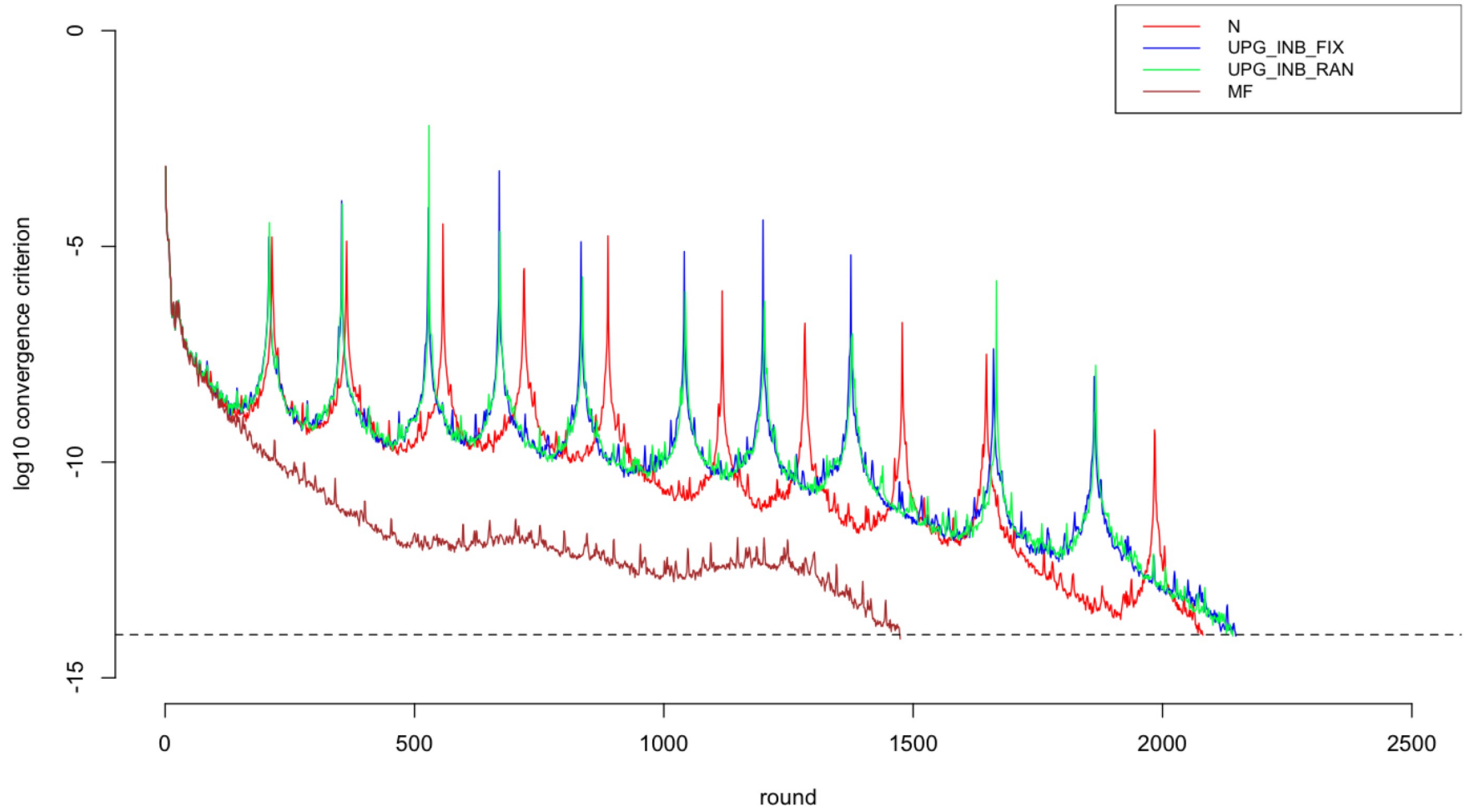
Modeling genetic differences of combined broiler chicken populations in single-step GBLUP

Matias Bermann,^{†,1} Daniela Lourenco,[†] Vivian Breen,[‡] Rachel Hawken,[‡] Fernando Brito Lopes,[‡] and Ignacy Misztal[†]

[†]Department of Animal and Dairy Science, University of Georgia, Athens, GA, USA, [‡]Cobb-Vantress Inc., Siloam Springs, AR 72761, USA





MFIX



Challenges: new information

- Omics
- Images
- Automatic feeders
- Are they useful?
- How to incorporate them?

Genetic evaluation including intermediate omics features

Ole F. Christensen ^{1,*} Vinzent Börner,¹ Luis Varona,² and Andres Legarra ³

The role of sensors, big data and machine learning in modern animal farming

Suresh Neethirajan

Estimation of direct and social effects of feeding duration in growing pigs using records from automatic feeding stations

Belcy K. Angarita,^{†,‡} Junjie Han,[‡] Rodolfo J. C. Cantet,[†] Sarah K. Chewning,[‡] Kaitlin E. Wurtz,[‡] Janice M. Siegford,[‡] Catherine W. Ernst,[‡] and Juan Pedro Steibel^{‡,||,1}

Challenges: complex models

- Threshold models
- Multi/cross breed

ORIGINAL ARTICLE

Journal of
Animal Breeding and Genetics | WILEY

Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: An application in chicken mortality

Matias Bermann¹  | Andres Legarra² | Mary Kate Hollifield¹ | Yutaka Masuda¹ | Daniela Lourenco¹  | Ignacy Misztal¹ 

ANIMAL GENETICS AND GENOMICS

Modeling genetic differences of combined broiler chicken populations in single-step GBLUP

Matias Bermann,^{†,1} Daniela Lourenco,[†] Vivian Breen,[‡] Rachel Hawken,[‡] Fernando Brito Lopes,[‡] and Ignacy Misztal[†]

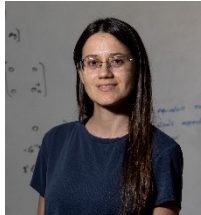
[†]Department of Animal and Dairy Science, University of Georgia, Athens, GA, USA, [‡]Cobb-Vantress Inc., Siloam Springs, AR 72761, USA



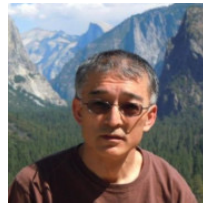
Conclusions

- Different computing strategies allow weekly single-step evaluations
- There is room for improving computational efficiency
- Anticipate next challenges and bottlenecks

The UGA team



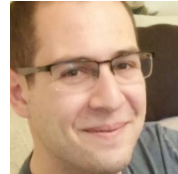
Daniela
Lourenco



Shogo
Tsuruta



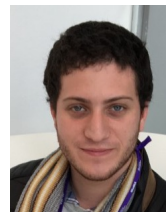
Ignacy
Misztal



Alberto
Cesarani



Jorge
Hidalgo



Matias
Bermann



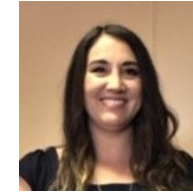
Natalia
Galoro



Sungbong
Jang



Jennifer
Richter



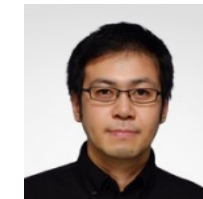
Taylor
Mcwhorther



Mary Kate
Hollifield



Yvette
Steyn



Yutaka
Masuda



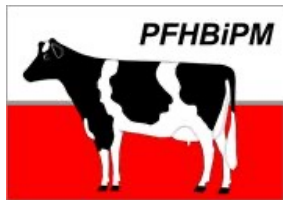
Andres
Legarra



Ignacio
Aguilar



Zulma
Vitezica



Questions?

mbermann@uga.edu



Genomic reliabilities (cont.)

- Core animals:

$$diag((\mathbf{W}_{cc} + \mathbf{G}^{cc} - \mathbf{G}^{cn}(\mathbf{W}_{nn} + \mathbf{M}_{nn}^{-1})^{-1}\mathbf{G}^{nc})^{-1})$$

- Non-core animals:

$$diag((\mathbf{W}_{nn} + \mathbf{M}_{nn}^{-1})^{-1} + (\mathbf{W}_{nn} + \mathbf{M}_{nn}^{-1})^{-1}\mathbf{G}^{nc}(\mathbf{W}_{cc} + \mathbf{G}^{cc} - \mathbf{G}^{cn}(\mathbf{W}_{nn} + \mathbf{M}_{nn}^{-1})^{-1}\mathbf{G}^{nc})^{-1}\mathbf{G}^{cn}(\mathbf{W}_{nn} + \mathbf{M}_{nn}^{-1})^{-1})$$

- $max(O(n_c^3), O(n_n n_c^2))$

