Single-step genomic evaluation of crossbreed dairy cattle in the US

Ignacy Misztal, D. Lourenco, S. Tsuruta, M. Bermann
University of Georgia
Alberto Cesarani **University of Sassari**
A. Legarra **INRA**
E.L. Nicolazzi **CDCB**
P. M. VanRaden **USDA**

UNIVERSITY OF GEORGIA

Introduction

- Genomic evaluation in US by multistep method
- Evaluation by crossbreds by SNP effects from purebreds
  - Accuracy slightly higher than parent average
- Crossbred genotypes not used for purebred evaluation
- Move to single-step models
  - Avoid preselection bias, allow more complex models, simplify pipelines

Goals

- Implement single-step evaluation for crossbreds
- Questions:
  - Are reliabilities for crossbreds higher than based on parent average?
  - Are PTA for purebreds negatively affected by crossbred data?
  - Is computing time reasonable?

Single step in dairy at UGA

- Original single-step paper (2010)
- ……many papers – inflation …….
- Holsteins – current CDCB data (2020)
  - High R^2 and low inflation
  - Data truncation before 2000 not affecting young animals
- 5 dairy breeds with 4 M genotypes (2021)
  - Similar accuracy for single- and multi-breed with choice of core animals

J. Dairy Sci. 93:743–752
doi:10.3168/jds.2009-2730

Hot topic: **A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score**[1]

I. Aguilar,*†[2] I. Misztal,* D. L. Johnson,‡ A. Legarra,§ S. Tsuruta,* and T. J. Lawlor#
*Animal and Dairy Science Department, University of Georgia, Athens 30602
†Instituto Nacional de Investigación Agropecuaria, Las Brujas 90200, Uruguay
‡Livestock Improvement Corp., Private Bag 3016, Hamilton 3240, New Zealand
§INRA, UR631 SAGA, BP 52627, 32326 Castanet-Tolosan, France
#Holstein Association USA Inc., Brattleboro, VT 05302-0808

ARTICLE IN PRESS

J. Dairy Sci. 104
https://doi.org/10.3168/jds.2020-19789

Genomic predictions for yield traits in US Holsteins with unknown parent groups

A. Cesarani,[1]* Y. Masuda,[1] S. Tsuruta,[1] E. L. Nicolazzi,[2] P. M. VanRaden,[3] D. Lourenco,[1] and I. Misztal[1]
[1]Department of Animal and Dairy Science, University of Georgia, Athens 30602
[2]Council on Dairy Cattle Breeding, Bowie, MD 20716
[3]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

J. Dairy Sci. 105
https://doi.org/10.3168/jds.2021-21505

Multibreed genomic evaluation for production traits of dairy cattle in the United States using single-step genomic best linear unbiased predictor

A. Cesarani,[1]* D. Lourenco,[1] S. Tsuruta,[1] A. Legarra,[2] E. L. Nicolazzi,[3] P. M. VanRaden,[4] and I. Misztal[1]
[1]Department of Animal and Dairy Science, University of Georgia, Athens 30602
[2]INRA, UMR1388 GenPhySE, Castanet-Tolosan, France, 31320
[3]Council on Dairy Cattle Breeding, Bowie, MD 20716
[4]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705

Data

Phenotypes from 2000 recorded in Holstein, Jersey, and their crosses

Phenotypes	47,417,185
Cows with records	20,367,132
Tot. animals	27,111,201
Genotypes	1,424,863

Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) 305-d yields recorded from January 2000 (Cesarani et al., 2021)

- Three-trait repeatability model $\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}$

herd × sire

herd × management, age × parity,
inbreeding, heterosis

Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) 305-d yields recorded from January 2000 (Cesarani et al., 2021)

- Three-trait repeatability model $\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}$

  - $\mathbf{Z}_h\mathbf{h}$: herd × management, age × parity, inbreeding, heterosis
  - $\mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a$: herd × sire

- **Complete data**: phenotypes up to August 2021
- **Reduced data**: phenotypes up to August 2017

Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) 305-d yields recorded from January 2000 (Cesarani et al., 2021)

- Three-trait repeatability model $\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}$

  - $\mathbf{Xb}$: herd × management, age × parity, inbreeding, heterosis
  - $\mathbf{Z}_h\mathbf{h}$: herd × sire

- **Complete data**: phenotypes up to August 2021
- **Reduced data**: phenotypes up to August 2017

- UPG (8 groups per breed) in $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}^{-1}$: difference by breed, YOB, and sex

Materials and Methods II

Validation cows = genotyped females with no phenotypes in the reduced dataset

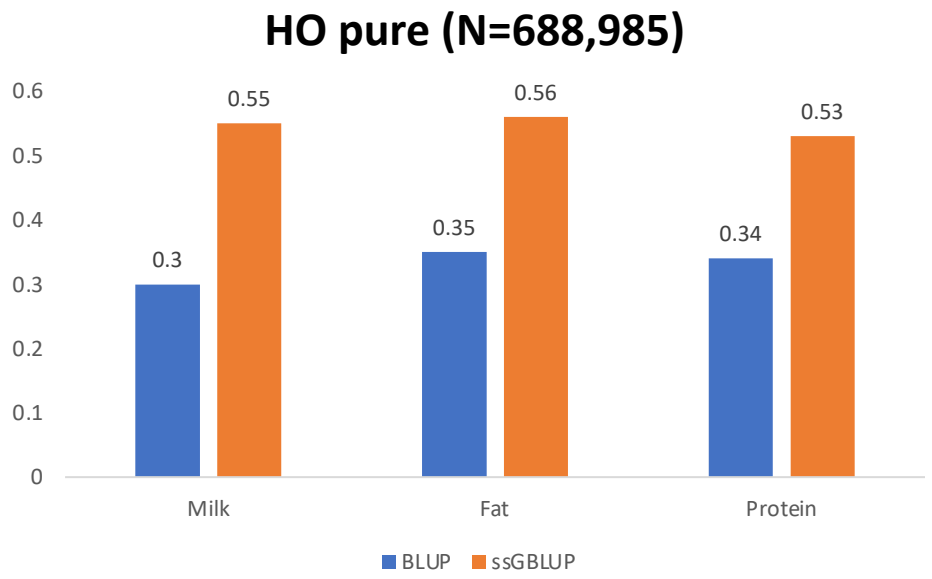
- **HO pure** = HO animals with both sire and dam HO
- **JE pure** = JE animals with both sire and dam JE
- **cross** = HO or JE animals with at least one parent of the opposite breed
- **cross F1** = cross animals with 100 heterosis

Validation method = predictivity based on adjusted phenotypes (correlation and b_1)

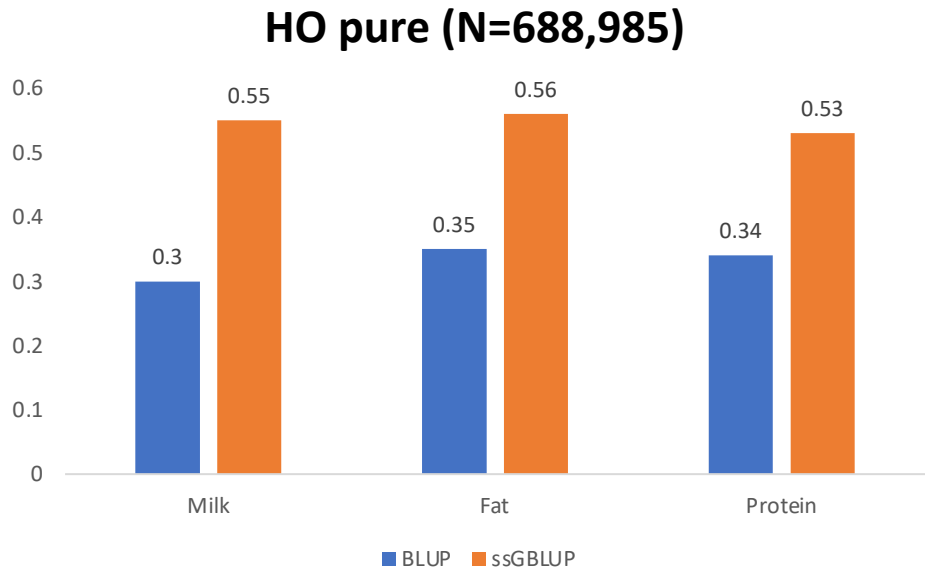
pred=corr(PTA, y-”fixed”)

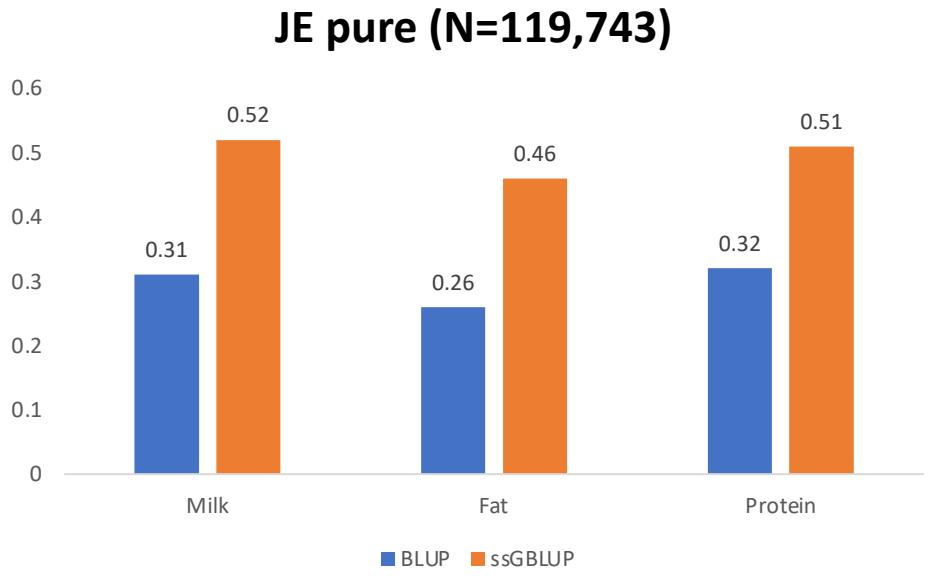
Predictive abilities for cows

Predictive abilities for cows

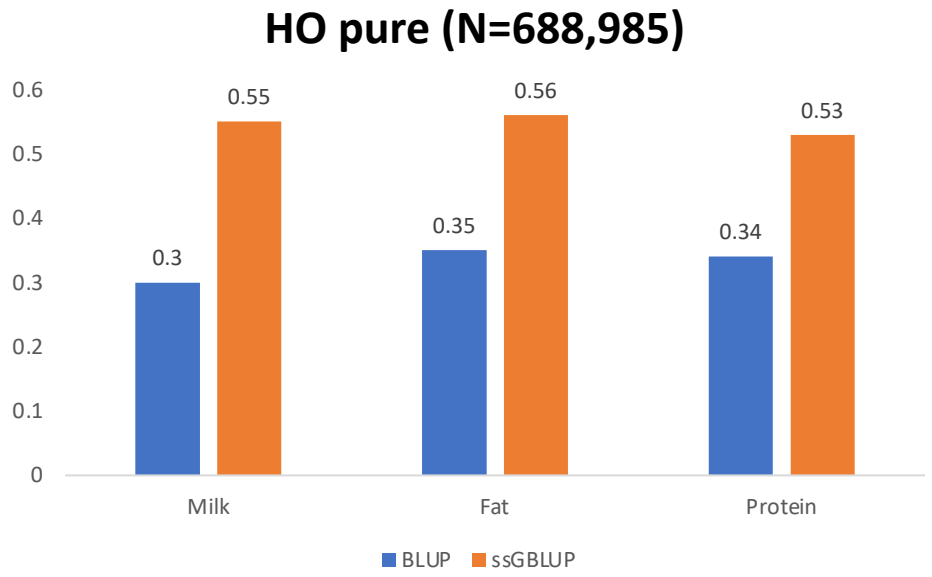


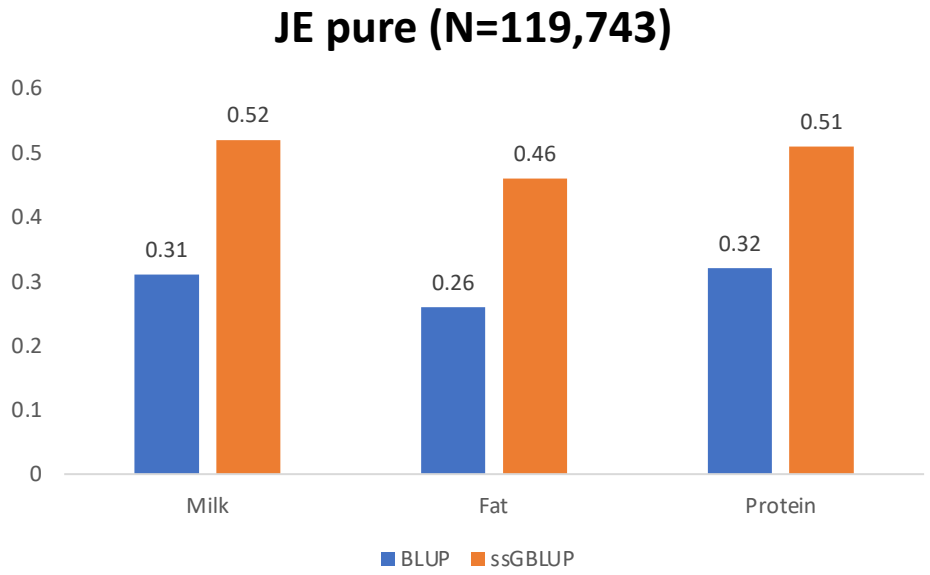
Predictive abilities for cows

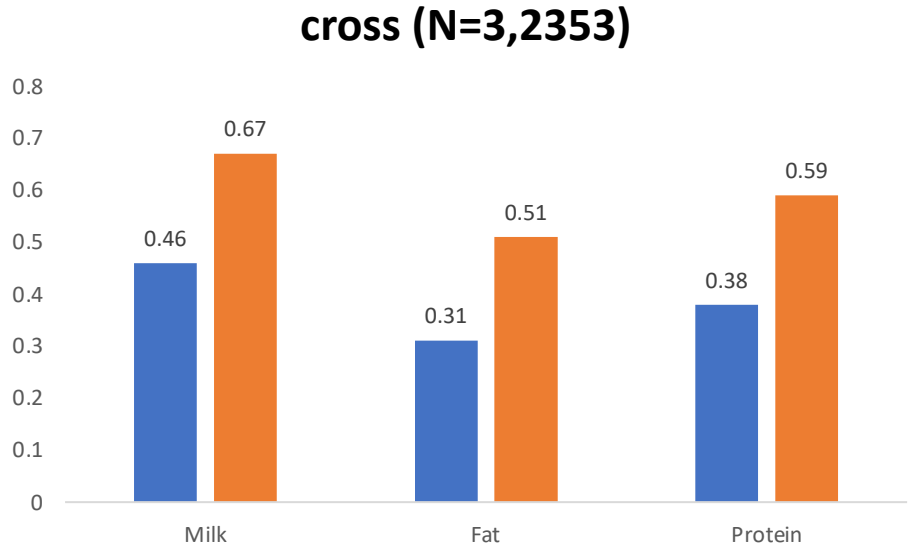




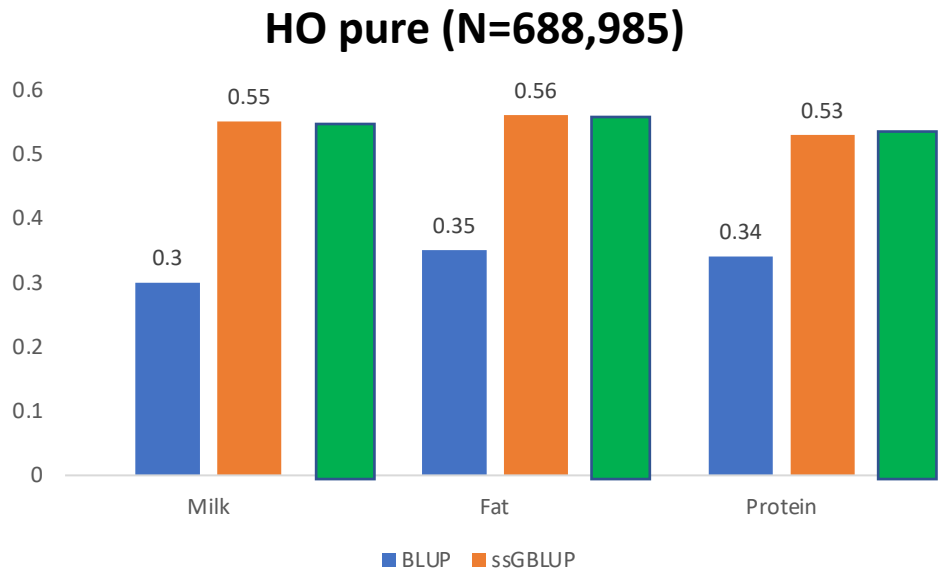
Predictive abilities for cows

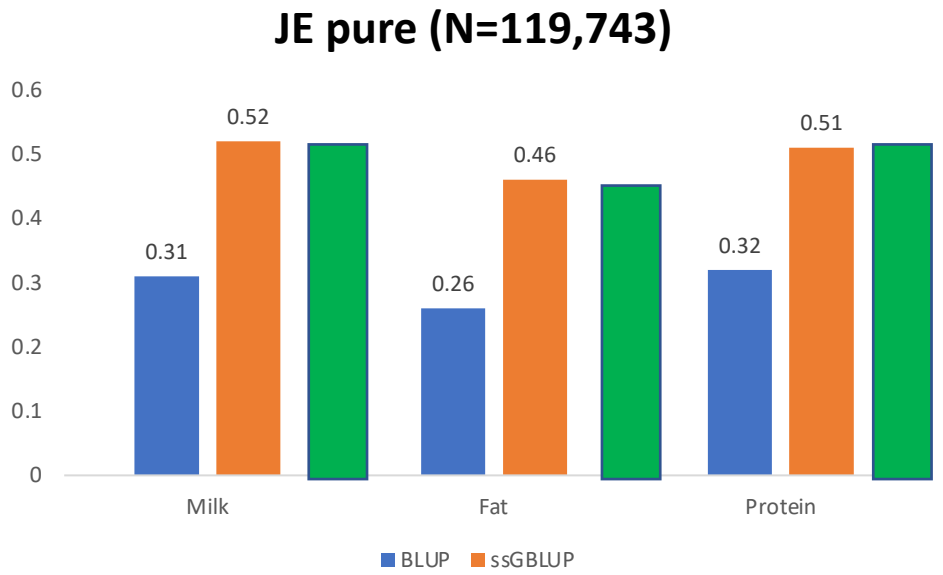


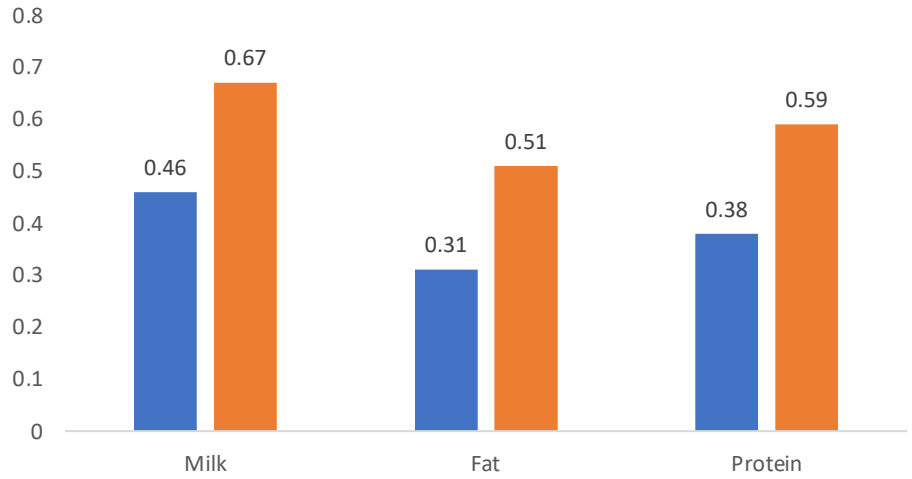




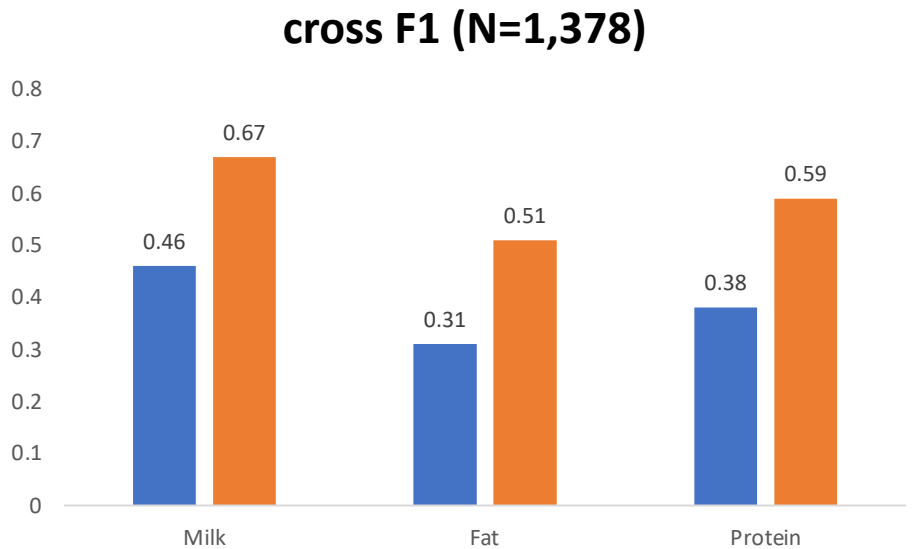
Predictive abilities for cows



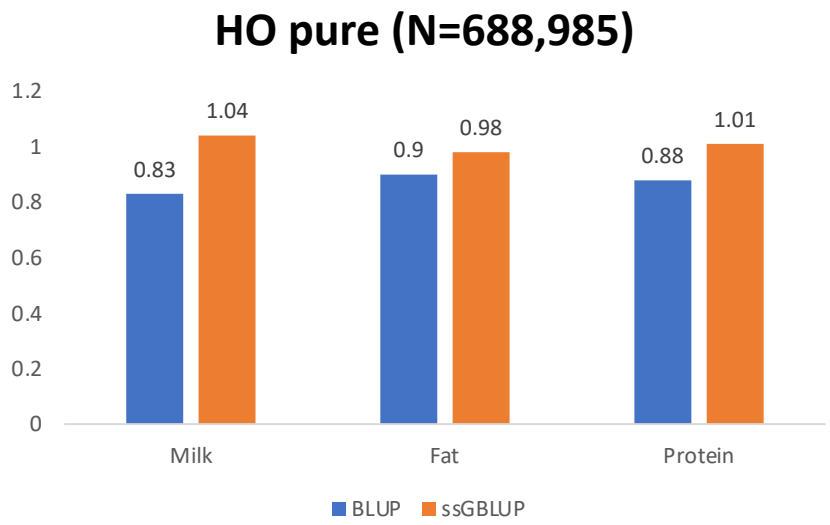




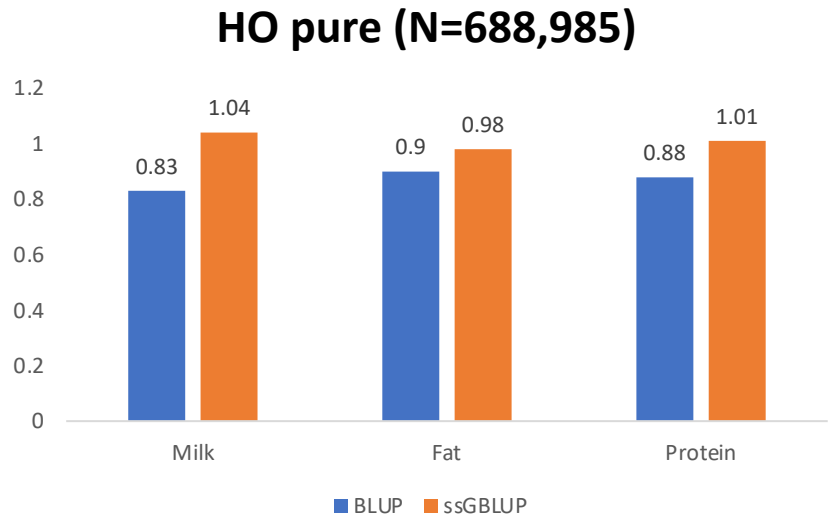
Single breed analyzes

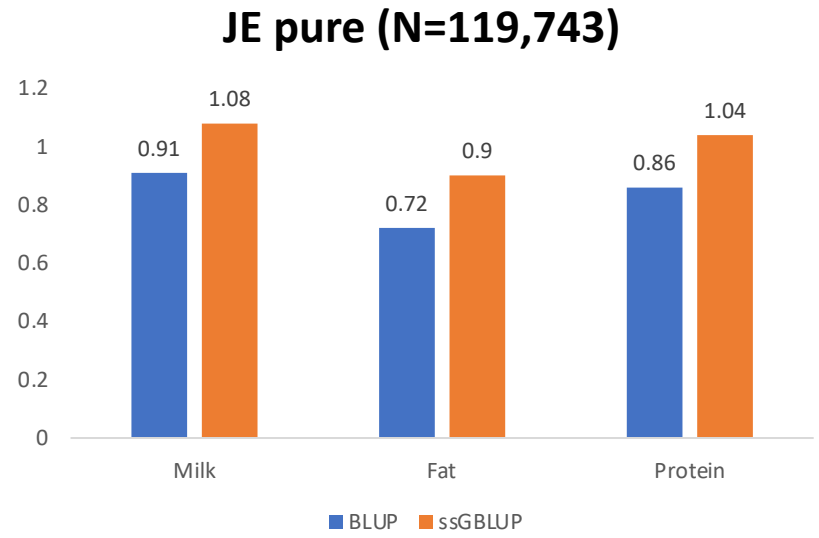


b1 values

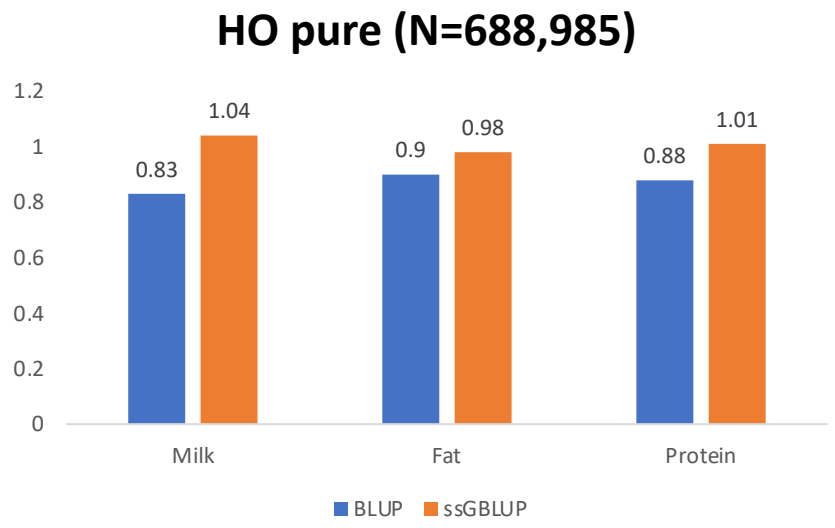

HO pure (N=688,985)
1.2
1
0.8
0.6
0.4
0.2
0
1.04
0.83
0.98
0.9
1.01
0.88
Milk
Fat
Protein
BLUP
ssGBLUP

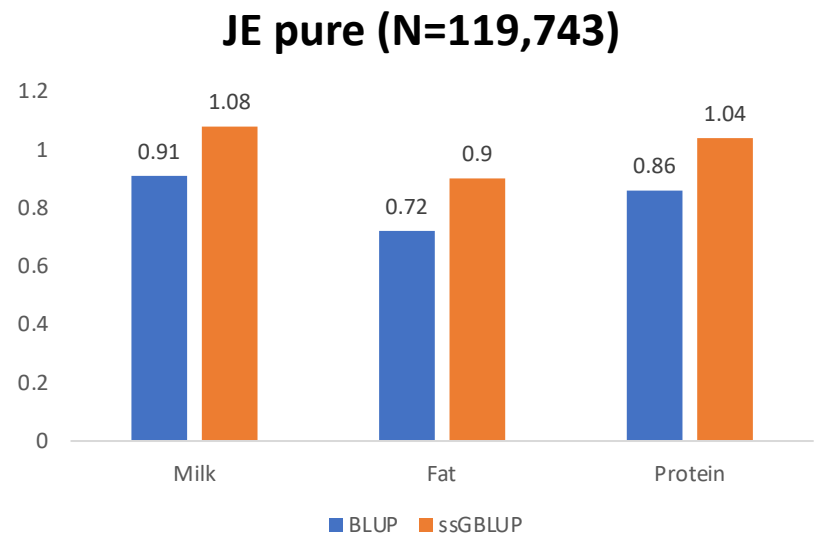
b1 values

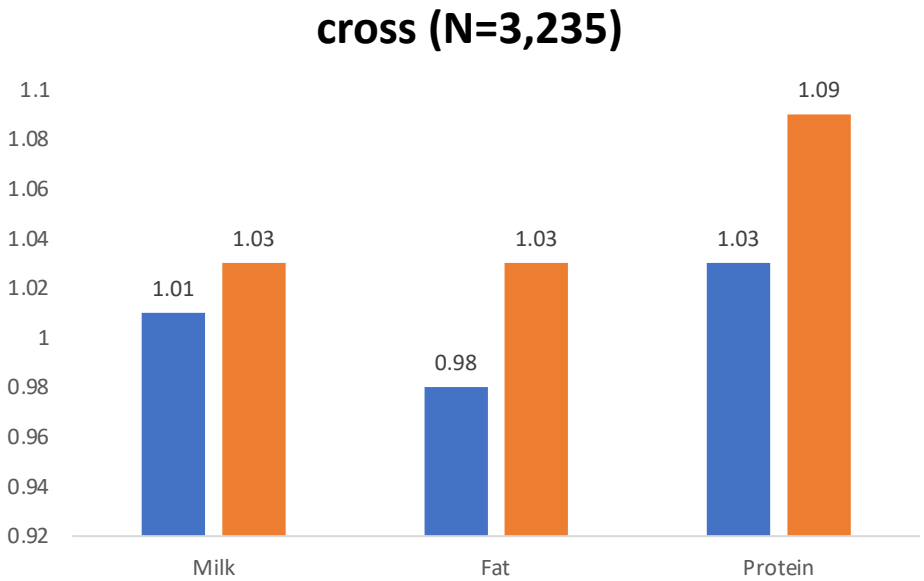

HO pure (N=688,985)
1.2
1
0.8
0.6
0.4
0.2
0
0.83
1.04
0.9
0.98
0.88
1.01
Milk
Fat
Protein
BLUP
ssGBLUP


JE pure (N=119,743)
1.2
1
0.8
0.6
0.4
0.2
0
0.91
1.08
0.72
0.9
0.86
1.04
Milk
Fat
Protein
BLUP
ssGBLUP

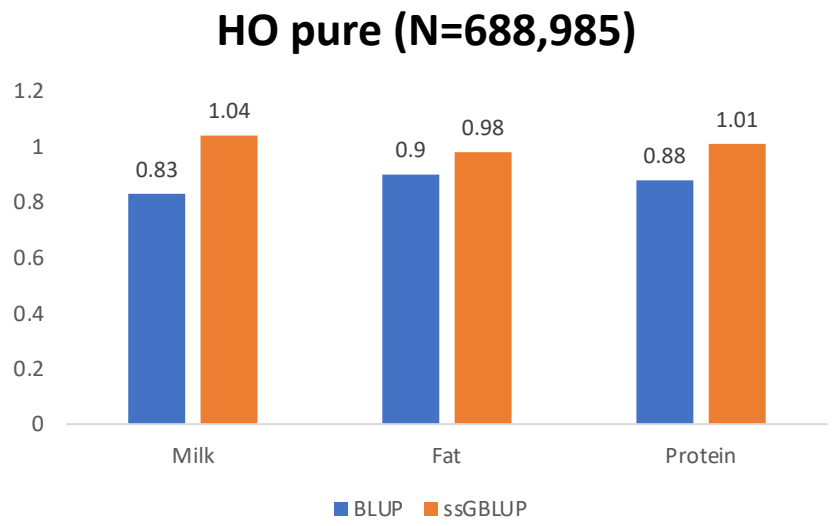
b1 values

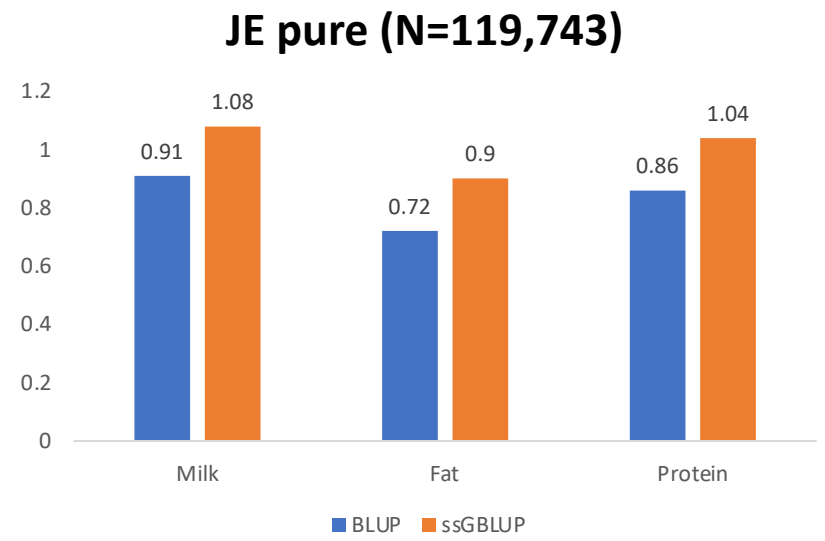

HO pure (N=688,985)
1.2
1
0.8
0.6
0.4
0.2
0
0.83
1.04
0.9
0.98
0.88
1.01
Milk
Fat
Protein
BLUP
ssGBLUP

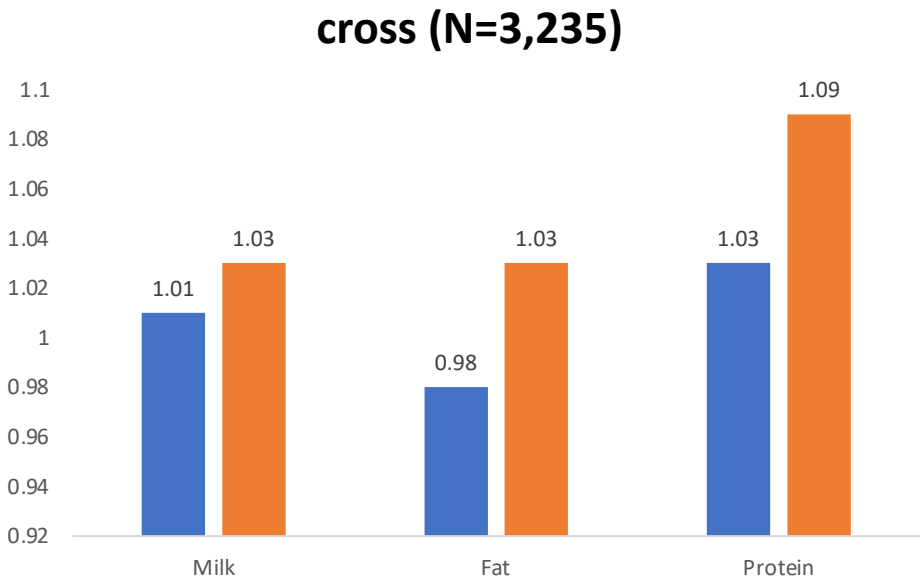

JE pure (N=119,743)
1.2
1
0.8
0.6
0.4
0.2
0
0.91
1.08
0.72
0.9
0.86
1.04
Milk
Fat
Protein
BLUP
ssGBLUP

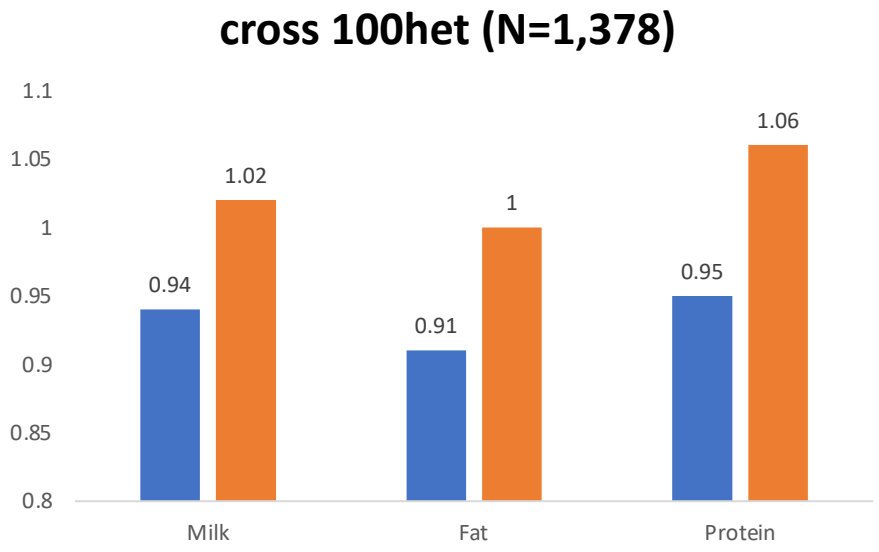

cross (N=3,235)
1.1
1.08
1.06
1.04
1.02
1
0.98
0.96
0.94
0.92
1.01
1.03
0.98
1.03
1.03
1.09
Milk
Fat
Protein

b1 values


HO pure (N=688,985)
1.2
1
0.8
0.6
0.4
0.2
0
0.83
1.04
0.9
0.98
0.88
1.01
Milk
Fat
Protein
BLUP
ssGBLUP


JE pure (N=119,743)
1.2
1
0.8
0.6
0.4
0.2
0
0.91
1.08
0.72
0.9
0.86
1.04
Milk
Fat
Protein
BLUP
ssGBLUP


cross (N=3,235)
1.1
1.08
1.06
1.04
1.02
1
0.98
0.96
0.94
0.92
1.01
1.03
0.98
1.03
1.03
1.09
Milk
Fat
Protein


cross 100het (N=1,378)
1.1
1.05
1
0.95
0.9
0.85
0.8
0.94
1.02
0.91
1
0.95
1.06
Milk
Fat
Protein

Predictivity and accuracy

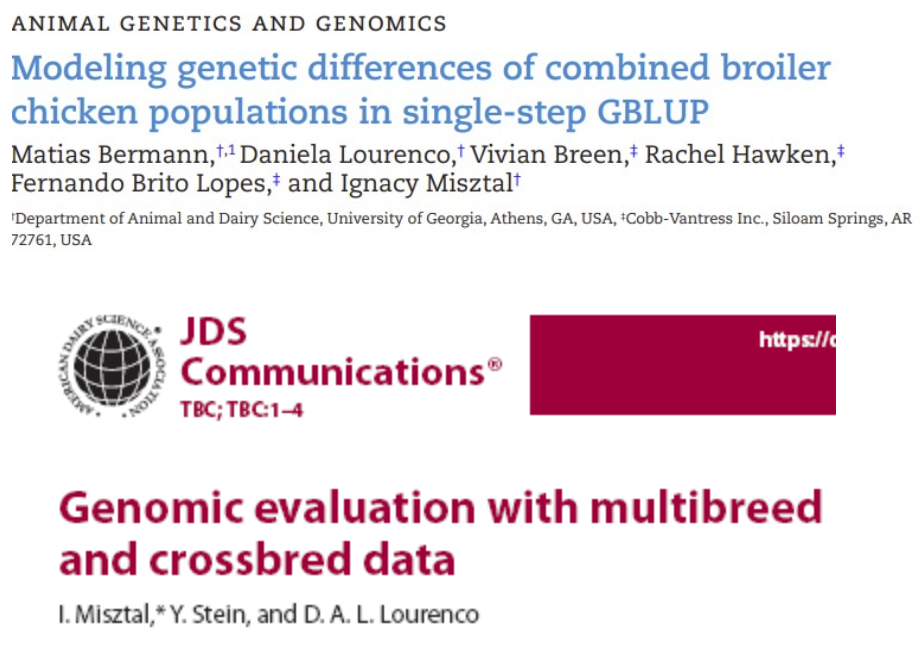
$$accuracy = \frac{pred}{h} = \frac{corr(PTA, y - \text{"fixed"})}{h}$$

For Holstein milk: pred = 0.55, acc ≈ 0.90, ➔ $h^2 \approx 0.22$

For crossbreds: pred = 0.67, assume $h^2 \approx 0.22$ ➔ acc = 1.10 ??????

Breed type(F1, F2, reciprocal…) ignored

hard to do from existing data



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

JDS Communications®

TBC; TBC:1–4

https://

Genomic evaluation with multibreed and crossbred data

I. Misztal,* Y. Stein, and D. A. L. Lourenco

Computing time

Step	Rounds	Sec / round	Time
preGSf90			~ 10 h
BLUP	671	88.98	~ 17 h
ssGBLUP	459	134.06	~ 17 h

Conclusions

- Predictivity for crossbreds higher than expected
  - Results superior if used for management
- Purebred evaluations unaffected by crossbred data
- Computations feasible

Acknowledgements