UNIVERSITY OF GEORGIA

Single-step genomic BLUP for national beef cattle evaluation in US:

from initial developments to final implementation

Daniela Lourenco

S. Tsuruta, B.O. Fragomeni, Y. Masuda, I. Aguilar

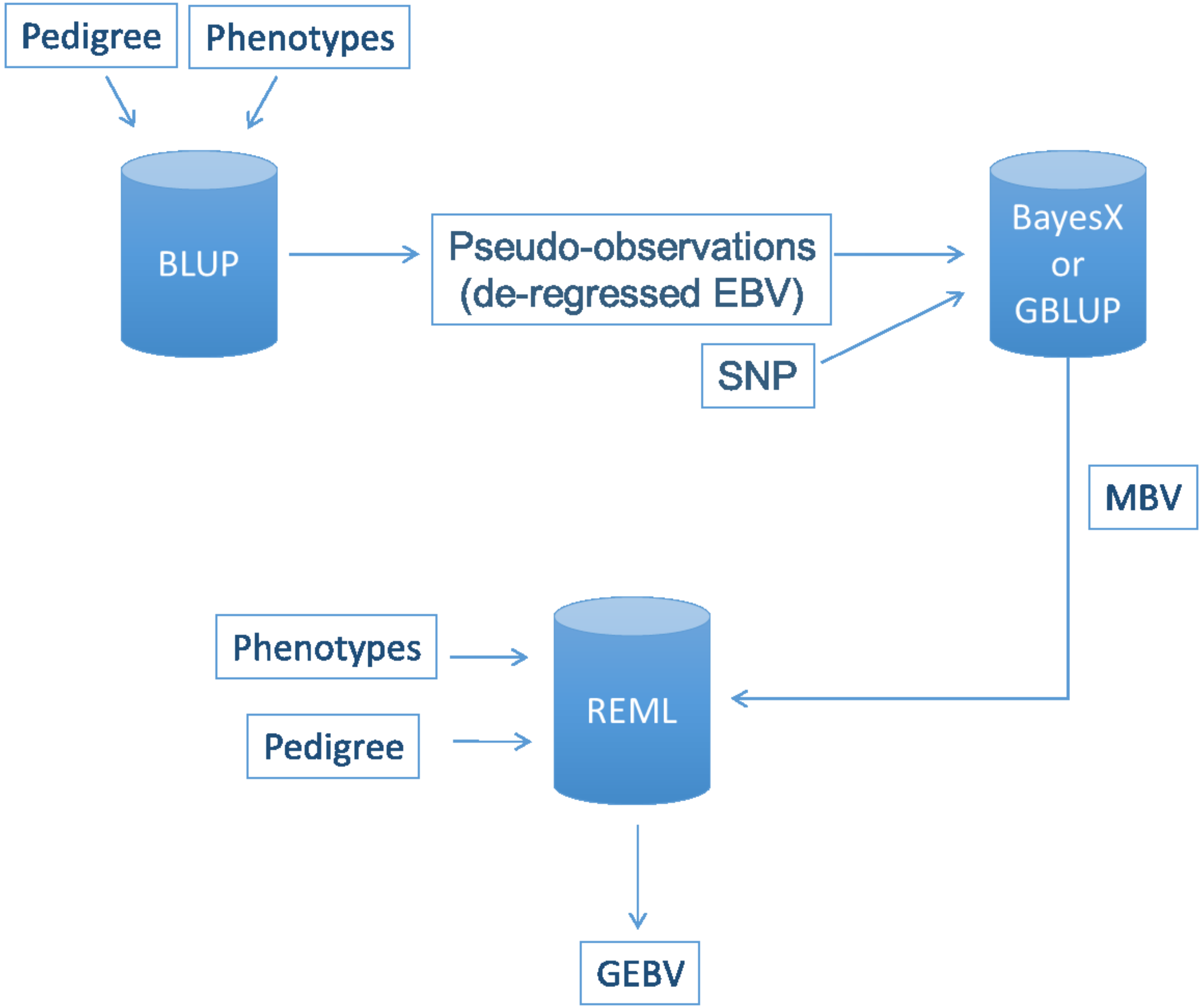
A. Legarra, S. Miller, D. Moser, I. Misztal

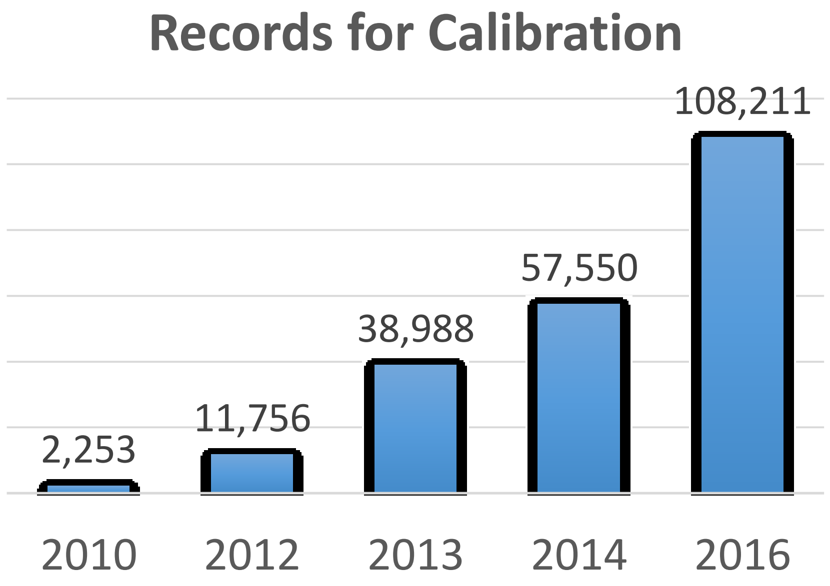
11th WCGALP 2018

Angus

- Main beef cattle breed in USA
- Genomic Selection since 2009

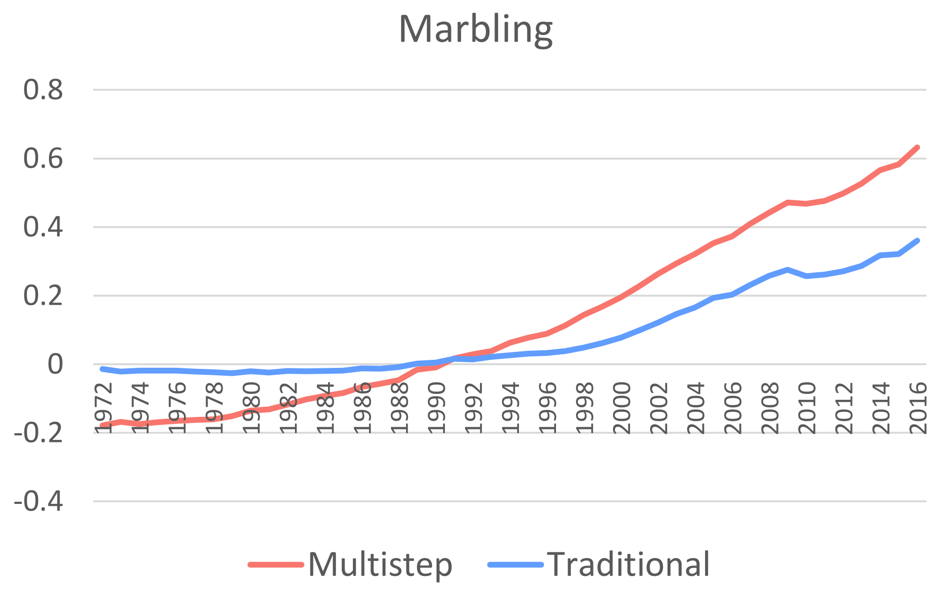
Multistep Genomic Evaluation



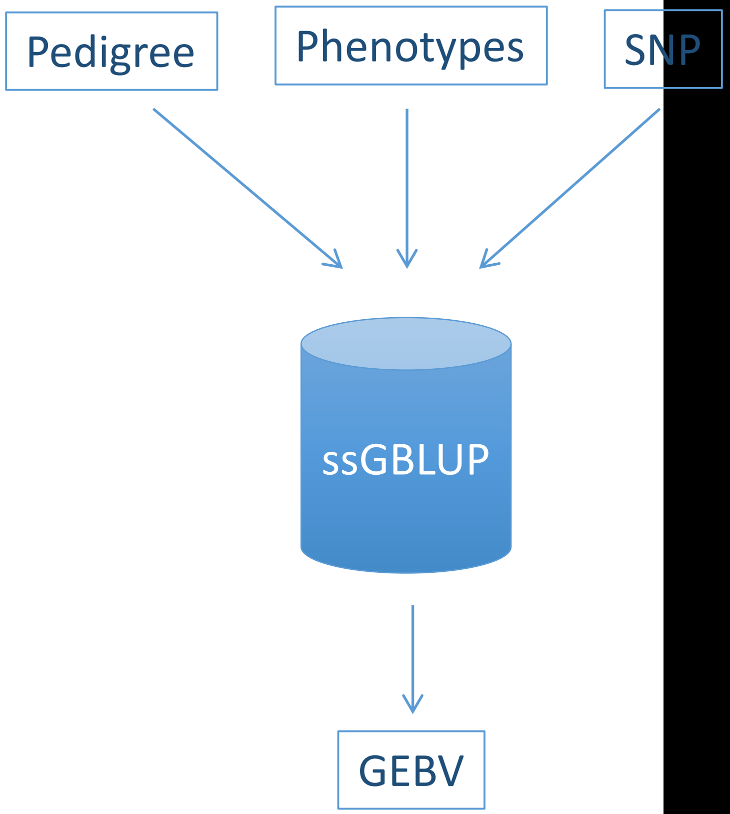


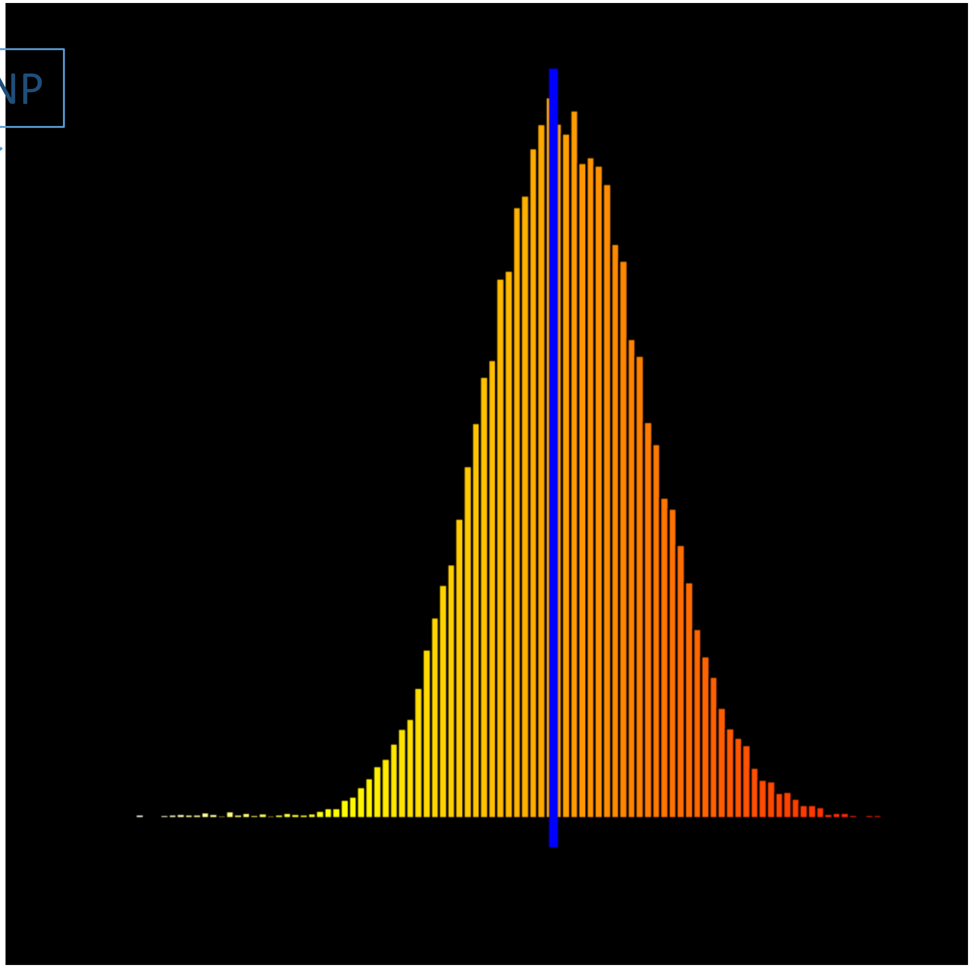
Problems with Multistep

- Big fluctuations in GEBV for new calibration
- Rank change for bulls with high accuracy
- Overfitted models – 2x the number of traits
- High genetic correlation between phenotype and MBV



Single-step genomic BLUP (ssGBLUP)



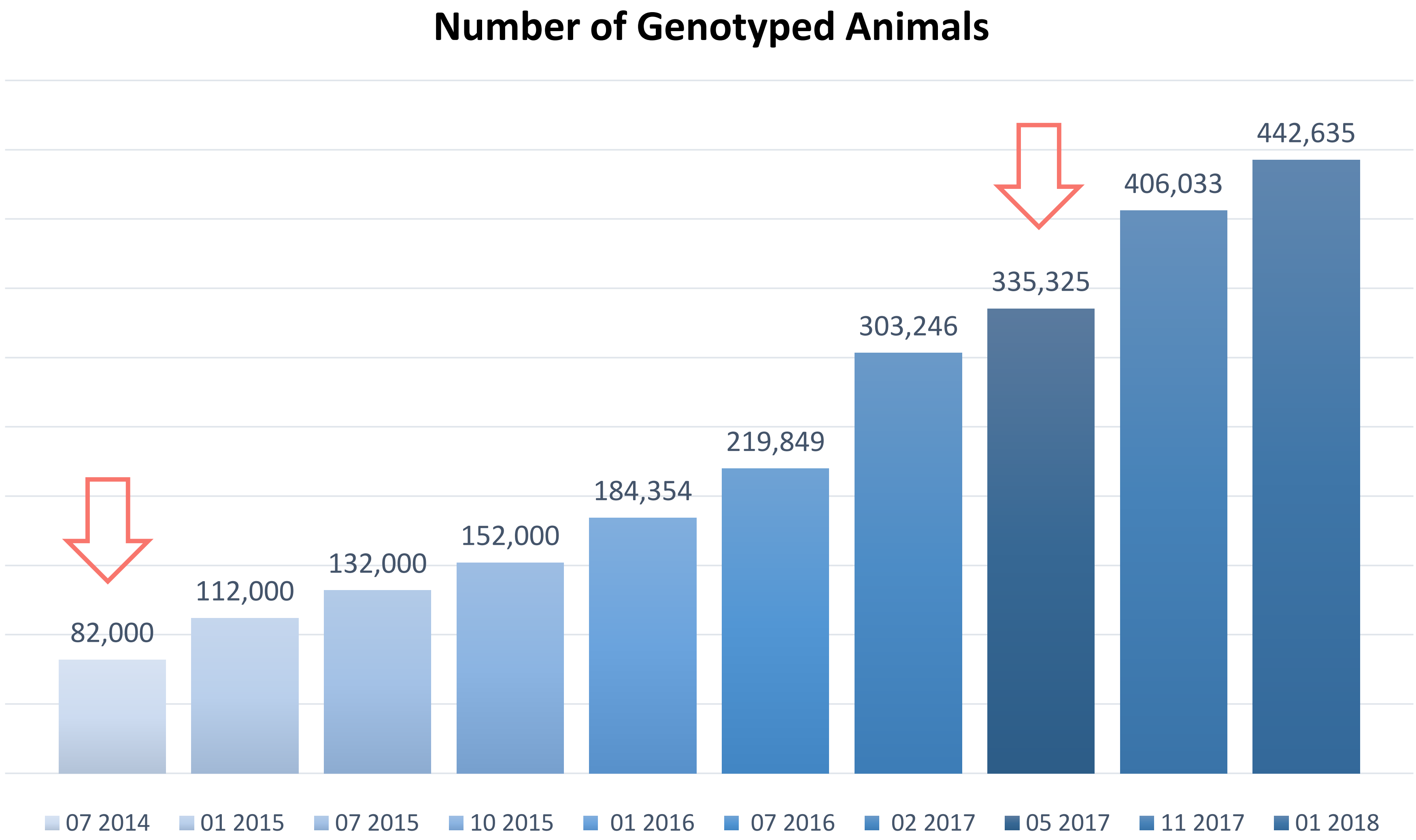


$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

Aguilar et al., 2010

UGA group
(2008 – now)

Initial tests of ssGBLUP for Angus



Ability to predict future performance

2014

- 8M animals in pedigree
- 6M BW and WW
- 3.4M PWG
- 52k genotyped animals
- 18.7k born in 2013

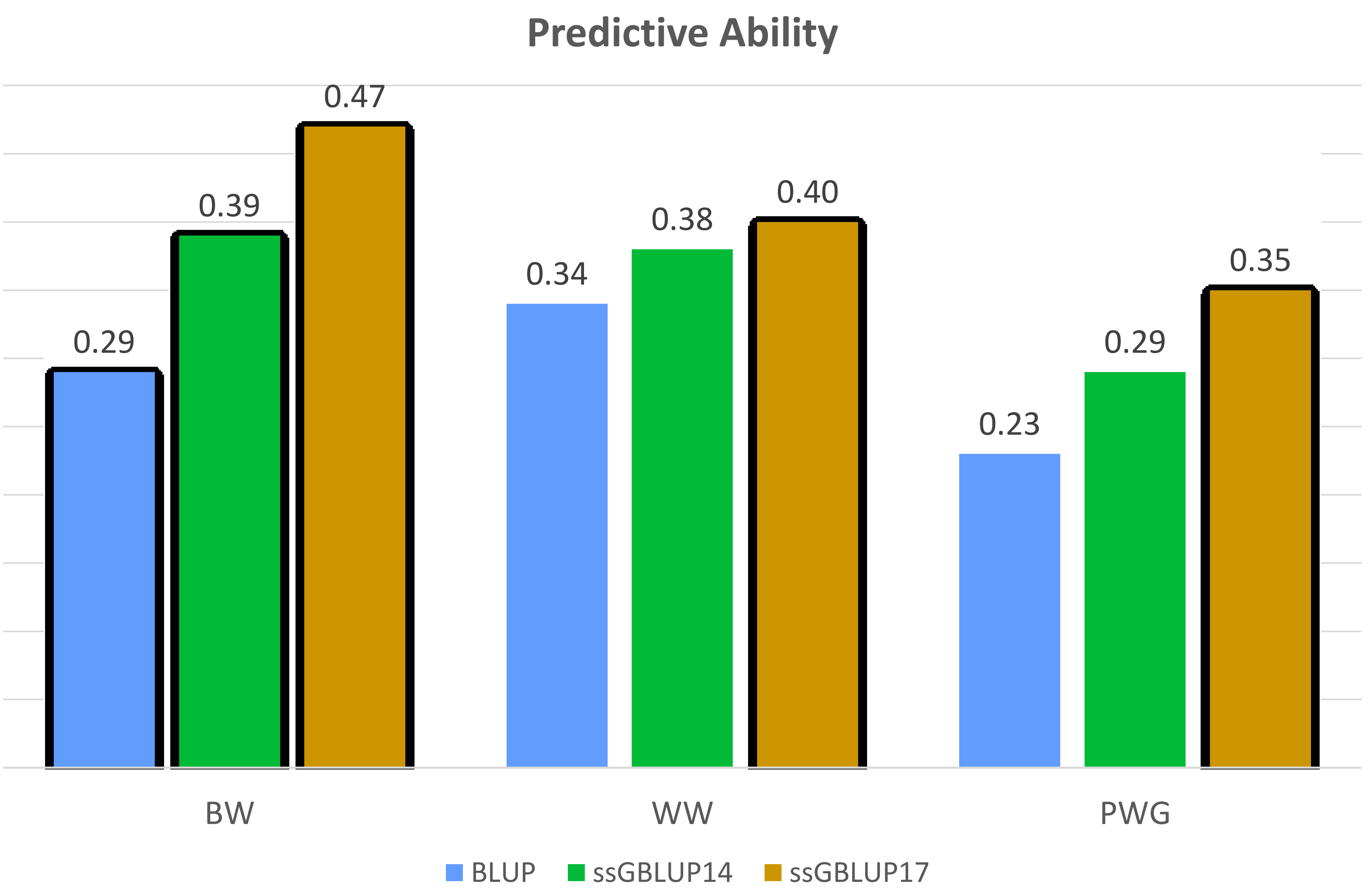
2017

- 10M animals in pedigree
- 8M BW and WW
- 4.2M PWG
- 335k genotyped animals
- 18.7k born in 2016

Predictive ability direct = COR(Y_adj, GEBV)

Predictive ability maternal = COR(Y_adj, total_maternal_GEBV)

Ability to predict future performance



Average Gain

Direct

2014 = 25%

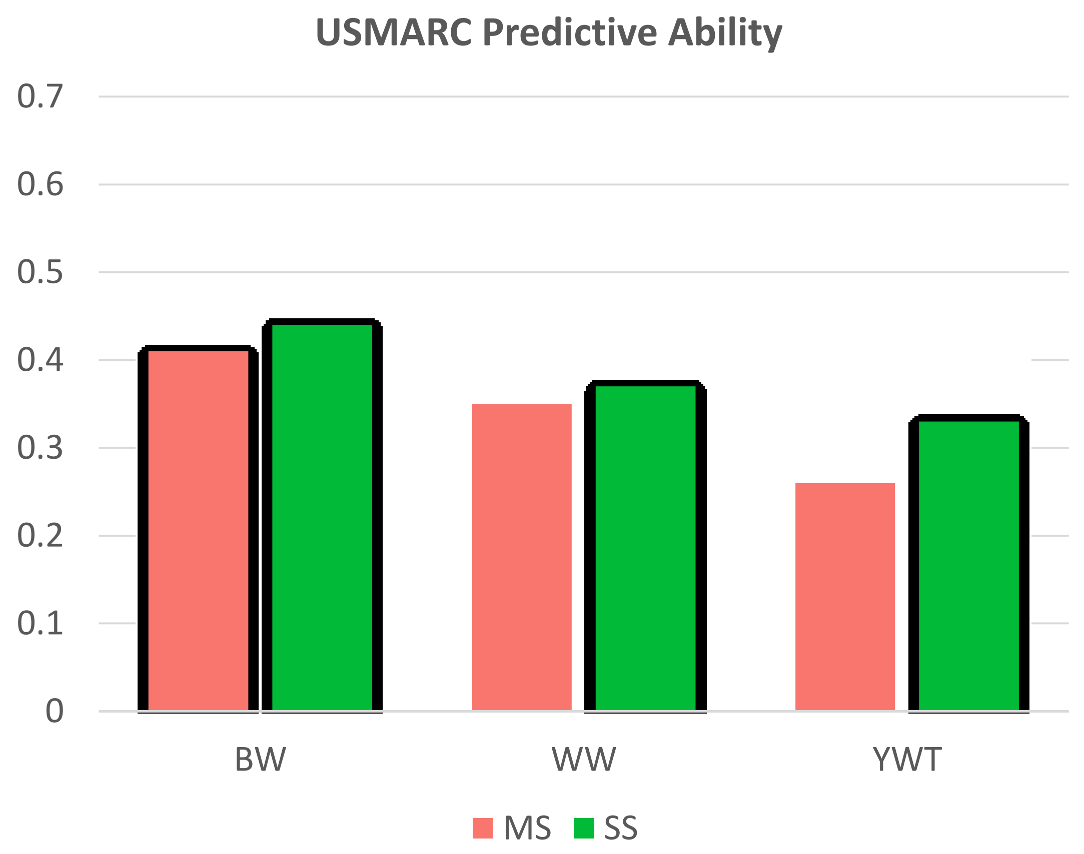
2017 = 36%

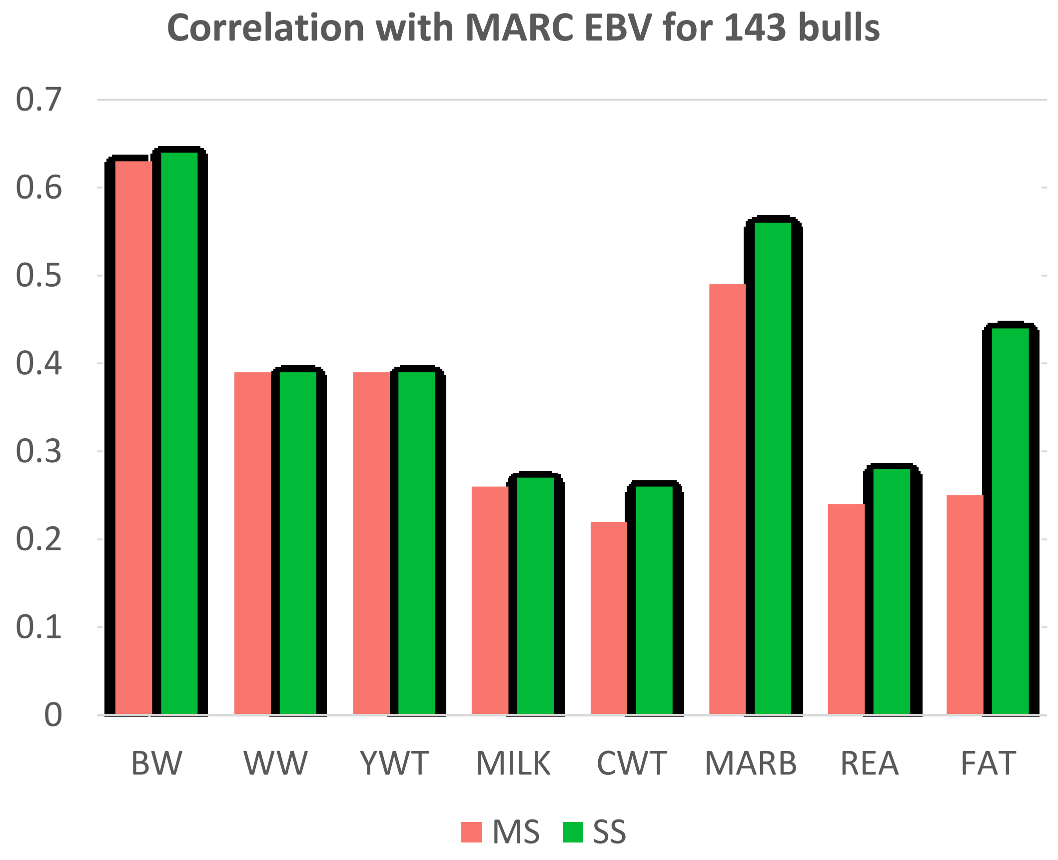
Maternal

2014 = 8%

2017 = 10%

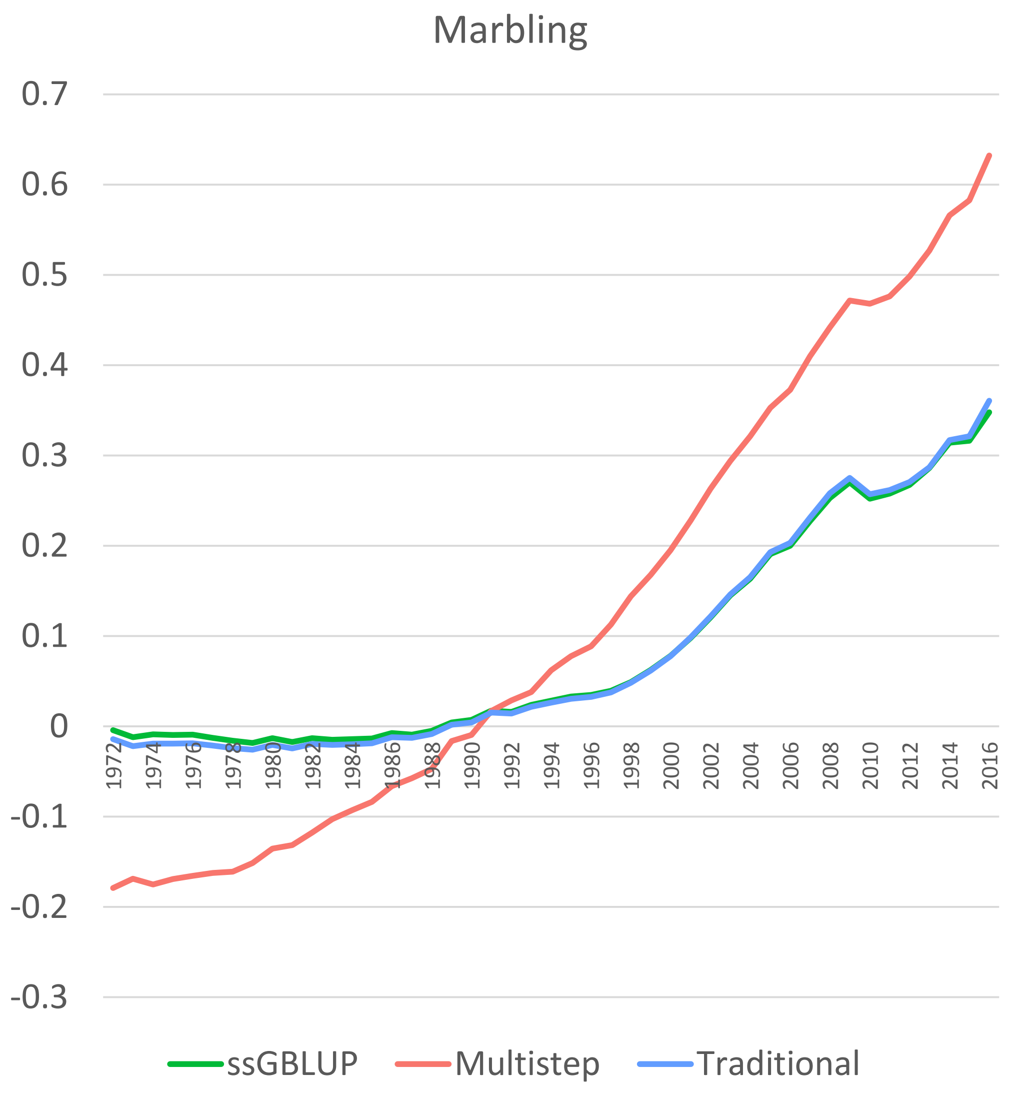
*US*MARC comparisons of ssGBLUP x multistep

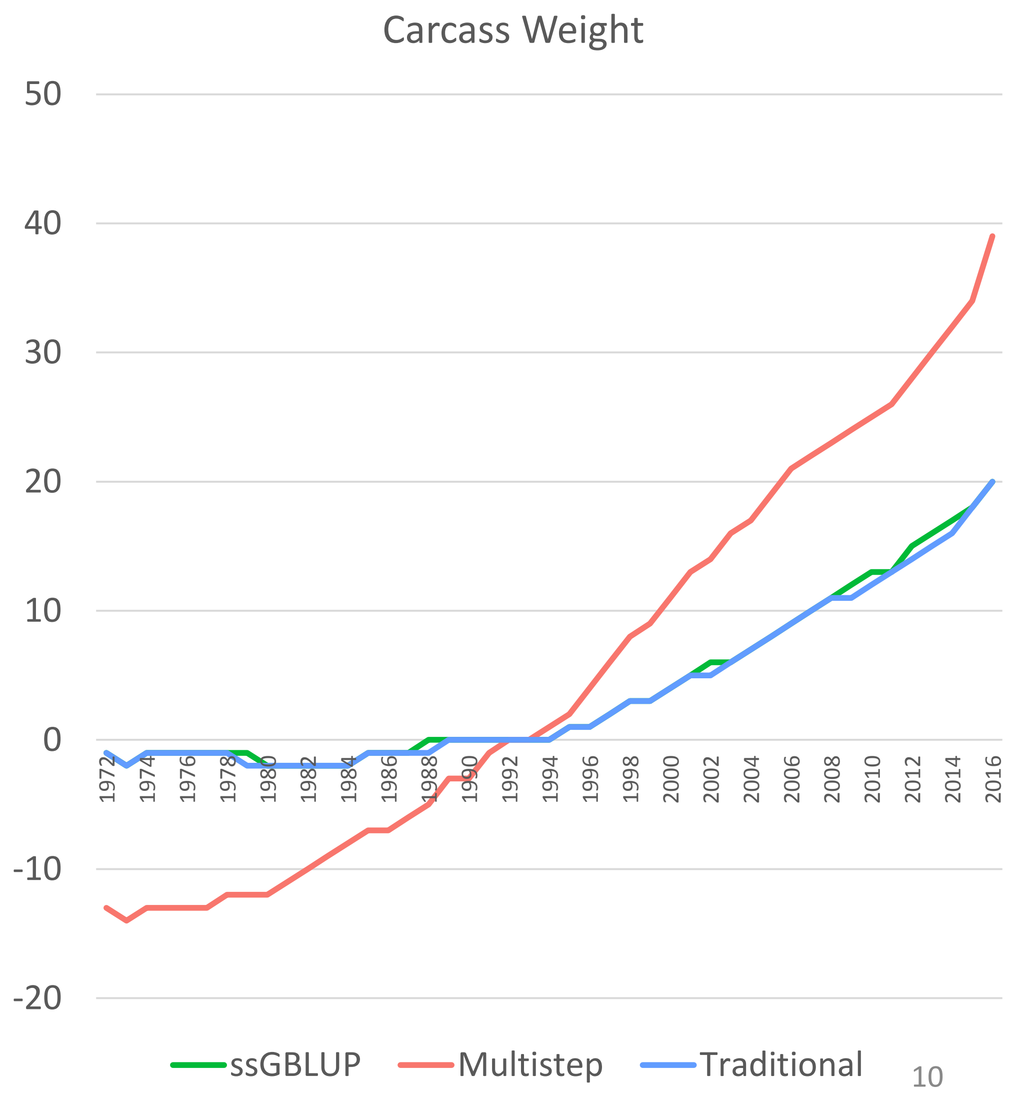




Kuehn et al., 2017

Genetic trends for carcass traits





Increasing number of genotyped animals

- Number of genotyped animals increased 5-fold from 2014 to 2018
  - 150,000
  - > 2 hours
  - > 700Gb RAM

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

- APY ssGBLUP
  - Borrowed from algorithm to construct A^{-1}
  - Core and Non-core

$$\mathbf{G}_{\mathrm{APY}}^{-1} = \begin{bmatrix} \mathbf{G}_{cc}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{G}_{cc}^{-1}\mathbf{G}_{cn} \\ \mathbf{I} \end{bmatrix} \mathbf{M}_{nn}^{-1} \begin{bmatrix} -\mathbf{G}_{nc}\mathbf{G}_{cc}^{-1} & \mathbf{I} \end{bmatrix} \qquad \mathrm{M}_{nn} = g_{ii} - g_{ic}G_{cc}^{-1}g_{ci}$$

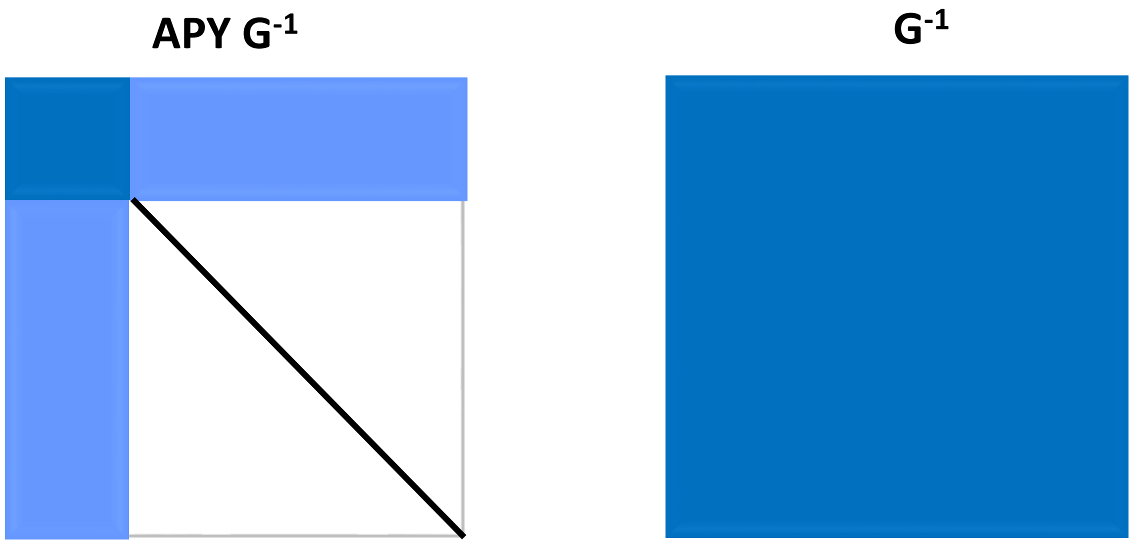
Misztal et al., 2014

APY ssGBLUP in 2014

```
AAA~~~8781383 1
AAA~~~8974207 1
AAA~~~9459638 1
AAA~~~9604587 1
AAA~~~9651921 1
AAA~~~9891499 1
AAA~~~9906195 1
AAA~~~9921007 1
AAA~~~10383408 2
AAA~~~10648572 2
AAA~~~10710606 2
AAA~~~10848637 2
AAA~~~10931840 2
AAA~~~10971449 2
AAA~~~10971485 2
AAA~~~11118769 2
AAA~~~11142393 2
AAA~~~11223766 2
AAA~~~11308904 2
AAA~~~11356568 2
AAA~~~11367940 2
AAA~~~11373742 2
AAA~~~11391800 2
AAA~~~11392490 2
AAA~~~11447335 2
AAA~~~11468795 2
AAA~~~11520398 2
AAA~~~11567326 2
```

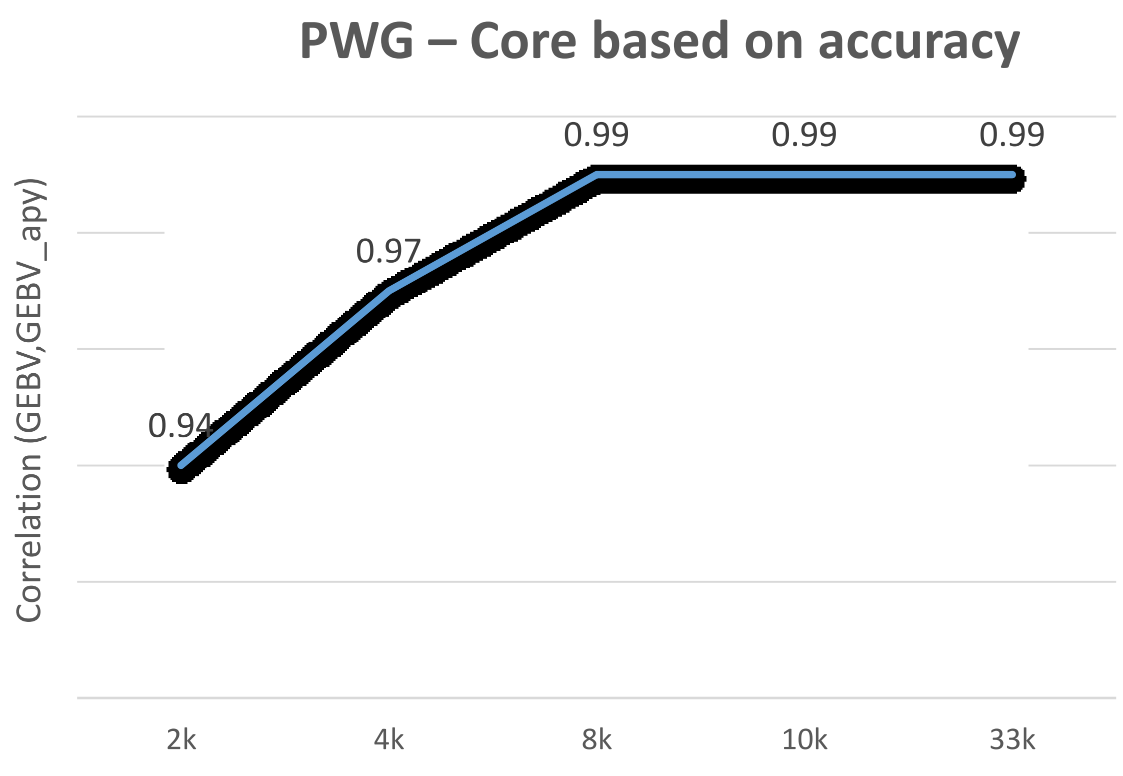
core

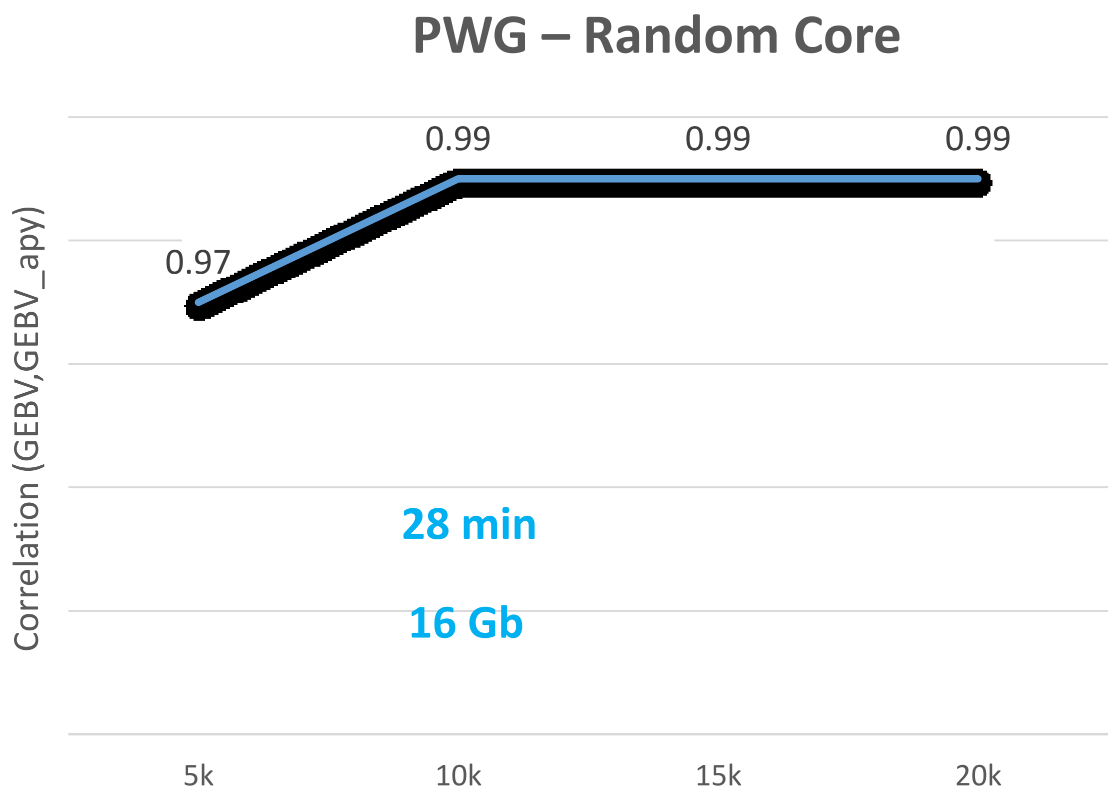
non-core



How to choose core animals?

APY ssGBLUP in 2014





Lourenco et al., 2015a
Lourenco et al., 2015b

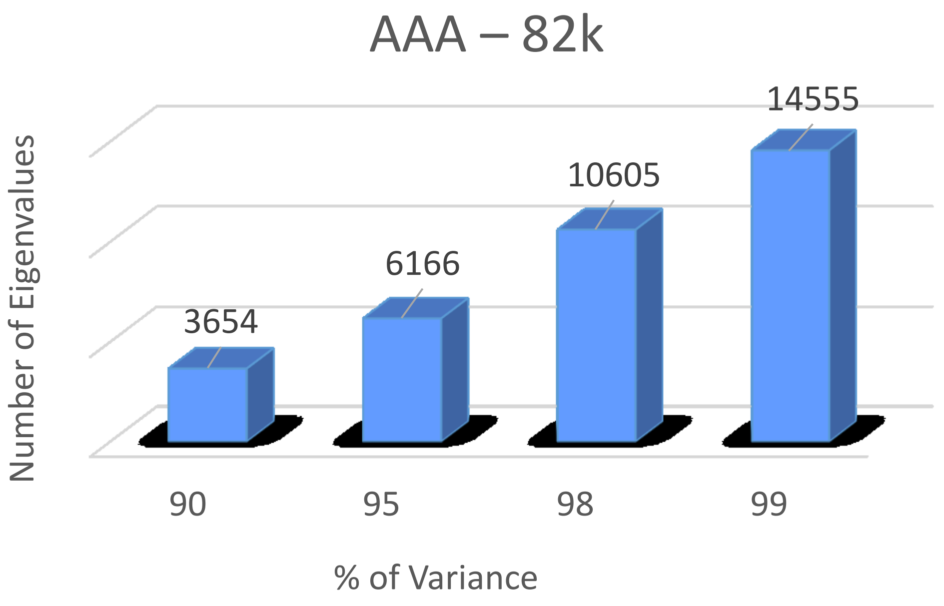
Regular inversion = 213 min

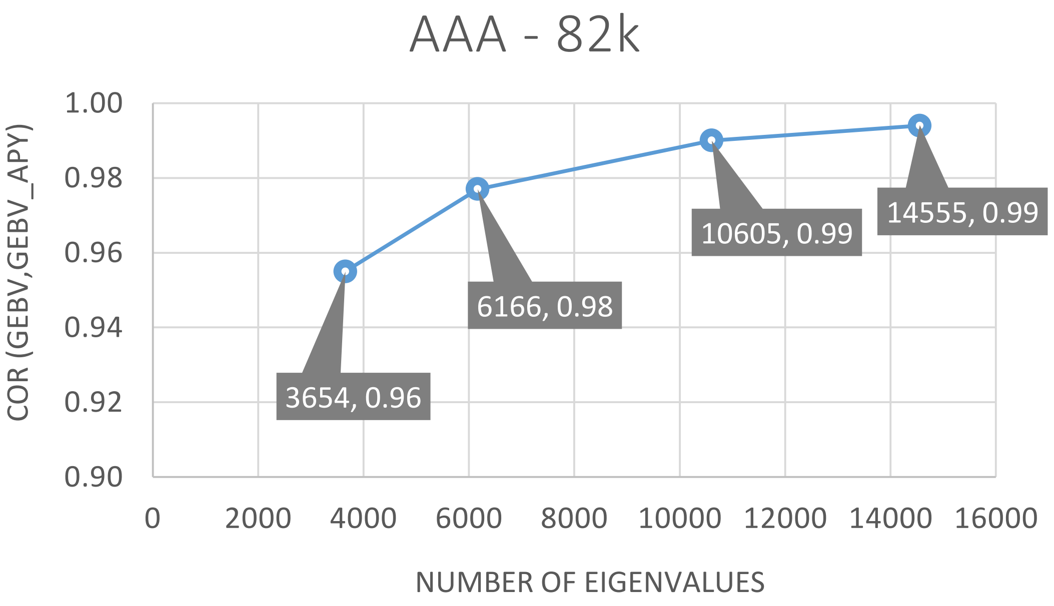
230 Gb

How to choose the number of core in APY?

- Ne, Me, ESM, Eigen of G
- Limited dimensionality

Pocrnic et al., 2016
Misztal, 2016





Additional features in ssGBLUP

- Commercial products
  - e.g. GeneMax for non-registered animals
  - Based on SNP effects
  - Accurate SNP effects with APY?

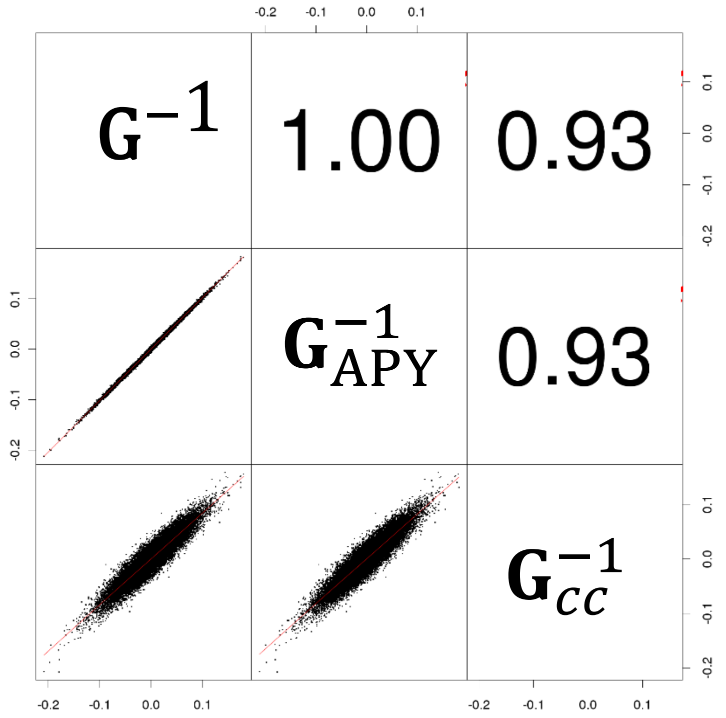
SNP effects in APY ssGBLUP

$$\hat{a}_{\mathbf{G}} = \lambda \mathbf{D}\ \mathbf{Z}'\mathbf{G}^{-1}\hat{u}$$

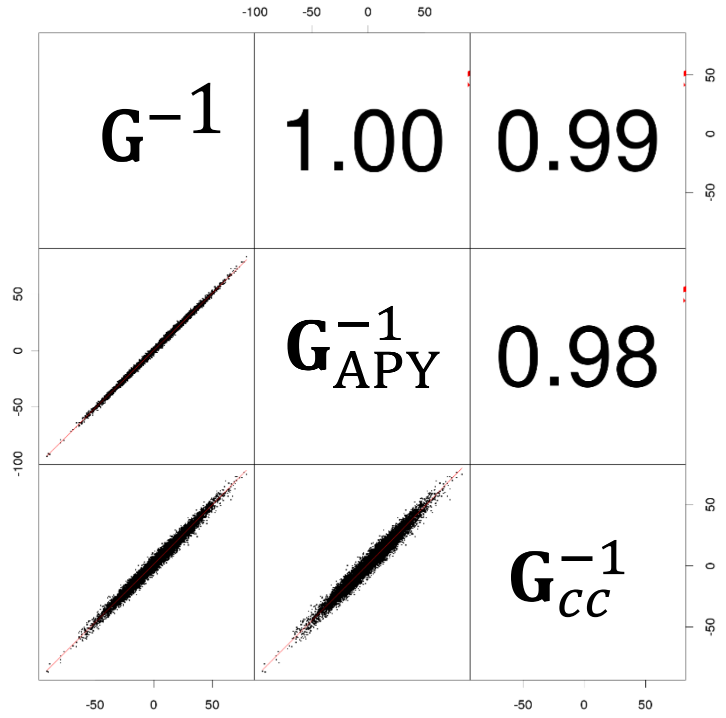
$$\hat{a}_{\mathbf{G}^{-1}_{APY}} = \lambda \mathbf{D}\ \mathbf{Z}'\mathbf{G}^{-1}_{\mathrm{APY}}\hat{u}_{\mathrm{APY}}$$

$$\hat{a}_{\mathbf{G}^{-1}_{cc}} = \lambda \mathbf{D}\ \mathbf{Z}'\mathbf{G}^{-1}_{CC}\hat{u}_{\mathrm{APY}}$$

$\hat{a}$



$\mathbf{Z}\hat{a}$



Additional features in ssGBLUP

- Interim evaluations
  - Indirect predictions
  - Quick evaluations between official runs
  - Should be comparable to GEBV

Indirect predictions for young animals

$$\left\{\mathbf{W}'\mathbf{W} + \alpha\,\mathbf{A}^{-1} + \alpha\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}\end{bmatrix}\right\}\hat{u} = \mathbf{W}'\mathbf{y}$$

parent average

yield deviation

progeny contribution

direct genomic value

pedigree prediction

$$\textbf{GEBV} = w_1\textbf{PA} + w_2\textbf{YD} + w_3\textbf{PC} + w_4\textbf{DGV} - w_5\textbf{PP}$$

$$\textbf{GEBV}_y = w_1\textbf{PA} + w_4\textbf{DGV} - w_5\textbf{PP}$$

$$\textbf{GEBV}_y \approx \textbf{DGV}$$

Lourenco et al., 2015

Problem with Indirect predictions

COR(GEBV,DGV) > 0.99

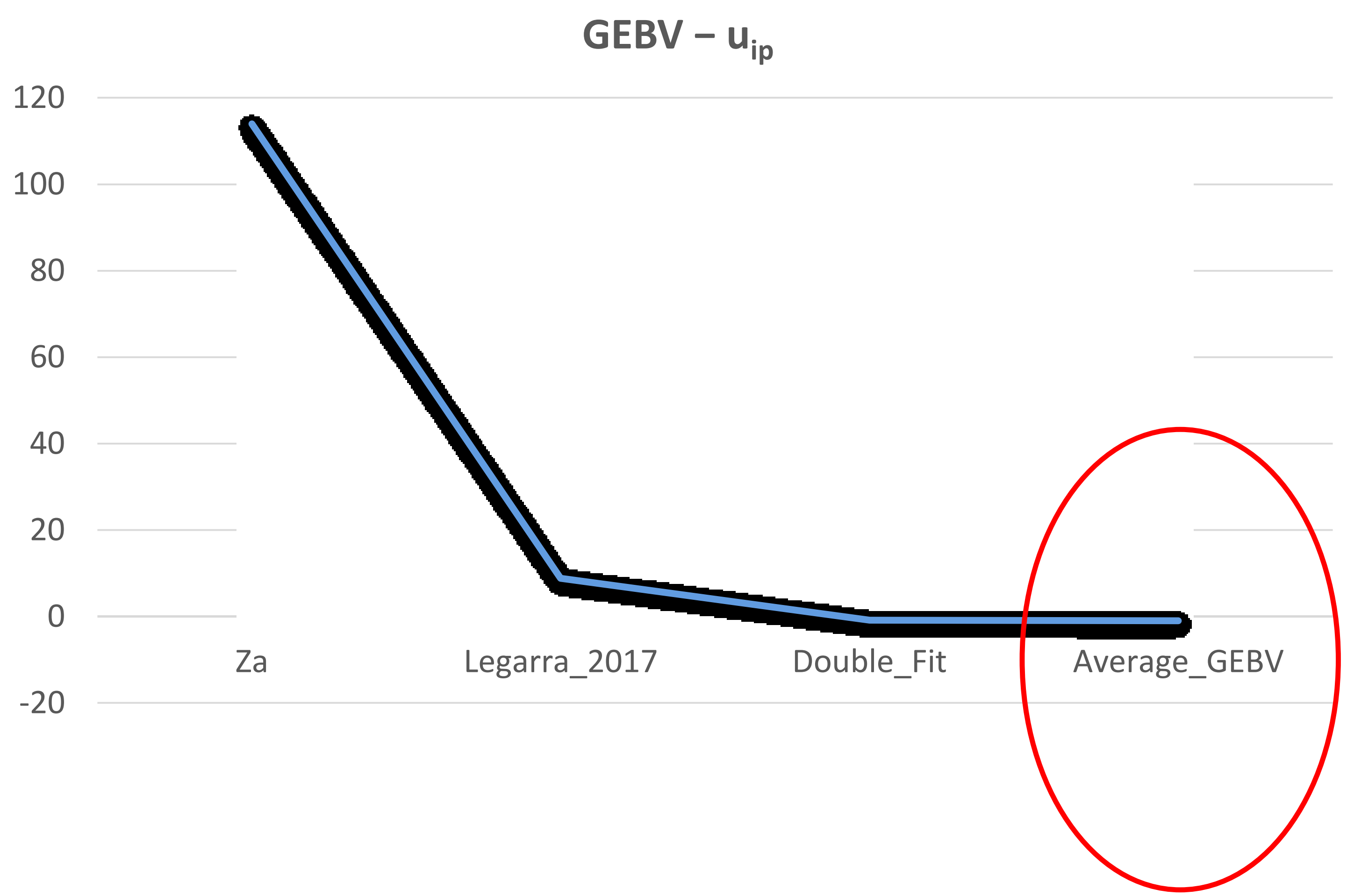
Lourenco et al., 2015

Avg(GEBV) ≈ 100 ≠ Avg(DGV) ≈ 0

- Base of SSGBLUP: modelled as a mean in genotyped animals
  - $p(\boldsymbol{u}_g) = N(\mathbf{1}\mu, \mathbf{G})$

    Vitezica et al. (2011)
  - μ = (Pedigree base) – (Genomic base)

Correcting for bias of indirect predictions



$$\mathrm{E}(\hat{\boldsymbol{u}}|\hat{\boldsymbol{a}}) = \mu + \mathbf{Z}\hat{\boldsymbol{a}}$$

$$\approx$$

$$\mathbf{DGV} = \overline{\mathrm{GEBV}} + \mathbf{Z}\hat{\boldsymbol{a}}$$

Lourenco et al., 2018

Issues in the implementation of ssGBLUP for Angus

1) Omega = 0.7 indicates inflation in GEBV

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

Inbreeding

Inbreeding

Inbreeding

Solution: adding inbreeding for A^{-1} removed inflation in GEBV

Omega = 1.0

Issues in the implementation of ssGBLUP for Angus

2) Inclusion of external EBV into growth evaluation

- 10k Red Angus EBV
- External EBV + genomics was not supported

$$\mathbf{H}^{*-1} = \begin{bmatrix} \mathbf{H}^{EE} + \mathbf{T}^{-1} - \mathbf{H}_{EE}^{-1} & \mathbf{H}^{EI} \\ \mathbf{H}^{IE} & \mathbf{H}^{II} \end{bmatrix}$$

- E = external
- I = internal
- T = PEV for E

Adapted from Legarra et al., 2007

Issues in the implementation of ssGBLUP for Angus

3) Calving ease evaluation was not quite easy

- BW + CE in linear-threshold model
- BLUP = 12 hours
- 152k genotyped animals
- APY ssGBLUP = 4.5 days

Scenario	Description of parameters			rounds	hours	correlation with genomic
	pcg rounds	alpha	beta			
traditional	40	-		60	12	-
genomic	40	0.9	0.1	488	108	-

Issues in the implementation of ssGBLUP for Angus

4) Accuracy of GEBV

- Large datasets
  - Impossible to invert

- d_i^r and d_i^p are approximated

(Misztal and Wiggans, 1988)

- Accuracy = 1 - LHS^{-1}

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Diag(C^{ZZ+}) = PEV

$$LHS_{uu}^{ii} = 1 / (\lambda + d_i^r + d_i^p)$$

Issues in the implementation of ssGBLUP for Angus

4) Accuracy of GEBV

$$LHS_{uu}^{ii} = 1 / (\lambda + d_i^r + d_i^p + d_i^g)$$

$$\left\{ \mathbf{Z}'\mathbf{Z} + \lambda \mathbf{A}^{-1} + \lambda \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \right\}$$

$\mathbf{Z}'\mathbf{Z} \rightarrow d_i^r$

$\lambda \mathbf{A}^{-1} \rightarrow d_i^p$

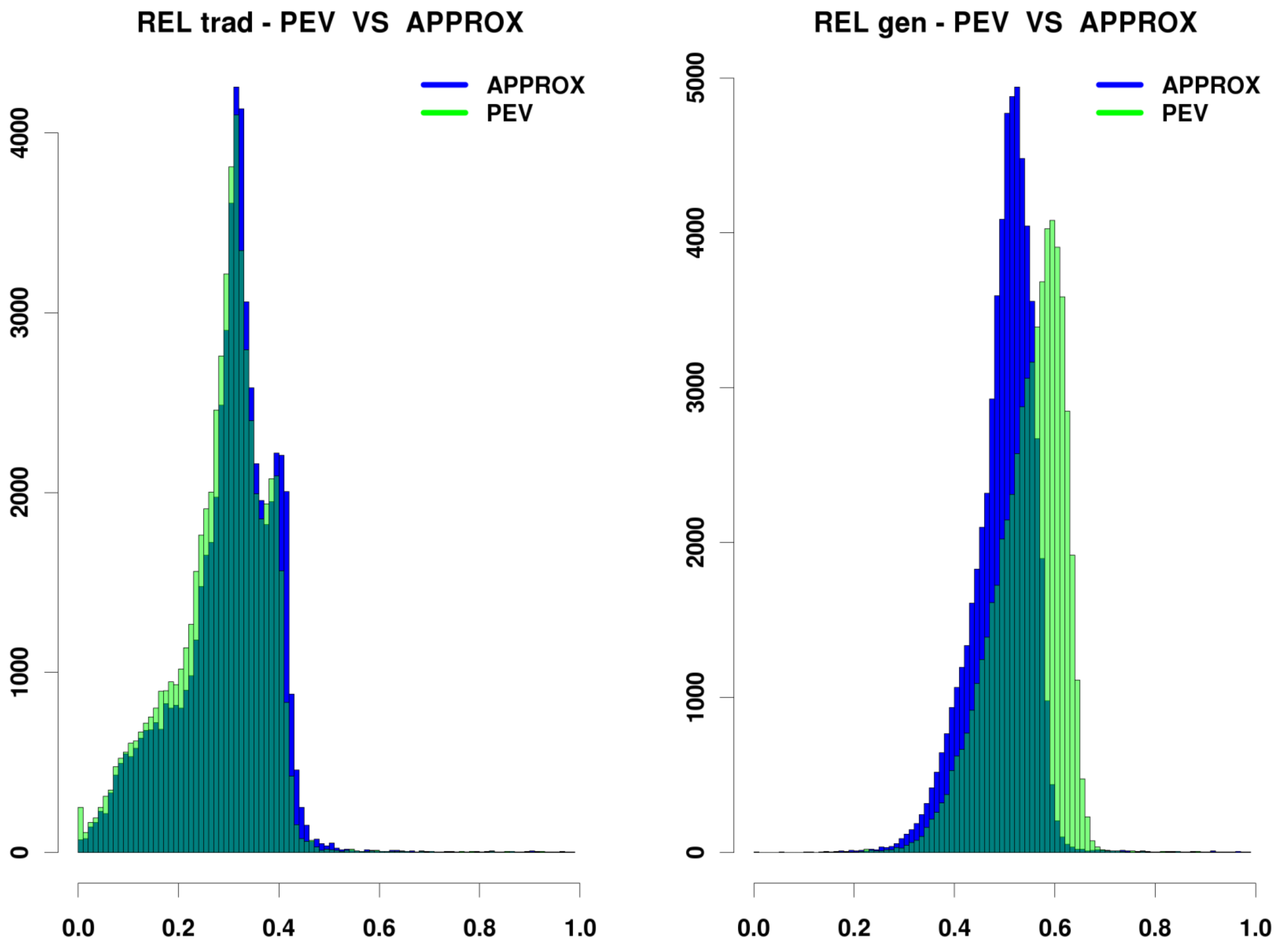
$\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \rightarrow d_i^g$

$$d_i^g = var\_ratio * [\overline{Rel} + (1 - g_{ii}) + zeta * \overline{Rel - Rel_{PA}}]$$

Issues in the implementation of ssGBLUP for Angus

4) Accuracy of GEBV

Cor = 0.87
Avg_True = 0.55
Avg_approx. = 0.50
MSE = 0.0035



Implementation of ssGBLUP on 7/7/2017

- Current Angus evaluation with ~ 450k
  - 19k core
  - Weekly evaluations
  - ~ 18 traits (maternal, categorical, external information)
  - Indirect predictions based on SNP effects $\hat{\boldsymbol{a}}_{\mathbf{G}_{cc}^{-1}} = \lambda \mathbf{D}\,\mathbf{Z}'\mathbf{G}_{CC}^{-1}\hat{\boldsymbol{u}}_{\mathrm{APY}}$

- Minimal changes for proven animals
- Considerable changes for young animals
- More variation among half- and full-sibs

Final Remarks

- ssGBLUP tests were extensive and took couple of years
  - More stable than multistep

- Implementation of ssGBLUP by Angus raised several issues
  - All solved
  - Successful weekly evaluations for 7 months
  - Evaluation with ~450k genotyped animals is possible with APY

- Implementation of ssGBLUP for Angus in 2017 set new standards for beef cattle evaluation in USA

Acknowledgements