

Issues in commercial application of single-step genomic BLUP for genetic evaluation in American Angus

Daniela Lourenco

S. Tsuruta, B. Fragomeni, Y. Masuda, I. Pocrnic, I. Aguilar, J.K. Bertrand,
D. Moser, I. Misztal

University of Georgia

July, 2016

ssGBLUP theory vs. practice

Theory

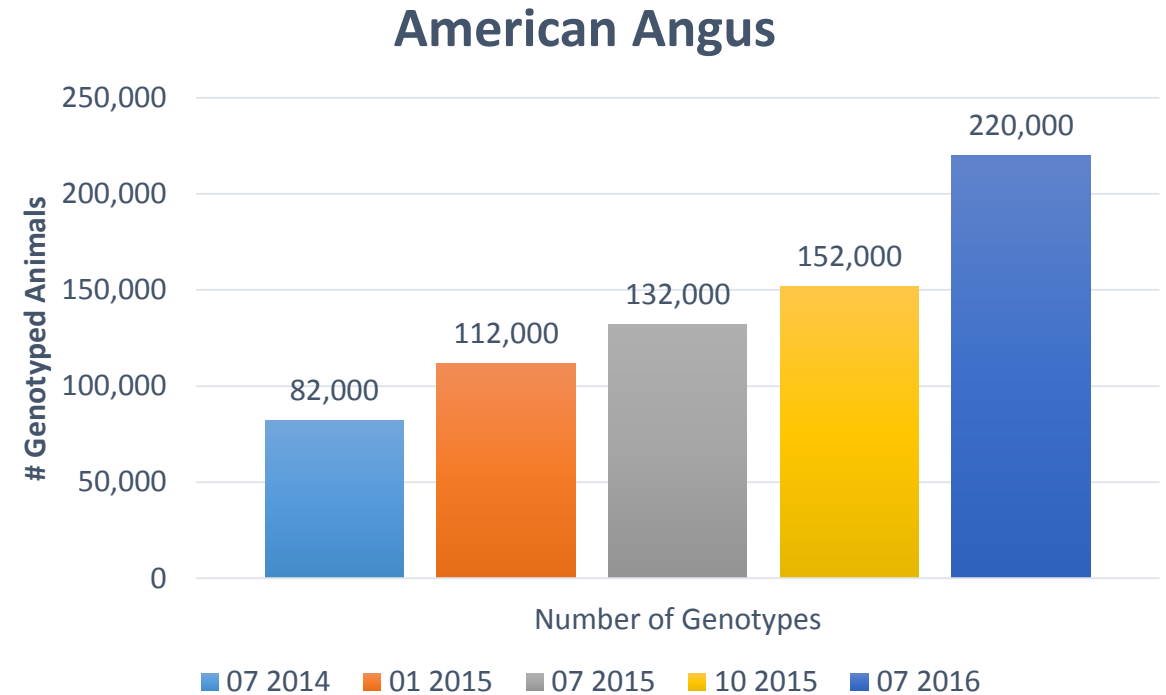
- Studies on several livestock species
- Simple

Practice

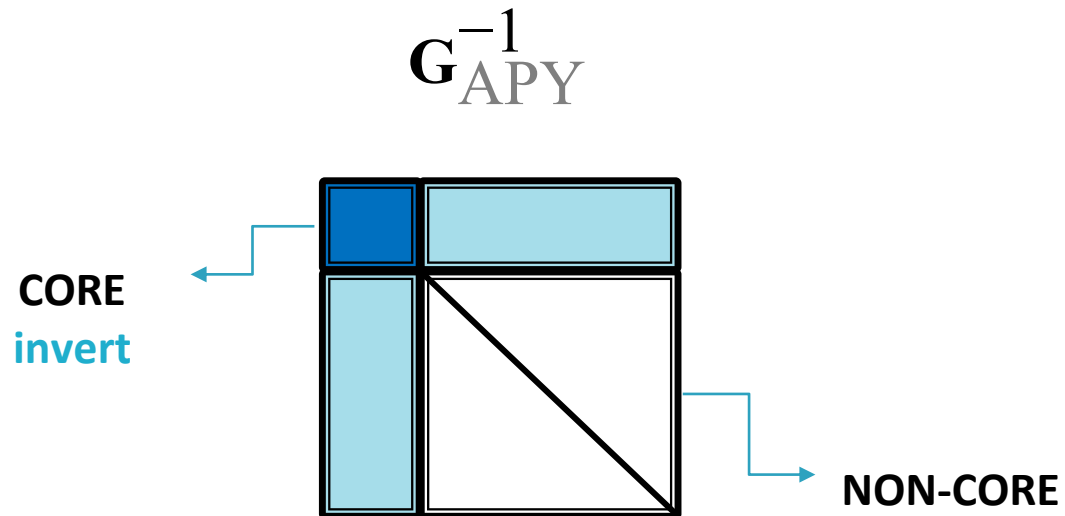
- Full capability
- Implementation for American Angus
- Challenges and problems

Angus Data

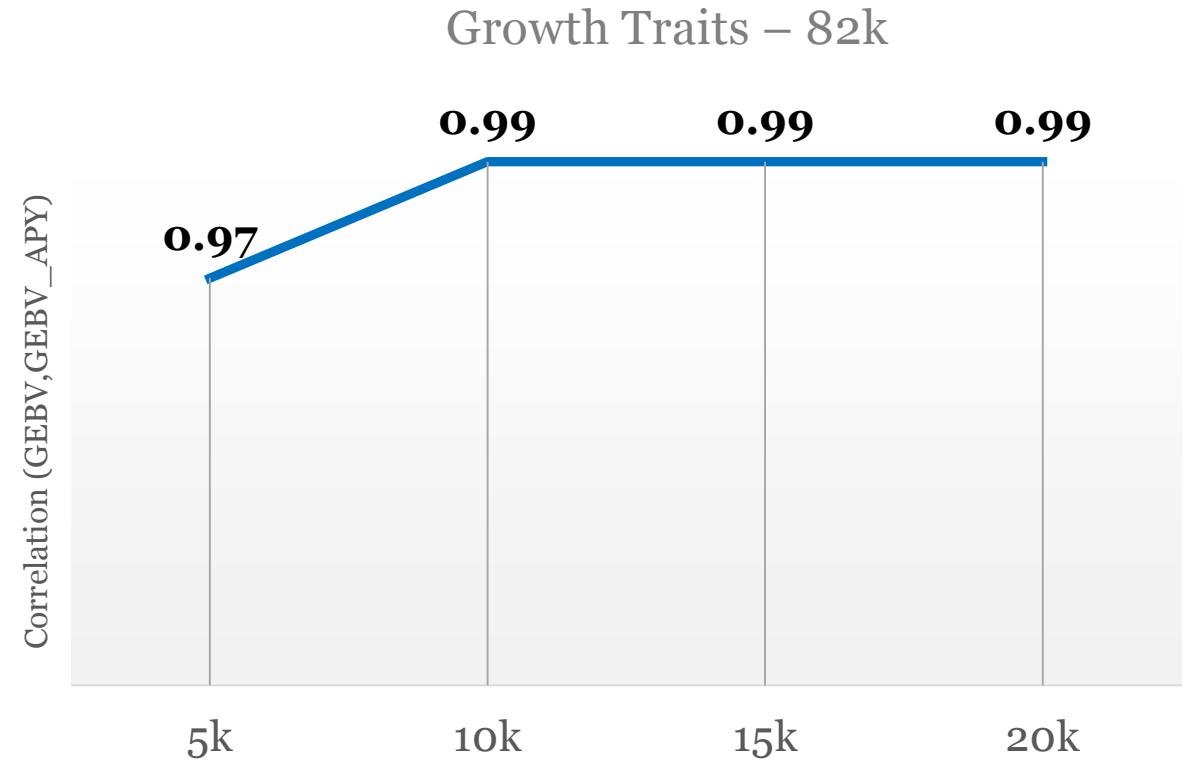
- 220,000 genotyped animals
- 9.7M pedigree
- BW, WW, PWG
- SC
- CE
- Docility
- Heifer pregnancy
- Yearling height
- Mature weight, height
- Carcass weight, marbling, ribeye area, fat thickness
- Dry matter intake



ssGBLUP - lots of genotyped animals



Misztal et al., 2014

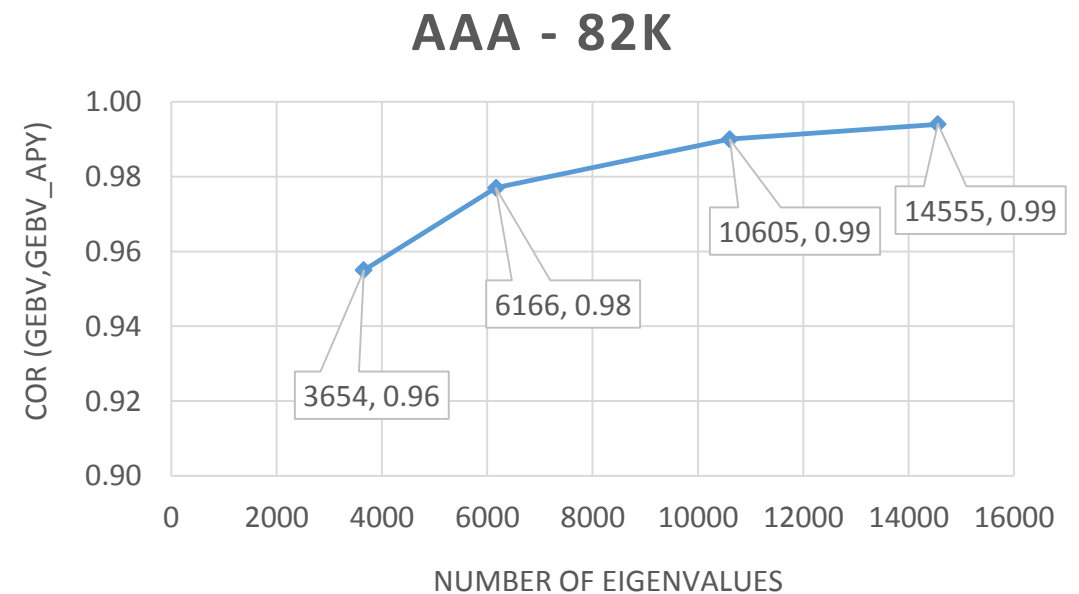
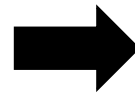
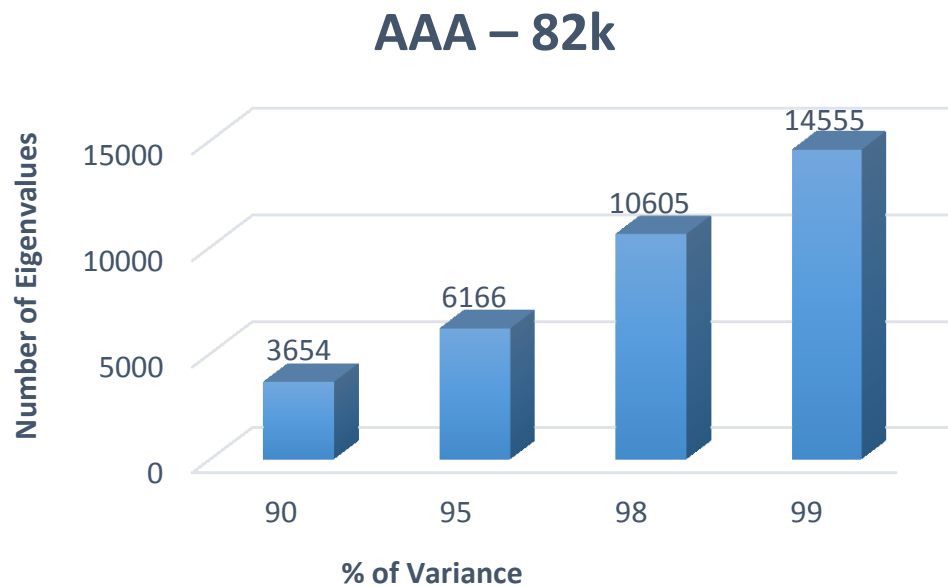


CORE animals randomly sampled from genotyped population

ssGBLUP - lots of genotyped animals

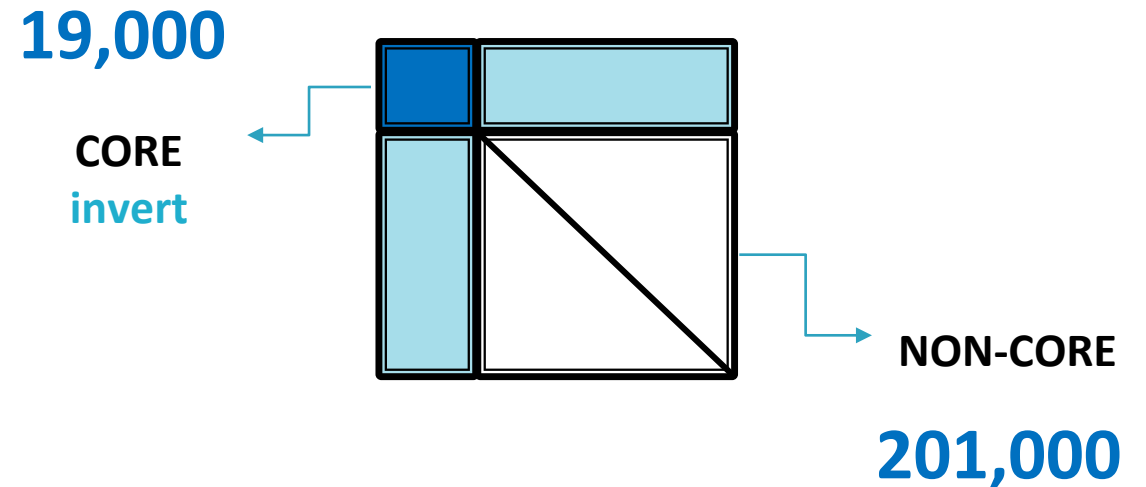
- How to choose number of core animals?
 - N_e , M_e , ESM, Eigen of G
 - Limited dimensionality

Pocrnic et al., 2016
Misztal, 2016



ssGBLUP – Set core animals for evaluation

- 220,000 genotyped animals
- 99% eigen
- Core = 19,000 high accuracy
- Keep core constant 1 year
- Add extra core
- Genomic set up = 30min



Growth Traits with external EBV

- External EBV from ~10k Red Angus
- 220,000 genotyped animals
 - 19,000 core
 - 201,000 non-core
- 9.7M pedigree
- 7.4M BW
- 8.1M WW
- 4M PWG
- Maternal effect for BW & WW

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

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

Adapted from Legarra et al., 2007

Computing time

BLUP = 8h

ssGBLUP = 12h

Calving Ease is categorical

- 220,000 genotyped animals
 - 19,000 core
 - 201,000 non-core
- 8.7M Pedigree
- 1.4M CE
 - 91% easy
 - 9% difficult
- 2-trait BW-CE linear-threshold
- BLUP = 12h
- **ssGBLUP = 4.5 days**

Calving Ease is categorical

- Cblup90iod2: 2 nested rounds

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

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	-
1	100	0.9	0.1	81	43	0.999
2	100	0.85	0.15	62	32	0.999
3	200	0.9	0.1	24	25	0.999
4	200	0.85	0.15	19	19	0.999

Working on OMP – 30% faster

Interim GEBV

- Weekly evaluation
- New genotypes daily
- 220,000 genotyped animals
 - 19,000 core
 - 201,000 non-core

- SNP effect

$$\hat{u} = \mathbf{DZ}' \mathbf{G}^{-1} (\mathbf{DGV})$$

 $d_i = \text{SNP weight} = \mathbf{I}$

- $\text{GEBV}_i = \mathbf{Z}\hat{u}$
- $\text{COR} (\text{GEBV}_{\text{I_CORE}}, \text{GEBV}_{\text{I_50k}}) = 0.98$

Accuracy of GEBV

- GEBV published with accuracy
- Measurement of precision

- Large datasets
 - Impossible to invert

Traditional

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



$$\text{Diag}(\mathbf{C}^{\mathbf{ZZ}+}) = \text{PEV}$$



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

- d_i^r and d_i^p are approximated

(Misztal and Wiggans, 1988)

- Accuracy = $1 - LHS^{-1}$

Accuracy of GEBV

Genomic

$$\left\{ \mathbf{Z}'\mathbf{Z} + \lambda \mathbf{A}^{-1} + \lambda \begin{bmatrix} 0 & \\ & \mathbf{G}^{-1} \\ 0 & -\mathbf{A}_{22}^{-1} \end{bmatrix} \right\} \longrightarrow LHS_{uu}^{ii} = 1 / (\lambda + d_i^r + d_i^p + d_i^g)$$

d_i^r d_i^p d_i^g

How to approximate d_i^g ?

Accuracy of GEBV

- Approximation 1

- h^2
- # effective SNP
- Average contribution from **G**
- Accuracy of PA

- Approximation 2

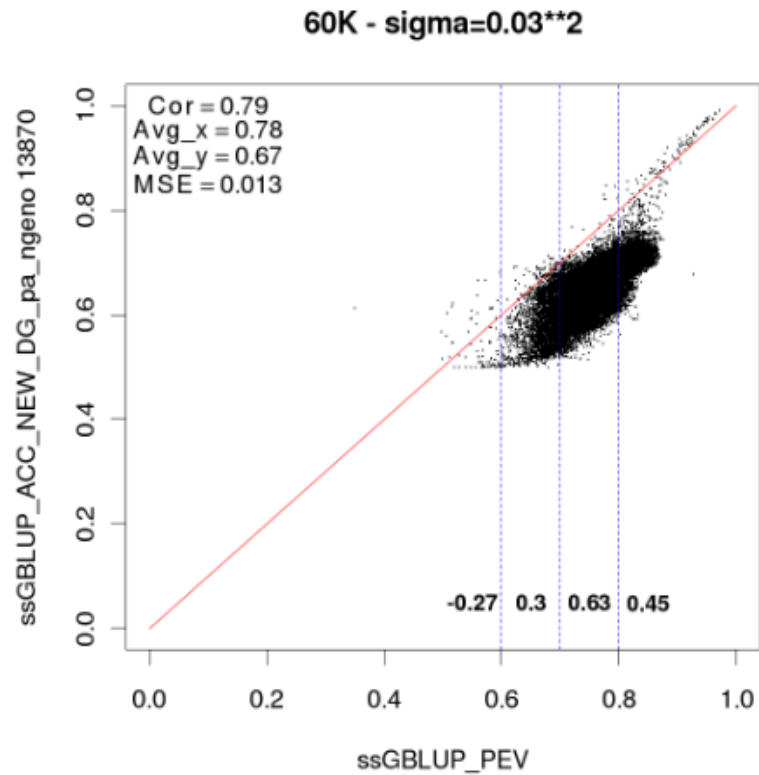
- d_i^r and d_i^p
- **G**
 - Minimum relationship (δ)
- Accuracy of PA

$$d_i^g = 1 + h^2 * ESM * \overline{G - A_{22}} * Acc_{PA}$$

$$d_i^g = \lambda \left[\frac{\sum_{i \neq j} Acc_{PA}}{count(\delta)} + \approx \frac{\sum_{i \neq j} \{g_{ij}(d_i^r + d_i^p)\}}{count(\delta)} \right]$$

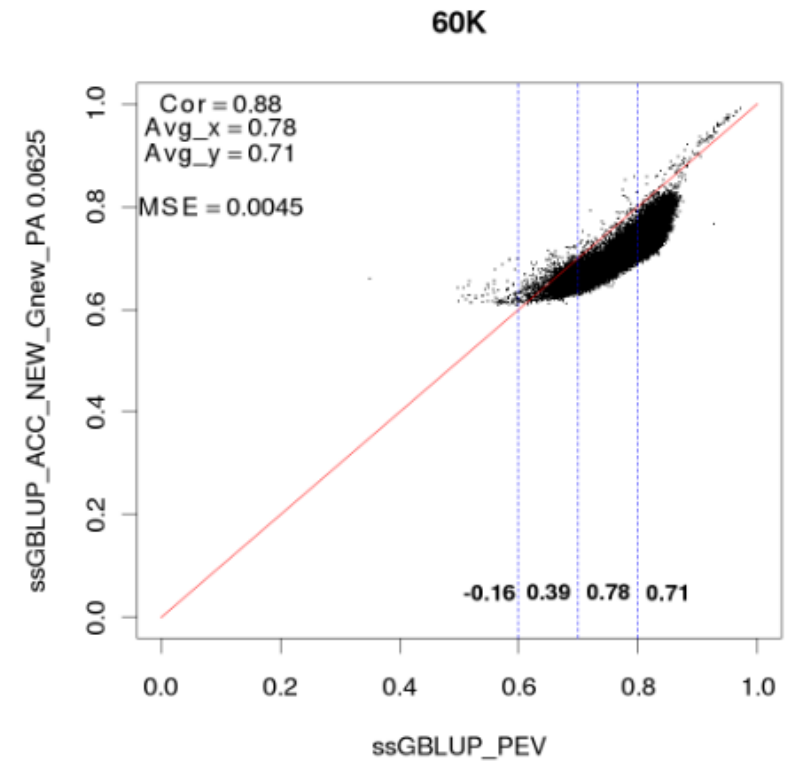
Accuracy of GEBV

Approximation 1



0'12"

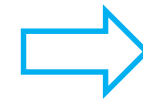
Approximation 2



3'24"

Considerations

Several problems and challenges



All solved

- multi-trait, categorical, maternal
- external info
- interim GEBV
- accuracy of GEBV

ssGBLUP ready for national Beef cattle evaluation

- Angus in 2016

Acknowledgements

