

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)



(2008 - now)

5

Initial tests of ssGBLUP for Angus

Number of Genotyped Animals



Ability to predict future performance

<u>2014</u>

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

<u>2017</u>

•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



USMARC comparisons of ssGBLUP x multistep



USMARC Predictive Ability



Correlation with MARC EBV for 143 bulls

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 $H^{-1}=A^{-1}+\begin{bmatrix} 0 & 0 \\ 0 & G^{-1}-A_{22}^{-1} \end{bmatrix}$
 - > 700Gb RAM
- APY ssGBLUP
 - Borrowed from algorithm to construct A⁻¹
 - Core and Non-core

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

APY ssGBLUP in 2014



APY ssGBLUP in 2014



Regular inversion = 213 min

230 Gb

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

How to choose the number of core in APY?

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

Pocrnic et al., 2016 Misztal, 2016



14

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

â

Zâ $\widehat{\boldsymbol{a}}_{\mathbf{G}} = \lambda \mathbf{D} \, \mathbf{Z}' \mathbf{G}^{-1} \widehat{\boldsymbol{u}}$ G^{-1} G^{-1} 0.93 0.99 1.00 1.00 $\widehat{\boldsymbol{a}}_{\mathbf{G}_{APY}^{-1}} = \lambda \mathbf{D} \mathbf{Z}' \mathbf{G}_{APY}^{-1} \widehat{\boldsymbol{u}}_{APY}$ 22 0.1 $\mathbf{G}_{\mathrm{APY}}^{-1}$ $\mathbf{G}_{\mathrm{APY}}^{-1}$ 0.93 0.98 0.0 0.1 20 9 $\widehat{\boldsymbol{a}}_{\mathbf{G}_{cc}^{-1}} = \lambda \mathbf{D} \, \mathbf{Z}' \mathbf{G}_{cc}^{-1} \, \widehat{\boldsymbol{u}}_{\mathrm{APY}}$ \mathbf{G}_{cc}^{-1} \mathbf{G}_{cc}^{-1}

0.0

-0.2 -0.1

-0.2

-0.1 0.0

0.1

50

Additional features in ssGBLUP

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

Indirect predictions for young animals

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

$$\mathbf{GEBV_y} = w_1 \mathbf{PA} + w_2 \mathbf{DGV} - w_5 \mathbf{PP}$$

$$\mathbf{GEBV_y} \approx \mathbf{DGV}$$

$$\mathbf{GEBV_y} \approx \mathbf{DGV}$$

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

$$\mathbf{Lourenco et al., 2015}$$

Problem with Indirect predictions

COR(GEBV,DGV) > 0.99

Lourenco et al., 2015

 $Avg(GEBV) \approx 100 \implies Avg(DGV) \approx 0$

• Base of SSGBLUP: modelled as a mean in genotyped animals

•
$$p(\boldsymbol{u}_g) = N(\boldsymbol{1}\mu, \mathbf{G})$$

Vitezica et al. (2011)

μ = (Pedigree base) – (Genomic base)

Correcting for bias of indirect predictions



1) Omega = 0.7 indicates inflation in GEBV



Solution: adding inbreeding for A⁻¹ removed inflation in GEBV Omega = 1.0

2) Inclusion of external EBV into growth evaluation

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

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

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

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
	pcg rounds	alpha	beta	rounus	nours	genomic
traditional	40	-		60	12	-
genomic	40	0.9	0.1	488	108	-

4) Accuracy of GEBV

- Large datasets
 - Impossible to invert

• d_i^r and d_i^p are approximated

(Misztal and Wiggans, 1988)

• Accuracy = $1 - LHS^{-1}$

Diag(C^{zz+}) = PEV
$$\longrightarrow$$
 $LHS_{uu}^{ii} = \frac{1}{(\lambda + d_i^r + d_i^p)}$

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

25

4) Accuracy of GEBV

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

$$\begin{cases} \mathbf{Z}'\mathbf{Z} + \lambda \mathbf{A}^{-1} + \lambda \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \\ \downarrow & \downarrow \\ d_i^r & d_i^p & \downarrow \\ \mathbf{d}_i^g &= var_ratio * [\overline{Rel} + (1 - g_{ii}) + zeta * \overline{Rel - Rel_{PA}}] \end{cases}$$



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

- 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







