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

Kachman, 2008
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

Number of Genotyped Animals
Ability to predict future performance

<table>
<thead>
<tr>
<th>2014</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 8M animals in pedigree</td>
<td>• 10M animals in pedigree</td>
</tr>
<tr>
<td>• 6M BW and WW</td>
<td>• 8M BW and WW</td>
</tr>
<tr>
<td>• 3.4M PWG</td>
<td>• 4.2M PWG</td>
</tr>
<tr>
<td>• 52k genotyped animals</td>
<td>• 335k genotyped animals</td>
</tr>
<tr>
<td>• 18.7k born in 2013</td>
<td>• 18.7k born in 2016</td>
</tr>
</tbody>
</table>

Predictive ability direct $= \text{COR}(Y\text{\_adj}, \text{GEBV})$

Predictive ability maternal $= \text{COR}(Y\text{\_adj}, \text{total\_maternal\_GEBV})$
Ability to predict future performance

Average Gain

Direct
2014 = 25%
2017 = 36%

Maternal
2014 = 8%
2017 = 10%
USMARC comparisons of ssGBLUP x multistep

Kuehn et al., 2017
Genetic trends for carcass traits

![Marbling Trend](chart1)

![Carcass Weight Trend](chart2)
Increasing number of genotyped animals

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

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

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

\[
G_{APY}^{-1} = \begin{bmatrix} G_{cc}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -G_{cc}^{-1}G_{cn} \\ I \end{bmatrix} M_{nn}^{-1} \begin{bmatrix} -G_{nc}G_{cc}^{-1} & 0 \\ I & 0 \end{bmatrix} M_{nn} = g_{ii} - g_{ic}G_{cc}^{-1}g_{ci}
\]

Misztal et al., 2014
APY ssGBLUP in 2014

<table>
<thead>
<tr>
<th>Animal ID</th>
<th>Core Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA-8781383</td>
<td>1</td>
</tr>
<tr>
<td>AAA-8974207</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9459638</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9604587</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9651921</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9891499</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9906195</td>
<td>1</td>
</tr>
<tr>
<td>AAA-9921007</td>
<td>1</td>
</tr>
<tr>
<td>AAA-10383408</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10648572</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10710606</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10848637</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10931840</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10971449</td>
<td>2</td>
</tr>
<tr>
<td>AAA-10971485</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11118769</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11142393</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11223766</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11308904</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11356568</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11367940</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11373742</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11391800</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11392490</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11447335</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11468795</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11520398</td>
<td>2</td>
</tr>
<tr>
<td>AAA-11567326</td>
<td>2</td>
</tr>
</tbody>
</table>

How to choose core animals?

APY $G^{-1}$

$G^{-1}$
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?

- $N_e$, $M_e$, 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

\[ \hat{a}_G = \lambda D Z' G^{-1} \hat{u} \]

\[ \hat{a}_{G_{APY}} = \lambda D Z' G^{-1}_{APY} \hat{u}_{APY} \]

\[ \hat{a}_{G_{CC}} = \lambda D Z' G^{-1}_{CC} \hat{u}_{APY} \]
Additional features in ssGBLUP

• Interim evaluations
  • Indirect predictions
  • Quick evaluations between official runs
  • Should be comparable to GEBV
Indirect predictions for young animals

\[
\{ \mathbf{W}' \mathbf{W} + \alpha \mathbf{A}^{-1} + \alpha \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \} \widehat{\mathbf{u}} = \mathbf{W}' \mathbf{y}
\]

\[
\text{GEBV}_y = w_1 \text{PA} + w_4 \text{DGV} - w_5 \text{PP}
\]

\[
\text{GEBV}_y \approx \text{DGV}
\]

\[
\text{GEBV} = w_1 \text{PA} + w_2 \text{YD} + w_3 \text{PC} + w_4 \text{DGV} - w_5 \text{PP}
\]

Lourenco et al., 2015
Problem with Indirect predictions

\[ \text{COR}(\text{GEBV}, \text{DGV}) > 0.99 \]

\[ \text{Avg}(\text{GEBV}) \approx 100 \quad \Rightarrow \quad \text{Avg}(\text{DGV}) \approx 0 \]

- Base of SSGBLUP: modelled as a mean in genotyped animals
  - \( p(u_g) = N(1\mu, G) \)
  - \( \mu = (\text{Pedigree base}) - (\text{Genomic base}) \)

Lourenco et al., 2015
Vitezica et al. (2011)
Correcting for bias of indirect predictions

\[ E(u|a) = \mu + Z\hat{a} \approx DGV = \overline{GEBV} + Z\hat{a} \]

Lourenco et al., 2018
Issues in the implementation of ssGBLUP for Angus

1) Omega = 0.7 indicates inflation in GEBV

\[ H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau G^{-1} - \omega A_{22}^{-1} \end{bmatrix} \]

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

\[ H^{-1} = \begin{bmatrix} H_{EE}^{EE} + T^{-1} - H_{EE}^{-1} & H_{EI}^{EE} \\ H_{IE}^{EE} & H_{II} \end{bmatrix} \]

- \( E \) = 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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description of parameters</th>
<th>genomic pcg rounds</th>
<th>alpha</th>
<th>beta</th>
<th>rounds</th>
<th>hours</th>
<th>correlation with genomic</th>
</tr>
</thead>
<tbody>
<tr>
<td>traditional</td>
<td></td>
<td>40</td>
<td></td>
<td>-</td>
<td>60</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>genomic</td>
<td></td>
<td>40</td>
<td>0.9</td>
<td>0.1</td>
<td>488</td>
<td>108</td>
<td>-</td>
</tr>
</tbody>
</table>
Issues in the implementation of ssGBLUP for Angus

4) Accuracy of GEBV

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

\[\text{Diag}(C^{zz+}) = \text{PEV}\]

- Large datasets
  - Impossible to invert

- \(d_i^r\) and \(d_i^p\) are approximated
  (Misztal and Wiggans, 1988)

- Accuracy = \(1 - LHS^{-1}\)

\[LHS_{uu}^{ii} = \frac{1}{(\lambda + d_i^r + d_i^p)}\]
Issues in the implementation of ssGBLUP for Angus

4) Accuracy of GEBV

\[
\left\{Z'Z + \lambda A^{-1} + \lambda \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A22^{-1} \end{bmatrix}\right\}
\]

\[
\downarrow \quad \downarrow \quad \downarrow \quad \downarrow
\]

\[d_i^r \quad d_i^p \quad d_i^g\]

\[
LHS_{uu}^{ii} = \frac{1}{(\lambda + d_i^r + d_i^p + d_i^g)}
\]

\[
d_i^g = \text{var}_\text{ratio} \times \left[\overline{\text{Rel}} + (1 - g_{ii}) + \text{zeta} \times \overline{\text{Rel}} - \overline{\text{Rel}_{PA}}\right]
\]
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
  \[ \hat{a}_{G_{cc}}^{-1} = \lambda D Z' G_{cc}^{-1} \hat{u}_{APY} \]

• 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