GEBV with GMACE correlations in the Czech Republic

Přibyl J.¹, Zavadilová L.¹, Pešek P.¹, Šplíchal J.², Bauer J.¹, Vostrý L.¹, Motyčka J.³, Fulínová D.², Čermák V.², Růžička Z.²,

¹Institute of animal science, Czech Republic
² Czech Moravian Breeding Corporation, Czech Republic
³ Holstein Cattle Breeders Association of the Czech Republic, Czech Republic
2011-2014 – several genomic approaches compared

• RRBLUP
• GBLUP
• Blending ssGBLUP
• ssGBLUP
• ssGBULP with combination of domestic production + Interbull DRPs
Correlation of GEBV (EBV) prediction 2008 to daughters results (DYD) 2012 for 140 young bulls

- 969,269 1st lactations 1991–2008, 1,762,905 in pedigree
- 1,185,225 1st lactations 1991–2012, in pedigree 1,958,139
- 98,037 INTERBULL EBVs through year 2008
- 1,605 genotyped bulls (1259 already proven in 2008)
Correlation of GEBV (EBV) prediction 2008 to daughters results (DYD) 2012 for 140 young bulls

<table>
<thead>
<tr>
<th>Method</th>
<th>Correlation</th>
<th>Validated reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>D – BLUP</td>
<td>0.47</td>
<td>0.29</td>
</tr>
<tr>
<td>D – ssGBLUP</td>
<td>0.61</td>
<td>0.48</td>
</tr>
<tr>
<td>I – RRBLUP</td>
<td>0.61</td>
<td>0.48</td>
</tr>
<tr>
<td>I – GBLUP</td>
<td>0.61</td>
<td>0.47</td>
</tr>
<tr>
<td>I – blending BLUP</td>
<td>0.53</td>
<td>0.36</td>
</tr>
<tr>
<td>I – blending ssGBLUP</td>
<td>0.65</td>
<td>0.54</td>
</tr>
<tr>
<td>D + I – BLUP</td>
<td>0.51</td>
<td>0.34</td>
</tr>
<tr>
<td>D + I – ssGBLUP</td>
<td>0.64</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Correlations of GEBVs (EBVs) predicted using different approaches to GMACE

- 1,257,462 1st lactations 1991-2013, 2,314,856 animals in pedigree
- 19,435,367 test-day records of 1,086,267 cows (calved 1995-2013), 2,142,354 animals in pedigree
- 112,880 INTERBULL EBVs
- 2,627 genotyped bulls (2,189 already proven)
Correlations of GEBVs (EBVs) predicted using different approaches to GMACE

96 young bulls not proven in the Czech Republic
Our choice:  
**Single step GBLUP**

Benefits:

- more accurate yet much simpler than multi-step methods
- avoiding bias in GEBV prediction
- increase of breeding value prediction reliability for ungenotyped animals
Correlations for production traits (RRTDM)

0.71 0.67 0.76 0.64 0.75

Milk kg  Protein kg  Protein %  Fat kg  Fat %  SCC

158 young bulls not proven in the Czech Republic
Correlations of GEBV to MACE for production traits (RRTDM)

73 bulls proven abroad but not in the Czech Republic
Correlations with GMACE for linear type traits – Body traits

43 young bulls not proven in the Czech Republic
Correlations with GMACE for linear type traits – Feet & Legs

43 young bulls not proven in the Czech Republic
Correlations with GMACE for linear type traits - Udder

43 young bulls not proven in the Czech Republic
Correlations with GMACE for linear type traits – overall characteristics

Body Condition Score: 72
Overall udder score: 54
Overall feet & leg score: 63
Overall conformation score: 50

43 young bulls not proven in the Czech Republic
Conclusion

Single-step Genomic Evaluation
as the best approach for the Czech Republic

• The stronger correlation with GMACE compared to multi-step approaches

• Medium to high correlations with GMACE for production traits as well as linear type traits
Thank You for Your attention
Additional notes
Weighted analysis

\[
\text{ERC} = \delta \left( \frac{\text{rel}}{1-\text{rel}} \right)
\]

\[
\delta = \left( \frac{1-h^2}{h^2} \right)
\]
Weighted analysis

Ridge Regression

\[ \text{DRP}_j = \mu + \sum \delta_i \cdot g_{ij} + e_j \]

Fixed effects
\( \mu \) – common constant

Random effects
\( g_{ij} \) – genotype of j-th bull in i-th locus
\( \delta_i \) – regression coefficient
\( e_j \) – residual

Heritability \( (h^2) = 0.25 \); Weight = ERC
**Weighted analysis**

**GBLUP**

\[ \text{DRP}_j = \mu + a_n_j + e_j \]

**Fixed effects**
- \( \mu \) – common constant

**Random effects**
- \( a_n_j \) – animal
- \( e_j \) – residual

Heritability \((h^2) = 0.25\); Weight = ERC
**BLUP / ssGBLUP, lactation model**

\[
milk_{ijkl} = HYS_i + \beta_1 \cdot ca_k + \beta_2 \cdot ca_k^2 + \beta_3 \cdot dol + \beta_4 \cdot dol^2 + an_j + e_{ijkl}
\]

**Fixed effects**
- **HYS** – contemporary group
- **\(\beta\)** – regression coefficients
- **\(ca_k; ca_k^2\)** – curvilinear regression on calving age
- **\(dol; dol^2\)** – curvilinear regression on days open

**Random effects**
- **\(an_j\)** – animal
- **\(e_{ijkl}\)** – residual

**Heritability** \((h^2) = 0.25;\)
ssGBLUP, RR-TDM, 3 lactations

\[ y_{ijn} = HTD_{in} + f_{fg,n} + f_{pe,n} + f_{an,n} + e_{ijn}, \]

\( y_{ijn} \) = test-day record of milk yield of cow in lactation \( n<1,2,3> \);

HTD\(_{in}\) = herd-test-day contemporary group \( i \) within a herd in lactation \( n \) (fixed effect);

\( f_{fg,n} \) = average LP of lactation curve according to groups of cows within management classes of systematic environment (Zavadilová et al., 2005b) (fixed effect);

\( f_{pe,n} \) = permanent environmental within lactation LP of lactation curve of cows, random effect with covariance matrix (Zavadilová et al., 2005a);

\( f_{an,n} \) = genetic within lactation LP of lactation curve of animal, random effect with covariance matrix;

\( e_{ijn} \) = random residual of test day records within lactation \( n \), reflecting changes of variability along the course of lactation.
Relationship

Ridge Regression...I
BLUP – AM ............A
GBLUP ..................G (VanRaden 2008)
ssGBLUP ..................H (Legarra et al., 2009)

20% $A_{22}$  G  80%

$G$ – normalised (aver. diag. = 1)  
(shifted (aver. $G = \text{aver. } A_{22}$))  
(Forni et al., 2011)

(Vitezica et al., 2011)
SNP editing:

- MAF,
- G-score,
- No. of loci per bull,
- No. of bulls per locus,
- Big error of prediction of old bulls in training set,
- Big discrepancy of relationship $A_{22} \times G$,
- Proportion of H.