Integration / blending of external information into genetic / genomic evaluations

J. Vandenplas\textsuperscript{1,2}, N. Gengler\textsuperscript{1}

\textsuperscript{1}University of Liège, Gembloux Agro-Bio Tech, Belgium
\textsuperscript{2}National Fund for Scientific Research, Brussels, Belgium
Introduction

- Most reliable EBV if estimated from all available sources
- Most situations
  - Multiple sources (e.g., dairy breeds)
    - Traditional genetic evaluations
    - International second step (Interbull)
  ➡ Animals with few (or no) local data: low accuracy
  - Development of genomic selection
    ➡ New genomic information sources

➡ Strategies for integration / blending of those multiple sources of external information
Introduction

• Single-step genomic evaluation (ssGBLUP)
  ➔ Genomic, pedigree and phenotypic information considered simultaneously

• Current limitation
  – Only available local records used

  ➔ Disadvantage compared to multi-steps methods using e.g. MACE-EBV as a way to access to external phenotypic information
Current situation

System A
- Performances recording
- Exchange of genetic material (e.g. frozen semen and embryos)

Breeding associations
- Phenotypic, pedigree and genomic information of local animals

Genetic/genomic evaluations
- Local EBV

System B
- External information of external animals

Breeders

System A → Breeding associations → Genetic/genomic evaluations
System B

External information of external animals

Exchange of genetic material (e.g. frozen semen and embryos)
Possible solution

- Performances recording
- Breeding associations
- Bayesian genetic/genomic evaluations
- Phenotypic, pedigree and genomic information of local animals
- Exchange of genetic material (e.g. frozen semen and embryos)

System A

System B

External information of external animals

Breeders
Aim

To integrate/blend multiple *a priori* known external information into a local genetic/genomic evaluation

→ Using a Bayesian approach

- Based on
  - Legarra et al. (2007)
  - Quaas and Zhang (2006)
  - Vandenplas and Gengler (2012a, 2012b)
  - Vandenplas et al. (2012a, 2012b)
External information

• Most situations
  – EBV and REL from other genetic evaluations
  – Information not taken into account by a local (ssG)BLUP
    ➔ Not an external genomic evaluation
    ➔ No double-counting between local and external evaluations
  – Only available for some animals
    • Having, or not, phenotypic information in the local (ssG)BLUP
    • Present in the pedigree of the local (ssG)BLUP
Regular (ssG)BLUP

- Mixed model equations

\[
\begin{bmatrix}
X'R^{-1}X & X'R^{-1}Z \\
Z'R^{-1}X & Z'R^{-1}Z + G^{*-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{u}_L
\end{bmatrix}
= \begin{bmatrix}
X'R^{-1}y_L \\
Z'R^{-1}y_L
\end{bmatrix}
\]

- \(G^{*-1}\) : Inverse of (combined genomic-) pedigree based (co)variances matrix
- \(y_L\) : vector of local observations
- \(\hat{\beta}_L\) : vector of estimated local fixed effects
- \(\hat{u}_L\) : vector of estimated local (G)EBV
- \(\rho(u_L) = MVN(0, G^*)\)
Integration of one source

• Assumption: Conditional distribution of $u_L$

$$
\rho(u_L|y_E) = \text{MVN}(\hat{u}_E, D)
$$

- $y_E$: unavailable vector of external observations
- $\hat{u}_E$: vector of external EBV
- $D$: prediction error (co)variances matrix of $\hat{u}_E$

• Issue: only available for some animals

$\Rightarrow \hat{u}_E$ and $D$: (partially) unknown
Integration of one source

- Estimation of $\hat{u}_E$
  - **Available**: External EBV of external animals ($\hat{u}^*_E$)
  - **Local animals**: prediction of external EBV ($\hat{u}_{EL}$)

\[
p(\hat{u}_{EL} | \hat{u}^*_E) = MVN(G_{LE}G_{EE}^{-1}\hat{u}^*_E, (G_{EE}^{-1})^{-1})
\]

$\hat{u}_E = \begin{bmatrix} \hat{u}_{EL} \\ \hat{u}^*_E \end{bmatrix}$

- Predicted external EBV
- Available external EBV

**Correct propagation of external information**
Integration of one source

- Estimation of $D$

$$D^{-1} = G^{-1} + \Lambda$$

- $G^{-1} = A^{-1} \otimes G_0^{-1}$: Inverse of genetic (co)variances matrix of $\hat{u}_E$

- $\Lambda = \text{block diag}(\Delta_i G_0^{-1} \Delta_i)$; $i = 1, \ldots, n$ animals

\[
\begin{cases}
\text{For external animals}: \Delta_i = \text{diag} (\sqrt{\text{REL}_{ij} / (1 - \text{REL}_{ij})}); j = 1, \ldots, t \text{ traits} \\
\text{For animals with only local information}: \Delta_i = 0
\end{cases}
\]
Integration of one source

- Integration of external EBV and REL

\[
\begin{bmatrix}
X' R^{-1} X & X' R^{-1} Z \\
Z' R^{-1} X & Z' R^{-1} Z + G^{*-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{u}_L
\end{bmatrix}
= 
\begin{bmatrix}
X' R^{-1} y_L \\
Z' R^{-1} y_L
\end{bmatrix}
\]

Least square parts of LHS of the hypothetical BLUP

RHS of the hypothetical BLUP
Integration of one source: simulation

• 100 replicates
• 2 populations
  – ±1000 animals/population
  – 5 generations
  – Random matings / cullings
  – Observations (Van Vleck, 1994)
    • Milk yield for the first lactation
    • Heritability : 0.25
  – Fixed effect
    • Random herd effect within population
Integration of one source: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Bayesian BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td></td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Integration of one source: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Bayesian BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Integration of one source: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Bayesian BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>50 external sires</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
</tr>
</tbody>
</table>
Integration of one source: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Bayesian BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>50 external sires</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Comparison with joint BLUP

- EBV of the 50 external sires used locally
Comparison with joint BLUP

- Rank correlations (r) and mean squared errors (MSE) expressed as a percentage of the local MSE

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Local animals</th>
<th>External sires</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r ± SD</td>
<td>MSE ± SD</td>
</tr>
<tr>
<td>Without external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local BLUP</td>
<td>0.95 ± 0.02</td>
<td>100.00 ± 33.5</td>
</tr>
<tr>
<td>With external information</td>
<td>&gt;0.99 ± 0.000</td>
<td>0.79 ± 0.52</td>
</tr>
</tbody>
</table>

→ Rankings more similar to those of the joint BLUP
Integration of multiple sources

Performances recording → Breeding associations → Bayesian genetic/genomic evaluations

- Phenotypic, pedigree and genomic information of local animals
- External evaluations of external animals

Exchange of genetic material (e.g. frozen semen and embryos)

- System A
- System B
- System C
- System D

Bayesian EBV → Breeders
Integration of multiple sources

- Assumption: *A priori* known information of $\mathbf{u}_L$
  - $n$ sources of external information:
    - $n$ vector of external EBV
      \[ \hat{\mathbf{u}}_{E_1}, \ldots, \hat{\mathbf{u}}_{E_i}, \ldots, \hat{\mathbf{u}}_{E_n} \]
    - $n$ prediction error (co)variances matrices
      \[ \mathbf{D}_{E_1}, \ldots, \mathbf{D}_{E_i}, \ldots, \mathbf{D}_{E_n} \]
  - Issue: only available for some animals

$\Rightarrow \hat{\mathbf{u}}_{E_i}$ and $\mathbf{D}_{E_i}$: (partially) unknown
Integration of multiple sources

- For each source $i$: Estimation of $\hat{u}_{E_i}$
  - Available: External EBV of external animals ($\hat{u}_{EE_i}^*$)
  - Local animals: prediction of external EBV ($\hat{u}_{EL_i}$)

$\rho(\hat{u}_{EL_i} \mid \hat{u}_{EE_i}^*) = MVN\left(G_{LE_i} G_{EE_i}^{-1} \hat{u}_{EE_i}^* , (G_{EE_i}^{EE})^{-1}\right)$

$\hat{u}_{E_i} = \begin{bmatrix} \hat{u}_{EL_i} \\ \hat{u}_{EE_i}^* \end{bmatrix}$

$\Rightarrow$ Correct propagation of external information

Predicted external EBV
Available external EBV
Integration of multiple sources

- For each source $i$: Estimation of $D_{E_i}$

$$D_{E_i}^{-1} = G^{-1} + \Lambda_{E_i}$$

$G^{-1} = A^{-1} \otimes G_0^{-1}$: Inverse of genetic (co)variances matrix of $\hat{u}_{E_i}$

$\Lambda_{E_i} = \text{block diag}(\Delta_j G_0^{-1} \Delta_j); i = 1, ..., n \text{ sources}; j = 1, ..., a \text{ animals}$

\[
\begin{cases}
\text{For external animals: } \Delta_j = \text{diag}(\sqrt{REL_{ijk}/(1-REL_{ijk})}); k = 1, ..., t \text{ traits} \\
\text{For animals with only local information: } \Delta_j = 0
\end{cases}
\]
Integration of multiple sources

- Integration of $n$ external information

\[
\begin{bmatrix}
X'\mathbf{R}^{-1}X & X'\mathbf{R}^{-1}Z \\
Z'\mathbf{R}^{-1}X & Z'\mathbf{R}^{-1}Z + \mathbf{G}^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{\mathbf{u}}_L
\end{bmatrix} =
\begin{bmatrix}
X'\mathbf{R}^{-1}y_L \\
Z'\mathbf{R}^{-1}y_L
\end{bmatrix}
\]

Sum of $n$ least square parts of LHS of $n$ hypothetical BLUP of $n$ sources of external EBV

Sum of $n$ RHS of $n$ hypothetical BLUP of $n$ sources of external EBV
Blending of multiple sources

- Blending of \( n \) external information
  - Assumption: no local records in \( y_L \)

\[
\begin{bmatrix}
G^{-1} + \sum_{i=1}^{n} \Lambda_{E_i}
\end{bmatrix} \hat{u}_L = \sum_{i=1}^{n} \left( D_{E_i}^{-1} \hat{u}_{E_i} \right)
\]
Issues

• Some issues arise before the implementation of a Bayesian evaluation like:
  – Compatibility between external and local traits
    • Same expression of external and local traits
      ➔ Need to translate external information (e.g., using conversion equations (e.g., Vandenplas et al., 2012a))
  – Double counting of information
    • Due to relationships among external animals
    • Among external and local information
    • Among multiple external information
Issues

• Double counting due to relationships
  – The animal’s EBV and REL combine contributions:
    • due to own records (if present)
    • due to relationships (Misztal and Wiggans, 1988)

→ Multiple considerations of parts of integrated external
  EBV for related external animals

→ Estimation of contributions due to relationships and due
  to own records through a two-step algorithm (Vandenplas and Gengler, 2012)

→ All $\Lambda_{E_i}$ only depend on contributions due to own records
Issues

• Double counting among external (and local) information
  – Each external information has to be free of local or other external information.

  – Ex.: Integration of MACE-EBV
    • Included local information
    ➔ Estimation of external information free of local information:

\[
D_E^{-1} \hat{u}_E = D_M^{-1} \hat{u}_M - D_L^{-1} \hat{u}_L
\]
Issues

• Integration of MACE-EBV

\[
\begin{bmatrix}
X'_L R^{-1}X_L \\
Z'_L R^{-1}X_L
\end{bmatrix}
\begin{bmatrix}
X'_L R^{-1}Z_L \\
Z'_L R^{-1}Z_L + G^{-1} + D^{-1}_M - D^{-1}_L
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{\mu}_L
\end{bmatrix}
= 
\begin{bmatrix}
X'_L R^{-1}y_L \\
Z'_L R^{-1}y_L + D^{-1}_M \hat{\mu}_M - D^{-1}_L \hat{\mu}_L
\end{bmatrix}
\]
Issues

• Integration of MACE-EBV

\[
\begin{pmatrix}
X'_L R^{-1} X_L & X'_L R^{-1} Z_L \\
Z'_L R^{-1} X_L & Z'_L R^{-1} Z_L + G^{-1} + D_M^{-1} - D_L^{-1}
\end{pmatrix}
\begin{pmatrix}
\hat{\beta}_L \\
\hat{\mu}_L
\end{pmatrix}
= \begin{pmatrix}
X'_L R^{-1} y_L \\
Z'_L R^{-1} y_L + D_M^{-1} \hat{\mu}_M - D_L^{-1} \hat{\mu}_L
\end{pmatrix}
\]

Inverse of (combined genomic -) pedigree based (co)variances matrix
Issues

• Integration of MACE-EBV

\[
\begin{bmatrix}
X' R^{-1} X_L \\
Z' R^{-1} X_L \\
Z_L R^{-1} Y_L + G^{-1} D_M^{-1} D_L^{-1} L \\
Z_L R^{-1} Y_L + D_M^{-1} L_L - D_L^{-1} \hat{u}_L \\
\end{bmatrix} \begin{bmatrix}
\beta_L \\
\hat{u}_L \\
\end{bmatrix} = \begin{bmatrix}
X' R^{-1} Y_L \\
Z_L R^{-1} Y_L + D_M^{-1} L_L - D_L^{-1} \hat{u}_L \\
\end{bmatrix}
\]

Inverse of (combined genomic -) pedigree based (co)variances matrix

Inverse of prediction error (co)variances matrix of MACE-EBV
Issues

• Integration of MACE-EBV

\[
\begin{bmatrix}
X_L' R^{-1}X_L \\
Z_L' R^{-1}X_L \\
X_L' R^{-1}Z_L \\
Z_L' R^{-1}Z_L
\end{bmatrix} + \begin{bmatrix}
G^{-1} \\
D_M^{-1} \\
D_L^{-1}
\end{bmatrix} \begin{bmatrix}
\hat{\beta}_L \\
\hat{u}_L
\end{bmatrix} = \begin{bmatrix}
X_L' R^{-1}y_L \\
Z_L' R^{-1}y_L + D_M^{-1}\hat{u}_M - D_L^{-1}\hat{u}_L
\end{bmatrix}
\]

Inverse of \textit{(combined genomic -) pedigree based (co)variances matrix}

Inverse of \textbf{prediction error (co)variances matrix of MACE-EBV}

Inverse of \textbf{prediction error (co)variances matrix of local EBV}
Issues

• Integration of MACE-EBV

\[
\begin{bmatrix}
X'_L R^{-1}X_L \\
Z'_L R^{-1}X_L
\end{bmatrix}
\begin{bmatrix}
X'_L R^{-1}Z_L \\
Z'_L R^{-1}Z_L
\end{bmatrix}
= 
\begin{bmatrix}
G^{-1} + D_M^{-1} & D_L^{-1} \\
D_M^{-1} & D_L^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{u}_L
\end{bmatrix} = 
\begin{bmatrix}
X'_L R^{-1}y_L \\
Z'_L R^{-1}y_L
\end{bmatrix}
\begin{bmatrix}
D_M^{-1} \hat{u}_M \\
D_L^{-1} \hat{u}_L
\end{bmatrix} - D^{-1}\hat{u}^*_L
\]

Inverse of \( \text{(combined genomic -) pedigree based} \) (co)variances matrix

Inverse of prediction error (co)variances matrix of MACE-EBV

RHS of an hypothetical BLUP of MACE-EBV

Inverse of prediction error (co)variances matrix of local EBV
Issues

• Integration of MACE-EBV

\[
\begin{bmatrix}
X'_L R^{-1} X_L \\
Z'_L R^{-1} X_L \\
Z'_L R^{-1} Z_L + G^{-1} D^{-1}_M D^{-1}_L \\
Z'_L R^{-1} Z_L
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_L \\
\hat{u}_L
\end{bmatrix}
= 
\begin{bmatrix}
X'_L R^{-1} y_L \\
Z'_L R^{-1} y_L + D^{-1}_M \hat{u}_M \\
D^{-1}_L \hat{u}_L
\end{bmatrix}
\]

Inverse of (combined genomic -) pedigree based (co)variances matrix

Inverse of prediction error (co)variances matrix of MACE-EBV

RHS of an hypothetical BLUP of MACE-EBV

RHS of an hypothetical BLUP of local EBV

Inverse of prediction error (co)variances matrix of local EBV
Bayesian ssGBLUP: simulation

- 100 replicates
- 2 populations
  - ±1000 animals/population
  - 5 generations
  - Random matings / cullings
- Observations (Van Vleck, 1994)
  - Milk yield for the first lactation
  - Heritability: 0.25
- Fixed effect
  - Random herd effect within population
Bayesian ssGBLUP: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local ssGBLUP</th>
<th>Bayesian ssGBLUP</th>
<th>Joint ssGBLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bayesian ssGBLUP: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local ssGBLUP</th>
<th>Bayesian ssGBLUP</th>
<th>Joint ssGBLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bayesian ssGBLUP: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local ssGBLUP</th>
<th>Bayesian ssGBLUP</th>
<th>Joint ssGBLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genotypes</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Bayesian ssGBLUP: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local ssGBLUP</th>
<th>Bayesian ssGBLUP</th>
<th>Joint ssGBLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local population</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>+ 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>50 external sires</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>External EBV and REL</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Comparison with joint ssGBLUP

• Rank correlations \((r+SD)\)

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Local animals</th>
<th>External sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local ssGBLUP</td>
<td>0.94 (\pm) 0.02</td>
<td>0.52 (\pm) 0.14</td>
</tr>
<tr>
<td>With external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian ssGBLUP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With double counting</td>
<td>(&gt;0.99) (\pm) 0.000</td>
<td>0.99 (\pm) 0.01</td>
</tr>
<tr>
<td>Without double counting</td>
<td>(&gt;0.99) (\pm) 0.000</td>
<td>0.99 (\pm) 0.01</td>
</tr>
</tbody>
</table>

 Rankings more similar to those of the joint ssGBLUP
### Comparison with joint ssGBLUP

- Mean squared errors (MSE+SD)
  - Expressed as a percentage of the local MSE

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Local animals</th>
<th>External sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local ssGBLUP</td>
<td>100.00 ± 32.8</td>
<td>100.00 ± 29.1</td>
</tr>
<tr>
<td>With external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian ssGBLUP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With double counting</td>
<td>2.37 ± 2.2</td>
<td>1.35 ± 1.2</td>
</tr>
<tr>
<td>Without double counting</td>
<td>1.97 ± 2.1</td>
<td>1.12 ± 1.2</td>
</tr>
</tbody>
</table>

⇒ Importance of double-counting
Blending: simulation

- 100 replicates
- 2 populations
  - ±1000 animals/population
  - 5 generations
  - Random matings / cullings
- Observations (Van Vleck, 1994)
  - Milk yield for the first lactation
  - Heritability: 0.25
- Fixed effect
  - Random herd effect within population
Blending: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Blending BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population + 50 external sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>External information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL (50 external sires)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local EBV and REL (all population)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Blending: simulation

- **Performed evaluations**

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Blending BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedigree</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population + 50 external sire</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>sires used locally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phenotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>External information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL (50 external</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>sires)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local EBV and REL (all population)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Blending: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Blending BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local population + 50 external sires used locally</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local observations</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>External information</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>External EBV and REL (50 external sires)</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Local EBV and REL (all population)</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
### Blending: simulation

- Performed evaluations

<table>
<thead>
<tr>
<th>Information</th>
<th>External BLUP</th>
<th>Local BLUP</th>
<th>Blending BLUP</th>
<th>Joint BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External population</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local population + 50 external sires used locally</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Phenotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External observations</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local observations</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>External information</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External EBV and REL (50 external sires)</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Local EBV and REL (all population)</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Comparison with joint BLUP

• Rank correlations (r+SD)

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Local animals</th>
<th>External sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local BLUP</td>
<td>0.95 ± 0.02</td>
<td>0.54 ± 0.11</td>
</tr>
<tr>
<td>Only external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blending BLUP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With double counting</td>
<td>0.99 ± 0.004</td>
<td>0.97 ± 0.01</td>
</tr>
<tr>
<td>Without double counting</td>
<td>&gt;0.99 ± 0.000</td>
<td>&gt;0.99 ± 0.001</td>
</tr>
</tbody>
</table>

 Rankings more similar to those of the joint BLUP
Comparison with joint BLUP

• Mean squared errors (MSE+SD)
  - Expressed as a percentage of the local MSE

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Local animals</th>
<th>External sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local BLUP</td>
<td>100.00 ± 26.7</td>
<td>100.00 ± 24.5</td>
</tr>
<tr>
<td>Only external information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blending BLUP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With double counting</td>
<td>21.20 ± 6.2</td>
<td>6.83 ± 1.9</td>
</tr>
<tr>
<td>Without double counting</td>
<td>0.48 ± 0.2</td>
<td>0.23 ± 0.1</td>
</tr>
</tbody>
</table>

➡ Importance of double counting
Conclusion

• Bayesian (ssG)BLUP
  ➔ Rankings most similar to those of a joint (ssG)BLUP
  – Importance of double counting among animals

➔ Bayesian procedure
  – Reliable integration/blending of multiple external information
  – Simple modifications of current programs
  – Applicable to multi-traits models
Practical example: jumping horses

- Belgium = crossroad of sport horse populations
- 2008: >50% approved foreign stallions
- Export of sport horses
- Limited Belgian population (~ 4500 foals/year)

→ Limited amount of information for reliable local EBV
Practical example: jumping horses

**Belgium**
- Jumping results recording
- Breeding associations
- Phenotypic and pedigree information
- Bayesian genetic evaluation (double counting)

**The Netherlands**
- 67 approved stallions

**France**
- 98 approved stallions

Admission and use of Dutch and French stallions

- EBV and REL of approved stallions
- Conversion
- Breeders

Bayesian EBV
Practical example: jumping horses

- Spearman rank correlations

<table>
<thead>
<tr>
<th>Group of horses</th>
<th>Nb</th>
<th>Spearman rank correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EBV_{AI} /EBV_{BI}</td>
</tr>
<tr>
<td>All horses</td>
<td>101,382</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>French stallions</td>
<td>98</td>
<td>0.87</td>
</tr>
<tr>
<td>Dutch stallions</td>
<td>67</td>
<td>0.95</td>
</tr>
</tbody>
</table>
### Practical example: jumping horses

#### Reliabilities

<table>
<thead>
<tr>
<th>Group of horses</th>
<th>Genetic evaluation(^1)</th>
<th>Datasets(^2)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>All horses</td>
<td>A</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
<td>0.21</td>
</tr>
<tr>
<td>French stallions</td>
<td>A</td>
<td>0.58</td>
<td>0.23</td>
<td>0.54</td>
<td>0.24</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.61</td>
<td>0.20</td>
<td>0.58</td>
<td>0.20</td>
<td>0.57</td>
</tr>
<tr>
<td>Dutch stallions</td>
<td>A</td>
<td>0.51</td>
<td>0.26</td>
<td>0.46</td>
<td>0.26</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.52</td>
<td>0.25</td>
<td>0.47</td>
<td>0.25</td>
<td>0.49</td>
</tr>
</tbody>
</table>

\(^1\)A = Belgian genetic evaluation; B = Bayesian evaluation

\(^2\)I = complete data; II = data for which all performances after December 31th, 2005 were assumed to be missing; III = data for which all French and Dutch stallions’ own performances were assumed to be missing
Practical example: jumping horses

- Reliabilities

<table>
<thead>
<tr>
<th>Group of horses</th>
<th>Genetic evaluation¹</th>
<th>Datasets²</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>All horses</td>
<td>A</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td>French stallions</td>
<td>A</td>
<td>0.58</td>
<td>0.23</td>
<td>0.54</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.61</td>
<td>0.20</td>
<td>0.58</td>
<td>0.20</td>
</tr>
<tr>
<td>Dutch stallions</td>
<td>A</td>
<td>0.51</td>
<td>0.26</td>
<td>0.46</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.52</td>
<td>0.25</td>
<td>0.47</td>
<td>0.25</td>
</tr>
</tbody>
</table>

¹A = Belgian genetic evaluation; B = Bayesian evaluation
²I = complete data; II = data for which all performances after December 31th, 2005 were assumed to be missing; III = data for which all French and Dutch stallions’ own performances were assumed to be missing
Practical example: jumping horses

• Reliabilities

<table>
<thead>
<tr>
<th>Group of horses</th>
<th>Genetic evaluation¹</th>
<th>Datasets²</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>All horses</td>
<td>A</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td>French stallions</td>
<td>A</td>
<td>0.58</td>
<td>0.23</td>
<td>0.54</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.61</td>
<td>0.20</td>
<td>0.58</td>
<td>0.20</td>
</tr>
<tr>
<td>Dutch stallions</td>
<td>A</td>
<td>0.51</td>
<td>0.26</td>
<td>0.46</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0.52</td>
<td>0.25</td>
<td>0.47</td>
<td>0.25</td>
</tr>
</tbody>
</table>

¹A = Belgian genetic evaluation; B = Bayesian evaluation

²I = complete data; II = data for which all performances after December 31th, 2005 were assumed to be missing; III = data for which all French and Dutch stallions’ own performances were assumed to be missing
Practical example: jumping horses

• Conclusions
  – External information partially to totally equivalent to 4 years of contemporary horses’ performances or to their own performances
  – Resulting Belgian ranking more similar to foreign rankings
  – Improvement of REL

-interesting for foreign stallions having no or few Belgian performances
Further work

• Integration of external genomic information

• Integration of correlated external information

  ➔ To improve the accuracy of (ssG)BLUP for only locally available traits (e.g., fatty acids)
Thank you for your attention!

Corresponding author’s e-mail: jvandenplas@ulg.ac.be

• **Acknowledgments for financial support**
  • National Fund for Scientific Research (FNRS), Belgium
  • Wallonie-Bruxelles International (WBI) and Government of the Republic of Slovenia (Slovenian Research Agency – ARRS) through bilateral cooperation agreement

• **Acknowledgments for hosting**
  • Dr Gregor Gorjanc, University of Ljubljana, Slovenia
  • Animal and Dairy Science Department, University of Georgia, Athens, USA


