How should we be predicting breed differences in multibreed genetic evaluation?

Larry Kuehn

Multibreed evaluation

• What is it?
• Depends on the goal
  – Genetic evaluation incorporating data from and generating predictions on crossbred animals
  – Evaluation that results in accurate comparisons of additive genetic merit across breeds/composites
  – Provide estimates of nonadditive effects such as heterosis/sire x dam breed interactions

Multibreed evaluation

• As implemented currently, primary goal is to include crossbred and composite records
  • Ultimate goal to be able to compare all animals of any breed as seedstock
    – Starting to occur more often
    – Relates to a ‘full’ multibreed analysis

Obstacles to full Multibreed

• Merger of multiple breed databases
  – Structures are often very different
  – IDs duplicated in several breeds (but not known as duplicates)
    • Difficult to resolve
    • Standardized ID system would help
  • Cooperation between database curators
    – Breed associations
    – Genetic prediction ‘centers’
    – Individual producers/commercial entities

Multibreed Obstacles

• Estimating population parameters
  – Direct and maternal heterosis
  – Direct and maternal additive breed effects
  – Field data usually not suitable
    • Contemporary groups structure
      – Will discuss further
    • Confounding between heterosis and breed
    • Less crossbred data relative to purebreed (depends on classification of ‘purebred’)
  • Research data useful here

Outline

• Review current multibreed methods
  – Contemporary group structure and importance

• Research data for breed differences
  – ABEPD update

• Update on 2,000 bull project prediction equations
Current Multibreed Methodology

Multibreed model

- Primarily descended from Arnold et al. (1992) animal model
  - Elzo et al., 1983, 1990 sire model
  
  \[ y = Xb + ZQg + Za + Ws + WT + e \]
  - Includes
    - Additive breed effects and heterosis
    - Additive animal effects
    - Animal x breed interaction (dominance)
      - Likely difficult to fit in most data sets

- Variance assumption
  - Scale additive/residual variance according to breed/breed percentage
  - Segregation variance not accounted for
    - (Lo et al., 1993; Cardoso and Tempelman, 2004)

- Field data generally not adequate
  - Rodriguez-Almeida et al. (1997)

Why is field data inadequate?

- Trying to predict multiple effects from crossbred progeny relative to purebreds
  - Heterosis, maternal breed effect, direct breed effect, maternal heterosis, etc.

  - All comparisons must take place within contemporary groups

  - Contemporary group definition is key

Review Contemporary Groups

- Formed to evaluate animals that have been in similar environmental conditions

<table>
<thead>
<tr>
<th>Herd</th>
<th>Year</th>
<th>Recording date</th>
<th>Management</th>
<th>Location (including different pastures)</th>
</tr>
</thead>
</table>

- All comparisons used to calculate EPDs, breed differences, and heterosis take place within contemporary groups

- Some of the factors in contemporary groups are intuitive
  - year, season, herd

- Want to be sure that environmental factors are not incorrectly attributed to genetics
  - Pasture: grass quality differences, creep
  - Weigh date: stomach contents (water, feed)
  - Pen: bunk space, animal hierarchy
Review Contemporary Groups

- Desirable properties
  - Reasonable size
    - If herd is small try to make 1 group
    - Compare as many animals as possible at once
    - Weigh on the same date
  - Multiple sires
    - Always try to use more than one sire
    - When only one sire is used, none of the progeny data contributes to that sire’s EPD
    - Generally, sires are our main goal in selection

Contemporary Groups (CG)

- Desirable properties
  - Connectedness – especially when size is limited (e.g. due to pasture, etc. in definition)

Across Herd EPDs

<table>
<thead>
<tr>
<th>Sire averages</th>
<th>Sire averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire A: 550 lb</td>
<td>Sire D: 660 lb</td>
</tr>
<tr>
<td>Sire B: 650 lb</td>
<td>Sire E: 650 lb</td>
</tr>
<tr>
<td>Sire C: 540 lb</td>
<td>Sire F: 600 lb</td>
</tr>
</tbody>
</table>

Which sire is the best?

Across Herd EPDs

<table>
<thead>
<tr>
<th>Herd 1 Sire averages</th>
<th>Herd 2 Sire averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>+10 Sire A: 550 lb</td>
<td>Sire D: 660 lb +60</td>
</tr>
<tr>
<td>+80 Sire B: 620 lb</td>
<td>Sire E: 640 lb -10</td>
</tr>
<tr>
<td>+0 Sire C: 540 lb</td>
<td>Same Bull</td>
</tr>
</tbody>
</table>

Which sire is the best?

Contemporary Groups

- In order to estimate breed differences from field data, we need contemporary groups that include purebreds of the same breed
  - Rarely occurs; often breeds are in different groups
  - Even when crossbreds and purebreds are in the same group, direct comparisons are not possible without adjusting for heterosis (requires good estimates of heterosis)

Estimation of Heterosis

- In order to estimate heterosis from field data, we need groups with crossbreds and purebreds of both parental breeds
  - Rarely occurs; usually crossbreds are in a different groups
  - Even when crossbreds and purebreds are in the same group, typically purebreds of only one of the breeds are present
Estimating breed differences

• Contemporary group
  - All mated to mature purebred Limousin cows
  - What was the cause of the different averages?
    - Heterosis
    - Sire breeding value
    - Breed differences
  
  | Sire 1: Limousin WW Avg: 650 lb |
  | Sire 2: Angus WW Avg: 675 lb |
  | Sire 3: Lim-Flex WW Avg: 670 lb |

  YES
  (and no)

Estimating breed differences

• Problem can be improved with more sires in group still other considerations that are difficult to address
  - Reciprocal matings
  - Biased sampling of sires from other breeds
  - Heterosis still difficult to separate from breed
  - Were calves really treated the same?

Multibreed model

• Prior estimates of breed effects and heterosis essentially required

• Source of information most likely from research data
  - Published studies
  - Current/ongoing projects

Research Data as Prior Information

Breed comparison research

• Too many studies to count...

• None compare all breeds that we are interested in (minimum 20 x 20)

• Breed specific heterosis impossible to obtain from literature for most crosses

• Data has to be mined/combined from multiple studies

Studies of combined data

• Williams et al. (2008)
  - Least squares means from studies published 1976 to 1996
  - Combined data by modeling fixed class effect (study) and percentage breed covariates

• Roughsedge et al. (2001)
  - Studies performed in 1970s and 1980s
  - Weighted analysis for sire sampling
Roughsedge et al. (2001)

- Weighing for sire sampling
  - Amer et al. (1992)
  \[
  SE(X_c) = \sqrt{\frac{h^2CV^2}{n_i} + \frac{3h^2CV^2 + (1-h^2)CV^2}{n_s}}
  \]
  - \(n_i\) is the number of sires sampled
  - \(n_s\) is the number of progeny per sire
  - Weight in analysis was 1/SE(Xc)

Studies of combined data

- Combined studies can be adjusted to account for information content
- These studies did not account for selection (genetic trend) in the breed
  - May still work if base is adjusted to study periods, but relies on several assumptions
- Breeds effects are not a constant

Adjusting data for trend

- Possible to rely on genetic trend in the model
  - Assumes that relationships, genetic variances in the multibreed model are accurate and predict trend
- Want current breed differences
  - Better to adjust study breed estimates
  - Use current EPD of sampled bulls

Adjusting sample for current EPD

- Across-breed EPD program
  - Estimate breed differences from GPE using a sire and dam model (Fₐ progeny records)
  - Adjust records for bull EPD
    \[
    B_i = \frac{\text{USMARC}_i}{b} + (\text{EPD}_{YY} - \text{EPD}_{\text{USMARC}})
    \]
  - \(\text{EPD}_{YY}\) is the breed average EPD (current)
  - \(\text{EPD}_{\text{USMARC}}\) Weighted average USMARC sire EPD
  - \(b\) is a scaling factor to convert USMARC solution to an industry scale

### Sire Breeds Used in the Germplasm Evaluation Program at the USMARC

<table>
<thead>
<tr>
<th>Cycle I</th>
<th>Cycle II</th>
<th>Cycle III</th>
<th>Cycle IV</th>
<th>Cycle V</th>
<th>Cycle VI</th>
<th>Cycle VII</th>
<th>Cycle VIII</th>
</tr>
</thead>
<tbody>
<tr>
<td>70-72</td>
<td>71-73</td>
<td>74-76</td>
<td>92-94</td>
<td>97-98</td>
<td>99-00</td>
<td>01-02</td>
<td></td>
</tr>
</tbody>
</table>

- Fₙ Crosses (Hereford or Angus dams) *a*
- Sire breeds mated to Angus and Hereford females, and Composite MARC III (1/4 Angus, Hereford, Red Poll and Pinzgauer) cows in Cycles V, VI, VII, and VIII.

### Genetic Trends for Birth Weight, lb

Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors
**Genetic Trends for Weaning Weight, lb**

- 1974: -20 lb
- 1984: 0 lb
- 1994: 25 lb
- 2004: 50 lb
- 2011: 75 lb

Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

**Genetic Trends for Yearling Weight, lb**

- 1974: -10 lb
- 1984: 5 lb
- 1994: 20 lb
- 2004: 40 lb
- 2011: 60 lb

Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

**Genetic Trends for Maternal Milk, lb**

- 1974: -15 lb
- 1984: 0 lb
- 1994: 15 lb
- 2004: 30 lb
- 2011: 45 lb

Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

**Genetic Trends for Yearling Weight, lb**

- 1974: -25 lb
- 1984: 5 lb
- 1994: 20 lb
- 2004: 40 lb
- 2011: 60 lb

Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

**GPE Target Population Structure**

- **AI Sires:** AN, HH, SM, CH, AR, LM, GV, SH, BN, BM, MA, BR, CL, SG, SA, BV
- **Dams:** AN, HH, SM, CH, AR, LM, GV, SH, BN, BM, MA, BR, CL, SG, SA, BV

**Current breed differences**

- Underlies need for continued research for breed differences
  - USMARC GPE not enough
  - More to the country than central Nebraska
  - Southern location most important
  - Use of common sires between locations will strengthen overall power
2,000 bull project predictions

- Whole genome selection pilot project in cooperation with breed associations
  - Over 2,000 industry bulls genotyped with Illumina BovineSNP50
- Formed prediction equations using USMARC GPE data as well as deregressed EPDs from the 2,000+ bulls

Proportion of variance explained

<table>
<thead>
<tr>
<th>Trait</th>
<th>MARC-trained</th>
<th>2000 Bull-trained</th>
</tr>
</thead>
<tbody>
<tr>
<td>BWT</td>
<td>0.134</td>
<td>0.168</td>
</tr>
<tr>
<td>WWT</td>
<td>0.038</td>
<td>0.102</td>
</tr>
<tr>
<td>YWT</td>
<td>0.096</td>
<td>0.066</td>
</tr>
<tr>
<td>RIB</td>
<td>0.098</td>
<td>0.058</td>
</tr>
<tr>
<td>MRB</td>
<td>0.053</td>
<td>0.058</td>
</tr>
<tr>
<td>HCW</td>
<td>0.081</td>
<td>0.048</td>
</tr>
</tbody>
</table>

*2,000 bull predictions excluded the sires of the MARC validation populations

2000 Bulls MBV

- Resulting MBV are being sent to breed associations today
  - Sent by Mark Thallman
- Contact us with any questions

Prediction Equations

- Equations are available at: https://www.ars.usda.gov/sp2UserFiles/Place/54380510/2000_Bull_Prediction_Equations.xlsx

  Link from www.marc.usda.gov -> Genetics & Breeding -> Documents

  - Special thanks to Kristina Weber

Questions