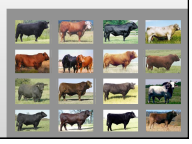
How should we be predicting breed differences in multibreed genetic evaluation?

Larry Kuehn

USDA ARS



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 purebred (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

$$\mathbf{y} = \mathbf{Xb} + \mathbf{ZQg} + \mathbf{Za} + \mathbf{WSd} + \mathbf{WT\delta} + \mathbf{e}$$

  - Includes
    - Additive breed effects and heterosis
    - Additive animal effects
    - Animal x breed interaction (dominance)
      - Likely difficult to fit in most data sets

Multibreed model

- 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
  - Rodriquez-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

Herd	Year
Recording date	Management
Location (including different pastures)	

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

Review Contemporary Groups

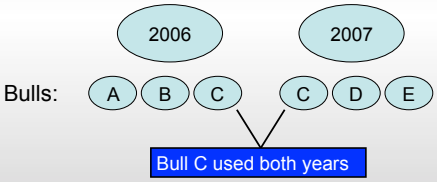
- Some of the factors in contemporary groups are intuitive
  - year, season, herd
- Want to be sure that environmental factors are not incorrectly attributed to genetics
  - Pasture: grass quality differences, creep
  - Weigh date: stomach contents (water, feed)
  - Pen: bunk space, animal hierarchy

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

Herd1 Sire averages	Herd 2 Sire averages
Sire A: 550 lb	Sire D: 660 lb
Sire B: 650 lb	Sire E: 650 lb
Sire C: 540 lb	Sire F: 600 lb

Which sire is the best?

Across Herd EPDs

Herd 1 Sire averages		Herd 2 Sire averages	
+10	Sire A: 550 lb	Sire D: 660 lb	+60
+80	Sire B: 620 lb	Sire E: 640 lb	-10
+0	Sire C: 540 lb	Same Bull → Sire C: 650 lb	+0

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

Sire 1: Limousin	Sire 2: Angus	Sire 3: Lim-Flex
WW Avg: 650 lb	WW Avg: 675 lb	WW Avg: 670 lb

  - All mated to mature purebred Limousin cows
  - What was the cause of the different averages?
    - Heterosis
    - Sire
    - Breed differences

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)

$$SE(Xc) \cong \sqrt{\frac{\frac{1}{4}h^2CV^2}{n_s} + \frac{\frac{3}{4}h^2CV^2 + (1-h^2)CV^2}{n_o}}$$

  - n_s is the number of sires sampled
  - n_a 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 (F_1 progeny records)
  - Adjust records for bull EPD

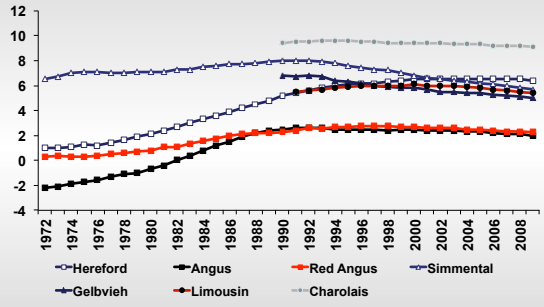
$$B_i = USMARC_i / b + (EPD_{i,YY} - EPD_{i,USMARC})$$

  - $EPD_{i,YY}$ is the breed average EPD (current)
  - $EPD_{i,USMARC}$ 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

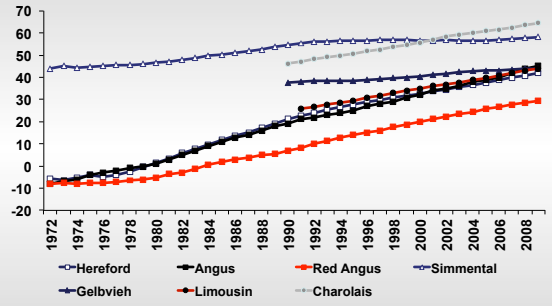
Cycle I 70-72	Cycle II 73-74	Cycle III 75-76	Cycle IV 86-90	Cycle V 92-94	Cycle VI 97-98	Cycle VII 99-00	Cycle VIII 01-02
			F_x Crosses (Hereford or Angus dams) [a]				
Hereford	Hereford	Hereford	Hereford	Hereford	Hereford	Hereford	Hereford
Angus	Angus	Angus	Angus	Angus	Angus	Angus	Angus
Jersey	Red Poll	Brahman	Longhorn	Tuli	Wagyu	Red Angus	Beefmaster
S. Devon	Braunvieh	Sahiwal	Salers	Boran	Norweg. Red	Limousin	Brangus
Limousin	Gelbvieh	Pinzgauer	Galloway	Belg. Blue	Sw. Red&Wh.	Charolais	Bonsmara
Simmental	Maine Anj.	Tarentaise	Nellore	Brahman	Friesian	Simmental	Romosinuano
Charolais	Chianina		Shorthorn	Piedmontese		Gelbvieh	
			Piedmontese				
	3-way crosses		Charolais				
Hereford	Hereford		Gelbvieh				
Angus	Angus		Pinzgauer				
Brahman	Brangus						
Devon	Santa Gertrudis						
Holstein							

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.



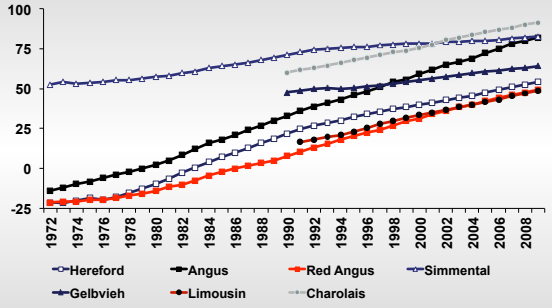
Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

Genetic Trends for Weaning Weight, lb



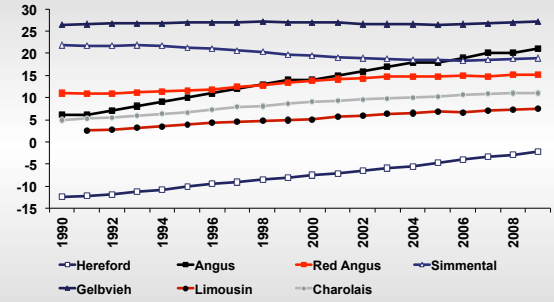
Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

Genetic Trends for Yearling Weight, lb



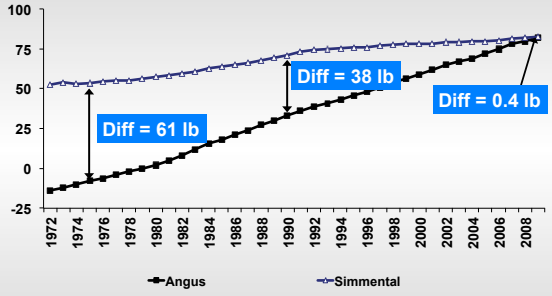
Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

Genetic Trends for Maternal Milk, lb



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

Genetic Trends for Yearling Weight, lb

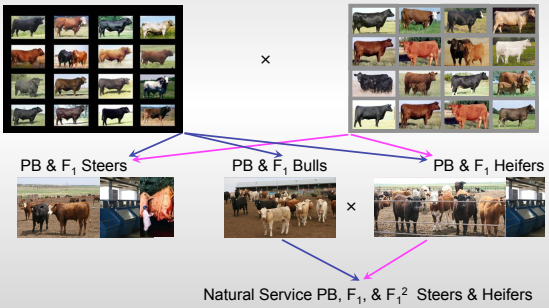


Adapted from Spring 2009 Genetic Trends from Breed Associations and 2011 AB-EPD factors

GPE Target Population Structure

AI Sires:
AN, HH, SM, CH, AR, LM, GV, SH, BN, BM, MA, BR, CI, SG, SA, BV

Dams:
AN, HH, SM, CH, AR, LM, GV, SH, BN, BM, MA, BR, CI, SG, SA, BV



Natural Service PB, F_1, & F_1^2 Steers & Heifers

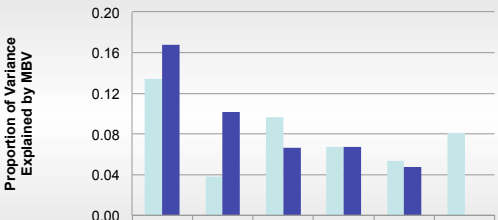
Current breed differences

- 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



	BWT	WWT	YWT	RIB	MRB	HCW
MARC-trained	0.134	0.038	0.096	0.068	0.053	0.081
2000 Bull-trained	0.168	0.102	0.066	0.068	0.048	

*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