

Multibreed Genetic Evaluations of Beef Cattle

E.J. Pollak and R.L. Quaas

Department of Animal Science, Cornell University, Ithaca, NY 14853 U.S.A.

INTRODUCTION

Interest in crossbreeding in commercial cattle herds in the latter half of the 20th century sparked a flurry of experiments characterizing breeds and breed crosses for additive and nonadditive gene effects. Analytical models for estimation of these effects naturally followed. A logical extension was to apply the methodology to estimation of genetic merit from records in populations representing multiple breeds, and this area of research has received increased interest as composite and other crossbred animals are finding their way into commercial herds as seedstock. This presentation will be a discussion of the models for multiple breed genetic evaluations and on experiences gained in the application of these models to beef cattle evaluations in the U.S. It will conclude with a description of the plan for expanding multiple breed evaluations being proposed by the National Beef Cattle Evaluation Consortium (NBCEC).

MODELING

The genesis of models to analyze data from crossbreeding experiments in livestock species, not surprisingly, can be found in the analysis of line-cross data in plants. Gardner and Eberhart (1966) described analysis of diallel crosses assuming diploid inheritance, two alleles per locus, and no epistasis. Robison *et al.* (1981) extended the model for use in analysis of crossbred dairy data. Their model has the following features:

- Observations are linear combinations of breed additive (both direct and maternal) effects and breed combinations (heterosis for direct, paternal, and maternal effects).
- Breed effects are regressed by the fraction of genes represented in an individual for that breed, with genes originating from the sire or the dam delineated separately.
- Maternal breed effects are included and regressed by fraction of genes in the dam representing each breed.
- Heterosis effects for the direct (maternal) expression of a trait due to dominance are regressed by the percentage of loci in the individual (dam) with alleles from different breeds.

These authors noted that a model containing breed-composition groups could be used (that is, a separate

group effect for each percent combination of breeds). These early models were devised with interest primarily in the estimation of genetic difference in strains (or breeds) and interactions between them (heterosis).

Paralleling investigations of multibreed analysis of data was research on grouping strategies in linear models for genetic evaluations. Models with group effects had been used in the earliest dairy sire evaluation procedures. In these early models, animals were assumed unrelated. There was renewed interest in research into grouping strategies when relationships were incorporated into the genetic evaluation models. Thompson (1979) described the concept of accumulated group effects in a sire model. His model incorporated ancestral group information by exploiting the structure of the numerator relationship matrix. Westell (1984, 1988) and Robinson (1986) independently extended this concept to an animal model. In addition, Westell (1984) used modified mixed model equations such that solutions of animal equations are genetic evaluations (combining the fixed group effects and random genetic components using the QP transformation [Quaas and Pollak 1981; Quaas 1988]). These studies were motivated by problems in evaluation of dairy cattle, but the results are general and provide an excellent mechanism to incorporate breed effects into a multibreed model.

Arnold *et al.* (1992) formulated the multibreed evaluation approach for a single-trait animal model. Using modified mixed model equations (QP transformed), they proposed the following model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{h} + \mathbf{e},$$

where \mathbf{y} is the vector of observation, $\boldsymbol{\beta}$ is a vector of fixed effects (typical of genetic evaluation problems, e.g. contemporary groups), \mathbf{u} is a vector of breeding values, \mathbf{h} is a vector of total nonadditive effects, and \mathbf{e} is a vector of random residuals. \mathbf{X} , \mathbf{Z} , \mathbf{W} are matrices relating effects to observations. The vector \mathbf{u} represented in the modified equation is $\mathbf{Q}\mathbf{g} + \mathbf{a}$, where \mathbf{g} is the vector of breed additive effects (group effects), \mathbf{a} is the vector of random additive genetic effects, and \mathbf{Q} is the matrix relating fractions of breeds represented in individuals to the breed group effects. Further, \mathbf{h} is modeled as $\mathbf{h} = \mathbf{S}\mathbf{d} + \mathbf{T}\boldsymbol{\delta}$, where \mathbf{d} is a vector of fixed heterosis effects (sire breed \times dam breed), and $\boldsymbol{\delta}$ is a

vector of random heterosis effects (e.g., sire \times breed of dam). They considered predicting the random heterosis component, δ , to be “problematic” and suggested that “initial applications of this model to existing cattle populations will most likely operate under the assumption that a fixed component of heterosis is sufficient.” This model is the foundation of the current multibreed evaluations systems in the U.S.

APPLICATIONS

Experimental results. An application of multibreed models is Rodriguez-Almeida *et al.* (1997). A simplified version of the Arnold model was fit for birth and weaning weights with direct and maternal effects for both traits. Of particular interest in this study was the ability of the procedure to separate effects in the model. Two data set were used:

1. Data set 1: Observations were from the base breeds and crosses used to create three composite breeds, MARC I, II, and III.
2. Data set 2: Data set 1 plus experimental data in which appropriate crosses and other matings provided better structure for estimation of effects in a multibreed analysis.

The results were sobering. Marked differences in estimates of direct and maternal breed effects were obtained from each data set. Results from MARC II for weaning weight are shown in the following adaptation of results in Rodriguez-Almeida *et al.* (1997):

Breed	Data set 1 (MARC II)			Data set 2 (MARC II)		
	Direct	Maternal	Sum	Direct	Maternal	Sum
Hereford	-4.7	-12.2	-16.9	0.9	-13.7	-12.8
Simmental	-42.2	94.6	52.4	29.8	26.0	55.8
Gelbvieh	-46.2	100.7	54.5	37.0	17.2	54.2

The authors pointed out that the sum of the direct and maternal estimates from both analyses were very similar while the components differed dramatically. This suggested that partitioning components in ill-designed data could be problematic. Data set 1 represents the nature and structure of data from the field. The authors concluded incorporation of experimental data into field data analysis, by some mechanism, is required to achieve satisfactory estimates of effects in multibreed models.

Multibreed model for national cattle evaluation: In the fall of 1997, the American (ASA) and Canadian Simmental Associations (CSA) published weight trait evaluations obtained from a joint analysis of three data sets that were previously analyzed separately. These populations were U.S. and Canadian Simmentals and U.S. Simbrahs. Because all three data sets included crossbred animals, a multibreed approach was used

(Klei and Quaas 1995). This multitrait model for birth, weaning weight, and postweaning gain has the following features:

1. Contemporary groups contain animals of different breed compositions.
2. Additive breed differences are accounted for by regression on breed of founder \times year groups.
3. Heterosis effects are accounted for assuming the expected heterosis is a result of the full heterosis between two breeds multiplied by the fraction of loci expected to contain one allele from each breed.
4. Heterogeneous variances by percent Simmental are used [Garrick *et al.* (1989)].
5. Age of dam effects are modeled with a 4th order polynomial (Bertrand *et al.*, 1994). Weighted averages of breed age curves are used for crossbred dams.
6. A nonlinear 205-day age of calf adjustments is used.

Bayesian procedures were used for both breed additive genetic effects and heterosis effects. Information from literature were used as prior means. Prior variances were chosen so that neither priors nor data dominate solutions, except for heterosis where variances were chosen such that most weight was on the priors.

EXPERIENCES IN APPLICATION

Application of models like that implemented by Cornell for the national evaluation of data from ASA and CSA provides valuable experiences, and lessons learned from this application will be discussed.

Combining breeds for breed, age of dam, and heterosis effects. The Simmental data set has animals representing over 60 breeds but is dominated by four breeds: Simmental, Hereford (foundation cows in early years), Angus (sires used on Simmental heifers in later years) and Brahman (through contributions to Simbrah). For these breeds, plus other well represented breeds (Charolais, Gelbvieh and Limousin), breed effects were fit by year. Breeds with limited information were combined into one of four categories: American, British, continental, dairy, plus two “catchall” groups, U.S. and Canadian. Ages of dam effects were fit for categories. For heterosis, the effects were for combinations of any breed represented in the following categories: British, Continental, Zebu, and other. Thus, for a Hereford-Angus calf, heterosis would be for British \times British; for a Simmental-Angus, the continental \times British value would be used, etc.

Autoregressive prior for year within breed effects. Breed (or breed grouping) of founder \times year effects

were included. A Bayesian approach was used to overcome the problem of little information for some subclasses. The prior's means for a breed were constant over years but were not assumed independent. An autoregressive covariance structure with a large between-year correlation (.95) was assumed. This effectively eliminated large year-to-year fluctuations but allowed the founder effects to change over time without specifying a functional form for that change. It also allowed estimation of a founder-year effect that has not yet been expressed, e.g. maternal founder effects for the current calf crop.

Founder effects. In articles discussing breed effects obtained from multibreed analysis, the effects were interpreted as the mean of a population defined by a "breed." For the Simmental data sets, we questioned whether this was a fair interpretation. Producers providing data on individuals with nonSimmental genes were either "grading up," producing crossbred animals for sale as seedstock, or creating composites. The "founders" coming from these breeds were not a representative sample of animals from that breed. Breed-year effects were not viewed as estimates that could be used to make statements about true breed differences, and we referred to them as founder effects; e.g., the trends we saw for various breeds were not necessarily the trends observed in the respective purebred populations.

Gametic versus genetic trends. Genetic trends are usually estimated as average EPDs (or BVs) for calves born by year. In a multibreed evaluation, these can be computed for a particular breed or breed combination. We calculated for each breed (group) a "gametic" trend. Within each year of birth, animals' EPDs were regressed on their breed compositions to partition the yearly average EPD among breed groups. The regression coefficients estimated the genetic merit of genes transmitted from each breed (group) to a year's calves. This trend used information from every animal born in a given year to measure the genes from the different breed of founder groups. Every animal with some fraction of Simmental breeding contributed to the Simmental gametic trend. Likewise, every animal with some fraction of Angus breeding contributed to the Angus gametic trends.

"True" breed fractions versus association designations. Arnold *et al.* (1992) pointed out the importance of correctly identifying breeds represented in individuals as their evaluation will be a function of the breed effects. Breed associations for breeds derived from upgrading or as composites often have rules for the designation of individuals resulting from various matings. The danger exists that the rules lead

to incorrect fractions of breeds represented. We recomputed breed fractions from all available pedigree information, which lead us to a problem of explaining evaluations of previously designated purebreds (100% Simmental) that now had fractions of other breed effects in their evaluations. We encountered an example of two bulls designated purebreds under ASA rules. They differed by the fraction of genes representing a second breed. With large numbers of progeny, the additive direct breed group effects had no significant impact on their evaluations. However, the heterosis effects differed when they were mated to purebred Simmental cows and as the number of progeny increased, the full expression of this difference was in the contrast between these bulls. This "change" in the contrast was difficult to explain given the bulls were designated by the association as purebreds.

Bases. In single-breed evaluation systems, the choice of a base usually revolves around which year establishes the base and then forcing solutions of animals born that year to be some constant, typically zero. In a multibreed analysis, base options are expanded, as now breed combinations can be included in the definition of the base. The base can be set by forcing solutions of a group of individuals to sum to zero or, as the ASA chose to do, the sum of the predominate four breeds' gametic values in 1991. A base is, of course, arbitrary and inconsequential to contrasts needed for ranking and selecting animals. However, EPDs are used in merchandising, and therefore, the choice of a base has economic ramifications.

Evaluations of animals from other breeds. The current Simmental multibreed system provided evaluations for animals of other breeds. These evaluations were based only on information in the Simmental data set. The evaluations and contrasts between them differed from official evaluations from their respective breeds. This was precisely the reason the ASA and CSA moved to a joint evaluation. As an example, we were challenged on the accuracy of the multibreed system based on contrasts reported between two prolific Angus bulls. Our initial contrasts between them (based on relatively few progeny) were quite different from the Angus analysis. As we have accumulated information, ours have moved to a reasonable reflection of the Angus evaluations. Providing results for these animals probably would not serve any real purpose. However, they will appear in pedigrees in our data set and so must be listed on official registration forms for those animals. To address this problem, the theoretical framework for incorporating "external EPDs" was developed (Quaas and Zhang, 2001) Incorporating

external EPDs is a process that uses EPDs and accuracy values obtained in a different evaluation system to supplement the information in the target data set. The need, obviously, does not exist for using external EPDs if the data for those breeds are included in the evaluation. Since the inception of the multibreed analysis, two breeds have added their data sets, Maine Anjou and Chianina.

Current status and future plans of multibreed evaluations: Since the first application of the multibreed model for weight traits in 1997, several other systems have evolved. Cornell/ASA now has a multibreed model for carcass data that includes information from ultrasound measures of breeding animals. The University of Georgia has developed a multibreed system for weight traits.

The NBCEC has recently developed and has begun the implementation of a strategy for expansion of multibreed evaluations. We are currently running a prototype analysis for a greatly expanded number of breeds for weight traits. This involves the development of a national pedigree file, which is being done at Cornell. This file will help maintain the unique identity of animals registered across several breed data sets. We are also developing a national data file for evaluations with each university and breeds providing information on performance. We have begun investigation into expanding the number of traits included in the multibreed system as well to include carcass quality measures and threshold traits such as calving ease, heifer pregnancy, and stayability.

Is an EPD an EPD? ASA has published a single EPD for each animal for each trait under the assumption that an EPD is an EPD. We were hoisted by our own petard of terminology. The EXPECTED PROGENY DIFFERENCE has a very definite meaning, and educational programs have emphasized this meaning. However, the additive values predicted from the multibreed evaluation do not predict progeny differences between prospective parents of different breed composition. We are on safer ground when the comparison is between two individuals of the same breed bred to similar mates, assuming the absence of

individual nonadditive differences. How will we deal with this in the future?

Colorado State University has taken the leadership in producing a decision-support web-based system to help producers assess the impact of selection decisions. This system uses existing EPDs to model the expected impact on total herd productivity. As such, phenotypic measures are generated based on herd information and the EPDs of prospective sires. This system will be expanded to incorporate information on cow breed composition and will incorporate heterosis into the predictions of phenotypic performance. It is unlikely that we will ever publish a "matrix" of EPDs showing the contribution of the genetic merit of the animal in question dependent on the breed composition of the mate. This decision-support system allows for the customizing the use of EPDs to each cattle operation.

REFERENCES

- Arnold, J.W., Bertrand, J.K and Benyshek, L.L. (1992) *J. Anim. Sci.* **70**:3322-3332.
- Bertrand, J.K., Nelson, A.H. and Middleton, B.K. (1994) Proc. 4th Genetic Pred. Workshop, Kansas City, MO.
- Gardner, C.O. and Eberhart, S.A. (1966) *Biometrics* **22**:439-452.
- Garrick, D.J., Pollak, E.J., Quaas, R.L. and Van Vleck, L.D. (1989) *J. Anim. Sci.* **67**:2515-2528.
- Klei, L. and Quaas, R.L. (1995) Proc.WCC-100 Reg. Coordinating Comm. Mtg., Brainard, MI.
- Quaas, R.L. (1988) *J. Dairy Sci.* **71**:1338-1345.
- Quaas, R.L. and Pollak, E.J. (1981) *J. Dairy Sci.* **64**:1868-1873.
- Quaas, R.L. and Zhang, Z. (2001) *J. Animal Sci.* **79** (Suppl. 1):342.
- Robinson, G.K. (1986) *J. Dairy Sci.* **69**:3106-3111.
- Robison, O.W., McDaniel, B.T. and Rincon, E.J. (1981) *J. Anim. Sci.* **52**:44-50.
- Rodriguez-Almeida, F.A., Van Vleck, L.D., and Gregory, K.E. (1997) *J. Anim. Sci.* **75**:1203-1212.
- Thompson, R. (1979) *Biometrics* **35**:339.
- Westell, R.A. (1984) Ph.D.Thesis, Cornell University.
- Westell, R.A., Quaas, R.L., and Van Vleck, L.D. (1988) *J. Dairy Sci.* **71**:1310-1318.