361 Field experience with gilt age and lactation length. L. V. Himmelberg*, *Danbred North America, Seward, NE*.

The objective of this presentation will be to relate field experiences regarding the impact of gilt age and lactation length on sow productivity. Over time the largest percentage of females to farrow within a herd are parity one and therefore overall herd productivity is greatly influenced by the results achieved by them. Gilt age and weight at first service have a direct impact on these results. Field observations show that increasing gilt age and weight by approximately 30 days and 20 kilograms respectively, over traditional targets of 210 days and 125 kilograms can improve first parity born alive by as much as .75 pig per litter. Early weaning programs designed to increase sow productivity and improve health have driven weaning averages below 21 days of age. Many producers have noted that weaning this early has a negative impact on subsequent parity born alive. To explore this, commercial production herds containing Danish genetics were examined. Born alive per litter was treated as a dependent variable and previous lactation length as an independent variable within each herd. The results of this analysis showed an average decrease in subsequent parity born alive of .12 pig per day for each day decrease in weaning average between 21 and 15.

Key Words: Gilt, Lactation

Breeding and Genetics: Genetics Methodology 2

362 Pedigraph 2.0: A software tool for the graphing and analysis of large complex pedigrees. J. R. Garbe* and Y. Da, *Department of Animal Science, University of Minnesota, St. Paul*.

Pedigree graphing and analysis are important tasks in animal breeding and genetics. However, pedigree graphing and analysis are difficult for large complex pedigrees. Pedigraph version 2.0 provides rapid graphical visualization and analysis of large, complex pedigrees with a variety of options and features. The program produces artistic pedigree drawings with options to draw the full pedigree, a summarized pedigree with three options, or an extracted pedigree for a selected individual. The user has control over drawing style, fill color, line color, page size, pedigree size, title of the pedigree, gap between two generations, display of trait status, genotype, or inbreeding coefficient, and the minimal number of offspring required for a family to be included in the pedigree. The program can also calculate inbreeding coefficients for all individuals and the coefficient of coancestry between each pair of individuals. The program has been tested using two large and complex populations for its capability of pedigree graphing and its computational feasibility to calculate inbreeding and coancestry coefficients. The program has not encountered limitations in sample size or complexity of the pedigree. The only limitation encountered so far is the ability to print a potentially huge pedigree drawing. These tests plus over a years trial indicate that Pedigraph could be a versatile and capable tool for pedigree visualization and analysis.

Key Words: Pedigree, Genealogy, Visualization

363 Multiple-breed genetic inference using a heavy-tailed structural model for heterogeneous residual variances. F. F. Cardoso, G. J. M. Rosa, and R. J. Tempelman*, *Michigan State University, East Lansing*.

Multiple-breed genetic models have been recently demonstrated to effectively specify the heterogeneous genetic variances that exist between different beef cattle breed groups. We extend these models to allow for heterogeneous residual variances that may be a function of fixed effects (e.g. sex, breed proportion, or breed group heterozygosity) and random effects such as contemporary groups (CG). We additionally specify the residual distributions to be either Gaussian or based on a heavy tailed alternative such as the Student t, in order to distinguish potential outliers from individuals in high variance environments. For either of these two distributions and their homoskedastic counterparts, we analyzed 22,717 post weaning gain records from a Nelore-Hereford population based on a Markov chain Monte Carlo animal model implementation. The heteroskedastic Student t error model (with estimated degrees of freedom 7.33±0.48) was clearly the best fitting model based on a Pseudo-Bayes factor criterion. Sex and breed group heterozygosity, but not breed proportion, appeared to be marginally important sources of residual heteroskedasticity. Specifically, the male residual variance was estimated to be 1.13±0.09 times that for females whereas, for example, the residual variance in F_1 animals was estimated to be 0.70±0.16 times that for purebreds. The CG effects were important random sources of residual heteroskedasticity; that is, the standard deviation of ratios of CG-specific residual variances relative to the baseline average was estimated to be 0.72±0.06. Purebred Nelores were estimated to have a larger genetic variance (124.87±21.75kg^2) compared to Herefords (40.89±6.70kg^2) under the heteroskedastic Student t error model; however, the converse was observed from results based on a homoskedastic Student t error model (46.24±10.90kg^2 and 60.11±8.54 kg^2, respectively). These results naturally have important implications for multiple breed genetic evaluations.

Key Words: Bayesian Inference, Genetic Predictions, Outlier Robustness

364 Genetic evaluation of male fertility using a threshold model with emphasis on accurate estimation of conception rate. G. Abdel-Azim*[1], S. Schnell[1], G. Gelbert[1], and H. Rycroft[2], [1]*Genex Cooperative, Inc., Shawano, WI,* [2]*Cooperative Resources International, Shawano, WI*.

A threshold model for male fertility evaluation was applied to a data set of size 2,233,377 records. The model included herd, year-season, bull age, cow parity, as fixed factors, and service bull, inseminated cow, and inseminator as random factors. Variance components estimated as ratios of the total phenotypic variance were 2.8, 11.1, and 5.2% for service bull, cow and inseminator, respectively. The main objective of the study was to transform bull solutions into percentages of conception rates, a more interpretable value to indicate male fertility. The transformation implemented avoids problems arising when linear combinations of estimates are used in a nonlinear parametric function, a practice that is currently in use to report national calving ease evaluations estimated by a threshold model. The transformation we implemented takes into account the accuracy of estimating breeding values and other factors in the model, hence, current practices that do not take accuracy into account were shown to consistently overestimate the probability of success in the binomial setting studied.

Key Words: Fertility, Conception Rate, Genetic Evaluation

365 Application of a random regression model to gene expression profiling. S. L. Rodriguez-Zas*, J. J. Loor, J. K. Drackley, and H. A. Lewin, *University of Illinois, Urbana*.

The patterns of gene expression recorded on individuals over a period can be studied using discrete or continuous representations of time. Within the later representation, the profile of expression can be modeled using common (fixed) and individual (random) polynomial coefficients in time. We evaluated the potential of random regression models to describe the fluctuations in the gene transcription levels recorded at successive time points. The data consisted of fluorescence intensities on more than 6000 unique genes recorded using spotted cDNA microarray technology. Liver samples were obtained at -65 d, -30d, -14d, +1d, +14d, +28d and +49d relative to calving on 8 Holstein cows. A reference design was implemented with each cow-day sample represented in two reverse-dye microarrays and each gene double spotted on each microarray. Fluorescence intensity measurements on 106 microarrays were filtered for weak signals and were normalized using a log2 transformation on the loess-adjusted values. The random regression model included linear to quartic polynomials on days and accounted for heteroscedasticity between days. Three percent of the genes had at least one significant ($P < 0.0001$) regression coefficient in days. The majority of these genes had significant quadratic trends alone or in combination with a significant quartic trend. Hierarchical and disjoint clustering of these coefficient estimates indicated the presence of 5 clusters. Four of these clusters were approximately characterized by significant (positive and negative) quadratic regression coefficients in combination with significant (positive and negative) quartic regression coefficient within each signed quadratic group. The last cluster was characterized by significant linear and cubic regression coefficients. Results from this study

indicate that random regression models are flexible to accommodate the variation in patterns that can be observed in genomic studies.

Key Words: Gene Expression, Random Regression, Clustering

366 Evaluation of three statistical methods for QTL analysis. J. Xu*, J. R. Garbe, N. R. London, Y. Mao, and Y. Da, *Department of Animal Science, University of Minnesota, St. Paul*.

Three statistical methods are widely used in QTL analysis: mixture model (MM) likelihood analysis, regression on genotypic probabilities (RGP) and the least squares method (LS). We evaluated these three methods for statistical power in QTL detection and for accuracy in parameter estimation using simulations. Additive and dominance effects were simulated separately under each method. Three significance levels were used for QTL detection under LS. Critical values for QTL detection under MM and RGP were determined by empirical type-I errors that are equivalent to those under the LS method to ensure fair comparison. Statistical power was measured by the percentage of significant results. Accuracy of parameter estimation was measured by the mean squared error (MSE) of the estimates. For detecting additive effect, LS had the highest power when the heritability of the trait is <0.1, and the three methods had similar power for higher heritabilities. For detecting dominance effect, MM and RGP had the same higher power and LS had the lowest power. For estimating marker-QTL recombination frequency based on the testing for additive effect, LS had the best accuracy (smallest MSE), followed by RGP, with MM having the worst accuracy. For estimating marker-QTL recombination frequency based on the testing for dominance effect, RGP had the best accuracy and MM had the worst accuracy when dominance heritability<0.1. The three methods had similar accuracy when dominance heritability is > 0.1. For estimating additive effect, LS had the best accuracy followed by RGP. MM had the same accuracy as RGP when heritability is >0.1 but its accuracy rapidly worsens as heritability decreases. For estimating dominance effect, the three methods had similar performance when dominance heritability is >0.1 but the accuracy of MM rapidly worsens as dominance heritability decreases. These results suggest that each of the three methods has its own strength and that the three methods could be used jointly to compliment each others strength.

Key Words: Mixture Model, Regression, Least Squares

367 Relationship between the choice of a random regression model and the possible shapes of the resulting variance function. S. D. Kachman*, *Department of Statistics, University of Nebraska, Lincoln*.

Random regression models provide a flexible means of modeling a trait, measured at variable points in time, using a relatively small number of covariance parameters. Random regression models have been applied in a variety of settings, including modeling lactation curves and growth curves. Within the class of random regression models, models using polynomials have proven to be very popular. The variance as a function of time can be obtained using the variance function, a linear transformation of a set of component variance functions. The component variance functions are not dependent on the variance components. In the case of polynomials the component variance functions are polynomials in time. An examination of the component variance functions, along with the restriction that a covariance matrix for the random regression coefficients must be positive semidefinite, results in the variance function having a lower bound formed by polynomials of even power. This results in a tendency for the variance to grow rapidly for extreme time points. For growth curve models without a fixed endpoint, a major consequence is large genetic and environmental variances for extreme ages. A secondary consequence is the tendency for variance functions to display a rounded wave pattern at intermediate time points. Regardless of the type of function used in a random regression model, the shapes of the component variance functions will determine the possible shapes of the variance function. Understating the component variance functions is important both in determining the suitability of a random regression model and in interpreting the resulting covariance matrices.

Key Words: Mixed Models, REML, Random Regression

368 Weighting of information when predicting breeding values using the standard or marker-based inverse of the numerator relationship matrix. A. Maiwashe*, D. J. Garrick, and R. M. Enns, *Colorado State University, Fort Collins*.

Best linear unbiased prediction uses information from relatives when evaluating the genetic merit of an animal by accounting for additive genetic relationships among animals. Relationships among animals are usually computed using pedigree information (A_p). Relationships among animals can now be computed using genetic markers (A_m). The objective of this study was to derive weights for an individual and its relatives when the A_p or A_m numerator relationship matrices were inverted and used in the mixed model equations. A hypothetical data set for a trait measured on females included one grandsire with half sib sons. The grandsire and his sons had several hundred daughters with yield deviations (DYD). Estimated breeding values (EBV) were obtained using a sire model. The EBV for each son was expressed as a linear function of his DYD if available and EBV of his relatives. The weights for a half-sib bull using A_p were always zero except for his DYD and the grandsire. The relative emphasis on the DYD increased as the number of daughters increased. The zero weight on half-sibs was a feature of the inverse of the A_p. In any non-inbred population, the A_p uses a relationship of 0.25 between all half-sibs. A bull with no daughters was evaluated as half its sires EBV. Its reliability is limited because there can be no accounting for Mendelian sampling. In contrast, the relationship between half sibs can vary from 0 to 0.5 in A_m. That resulted in an inverse with more non-zero coefficients than existed when A_p was used. Accordingly, the evaluation of any particular half-sib had non-zero weights for all half-sibs with the weight on each half-sib varying with the proportions of alleles shared in common. A bull with no daughters could be more reliably assessed using A_m rather than A_p if he had half sibs with daughters because the inferiority or superiority of his Mendelian sampling could be assessed to some extent. The use of relationships based on markers may be advantageous for young bulls.

Key Words: Relationship Matrix, Marker-Assisted Selection, BLUP

369 A practical longitudinal model for evaluating growth in Gelbvieh cattle. K. R. Robbins*, I. Misztal, J. K. Bertrand, A. Legarra, and S. Tsuruta, *University of Georgia, Athens*.

Using Gelbvieh records spanning 1972-2001, the practicality of using a longitudinal model to evaluate growth in beef cattle was examined by subjecting the data to multiple-trait (MT) and random regression (RR) analysis. Approximately 15% of the animals in the dataset had at least one record measured outside of the accepted MT age ranges for weaning weight (WW) and yearling weight (YW). Fourteen percent of WW records and 19% of YW records were measured outside of the accepted ranges for MT analysis. This suggested that a longitudinal model could provide additional information for breeding value (BV) prediction. Three RR evaluations were performed using cubic Legendre Polynomials. The first (RR1), included only records measured within MT ranges. The second (RR2), used records as in RR1, but ages were set to 205 days and 365 days for WW and YW respectively. This was done to determine the extent to which the spread in days at weighing affected the predicted BVs. Finally, the third evaluation (RR3) utilized all available records. Due to convergence problems the random regression parameters were diagonalized via eigenvalue decomposition. After diagonalization, it was found that all models converged in a similar number of iterations. Correlations were computed between model predictions for fixed, direct, maternal, and maternal permanent environmental effects. Correlations between the MT, RR1, and RR2 evaluations were high for all effects. Correlations of MT with RR1 and RR2 for the direct effect were #8805 .99 for all three traits, leading to the conclusion that both models were equivalent when predicting BVs from data containing records measured within the MT age ranges. Correlations between MT and RR3 were .992, .958, and .951 for BW, WW, and YW. These correlations show that the added information does affect BV prediction. The results of this study indicate that the RR model has increased capability at a cost similar to MT.

Key Words: Legendre Polynomial, Random Regression, Multiple Trait