BARD

FINAL REPORT

PROJECT NO. US-805-84

Development of Methods for Genetic Analysis of Discrete Traits of Economic Importance in U.S. and Israeli Dairy Cattle Populations

**D. Gianola, R. L. Fernando, C. R. Henderson (deceased),
R. Bar-Anan (deceased), J. Weller**

10

630[#]72
BAR/GIA
2nd copy

1989

32139

000000127363

Standard BARD Cover Page for Scientific Reports

Date: October 31, 1989

BARD
P.O. BOX 6
Bet Dagan, ISRAEL

BARD project No. US-805-84

Title

Development of Methods for Genetic Analysis of Discrete Traits of Economic
Importance in U.S. and Israeli Dairy Cattle Populations

Investigators' Names (Principal listed first)

D. Gianola

R. L. Fernando

C. R. Henderson (deceased)

R. Bar-Anan (deceased)

J. Weller

Investigators' Institutions

University of Illinois, Urbana

University of Illinois, Urbana

University of Illinois, Urbana

A.R.O., The Volcani Center

A.R.O., The Volcani Center

Project's starting date: January 1, 1986

Type of Report: 1st Annual _____ 2nd Annual _____ Final ✓

Signature
Principal Investigator

Signature
Institution's Authorizing Official

RECEIVED
BARD
10-31-89

I. TABLE OF CONTENTS

	<u>Page</u>
II. ABSTRACT	2
III. ORIGINAL OBJECTIVES OF RESEARCH	3
IV. INTRODUCTION	3
V. PREDICTION OF BREEDING VALUES AND ESTIMATION OF PARAMETERS: THEORETICAL ASPECTS	4
VI. INFERENCES UNDER SELECTION	9
VII. COMPUTING TECHNIQUES	10
1. Linear models	10
2. Threshold models	13
VIII. BASIC QUANTITATIVE GENETICS	14
IX. DAIRY CATTLE BREEDING	15
1. Analysis of lactation curves	15
2. Health and disease	16
3. Causes of cow culling in Israel	17
4. Methods of cow evaluation	17
5. Calving traits	18
6. Genetic analysis of twinning rate in Israel	19
7. Genetic and environmental relationship between fertility and production in Israel	19
8. Genetic, environmental and phenotypic trends for production and secondary traits	24
9. Genetic and environmental parameters of milk composition in Israel	25
10. Strategies for genetic mapping of quantitative trait loci in dairy cattle	26
11. Heterogeneity of variances across production levels	27
12. Optimization of the Israeli breeding program	27
X. DESCRIPTION OF THE COOPERATION	30
XI. EVALUATION OF RESEARCH	30
XII. PUBLICATIONS	33

II. ABSTRACT

The project studied problems in dairy cattle breeding, ranging from the implementation and computation of large linear and nonlinear models to selecting bulls for twining rate of their daughters. Programs were written to calculate BLUP sire evaluations for production, fertility, calving, and growth traits, and for mastitis susceptibility and cow survival. A fairly general computer program was developed to carry out variance component estimation and sire evaluation by threshold models. Theoretical problems studied resulted in the development of more suitable methods for the analysis of threshold characters and in an increased understanding of why likelihood based procedures can account for certain forms of selection. Methods were developed to adjust production records for past reproductive performance, calving age and month, and to predict total production from incomplete lactations. Genetic evaluations computed from adjusted records were slightly more accurate than non-adjusted records. The adjustment factors developed in Israel were incorporated into the standard sire and cow evaluation procedures. Methods to account for heterogeneous variance components among herds were developed. This problem is probably not important for sire evaluation in Israel, because sires are evaluated over many herds. It may be of major significance for evaluation of US sires, and bull dams in both countries. Preliminary results on the relationship between bacterial infection and somatic cell counts have been presented to the industry, and will serve as the basis for a more extensive Israeli study. Phenotypic, genetic, and environmental trends in Israel for production and reproductive traits were estimated. Genetic trends were economically favorable for all traits studied, except fat percent and days to first breeding. Methods were developed to optimize the Israeli breeding program. It was found that genetic progress could be increased if more young sires were tested each year, and if a larger percentage of cows were mated to young sires. MOET for bull dams in Israel should increase genetic progress by 6-7 kg per year, and will reach the economic break-even point in about 15 years. Major problems were solved and new problems were more clearly defined and studied.

III. ORIGINAL OBJECTIVES OF RESEARCH

1. To develop statistical models aimed to describe variation of categorical traits in Israeli and U.S. populations of Holstein cattle.
2. To develop methods for estimation of variance and covariance components in nonlinear models suitable for categorical traits.
3. To assess robustness of methods of estimation of variance and covariance components and of sire evaluation using linear models, to departures in the assumptions required for their use when applied to categorical responses.¹
4. To study, using field records and computer simulation, the relative cost-effectiveness of nonlinear vs. linear methods of sire evaluation for categorical traits.
5. To examine, by analysis and simulation, the relative ability of linear versus nonlinear methods to reduce (or eliminate) bias due to different types of selection protocols.
6. To develop methods of genetic evaluation for mixtures of economically important quantitative and categorical traits, and merit functions aimed to describe breeding goals at the national level and to provide guidance for mating programs.
7. To develop and implement computer programs and algorithms which could be general in power and scope so as to accommodate reasonably large data sets.

IV. INTRODUCTION

The project dealt essentially with problems posed by the analysis of discrete attributes in dairy cattle breeding. Many traits of economic importance in dairy cattle production such as fertility, survival, calving ease and liability to disease are frequently measured as a discrete or categorical outcome.

The objectives were researched using theoretical, computer simulation and data-analytic approaches. Not all objectives were researched with equal intensity, and the broad nature of dairy cattle breeding led to incursions into other areas as well. Hence, this report will be organized by area of work rather than by initial project objective.

¹A Ph.D. thesis presented at Hohenheim University, Germany, in 1986 (Ina Hoschele, "Estimation of breeding values and variance components with quasi-continuous data") partially addressed this problem. In view of her results and of the complexity of the problem (in view of all the objectives of the project), the topic was not pursued.

V. PREDICTION OF BREEDING VALUES AND ESTIMATION OF PARAMETERS: THEORETICAL ASPECTS

Gianola et al. (1986) studied prediction of breeding value when dispersion parameters, e.g., heritability, are not known, so these must be estimated using the same body of data from which predictions are made. The joint distribution of breeding values and of records usually depends on unknown parameters such as means, variances and covariances in the multivariate normal distribution. If the objective of the analysis is to make selection decisions, these parameters should be considered as "nuisances". When their values are unknown, the state of uncertainty can be represented via a prior probability distribution. This is combined with the information provided by the data into a posterior distribution from which the needed predictors are calculated after integrating out the "nuisances". Prediction under alternative states of knowledge was studied and the corresponding solutions were developed. It was shown that modal estimators of variance parameters provide a reasonable answer to the problem. Because a Bayesian framework was adopted, the estimates so obtained are non-negative. Under multivariate normality and with completely vague prior knowledge about means and variances, it was shown that approximately "optimal" (in the sense of maximizing expected merit of the selected candidates) predictors are those obtained using the "mixed model" equations with the unknown variances replaced by restricted maximum likelihood estimates. This leads to empirical Bayes predictors of breeding value.

Foulley and Gianola (1986) described a procedure of sire evaluation for multiple binary responses when information is missing for some traits. The genetic model assumes a conceptual underlying distribution rendered discrete by abrupt thresholds. Statistical inferences are made from a posterior distribution combining a multivariate normal prior and several conditionally independent likelihood functions, each corresponding to a particular class of information available. Point estimators and predictors of fixed and random effects are the values that maximize the posterior distribution conditionally on heritabilities, genetic and residual correlations. The methodology provides a potential means for removing bias due to selection for categorical traits.

The principles used by Gianola et al. (1986) were employed to develop estimators and predictors for n polygenic binary traits, e.g., calving difficulty, calf mortality and calving "preparation" (Foulley et al. 1987). The conditional probability of an observation in a subpopulation i (a combination of levels of explanatory variables) falling into one of 2^n mutually exclusive and exhaustive categories was modelled using a normal integral in n -dimensions. The mean of the subpopulation was written as a linear combination of an unknown vector which can include "fixed" and "random" effects such as additive genetic value. Conditionally on θ , the normal integral depends on an unknown matrix R comprising residual correlations in a multivariate standard normal conceptual scale. The random variables in θ have a dispersion matrix $G * A$ where usually A is a known matrix of additive genetic relationships, and G is an unknown

matrix of genetic variances and covariances . It was assumed that a priori θ follows a multivariate normal distribution $f(\theta|G)$ which does not depend on R, and the likelihood function was taken as product multinomial. The point estimator of θ was the mode of the posterior distribution $f(\theta|Y, G=G^*, R=R^*)$, where Y is data, and G^* and R^* are the components of the mode of the marginal posterior distribution $f(G, R|Y)$ using "flat" priors for G and R. The matrices G^* and R^* correspond to the marginal maximum likelihood estimators of the corresponding matrices, and the point estimator of θ is of the empirical Bayes type. Computations involve solving three nonlinear systems in θ , G and R. G^* can be computed with an expectation-maximization type algorithm, and R^* can be estimated by a full maximum likelihood procedure.

Im and Gianola (1988) studied mixed nonlinear probit and logit models with independent and normally distributed random effects in a two-way nested layout. These models arise often in animal breeding and were motivated by lamb mortality data. The objective was to estimate "fixed" effects and variance components, and to predict random effects. Estimation was by maximum likelihood, with the integrations carried out by Gaussian quadrature. The simplex method, a derivative free algorithm was used as an alternative to the EM algorithm in computing maximum likelihood estimates. Random effects were predicted by their conditional means, given the data. The data set used to illustrate the procedures consisted of 1,119 mortality (live vs. dead) records on lambs from 491 dams mated to 37 sires. Logit and probit models essentially gave the same heritability estimates (.176) and sire rankings; however, the probit model required more computing time. The simplex method was preferred over the EM algorithm because it can be adapted to give the asymptotic covariance matrix, and it is less expensive to compute.

Foulley, Gianola and Planchenault (1987) developed a sire evaluation procedure for situations in which there is uncertainty with respect to the assignment of progeny to sires. The method requires specifying prior probabilities p_{ij} that progeny i is out of sire j . Inferences about "fixed" environmental and group effects, and "random" transmitting abilities of sires are based on Bayesian statistical procedures. Modal values of the posterior distribution of these parameters are employed as point estimators. Finding these modes entailed solving a nonlinear system of equations and several algorithms were suggested. The methodology was described for univariate evaluations obtained from normal or binary traits, and estimation of variances was also addressed.

The importance of cost saving measures in cattle production has increased on a worldwide basis because of, e.g., milk quotas in Europe, and the need to put a more competitive product in the consumer's market. Hence, secondary traits such as calving ease, time between first and last insemination within a lactation, productive lifespan and liability to diseases (e.g., mastitis, ketosis, milk fever) should be taken into consideration when formulating breeding objectives for dairy populations. The above auxiliary traits present a discrete distribution of phenotypes while their inheritance is assumed polygenic so they are referred to as quasi-continuous. The variance components needed to implement recently

developed methodology for analysis of such traits may not be known, in general. Hoschele, Gianola and Foulley (1987) studied Bayesian approaches to variance component estimation in threshold models for ordered categorical data. The methods are based on finding the mode of marginal or joint posterior distribution of variance parameters using flat or informative priors. The estimates correspond directly to parameters in the "underlying" scale. A simulation study was conducted to assess bias and mean squared error properties of the proposed procedures using alternative models to describe the data. The procedure based on the marginal posterior distribution and flat priors gave satisfactory estimates of heritability except when incidences were larger than 90% and subclass sizes were less than 2, in which case estimates were biased. Informative priors reduced considerably mean squared error of the estimates.

The new methods of genetic evaluation studied in this project use the threshold and liability concepts and assume binomial or multinomial sampling. However, there are situations where distributions other than the multinomial describe the data more suitably. For example, the number of veterinary treatments during lactation in a dairy cow, ovulatory response to exogenous gonadotropin and sibship size cannot be viewed as categorical responses, and are probably well described by a Poisson distribution. Further, when such traits are analyzed jointly with other discrete characters, e.g., calf survival, the resulting contingency tables show specific patterns, i.e., they are structurally "incomplete". Foulley, Gianola and Im (1987) developed a procedure of genetic evaluation for fitness traits under the assumption of polygenic inheritance. Using the concept of generalized linear models, logarithmic (ovulatory response) and probit (calf survival) functions are described by linear combinations of environmental and genetic effects; the liability of survival is expressed conditionally to the logarithm of ovulatory response. Inferences on parameters are based on the mode of their joint posterior density assuming a prior multivariate normal distribution. A method for estimating the dispersion parameters was also developed. A "truncated" Poisson distribution can be used to account for missing records on ovulatory response.

Hoschele, Rothschild and Gianola (1988) used Bayes theorem to develop a recursive procedure for estimating breeding values with threshold characters. The likelihood of the data at stage k was combined with the posterior distribution of the parameters conditional on the data after stage $k-1$, or prior for stage k . A second order Taylor series approximation to the prior density was used. The mode and the matrix of second derivatives of the posterior density at stage $k-1$ were used to approximate the prior mean and dispersion matrix at stage k . Hence, the recursive procedure only approximates the estimates that would be obtained from analyzing all data jointly. Under normality, the results are exactly the same. A numerical example was employed to illustrate the method. The proposed procedure would be useful in within-herd evaluations for calving ease, survival of the calf, liability to certain diseases and other categorical traits of economic importance, where only data on live animals

need be stored, provided that the size of the estimation equations permits inverting the coefficient matrix.

The assumption of linearity of the regression function is usually accepted by animal breeders. In the usual normal model, the regression of offspring on parent, on mid-parent, or on the two parents is linear. When observations are discrete, the normal model obviously does not hold. Im and Gianola (1988) studied offspring-parent regression for a binary trait (e.g., "healthy" vs. "diseased") without invoking the existence of an underlying normal "liability". The regression of offspring on a single parent is always linear; the maximum likelihood estimator of the slope being $n_{11}/(n_{10}+n_{11})-n_{01}/(n_{00}+n_{01})$, where n_{01} is the number of "healthy" offspring out of "diseased" parents, and so on. However, the regression of offspring on both parents is non-linear, the function being

$$E(Y|X_1=x_1, X_2=x_2) = \pi_{00} + (\pi_{10} - \pi_{00})x_1 + (\pi_{01} - \pi_{00})x_2 + (\pi_{11} - \pi_{10} - \pi_{01} + \pi_{00})x_1x_2$$

where Y is the score on the offspring and x_i is the realized value of the score on parent i ($i=1,2$). Further, the π 's are the conditional probabilities that the offspring has the attribute given the scores of the parents. Unless $\pi_{10} = (\pi_{00} + \pi_{11})/2$ (assuming $\pi_{01} = \pi_{10}$), the regression function contains an interaction between sire and dam. In a log-linear model without interaction between parents, the regressions are always linear on a logistic scale, and the slope can be estimated by maximum likelihood.

A Bayesian and a maximum-quasi-likelihood approach of nonlinear sire evaluation for binary threshold data were compared by simulation (Hoschele and Gianola, 1989). Using a protocol incorporating selection, the two methods were compared in terms of mean squared error of estimates of fixed group effects and of heritability, and mean true transmitting ability of the sires eventually selected at three intensities. The two procedures tended to overestimate true differences between groups, but not significantly so; the Monte-Carlo variance and mean squared error were smaller for the Bayes procedure in all cases examined. Heritability in the underlying scale was overestimated significantly by the two procedures, but estimated variance and mean-squared error were somewhat smaller for the Bayesian method. The two methods differed negligibly in true mean value of selected sires at the three levels of selection intensity.

Gianola, Im and Macedo (1990) discussed prediction of breeding value in a Bayesian framework under alternative states of knowledge. Using the concept of a loss function, it was argued that uniformly best predictors cannot be obtained in a classical framework, although an unequivocal Bayesian solution exists. Procedures to choose among alternative models, e.g., grouping strategies in dairy cattle breeding, and to predict future observations were also considered.

Gianola et al. (1990) extended the Box-Cox theory of transformations to univariate mixed models such as those arising in animal breeding applications. The objective of the transformation is to obtain a model where additivity, normality and constancy of variance hold so that the

analysis is simpler, and fewer parameters need to be estimated. The transformation parameter (λ) can be estimated from its marginal posterior distribution, from the joint distribution of λ and the variance ratio, or from the joint distribution of variance components and λ . Finding a model satisfying the features listed above is of current importance in dairy cattle breeding, e.g., selection of bull dams, because production testing occurs in a range of environments with heterogeneous genetic and residual variances. The transformed model can potentially reduce the extent of "unfair" selection decisions caused by differential variability among strata of the population.

As stated earlier, theory suggests that in the absence of information about variance parameters, breeding values should be predicted using BLUP with variances replaced by their REML estimates (Gianola et al., 1986). This strategy approximates the "best predictor" of breeding value because whereas uncertainty on fixed effects is taken into account, variances are approximately integrated out. This approximation is likely to be unsatisfactory when the restricted likelihood is "flat". Gianola and Macedo (1988) showed that the variances can be exactly integrated out, which solves an important problem in univariate estimation and prediction. Using a Bayesian model with flat priors for the fixed effects and inverted gamma distributions for the variance components they showed that the joint posterior distribution of fixed effects and breeding values is in the form of a poly-Student distribution. The marginal distribution of the random effects is also in the form of a poly-t. Algorithms for computing the modes of these distributions were developed; these algorithms involve an "extended" form of the mixed model equations. If "flat" priors are used for the variance components, the problem can degenerate because this prior gives too much weight to zero. On the other hand, the inverted chi-square distribution is very flexible and its hyper-parameters can be specified so as to convey "little" prior knowledge (if this is the case) or a sharp prior assessment of probabilities. Macedo and Gianola (1987) also addressed this problem using a precision component (the reciprocal of a variance) parameterization, and several such components. Distributions and point estimators were presented for joint, conditional and marginal estimation of fixed effects, random effects and variance components.

Weller and Ron (1987) studied the effect of missing data on estimates of genetic correlation. Various studies have estimated covariance components as half the difference between the variance component of the sum of the variable values for each observation and the sum of the corresponding variable variance components. Although the variance components for the separate variables can be estimated using all available data, the variance components of the sum can be computed only from those observations with records for both variables. Previous studies have suggested eliminating observations with missing data, because of possible selection bias. The effect of missing data on estimates of covariance components and genetic correlations was tested on sample beef cattle data and simulated data by randomly deleting differing proportions of records of one variable for each pair of variables analyzed. Estimates of genetic correlations obtained with observations with missing data eliminated were

more accurate than estimates computed using all available data. Furthermore, when observations with missing data were included, estimates of genetic correlation far outside the parameter space were common. Therefore, this method should be used only if observations with missing data have been eliminated, provided data are missing at random.

VI. INFERENCES UNDER SELECTION

Prediction of breeding value by maximization of the joint density of the random effects and of the data (M) was studied in a conditional selection scheme proposed by Henderson (Gianola, Im and Fernando, 1988a). Dispersion parameters were assumed known. A condition for being able to ignore selection when using such predictors is that the distribution of the selection variable must not depend on the location parameters being estimated; the same applies to the conditional distribution of the selection variable, given the data. Selection can be ignored when culling is based on linear or nonlinear functions of the data that are not a function of the fixed effects. Contrary to the results given by Henderson for best linear unbiased estimation and prediction, when using M predictors, selection can be ignored when it is based on linear or nonlinear functions of the residuals, or of the random effects. Implications of these findings on expected rate of genetic progress are discussed.

Henderson's treatment of selection was also revisited from the viewpoint of parameter estimation by maximum likelihood (Gianola et al., 1988). The likelihood function (L) in his conditional selection scheme is given by the ratio between L without selection and the probability of selection (P). Hence, a sufficient condition for ignorability of selection is that P does not depend on the parameters being estimated. For example, when estimating fixed effects (β) with known dispersion parameters, the condition is $dP/d\beta=0$. If selection is based on functions (linear or non-linear) of the data available for analysis, this is satisfied when the selection function is translation invariant. However, selection based on functions of the random effects or of the residuals can be ignored for this purpose. On the other hand, when estimating fixed effects and dispersion parameters, selection cannot be ignored when the distribution of the selection function depends on any of these. In general, selection cannot be ignored for parameter estimation in Henderson's scheme, although it is possible to do so in unconditional selection models.

The conditions for being able to ignore selection are dependent on the way that selection is modelled, as shown by Fernando et al. (1988). Consider a simple example where the better of two cows is selected to have a second lactation. Here, selection is based on $y_0=(y_{11}, y_{21})$ where y_{ij} is lactation record j of cow i . The data vector available for estimation is $y_{s1}=(y_{11}, y_{21}, y_{12})$ when $y_{11} > y_{21}$, or $y_{s2}=(y_{11}, y_{21}, y_{22})$ when $y_{11} \leq y_{21}$. In constructing the likelihood $f(y_{sk}|\theta)$, if k is treated as the realized value

of a random variable, the likelihood equations are identical to those ignoring selection. Note that with this formulation of selection, y_0 is always sampled, and the distribution of y_0 is not altered by selection. However, some authors have used the conditional likelihoods $f(y_{s1}|y_{11}>y_{21})$, or $f(y_{s2}|y_{11}\leq y_{21})$, instead of $f(y_{sk}|\theta)$, for estimating θ . Working with conditional likelihoods gives different results, because using these implies that the distribution of y_0 is altered by selection. However, it should be clear from the example that y_0 is always sampled at random so its distribution should not change. Therefore, the use of conditional likelihoods with this type of selection is not justified.

A reasonable framework to study parameter estimation by maximum likelihood under selection is that provided by what is called "missing data" theory (Im et al., 1989). The data vector y is partitioned into observed values, y_0 and missing values y_m . Let $f(y|\theta) = f(y_0, y_m|\theta)$ be the probability density of y , and θ be an unknown parameter vector. Define r to be a vector with the i th element taking value 1 if y_i is observed and 0 if y_i is missing. In analysis of selected data, the pattern of records available for analysis, characterized by the value of r , should be considered as part of the data. The density of y and r can be written as $f(y, r|\theta, \psi) = f(y|\theta) \cdot f(r|y, \psi)$ where $f(r|y, \psi)$ is the conditional density of r given the complete data vector and ψ is a parameter of this density. Now, the density of y_0 and r is $f(y_0, r|\theta, \psi) = \int f(y_0, y_m|\theta) f(r|y_0, y_m, \psi) dy_m$. It was shown that working with this density is equivalent to working with the marginal density of y_0 ignoring selection when 1) the missing data are missing at random, i.e., $f(r|y_0, y_m, \psi) = f(r|y_0, \psi)$ for all ψ and y_m evaluated at the observed values of r and y_0 , and 2) the parameters θ and ψ are distinct, in the sense that the joint parameter space of (θ, ψ) is the product of the parameter space of θ and the parameter space of ψ . When selection is based on y_0 , $f(r|y_0, y_m, \psi) = f(r|y_0, \psi)$ and thus, selection can be ignored.

VII. COMPUTING TECHNIQUES

1) Linear models

Wiggans and Misztal (1987) computed evaluations with an animal model using 110,278 Ayrshire first lactations. The model included fixed herd-

year seasons (20,322), genetic groups (12) and random cows and residual effects. Sires (7641) were included through the relationship matrix. Groups were defined by year of birth and presence of dam in the data. Best linear unbiased predictions were obtained without forming the mixed model equations. Solutions for herd-year-seasons (HYS) were by Gauss-Seidel iteration; other solutions were by second order Jacobi. This method does not require sorts by effects other than HYS. Iteration was on a CRAY XMP-24 supercomputer. Computing time was minimized by adapting DO loops for vector processing, by reading records in groups of 500, by coding data so that animal identification was obtained from record number, and by including an indicator of last record in HYS in the group variable. Relaxation factors from .3 to .9 were tested on 100 rounds of iteration with an optimum .83. Then, 400 rounds of iteration were performed requiring about 1 sec of central processing unit time per round. This suggests that the method may be practical for larger data sets. Nonlinear adjustment was tested, over a range of relaxation factors, after 100 rounds. This adjustment achieved accuracy that normally would occur 50-80 rounds later. At some relaxation factors, 125 to 150 rounds were required before nonlinear adjustment was beneficial. Results on optimum relaxation factors and nonlinear adjustments may differ for other data. Additional work is planned to account for repeated records, missing information on parents of bulls, sire by herd interaction, and larger data sets.

Large scale genetic evaluation of animals by best linear unbiased prediction and nonlinear procedures can have a high computational cost. This is partly due to the need to set up mixed model equations, which are then solved in an iterative way. Solutions can also be obtained by successive averaging data without setting the mixed model equations directly. Misztal and Gianola (1987) developed formulae for a class of models with fixed and random factors, including an additive relationship matrix. Two iterative procedures were investigated, Gauss-Seidel and Jacobi. With a balanced data set, putting restrictions on fixed effects is not effective for improving convergence rates in Gauss-Seidel, but essential in Jacobi. Computing techniques needed to implement the indirect procedures were discussed.

Misztal, Gianola and Schaeffer (1987) simulated a population of animals to study convergence rate of iterative methods for solving mixed model equations for an animal model. The solutions drifted for many rounds, and their true accuracy was far from that suggested by criteria such as the difference between right and left-hand sides, or a modified difference between consecutive solutions. The drift in later rounds of iteration closely followed an exponential curve. Formulae were presented to estimate the true solutions, the true accuracy of solutions in a given round, and the number of rounds needed to increase the accuracy of solutions by a given factor.

Wiggans, van Dijk and Misztal (1988) applied an animal model to evaluate 120,073 Alpine, LaMancha, Nubian, Saanen and Toggenburg bucks and does. Parities higher than six were excluded. The model included fixed herd-year-season and a random sire x herd interactions, breeding value,

permanent environmental effect and a residual. Breeding value was the sum of an accumulated group effect and of additive genetic value. Records made by a doe in different herds were treated by predicting a separate permanent environmental effect for each herd. Lactations were weighted according to their length, and does of all breeds were used as contemporaries. Breed differences were accounted for with a grouping strategy that grouped unknown parents by breed, sex of parent, sex of progeny and birth-year of progeny. Univariate evaluations for milk and fat were obtained. One hundred iterations were carried out to achieve a 3-digit accuracy; each iteration required slightly over 4 seconds of CPU in a CRAY XMP/48 computer. Predicted producing ability was the sum of predicted breeding value, permanent environmental effect and herd x sire interaction. This study served as a test for the animal model evaluation system being developed by the USDA for dairy cattle.

Wiggans, Misztal and Van Vleck (1988) studied an animal model evaluation of milk yield in Ayrshires. The model included 32,159 fixed herd-year-seasons, random herd x sire interactions (38,287), permanent environmental effects (119,541), animal and residual effects. The animals evaluated consisted of 119,541 cows with 301,799 records, 5762 sires and 11,893 dams without records. Genetic groups (36) were defined to represent unknown ancestors, and ancestors not contributing "ties" or records. Groups were defined by sex of ancestor, and birth year and sex of progeny of ancestor. These distinctions allowed groups to account for trend and differing selection intensities for ancestors of bulls and cows. The animal effect included combinations of these group effects derived from tracing each path in the pedigree back to an unknown parent group. Iteration for herd-year-season, permanent environment and herd x sire was by Gauss-Seidel, while that for other factors was by Jacobi. Iteration was conducted without forming mixed model equations; instead, a copy of the data sorted by herd and sire was read each round. To evaluate computer time requirements, solutions were also calculated with and without herd x sire effects, and with first versus all lactations included. About 3.5 seconds of CPU (CRAY XMP/48 computer) were required per round of iteration; herd x sire effects and later lactations accounted for .7 and 1.6 seconds, respectively. The study demonstrated the feasibility of carrying out genetic evaluations at a national level with an animal model and available computer resources.

Computation of prediction error variance (PEV) of genetic evaluations obtained by mixed model methodology requires inversion of the coefficient matrix. It is not feasible to do this in large data sets. Methods have been developed for estimating PEV in sire models, but these are not suitable for animal models because effects of the relationship matrix are not taken into account. Misztal and Wiggans (1988) developed an approximation to PEV in animal models that offers considerable promise. The reciprocal of the PEV for each animal is approximated by a linear function of contributions due to production records (if any) and due to relationships. The contribution due to records is a weighted number of records, where the weights reflect the distribution of records across factors; weighing techniques developed for sire models can be used. The contribution due to relationships is a sum of contributions from parents

and from offspring. An iterative algorithm was developed that combines both kinds of contributions, and PEV is estimated for both producing and non-producing animals. The accuracy of the algorithm was investigated using a simulated data set consisting of 3 generations of animals comprising 1000 cows, 40 sires, 100 herd-year-seasons and 2315 records. In addition to additive genetic values, the model included herd-year-season and permanent environmental effects. The correlation between "repeatability" estimates obtained by the algorithm and by inversion was .996. Storage requirements of the algorithm are 3 variables per animal. Iteration involves reading the file with records once, and reading the relationship file once per round (7 rounds were required in the simulation).

Da et al. (1988) studied computation of prediction error variance and of restricted maximum likelihood estimates of variance in an animal model with Westell-Quaas grouping techniques. Algorithms using a transformed coefficient matrix seem easier to implement because of easier computation of traces; sparse matrix techniques can reduce computations further.

2) Threshold models

The threshold model equations were shown to correspond to a "pseudo-linear" model which is simple to program and to test (Misztal et al., 1987, 1989). A modification of Jacobi iteration "on data" was developed that does not require setting up the system of equations explicitly.

For some models it has the convergence properties of Gauss-Seidel iteration but it does not require sorted files and it is simpler. Two threshold model programs were developed: an iterative and an absorption-inversion version. The programs, using the "pseudo-linear" concept, were applied to a set of data (5000 observations with 3 effects and 3 categories) to investigate convergence properties and several computing options. Runs did not terminate normally when the data contained levels of fixed effects with only observations in an extreme category. This was caused by numerical errors in computing differences between normal integrals, which appear in denominators of several formulae in threshold model equations. The problem was solved by not allowing such differences to be smaller than a certain level. The absorption-inversion program behaved erratically when the model was not of full rank, even though a generalized inverter was used; this occurred because numerical accuracy problems did not allow the inversion program to detect dependencies between thresholds and other effects. The program using inversion needed 4 rounds of iteration to reach approximately a two-decimal digit accuracy in sire solutions. The iterative version was fastest when only one linear round was performed for each non-linear round, requiring 6 rounds for the same accuracy. In either case, setting the first threshold to zero had no negative influence on convergence. A round of iteration for the program using inversion required 837, 11.8 and 1.24 s of CPU on a Personal Computer AT, an IBM 3081 "mainframe", and a CRAY X-MP/48 supercomputer, respectively. Corresponding times for the iterative program were 55.2, 2.6 and .57 seconds. In the CRAY, relatively little time was spent on inversion while reading and writing data was relatively efficient in the

Personal Computer AT. This study provided the bases for application of the programs to field data.

Misztal et al. (1988) were able to relax the assumptions of normality and constant variance in some threshold models. Estimation equations in threshold models were originally derived assuming that the residuals of the underlying variable are normally and identically distributed. This may not be true in general. For example, in sire-maternal grand-sire models with missing identification the residual variance is not homogeneous. The threshold model can be modified to account for heterogeneous variance in two steps. First, for observation j , arguments in normal distribution functions are divided by the standard deviation of the residual (σ_j). Second, the "contributions" of each observation to the coefficient matrix and right-hand side are divided by σ_j^2 and σ_j , respectively. This differs from the linear model, where "contributions" are divided by the same quantity, the residual variance. When the underlying variable is not normal, it was shown that the threshold model still holds, provided that liability can be transformed to achieve normality, even if the transformation is not known. Departure from normality causes shifts in the values of thresholds, but solutions for fixed and random effects can be calculated as if normality held. The essential aspect is that the same probability distribution is generated irrespective of changes in the point estimates.

VIII. BASIC QUANTITATIVE GENETICS

Chang et al. (1988a) noted and resolved a peculiarity of the tabular method for computing additive relationships between individuals of different generations.

Inverses of dominance and epistatic relationship matrices are needed in order to predict non-additive genetic values efficiently in outbred populations of livestock. For field data sets, these inverses cannot be computed by direct procedures because of the large dimension of the matrices involved. Chang et al. (1988b) developed a computationally efficient method for inverting an additive x additive relationship matrix. The additive x additive (AxA) epistatic value of an individual can be expressed in matrix notation as $u=1/4Pue$, where u and e are AxA effects and residuals, respectively, and P is a transition matrix. The i th row of P contains at most two non-zero elements corresponding to the known parents of individual i . These elements are $p_{is}=1+2a_{sd}a_{dd}[a_{ss}a_{dd}+a_{sd}^2]^{-1}$ $p_{id}=1+2a_{sd}a_{ss}[a_{ss}a_{dd}+a_{sd}^2]^{-1}$ for sire and dam, respectively. Then $\text{Var}(u)=(I-1/4P')^{-1}R(I-1/4P')^{-1}\sigma^2=G\sigma^2$ where $R\sigma^2=\text{Var}(e)$. Thus $G^{-1}=(I-1/4P')R^{-1}(I-1/4P)$. Whereas R^{-1} is required for computing G^{-1} , calculating R^{-1} is easier because R is sparse. Without inbreeding, the off-diagonals of R are null except for elements corresponding to individuals related through two lines of descent. The i th diagonal of R is: $r_{ii}=1$ when both parents

of i are unknown, r_{ii} when one parent is known, or $r_{ii}=.875$ when both parents are known. In general, for individuals i and j with parents (x,y) and (w,v) , respectively, the elements of R are: $r_{ii}=(1+1/2a_{xy})^2-.0625[(p_{ix}a_{xx})^2+(p_{iy}a_{yy})^2+2p_{ix}p_{iy}a_{xy}^2]$, and $r_{ij}=(a_{ij})^2-1/4[p_{ix}a_{jx}^2+p_{iy}a_{jy}^2+p_{jw}a_{iw}^2+p_{jv}a_{iv}^2]+.0625[p_{ix}p_{jw}a_{xw}^2+p_{ix}p_{jv}a_{xv}^2]$. Because R is sparse iterative methods can be used to invert it. However, inverting A/A is, in general, much more difficult than inverting A .

Chang et al. (1988c) considered the problem of ranking sires for a trait exhibiting epistatic genetic variance. The point of departure was to use the result that genetic merit in the progeny of a sample of selected animals is maximized by ranking sires with the conditional means of future progeny performance. Under normality, these conditional means can be obtained solving the mixed model equations. With an additive trait, the conditional mean of the genetic value of a future offspring, given a sire evaluation (u_a) is: $(1+a_{sd})u_a$, where a_{sd} is the relationship between sire and dam. With additive by additive epistasis the corresponding conditional mean of total genetic value is:

$$(1+a_{sd})u_a + [1+a_{dd}k_o + (1+a_{ss}k_o)a_{sd}^2]u_{aa}$$

where $k_o=2a_{sd}[a_{ss}a_{dd}+a_{sd}^2]^{-1}$ and u_{aa} is the mixed model solution for 1/4 of the epistatic effect of the sire. In the additive case, ranking of sires is "free" of a_{sd} when this is constant across sires. With epistasis, ranking may depend on a_{ss} , a_{dd} and a_{sd} . A case where this does not occur is when $a_{sd}=0$ because the conditional mean of the genetic value of an offspring reduces to u_a+u_{aa} .

IX. DAIRY CATTLE BREEDING

1. Analysis of lactation curves

To maximize genetic progress, environmental influences that reduce accuracy of selection should be eliminated. Methods were developed to adjust production records for each cow's previous reproductive history. Adjustment factors for Israeli Holsteins were estimated by Ezra, Weller, and Bar-Anan (1986) for calving age and month. Factors to extend partial lactations were developed by Bar-Anan, Weller, Wiggans, and Osterkorn (1986). Weller (1988) developed a method to differentially weigh lactations in progress in the computation of genetic evaluations; it is only slightly more complicated than procedures currently in use.

A multiphasic lactation function was fitted by De Boer et al. (1989) to first through third parity data for milk and fat yields of Israeli Holsteins. Cows were grouped by yield and by days open classes. A triphasic function, the sum of three logistic functions, was fitted to

milk yield, whereas a diphasic function, the sum of two logistic functions, was fitted to fat yield. For first parity milk, about 5% of yield occurred in first phase, about 25% in second phase, and about 70% in third phase. For first parity fat, about 14% of yield occurred in first phase and about 86% in second phase. Effect of yield was significant for initial, peak, and 305-d yields, generally. Effect of days open tended to be significant for time of peak yield and duration of each phase.

Weller and Forman (1989) studied the effect of days open and of days to first breeding on cumulative calf and milk production in the current and following lactation. The analysis was based on 121,627 first through third parity annualized economically fat-corrected lactations. Optimum first parity days open ranged from 110 for low calf value (500 kg milk) to 91 for high calf value (4000 kg milk). Optimum second parity days open ranged from 91 for low calf value to 40 for high calf value. Expected production as a function of days to first breeding was calculated as a function of the probability of conception, which was varied from .4 to .6, and of estrus detection, which was varied from .5 to .7. Optimum days to first breeding as a function of calf value and reproductive management level ranged from 95 to 65 days for first parity, and from 77 to less than 40 for second parity. Optimum days to first breeding was higher with lower calf value, and better reproductive management. Expected losses from early first breeding (40 days in milk), as compared with the optimum, ranged up to 780 kg milk for first parity cows, while expected losses from late first breeding (120 days in milk) ranged up to 1000 kg for second parity cows.

2. Health and disease

Dr. O. Distl, while on sabbatical leave from the Munich Veterinary Institute (Federal Republic of Germany), analyzed data on six reproductive syndromes from 102 Kibbutz herds with 76,170 recorded calvings. The syndromes were: retained placenta, lochia (vaginal discharge during the first two weeks after parturition), endometritis, postpartum anestrus, ovarian cysts with anestrus and ovarian cysts with nymphomania. Frequency was highest for anestrus postpartum, about .35. The frequency of all other syndromes was under .10. The same model of analysis was used for all syndromes and for two additional traits: all diseases, and number of disease occurrences during lactation. The effects of herd-year-season, sire, cow's age, days pregnant, and an interaction between sex of the previous calf and the difficulty of the calving were included in the model. Only the sire effect was assumed random. Coefficients of determination ranged from .05 for retained placenta to .23 for anestrus postpartum. Herd-year-season explained more of the variation than any of the other effects in the model. All heritabilities were under .04. Genetic correlations among retained placenta, lochia and endometritis were larger than .6. Genetic correlations among other syndromes were lower. Although recording in this data set was more reliable than in most data sets used previously to estimate genetic parameters for disease traits, heritabilities were not higher than previous estimates.

Weller (1988c) analyzed environmental effects on somatic cell concentration (SCC) in Israel using 30,000 lactation records with at least one SCC estimate during lactation. Mean SCC was about 450,000 cell/ml. SCC was lowest during the second month postpartum, and increased during the remainder of the lactation. SCC increased with parity up to the sixth. Correlations between the log of SCC in consecutive months were from .45 to .63. Correlations between log SCC four months apart were from .32 to .53. Correlations among actual SCC were lower. A sample of 7047 cows in 24 herds were scored for SCC and tested bacteriologically for udder infection. Log SCC scores before and after bacteriological assay, the mean of these scores, and the mean of all SCC scores during lactation were analyzed as functions of herd, parity, and infection status, i.e., presence or absence of bacterial growth in the sample. Days in milk when SCC was scored was included in all analyses, except the last. All effects included in the models were significant. SCC was higher for infected cows, as expected. The coefficient of determination was highest (.35) for the mean of the two SCC scores nearest in time to the bacteriological test. The difference between infected and non-infected cows for this analysis was 0.31 on the log scale, which corresponds to a factor of two on the linear scale. Less than 5% of cows with an SCC of less than 160,000 in the milk sample taken after the bacteriological assay were infected, while 25% of cows with an SCC above 400,000 were infected.

3. Causes of cow culling in Israel

Studies on criteria for cow culling in Israel centered on the assessment of genetic relationships that may exist between cow culling and its main causes: low production, infertility and mastitis. Sires were evaluated for first and second parity daughter culling. Each parity was analyzed separately. Sires were grouped by year of birth, but the analyses were conducted on sire solutions, rather than on the sum of sire and group effects. This was done to obtain a more meaningful comparison between younger and older sires. Because there is a positive genetic trend for production, older sires would tend to have a higher culling rate than younger sires. Correlations between sire solutions for culling and sire evaluations for production, fertility and mastitis were used to estimate genetic correlations between these traits. First parity culling was found to be highly genetically associated with low production. Genetic correlations between second parity culling, first parity culling and the other traits were all low, but in the expected directions; that is, daughters of sires with higher culling rates have lower milk production and fertility and higher mastitis rate. Further study is necessary to find useful criteria for between-cow selection.

4. Methods of cow evaluation

There are three main objectives in cow evaluation: selection of superior cows to be used as bull dams, selection of cows to serve as embryo donors, and identification of inferior cows to be culled from the population. Preliminary results indicate that under Israeli conditions, criteria for selection for these objectives may differ.

In most advanced breeding programs, bull dams are selected based on their genetic evaluations for production and conformation traits. In Israel, selection of bull dams based only on genetic evaluations may lead to a significant increase in inbreeding and to reduction of the genetic base of the population. This is because in the relatively small Israeli population, most cows with the highest genetic evaluations will be daughters of only a few bulls. It is, therefore, suggested that selection of bull dams be based on an index of the cows' genetic and phenotypic evaluations for production and conformation traits.

A method for more accurate genetic evaluation of cows was developed by Weller, Wiggans and Lawlor (1987). This method computes evaluations as back solutions from the sire evaluations, but includes the genetic level of the dam in the final cow evaluations. The method was tested on the complete Israeli populations, by comparing bull dam evaluations to their sons' genetic evaluations. Correlations were near the expected values based on the dam and son repeatabilities.

5. Calving traits

Ron, Bar-Anan, and Weller (1986); Weller, Ezra, and Bar-Anan (1986); Weller, Misztal, and Gianola (1988) and Weller and Gianola (1989) studied calving traits in the Israeli Holstein population. Calvings of 106,751 Israeli heifers were analyzed for dystocia and calf mortality, scored dichotomously, and a composite trait scored trichotomously. Dystocia was also studied with 146,973 second and third parity records. Data were analyzed by both linear (LM) and threshold models (TM). Models fitted included herd-year-season (HYS), sex of calf, calving age, calving month, sire of cow, sire of calf, and groups of sire of cow and of calf. HYS, sire of cow and calf, and residuals were random, with diagonal variance-covariance matrices. The HYS variance was assumed to be 10% of the residual component. Other variance components were estimated by REML for the LM analyses and by the counterpart of REML for the TM analyses. Heritability estimates were two to five times greater in threshold than in linear models, but correlations between corresponding sire evaluations were all larger than .9. LM sires evaluations were skewed positively, whereas TM evaluations had symmetrical distributions. Heritability for dystocia was greater in first than in later parities. Correlations between first and later parity sire evaluations were lesser than .5. Thus, the genetic control of dystocia seems to be different for heifers and cows. Correlations between sire of cow and calf evaluations were lesser than .3. Correlations between dystocia and calf mortality evaluations were about .7.

Dystocia and calf mortality were also analyzed without including the herd-year-season effect, and dystocia was analyzed separately for male and female calvings. Correlations between evaluations computed with and without herd-year-season effects were .9 for sire of calf evaluations for dystocia and greater than .97 for all other evaluations. Correlations between sire evaluations computed separately for male and female calvings were between .3 and .7. Thus, calving difficulty is expressed differently

in male and female calvings. Genetic trends were unfavorable for dystocia, but favorable for calf mortality. Phenotypic trends were curvilinear, with maxima in 1982 and 1983 for calf mortality and dystocia, respectively.

6. Genetic analysis of twinning rate in Israel

Twinning rate was analyzed in 221,348 second and third parity calvings of Israeli Holsteins, daughters of 473 sires. Mean rate of twinning was 2.4% and heritability was 2.4%. The range of BLUP sire evaluations was 8.7%. Evaluations were positively skewed, while log-transformed evaluations were symmetrically distributed. Correlations between sire and maternal grandsire evaluations were near the expected value of .5. Thus, there is no evidence for segregating major genes, or for nonadditive inheritance for this trait. Correlations between sire evaluations for twinning rate and milk production, fat percent, dystocia, and carcass gain were .14, -.13, -.28, and -.34, respectively. Thus, there are no indications of economically important adverse genetic correlations with twinning rate. Breeding for increased twinning rate may be feasible, by selection of high-repeatability sires with positive evaluations for twinning rate.

7. Genetic and environmental relationships between fertility and production in Israel.

Despite the economic importance of fertility in dairy cattle there has been little emphasis on breeding for either male or female fertility, because heritability for most measures of fertility is low. Because field data generally available consist of insemination dates, most studies have assessed female fertility by non-return rates, the number of times a cow is bred during a lactation, or by days open. Several studies found negative genetic correlations between fertility and production traits, while others found no significant relationships. Fertility traits pose difficulties in genetic analysis. Cows that do not conceive do not have a value for days open. Therefore, studies of days open have either deleted cows that did not conceive from the analysis, or used some arbitrary values such as the date of last insemination. All of these options will tend to cause bias. The first approach is undesirable because cows with the most obvious fertility problems are eliminated from the analysis. Among cows that do not conceive readily, farmers are more likely to re-inseminate those with superior milk production, or cows that are daughters of elite sires. Thus, cows with long open periods will not be a random sample with respect to either production or pedigree. Furthermore, since data on conception are not generally available, neither number of breedings nor non-return rate distinguish whether or not the cow conceived on a given insemination. Thus, for cows that did not conceive, a low number of inseminations is not necessarily indicative of high fertility.

Field data collected in Israel are unique in that pregnancy status is diagnosed by veterinary examination for all cows that do not display estrus within 60 days of insemination. Weller (1989) estimated genetic

and phenotypic parameters of fertility using these data, and genetic and environmental trends for fertility and production traits in the Israeli Holstein population.

Data were first through third parity insemination and production records of Israeli Holsteins inseminated between January 1, 1980, and July 31, 1986. Records were eliminated if: 1) no insemination was recorded prior to 153 days in milk; 2) days in milk were less than 34, or number of monthly tests was less than two; 3) previous parity records of the cow were missing; 4) fewer than 20 daughters per sire; or 5) otherwise abnormal. The edited data set consisted of 91,770 cows, daughters of 296 sires. This sample included daughters of the 26 sires in the international sire comparison experiment (Bar-Anan et al., 1987, Livest. Prod. Sci. 17:305).

Fertility traits analyzed were conception status (CS) [1/number of inseminations to conception], and days from calving to first breeding (DFB). Cows that were inseminated but did not conceive, as determined by veterinary examination, were scored as zero for CS. Cows with more than five inseminations were given the same score as those with five inseminations. Fertility traits were adjusted for the effect of parity, using additive adjustment factors. A frequency table of fertility traits for first parity cows is given in Table 1.

Table 1. Frequency table for fertility traits of 91,770 first parity cows.

No. of inseminations to conception	Conception status	Frequency		Days to first insemination	Frequency	
		no.	%		no.	%
1	100	43993	48	<50	1617	2
2	50	19603	21	51-70	25594	28
3	33	9999	11	71-90	34570	38
4	25	5205	6	91-110	19238	21
>4	20	6115	7	111-130	7155	8
No conception	0	6855	7	>130	3596	4

For sire evaluation, it was assumed that the sire, cow, and residual variance components were 1, 9, and 90% of the total variance for CS and DFB; and 5, 45, and 50% of the variance for milk and fat production. Groups were formed on the basis of year of birth for the general population, and by country of origin for the 26 test sires. HYS effects were absorbed, while sire and group solutions were obtained by iteration. Sires were assumed to be unrelated. Sire evaluations were computed as the sum of each sire's group effect plus the sire solution.

Variance components were estimated separately for first and second parity by Henderson's method 3 (HM3). The model was the same as described above, except that a cow effect was not included because parities were

analyzed separately. In the first parity analyses, only cows with records on all four traits were included.

First parity (co)variance components for the four traits are given in Table 2, and the corresponding heritabilities and the genetic and environmental correlations are in Table 3. All the genetic and environmental correlations listed are within the range of .3 to -.3, except for the correlations between milk and fat production, which are similar to previous results of Weller, Ron, and Bar-Anan (1986) for these traits. The next highest genetic correlation is that between milk production and DFB. All the remaining correlations are within the range of .2 to -.2. The genetic correlation between DFB and production traits is in agreement with those reported in other populations. The virtually nil genetic correlation between CS and production traits is probably not significantly different from the value of -.2 found by Hermas et al. (J. Dairy Sci. 70:1252) on a small sample of Guernsey cows. However, the definition of CS was different in that study. Environmental correlations between fertility and production traits were also close to zero, similar to results found for annualized yields, and 305-day yields corrected for days open. Most previous studies did not correct production traits for the effect of days open. Heritabilities for fertility traits were similar to those from previous studies in different populations. Heritabilities for yield traits were lower than most results for Holsteins, but consistent with results of Weller, Ron and Bar-Anan (1986).

Table 2. First parity (co) variance components for conception status (CS), days to first breeding (DFB) and production traits.

Variance component	Trait	Conception status	Days to first insemination	Milk	Fat
Sire	CS	10.66	-.18	-.14	-.81
	DFB		4.38	1.32	2.29
	Milk			5.32	10.57
	Fat				32.03
Residual	CS	1220.8	18.0	10.4	57.2
	DFB		362.9		9.4
	Milk			152.5	354.8
	Fat				1116.0

Table 3. Heritabilities (on the diagonal), genetic correlations (above the diagonal), and environmental correlations (below the diagonal) among first parity fertility and production traits.

	CS	DFB	Milk	Fat
Conception status	.035	-.026	-.019	-.044
Days to 1st breeding	.027	.048	.271	.194
kg milk	.024	.045	.135	.810
kg fat	.049	.015	.860	.112

First and second parity heritabilities for fertility traits, and genetic and environmental correlations between first and second parities for these traits are in Table 4. Heritabilities were slightly lower in second parity. Similar to previously published results, genetic correlations were lower than .7 while environmental correlations were lower than .1. Although genetic control of different parities is not identical, it is similar, and the assumption of a common sire component of variance for all parities included in the model seems reasonable. Conversely, since both heritability and the environmental correlation are low, and selection is based on actual performance, second parity results should not have been appreciably biased by selection.

The frequency distribution of the sire evaluations for CS and DFB are given in Table 5. The 26 sires evaluated in the international sire comparison were deleted from this table. The range of repeatabilities was from .22 to .98, with a mean of .56. Even though the heritability of CS was less than .04, more than 10% of the sires were outside the range of -4 to 4% for CS. Of these sires, 16 (53%) had repeatability larger than .5. Thus, real genetic differences do exist for these traits, as previous studies have indicated.

Table 4. Heritabilities and genetic and environmental correlations for first and second parity fertility traits.

	Conception status	Days to 1st breeding
Heritability, parity 1	.035	.048
Heritability, parity 2	.022	.031
Genetic correlation	.792	.732
Environmental correlation	.050	.061

Table 5. Frequency distribution of 273 sire evaluations for fertility traits.

<u>Conception status</u>			<u>Days to first insemination</u>		
Range	<u>Frequency</u>		Range	<u>Frequency</u>	
	No.	%		No.	%
<-4	14	5	<-2	17	6
(-4)-(-2)	38	14	(-2)-(-1)	57	21
(-2)-0	91	33	(-1)-0	71	26
0-2	76	28	0-1	58	21
2-4	38	14	1-2	54	20
>4	16	6	>2	16	6

Holstein strain differences for the Israeli, US and Swedish sires from the international sire comparison are given in Table 6. Canadian and New Zealand sires were not significantly different from Israeli sires for CS so, therefore, these data are not presented. Swedish sires were most favorable for fertility, but low for production traits. Israeli sires were highest for production, but intermediate for fertility, while US sires were lowest for fertility, and intermediate for production. Although differences among strains were significant ($p < .01$), the strain effects are based only on 5-6 sire elite sires from each country. These differences may reflect different national selection criteria. In Israel, breeding has been for annualized yield adjusted for days open (Bar-Anan, J. Sci. Agri. Soc. Finland, 55:489, 1984), while in the US breeding has generally been for 305-day yield, unadjusted for days open (Dickinson et al., Agric. Res. Serv., USDA Prod. Res Rep. No. 165:18). Because the genetic correlation between 305-day yields and days open is positive, selection in the US for high yields may have led to selection for low conception rate.

Table 6. Holstein strain differences for production and fertility.

Traits	<u>Strains</u>		
	Israel	U.S.A.	Sweden
Conception status	3.32	-2.24	9.44
Days to 1st insemination	1.24	-.60	-1.25
Kg milk	817.9	432.1	-103.4
Kg fat	16.5	6.8	2.4

8. Genetic, environmental and phenotypic trends for production and secondary traits.

Weller, Ron, and Bar-Anan (1986); Weller (1989); and Weller and Ron (1989) estimated phenotypic, genetic, and environmental trends for production and fertility traits in the Israeli Holstein population. Phenotypic trends for production and fertility traits in Israel were computed as the regression of each cow's first parity record on her first parity calving date. Genetic trends were estimated as twice the regression of each cow's sire evaluation on the cow's first parity calving date. Environmental trends were estimated as the difference between the phenotypic and genetic trends. Trends for milk and fat production were estimated for two time periods: 1976 through 1985, and 1980 through 1986. The procedure of doubling sire evaluations, which assumes equal genetic progress on the dam side, although commonly used for production traits, may lead to overestimating genetic trend because progress in the dam line tends to lag behind progress in the sire line.

The genetic, environmental and phenotypic trends are in Table 7. Genetic trends for milk and fat production were larger during the later time period. Environmental trends were positive during the earlier time period, and negative during the later. Genetic trends were economically favorable for milk, fat, and economically fat corrected milk (FCM) production, and for conception status, but negative for fat percent and days to first breeding. If the genetic correlation between production and conception status were negative, and the genetic trend for production positive, then the genetic trend for conception status should have been negative. Although the number of cows and sires was relatively large, the number of progeny per sire varied greatly, and a few sires with many daughters could have had major effects on the values presented. The reason for negative environmental trends during the later period may be due to stricter implementation of production quotas during the later years.

Table 7. Genetic, environmental and phenotypic trends in the Israeli dairy population from 1980 through 1986.

Period	Traits	Trends		
		Genetic	Environmental	Phenotypic
1976-85	Milk (kg/yr)	89.00	84.00	173.00
	Fat (kg/yr)	2.16	4.48	6.64
	Fat (%/yr)	-.01	.23	.02
	FCM (kg/yr)	81.00	102.00	183.00
1980-86	Milk (kg/yr)	154.00	-38.00	116.00
	Fat (kg/yr)	3.03	-.45	2.58
	Conception status (%/yr)	0.96	-1.01	-0.05
	Days to 1st breeding (d/yr)	0.09	-0.72	-0.63

FCM: fat-corrected milk

9. Genetic and environmental parameters of milk composition in Israel.

The laboratory of the Israeli Dairy Union began extensive testing of milk protein concentration of individual cows in 1985. The data are automatically recorded on computer diskettes and have been made available for analysis. Ezra, Raz, and Weller (1988); Ezra, Weller and Drori (1987); and Ron, Ezra, and Drori (1988) studied environmental and genetic factors affecting protein production and concentration. A total of 33,185 first through third parity lactation records from 105 herds that passed edits were analyzed. Means of fat and protein monthly tests were 3.2% and 3.0%, respectively. Standard deviations were .6% and .3%. Fat and protein concentration was lowest during the peak production postpartum period, and increased with cows' parity, in agreement with results from other countries.

Mean annualized protein production was 282 kg for Kibbutz (communal) herds, and 257 kg for Moshav (family) herds. The standard deviations were 41 and 35 kg, respectively. Means for protein concentration were $3.03 \pm .12$ and $3.17 \pm .20$ for Kibbutz and Moshav herds, respectively. Ratios of protein to fat production were .965 and .977, respectively. After correction for milk production level, fat percent was equal at both farm types, while protein production was still higher at Moshav farms. Significant differences among farms were found for both fat and protein production.

Variance components (Henderson's method 3) and genetic evaluations were obtained for a subset of 9815 cows, daughters of 85 sires from 382 herd-years. Estimates of heritabilities, genetic and environmental correlations among milk, fat, protein, and fat and protein concentration

are presented in Table 8. These estimates are similar to results of several studies for the US Holstein population. Heritabilities for milk and fat production were slightly higher than the estimates of Weller, Ron, and Bar-Anan (1986) for the Israeli population. The genetic correlations between fat percent, and milk and fat production were higher than in the previous study, which was conducted on a more extensive sample. These estimates were used to predict future genetic trends for the Israeli population.

Table 8. Heritabilities (on the diagonal), genetic correlations (above the diagonal), and environmental correlations for production traits in the Israeli population.

	Milk (kg)	Fat (kg)	Protein (kg)	Fat (%)	Protein (%)
Milk (kg)	.17	.63	.76	-.13	.44
Fat (kg)	.69	.25	.67	.68	.05
Protein (kg)	.92	.76	.16	.0	.25
Fat (%)	-.28	.47	-.05	.29	.40
Protein (%)	-.35	.05	.05	.51	.46

Under Israeli conditions of surplus fat production and importation of powdered skim milk, protein production is more important economically than butterfat. It is, therefore, recommended that protein be included as a major breeding objective once reliable sire evaluations become available.

10. Strategies for genetic mapping of quantitative trait loci in dairy cattle

Weller, Kashi, and Soller (1988, 1989) studied strategies for genetic mapping of quantitative trait loci (QTL) in the Israeli and US dairy cattle populations. Methods are based on the observation that progeny of a sire that is heterozygous for both a genetic marker and a linked QTL inheriting different alleles for the marker, will have different quantitative trait means. Power to detect QTL, as a function of QTL effect, heritability of the trait, and number of animals tested was determined for (1) "Daughter design" (marker genotype and quantitative trait values assessed on daughters of sires heterozygous for the markers), and (2) "Granddaughter design" (marker genotype determined on sons of heterozygous sires and quantitative trait values measured on daughters of the sons). For an equal number of assays, power increased with the number of daughters/sire (design 1) and sons/grandsire (design 2). For equal power and heritability less than or equal to .2, design 2 required half as many marker assays as design 1. For example, with a heritability of .2, a QTL effect of .2 standard deviation units, and a Type I error of .01 power is .70 if 400 daughters of each of ten sires are assayed for the

markers; on the other hand, power is .95 if markers are assayed on 100 sons of each of 20 sires, with 50 granddaughters/son.

11. Heterogeneity of variances across production levels

A number of studies have found that sire and residual variance components and heritability increase with herd production level. Weller, Ron, and Bar-Anan (1987) stratified herd-years of Israeli-Holsteins into three groups by two criteria: 1) mean annual milk production, and 2) mean persistency, estimated as the ratio of daily milk production at the fifth month postpartum to daily production during the postpartum peak period. As found in previous studies, sire and residual variance components and heritability as estimated by Henderson's method 3, increased with increased production level. This was true for both primiparous and multiparous cows, but the effect was lower for multiparous cows. Even though production and persistency were correlated, no clear trends were evident when stratification was on the basis of persistency. Genetic correlations among stratification groups were between .7 and .9 for persistency, and between .6 and .86 for production. The lower correlations for production are indicative of a possible sire x production level interaction.

Gianola, Foulley, Fernando and Henderson (1989) developed a new procedure for estimation of components of variance when there is heterogeneity of dispersion parameters with respect to some criterion of classification such as herd or production level. Genetic and residual variances are viewed as following distributions with either known or unknown parameters. The techniques are based on Bayesian ideas but the procedures can be given a frequentist interpretation, with the parameters of the prior distributions estimated from the data at hand. The estimates obtained are weighted averages of the corresponding parameter and of a statistic. Techniques are described for situations where animals are related or unrelated across herds. The estimators proposed may have better mean squared error properties than those obtained by grouping observations in some way, and then applying REML within each group.

12. Optimization of the Israeli breeding program.

Weller and Ezra (1989) studied the effect of varying the number of sires progeny tested, the number of cows mated to young sires, and the mean number of years of service of proven sires on genetic progress for milk production, with and without multiple ovulation and embryo transplant (MOET) for bull dams. The following assumptions, which approximate the conditions of the Israeli breeding program, were employed: 65,000 cows are milk-recorded out of a population of 100,000; six elite sires are used as bull sires each year; all cows not mated to young sires are mated to 20 proven sires; the phenotypic standard deviation and heritability of milk production are 1400 kg and .2, respectively; 17.5% of cows are culled due to low production; the number of bull dams is equal to two and six times the number of young sires with and without MOET, respectively; and sires used in the general population are selected after results of their

daughters' first lactation, while sires of sires are selected after their daughters' second lactation.

Genetic progress was greatest when proven sires were used four years. Without MOET, maximum genetic progress, 132.9 kg/yr, was obtained when 45% of all cows were mated to 190 young sires each year. Expected genetic progress was 2% less if 25% of the cows were mated to 100 young sires, the current situation. MOET for bull dams increased genetic progress consistently by 6 to 7 kg/yr. Assuming 1) the value of an additional kg of milk is \$.2, 2) a discount rate of 5%, 3) the cost of an additional sire stall is \$3557, 4) the cost of sire maintenance is \$876/yr, 5) the cost of establishment of a MOET unit is \$300,000, and 6) the cost of each ET is \$500; time to achieve the economic break even point for increasing the number of sires tested from 40 to 60 was 23 yr, and for the MOET scheme was 15 yr.

The expected correlated responses to two generations of index selection for milk, fat, and conception status were estimated by Weller (1989) as follows:

$$CR_t = h \times I \times r_t \times \sigma_t$$

where CR_t is the correlated response for the trait t , I is the selection intensity, h is the accuracy of the index, r_t is the genetic correlation between the selection index and the t th trait, and σ_t is the square root of the genetic variance for the t th trait; I was assumed to be equal to 1.5, which is equivalent to selection of the best 17%. The parameters h and r_t were calculated as follows:

$$h = \frac{[a'GRa]^{.5}}{[a'Ga]^{.5}}$$

$$r_t = \frac{a'Gb}{\{[a'Ga][b'Gt]\}^{.5}}$$

where a is the vector of economic weights; G is the genetic variance-covariance matrix, as estimated from the method 3 analysis; B is a matrix of order equal to G with mean sire repeatabilities on the diagonal and zeroes elsewhere; and b is a vector of order equal to a with an element equal to one for trait t , and all other elements equal to zero. Mean repeatabilities were computed as follows:

$$MRPT_t = \frac{100}{(K_t) + 100}$$

where $MRPT_t$ is the mean repeatability for trait t , and K_t is the ratio of residual to sire variance for trait t . The mean number of effective daughters was assumed to be equal to 100, similar to the actual situation for young sires in Israel.

CR_t is dependent only on I, h, a and G, and I and h were constant for all traits. Thus, although I was chosen arbitrarily, and h was computed for an arbitrary number of daughters, the relative values of CR_t did not differ. The following two indexes were compared:

$$\begin{aligned}\text{INDEX 1} &= .51(\text{kg milk}) + 14(\text{kg fat}) \\ \text{INDEX 2} &= .51(\text{kg milk}) + 14(\text{kg fat}) + 20(\% \text{ CS})\end{aligned}$$

INDEX 1 is proportional to the US PD\$. With the coefficients given, a kg of 3.5% milk will have a value of unity. Therefore, if the profit from an additional kg of 3.5% milk is \$.1, then an increase of CS of 1% is assumed to have a value of \$2.0.

Expected correlated responses of milk, fat, and CS to selection by INDEX 1 and INDEX 2 are presented in Table 9. Accuracy is slightly lower for INDEX 2, due to the lower mean repeatability for CS, as compared to the production traits. Although response for milk and fat are lower by 8% and 9% in INDEX 2, which gives an economic weight to CS, progress for CS will be 5.1%. If an economic value of zero is assumed for CS, then the genetic decline for this trait will be .52. However, even if fertility is not included in the selection index, there may still be positive selection for fertility, due to culling of cows that do not conceive. The correlated responses listed in this table were computed for two generations of sire selection, which is roughly equal to twelve years. Although smaller than the values presented in Table 7, the expected responses are similar to estimates of genetic trends for this population.

Table 9. Expected correlated responses of component traits and accuracies of sire evaluations for two selection indexes

		INDEX 1	INDEX 2
Traits			
Economic weights	Milk (kg)	.51	.51
	Fat (kg)	14.00	14.00
	Conception status (%)	0.00	20.00
Correlated response	Milk (kg)	1175.29	1090.92
	Fat (kg)	27.69	25.46
	Conception status (%)	-.52	5.10
Accuracy of index		.88	.86

Expected response after two generations of selection to changes in the economic weights for production traits (kg milk, fat and protein) were estimated by Ezra, Raz, and Weller (1988). Expected response in protein production varied from 18.1 kg in the current system (no price differential for protein) to 23 kg for direct selection on protein. Under the current pricing system, the expected genetic reduction in the sum of

protein and fat percent is .21%. Under the Dutch pricing system (a negative price for fluid milk), the sum of protein and fat percent would increase by .05%. The expected ratio of fat to protein varied from .94 in the current system to .98 for selection directly on protein.

X. DESCRIPTION OF THE COOPERATION

A preparatory meeting between Daniel Gianola and Reuven Bar-Anan was held in September 1985 in the 35th Annual Meeting of the European Association for Animal Production, Kallithea, Greece. The objective was to exchange ideas and to plan future work. This was followed by another meeting with Joel Weller in July 1986 at the Third World Congress of Genetics Applied to Livestock Production, Lincoln, Nebraska. At that time, we discussed the Israeli data and planned several analyses.

Communication between the University of Illinois and the Volcani Center was greatly facilitated by the computer network BITNET. This permitted exchanging electronic mail, files, data and ideas, without the delays always arising when regular mail service is used. BITNET became a primary vehicle for interaction and was used routinely by both parts in connection with the project.

Dr. Joel Weller from the Volcani Center visited the University of Illinois from September, 1987 through mid-July 1987. During his visit, further work on implementing the threshold model was completed and several analyses were completed. Interaction was effective and several areas for further cooperation in the future were identified, including the submission of a new BARD proposal. Dr. Weller visited again at the University of Illinois for a few days in August 1989 in connection with the preparation of such proposal.

The project was affected by the deaths of the Israeli principal investigator, Dr. Reuven Bar-Anan, in November, 1986, and of Dr. C. R. Henderson in March, 1989. Much more could have been accomplished if Dr. Bar-Anan had remained with us, and Dr. Henderson was always a most valuable resource person in discussions related to methodological developments.

XI. EVALUATION OF RESEARCH

The project provided a framework for studying a wide range of problems in dairy cattle breeding, ranging from the implementation and computation of large linear and nonlinear models to examining the feasibility of selecting bulls for twinning rate of their daughters.

Dairy sire and cow evaluation procedures have been enhanced by incorporating new methodology. Programs were written to calculate BLUP sire evaluations for production, fertility, calving, and growth traits, and for mastitis susceptibility and cow survival. These programs are now part of the standard Israeli sire evaluation program. Since 1988 Israeli sires are routinely evaluated for protein production. Procedures for selecting bull dams in Israel have been improved, based on the results of

this project. Nearly all Israeli heifers are bred to bulls with positive evaluations for calving traits.

A major breakthrough was the development of a fairly general computer program to carry out variance component estimation and sire evaluation by threshold models. The program runs on a series of computers, ranging from PC's to supercomputers. Because it is designed for small or large scale data analyses, it can be used for many purposes. Israel has now the elements needed to implement sire evaluations for dystocia and calf mortality using the modern nonlinear techniques developed primarily at the University of Illinois. The program continues being expanded, a future project being the analysis of "sire plus maternal grand-sire" models. This will facilitate the study of categorical traits influenced by maternal effects.

The theoretical problems studied at Illinois resulted in the development of more suitable methods for the analysis of threshold characters and in an increased understanding of why likelihood based procedures can account for certain forms of selection. It is felt that the application of statistical methodology to dairy cattle breeding is on a much firmer theoretical basis.

Methods were developed to adjust production records for past reproductive performance, calving age and month, and to predict total production from incomplete lactations. Genetic evaluations computed from adjusted records were slightly more accurate than non-adjusted records. The adjustment factors developed in Israel were incorporated into the standard sire and cow evaluation procedures.

Methods to account for heterogeneous variance components among herds were developed. This problem is probably not important for sire evaluation in Israel, because sires are evaluated over many herds. It may be of major significance for evaluation of US sires, and bull dams in both countries.

The results of the genetic analysis of reproductive diseases in Israel were presented to the industry. Although the data set was more reliable than most previous sets analyzed, heritabilities for all disease traits were still low. The preliminary results on the relationship between bacterial infection and somatic cell counts have been presented to the industry, and will serve as the basis for a more extensive Israeli study.

Phenotypic, genetic, and environmental trends in Israel for production and reproductive traits were estimated. Genetic trends were economically favorable for all traits studied, except fat percent and days to first breeding. Prediction of future trends based on different selection goals have been presented to the Israeli dairy industry.

Methods were developed to optimize the Israeli breeding program. It was found that genetic progress could be increased if more young sires were tested each year, and if a larger percentage of cows were mated to young sires. MOET for bull dams in Israel should increase genetic

progress by 6-7 kg per year, and will reach the economic break-even point in about 15 years.

Under this BARD project, major problems were solved and new problems were more clearly defined and studied. We feel that this project has already been and will be of significant value to the US and Israeli Dairy Industries.

XII. PUBLICATIONS

1. Refereed articles

- Bar-Anan, R., J. I. Weller, G. R. Wiggans, and K. Osterkorn. 1986. Prediction of annualized lactation yield from partial lactations. J. Dairy Sci. 69:1897.
- Chang, H. L., R. L. Fernando and D. Gianola. 1989. On the inverse of a matrix of additive x additive epistatic relationships. Genet. Sel. Evol. (submitted)
- Chang, H. L., R. L. Fernando and D. Gianola. 1989. Inverse of an additive relationship matrix due to sires and maternal grandsires. J. Dairy Sci. (In Press)
- Da, Y., M. Grossman, and I. Misztal. 1989. Prediction error variance and REML algorithm for animal model with relationship grouping and transformation. J. Dairy Sci. (Accepted)
- De Boer, J. A., J. I. Weller, T. A. Gipson, and M. Grossman. 1989. Multiphasic analysis of milk and fat yield curves of Israeli Holsteins. J. Dairy Sci. 72:2143.
- Foulley, J. L. and D. Gianola. 1986. Sire evaluation for multiple binary responses when information is missing on some traits. J. Dairy Sci. 69:2681.
- Foulley, J. L., S. Im, D. Gianola and I. Hoschele. 1987. Empirical Bayes estimation of parameters for n polygenic binary traits. Genet. Sel. Evol. 19:197.
- Foulley, J. L., D. Gianola and Plachenault. 1987. Sire evaluation with uncertain paternity. Genet. Sel. Evol. 19:83.
- Foulley, J. L., D. Gianola and S. Im. 1987. Genetic evaluation of traits distributed as Poisson-binomial with reference to reproductive characters. Theor. Appl. Genet. 73:870.
- Gianola, D., J. L. Foulley and R. L. Fernando. 1986. Prediction of breeding values when variances are not known. Genet. Sel. Evol. 18:485.
- Gianola, D., S. Im, and R. L. Fernando. 1988. Prediction of breeding value under Henderson's selection model: a revisitation. J. Dairy Sci. 71:2790.
- Gianola, D., J. L. Foulley, R. L. Fernando and C. R. Henderson. 1989. Estimation of heterogeneous variances using empirical Bayes methods: theoretical considerations. J. Dairy Sci. (submitted)

- Gianola, D., R. L. Fernando, S. Im and J. L. Foulley. 1989. Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome* (In press)
- Hoschele, I., D. Gianola and J. L. Foulley. 1987. Estimation of variance components with quasi-continuous data using Bayesian methods. *Z. Tierzuchtg. Zuchtgsbiol.* 104:334.
- Hoschele, I., M.F. Rothschild and D. Gianola. 1988. Recursive estimation of breeding values with categorical data. *J. Dairy Sci.* 71:452.
- Hoschele, I. and D. Gianola. 1989. Bayesian versus maximum quasi-likelihood methods for parameter estimation and sire ranking with binary threshold variables. *J. Dairy Sci.* 72:1569.
- Im, S. and D. Gianola. 1988. Mixed models for binomial data with an application to lamb mortality. *Appl. Stat.* 37:196.
- Im, S. and D. Gianola. 1988. Offspring-parent regression for a binary trait. *Theor. Appl. Genet.* 75:720.
- Im, S., R. L. Fernando and D. Gianola. 1989. Likelihood inferences in animal breeding under selection: a missing data theory viewpoint. *Genet. Sel. Evol.* (In Press).
- Misztal, I. and D. Gianola. 1987. Indirect solution of mixed model equations. *J. Dairy Sci.* 70:716.
- Mistal, I., D. Gianola and L. R. Schaeffer. 1987. Convergence rates in animal model solutions. *J. Dairy Sci.* 70:2577.
- Mizstal, I., D. Gianola and I. Hoschele. 1988. Note on threshold models with heterogeneous residual variance due to missing information. *Genet. Sel. Evol.* 20:511.
- Misztal, I. and G. R. Wiggans. 1988. Approximation of prediction error variance in large-scale animal models. *J. Dairy Sci.* (Submitted)
- Misztal, I., D. Gianola and J. L. Foulley. 1989. Computing aspects of a nonlinear method of sire evaluation for categorical data. *J. Dairy Sci.* 72:1557.
- Ron, M., R. Bar-Anan, and J. I. Weller. 1986. Sire and maternal grandsire effects on calving difficulty and calf mortality in Israeli Holsteins. *J. Dairy Sci.* 69:243.
- Weller, J. I. 1986. Comparison of multi-trait and single-trait multiple parity evaluations by Monte Carlo simulations. *J. Dairy Sci.* 59:493.
- Weller, J. I. and M. Ron. 1987. A note on the effect of observations with missing data on genetic correlation estimates. *Theor. and Appl. Genet.* 74:549.

- Weller, J. I., M. Ron and R. Bar-Anan. 1987. The effect of persistency and production level on the genetic parameters of milk and fat yield in the Israeli-Holstein population. J. Dairy Sci. 70:672.
- Weller, J. I., G. R. Wiggans and T. J. Lawlor. 1987. Genetic evaluation of dairy goat does for milk and fat as an extension of buck evaluations. J. Dairy Sci. 70:681.
- Weller, J. I. 1988. Inclusion of partial lactations in the genetic analysis of yield traits by differential weighting of records. J. Dairy Sci. 71:1873.
- Weller, J. I., I. Misztal and D. Gianola. 1988. Genetic analysis of dystocia and calf mortality in Israeli-Holsteins by threshold and linear models. J. Dairy Sci. 71:2491.
- Weller, J. I. 1989. Genetic analysis of fertility traits in the Israeli daily cattle population. J. Dairy Sci. (In press)
- Weller, J. I. and Y. Folman. 1989. Effects of calf value and reproductive management of optimum days to first breeding. J. Dairy Sci. (In press)
- Weller, J. I. and D. Gianola. 1989. Models for genetic analysis of dystocia and calf mortality. J. Dairy Sci. (In press)
- Weller, J. I., Y. Kashi, and M. Soller. 1989. Power of "daughter" and "granddaughter" design for genetic mapping of quantitative traits in dairy cattle using genetic markers. Submitted to J. Dairy Sci.
- Wiggans, G. R. and I. Misztal. 1987. Supercomputer for animal model evaluation of Ayrshire milk yield. J. Dairy Sci. 70:1906.
- Wiggans, G. R., J. W. J. van Dijk and I. Misztal. 1988. Genetic evaluation of dairy goats for milk and fat yield with an animal model. J. Dairy Sci. 71:1330.
- Wiggans, G. R., I. Misztal and L. D. Van Vleck. 1988. Animal model evaluation of Ayrshire milk yield with all lactations, herd-sire interaction and groups based on unknown parents. J. Dairy Sci. 71:1319.
- Wiggans, G. R., I. Misztal and L. D. Van Vleck. 1988. Implementation of an animal model for genetic evaluation of dairy cattle in the United States. J. Dairy Sci. 71 (Suppl. 2):54.

2. Chapters in books

- Gianola, D., S. Im, R. L. Fernando and J. L. Foulley. 1990. Mixed model methodology and the Box-Cox theory of transformations: a Bayesian approach. In D. Gianola, K. Hammond (eds.) "Advances in statistical methods for the genetic improvement of livestock" (Springer Verlag, In Press)
- Gianola, D., S. Im and F. W. Macedo. 1990. A framework for prediction of breeding value. In D. Gianola, K. Hammond (eds.) "Advances in statistical methods for the genetic improvement of livestock" (Springer Verlag, In Press)
- Foulley, J. L., D. Gianola and S. Im. 1990. Genetic evaluation of discrete polygenic traits in animal breeding. In D. Gianola, K. Hammond (eds.) "Advances in statistical methods for the genetic improvement of livestock" (Springer Verlag, In Press)

3. Abstracts and Symposia

- Chang, H. L., R. L. Fernando and D. Gianola. 1988a. A note on computing additive relationships. J. Anim. Sci. 66(Suppl. 1):99.
- Chang, H. L., R. L. Fernando and D. Gianola. 1988b. A computationally efficient method to invert an additive x additive relationship matrix. J. Anim. Sci. 66(Suppl. 1):220.
- Chang, H. L., R. L. Fernando and D. Gianola. 1988c. A note on ranking sires for a trait with epistatic variance. J. Anim. Sci. 66(Suppl. 1):220.
- Da, Y., M. Grossman and I. Misztal. 1988. Prediction error variance and REML algorithm for animal model with relationship grouping and transformation. J. Dairy Sci. (Suppl. 1):264.
- Fernando, R. L., D. Gianola and S. Im. 1988. Ignorability of selection in estimation of parameters by maximum likelihood. J. Dairy Sci. 71(Suppl. 1):261.
- Gianola, D., S. Im, R. L. Fernando and J. L. Foulley. 1987. Transformations in variance component estimation and sire evaluation. J. Dairy Sci. 70(Suppl. 1):123.
- Gianola, D., J. L. Foulley and S. Im. 1987. Quantitative genetic analysis of count variates. J. Anim. Sci. 65(Suppl. 1):91.
- Gianola, D., J. L. Foulley and S. Im. 1987. A model for joint genetic evaluation of prolificacy and mortality (survival) in a polytocous species. J. Anim. Sci. 65(Suppl. 1):92.

- Gianola, D., R. L. Fernando and S. Im. 1988. Maximum likelihood estimation of parameters under two selection models. *Genome* 30(Suppl. 1):25.
- Gianola, D., R. L. Fernando, S. Im and J. L. Foulley. 1988. Maximum likelihood estimation of parameters in a mixed model under Henderson's conditional selection scheme. *J. Dairy Sci.* 71(Suppl.1):261.
- Gianola, D. and F. W. Macedo. 1988. Remarks on prediction of breeding value with unknown variances. *J. Anim. Sci.* 66(Suppl. 1):103.
- Gianola, D., J. L. Foulley, R. L. Fernando and C. R. Henderson. 1989. A Bayesian model for estimation of heterogeneous variance. *Summaries* 1:245. 40th Annual Meeting of the European Association for Animal Production.
- Hoschele, I., D. Gianola and J. L. Foulley. 1987. Estimation of variance components with quasi-continuous data using Bayesian methods. *J. Dairy Sci.* 70(Suppl. 1):124.
- Hoschele, I., M. F. Rothschild and D. Gianola. 1987. A Bayesian approach to recursive estimation of breeding values with categorical data. *J. Anim. Sci.* 65(Suppl. 1):97.
- Hoschele, I. and D. Gianola. 1988. Comparison of maximum quasi-likelihood and approximate Bayes estimation with categorical data. *J. Dairy Sci.* 71(Suppl. 1):263.
- Im, S., J. L. Foulley and D. Gianola. 1987. A linear model for genetic evaluation of categorical traits. *J. Dairy Sci.* 70(Suppl. 1):124.
- Im, S. and D. Gianola. 1987. Mixed models for binomial data with an application to lamb mortality. *J. Anim. Sci.* 65(Suppl. 1):32.
- Im, S. and D. Gianola. 1988. Nonlinearity of offspring-parent regression with binary observations. *J. Anim. Sci.* 66(Suppl. 1):103.
- Im, S., R. L. Fernando and D. Gianola. 1988. Likelihood inferences in animal breeding under selection: a missing data theory viewpoint. *J. Dairy Sci.* 71(Suppl. 1):263.
- Macedo, F. W. M. and D. Gianola. 1987. Bayesian analysis of univariate mixed models with informative priors. *European Association for Animal Production. 38th Annual Meeting. Lisbon, Portugal. 35 pp.*
- Misztal, I. and D. Gianola. 1987. Computing aspects of a nonlinear method of sire evaluation for categorical data. *J. Dairy Sci.* 70(Suppl. 1):124.
- Misztal, I. and D. Gianola. 1987. Solving mixed model equations by iterating on data. *J. Anim. Sci.* 65(Suppl. 1):96.

- Misztal, I., D. Gianola, and L. R. Schaeffer. 1987. Extrapolation and convergence criteria with Jacobi and Gauss-Seidel iteration in animal model. J. Dairy Sci. 70:2577.
- Mistal, I. and G. R. Wiggans. 1988. Approximation of prediction error variance in large-scale animal models. J. Dairy Sci. 71(Suppl. 1):27.
- Misztal, I., D. Gianola, I. Hoschele and S. Im. 1988. Relaxing the assumptions of normality and constant variance in threshold models. J. Anim. Sci. 66(Suppl. 1):103.
- Ron, M., J. I. Weller and E. Ezra. 1987. Estimation of between country conversion factors for sire evaluations based on common sires and sons of common sires. Proc. Interbull Meeting. Lisbon, Portugal.
- Weller, J. I., E. Ezra and R. Bar-Anan. 1986. Studies on the model of choice for genetic analysis of calving traits. J. Dairy Sci. 69(Suppl. 1):124.
- Weller, J. I., M. Ron and R. Bar-Anan. 1986. Multilactation genetic analysis of the Israeli dairy cattle population. Third World Congress on Genetics Applied to Livestock Production. Lincoln, NE:202.
- Weller, J. I. 1986a. Inclusion of partial lactations in the genetic analysis of yield traits by differential weighting of records. J. Dairy Sci. 69(Suppl. 1):99.
- Weller, J. I. 1987. Genetic analysis of fertility traits in the Israeli dairy cattle population. Annual Congress of the Israeli Genetic Association. Jerusalem, Israel.
- Weller, J. I., Y. Kashi and M. Soller. 1988. Estimation of sample size necessary for genetic mapping of quantitative traits in dairy cattle using genetic markers. J. Dairy Sci. 71(Suppl. 1):142.
- Weller, J. I., I. Misztal and D. Gianola. 1988. Genetic analysis of dystocia in the Israeli-Holstein population by threshold and linear models. J. Anim. Sci. 66(Suppl. 1):99.
- Weller, J. I., I. Misztal and D. Gianola. 1988. Genetic analysis of dystocia and calf mortality in Israeli-Holsteins by threshold and linear models. J. Anim. Sci. 66(Suppl. 1):101.
- Weller, J. I. 1988a. Genetic analysis of fertility traits in the Israeli dairy cattle population. J. Anim. Sci. 66(Suppl. 1):101.
- Weller, J. I. and E. Ezra. 1989. Optimization of a national dairy breeding program with multiple ovulation and embryo transplant. J. Dairy Sci. 72(Suppl. 1):65.

Weller, J. I. and M. Ron. 1989. Trends in secondary traits in the dairy cow population in Israel. 40th Annual Meeting of the EAAP. Dublin, Ireland. p 49.

Wiggans, G. R. and I. Misztal. 1987. Investigation of an animal model for genetic evaluation of milk yield. J. Dairy Sci. 70(Suppl. 1):125.

Wiggans, G. R., I. Misztal and L. D. Van Vleck. 1988. Implementation of an animal model for genetic evaluation of dairy cattle in the United States. J. Dairy Sci. 71(Suppl. 2):54.

Wiggans, G. R., J. W. J. Van Dijk and I. Misztal. 1988. Genetic evaluation of dairy goats for milk and fat yield with an animal model. J. Dairy Sci. 71:1330.

4. Other publications

Bar-Anan, R. 1986. Breeding and judging cattle in Israel - To where? Meshek Habakar Vehahalav 202:14. (In Hebrew)

Bar-Anan, R. 1986a. Heifer milk production and feeding during the growth period. Meshek Habakar Vehahalav 200:40. (In Hebrew)

Bar-Anan, R. and E. Ezra. 1986. Judging at "On". Meshek Habakar Vehahalav 203:9. (In Hebrew)

Bar-Anan, R. and M. Ron. 1986. Genetic implications of embryo transfer in Israel. Meshek Habakar Vehahalav 201:33. (In Hebrew)

Bar-Anan, R., M. Ron, and J. I. Weller. 1987. Dairy cattle improvement - the Israeli experience. Acta Agronomica Hungarica 36:173.

Ezra, E. and J. I. Weller. 1989. Can the Israeli breeding program be improved. Meshek Habakar Vehahalav 210:13.

Ezra, E., A. Raz and J. I. Weller. 1988. Milk protein and fat in the Israeli dairy cattle population. Publication of the Israeli Cattle Breeders' Association. 40 pp. (In Hebrew)

Ezra, E., J. I. Weller, and R. Bar-Anan. 1986. The effects of age and calving month on milk production. Heker Umas 8:7. (In Hebrew)

Ezra, E., J. K. Weller and D. Drori. 1987. Estimation of environmental effects on milk protein content. Heker Umas 9:31. (In Hebrew)

Raz, A., E. Ezra and R. Bar-Anan. 1986. Preliminary results of analyses from the Beitan Aaharon milk laboratory. Heker Umas 7:19. (In Hebrew)

Ron, M., E. Ezra, and D. Drori. 1988. Genetic parameters for milk protein quantity and concentration in the Israeli dairy herd. Heker Umas 10:9. (In Hebrew)

- Ron, M., E. Ezra and J. I. Weller. 1987. Genetic differences in growth rate of male calves from the dairy herd. Heker Umas 9:23. (In Hebrew)
- Ron, M., J. I. Weller and E. Ezra. 1988. Comparison of the genetic value of bulls from different countries. Meshek Habakar Vehahalav 212:13. (In Hebrew).
- Weller, J. I. 1988b. Dairy cattle improvement in Israel. Proc.Poultry Breeders Roundtable. St. Louis, MO. pp 50.
- Weller, J. I. 1988c. Preliminary analysis of somatic cell concentration in the Israeli dairy herd. Meshek Habakar Vehahalav 212:39. (In Hebrew)