

Prediction of Dairy Bull Fertility from Field Data: Use of Multiple Services and Identification and Utilization of Factors Affecting Bull Fertility

M. T. Kuhn¹ and J. L. Hutchison

Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

The objectives of this research were to assess the utility of multiple services, rather than first service only, and an expanded service sire term for prediction of bull conception rate (CR) by artificial insemination in the United States. The intent with the expanded service sire term was to determine whether accuracy could be improved by estimating factors affecting the bull's CR explicitly in the model and then formulating the bull's prediction as the sum of his own service sire solution along with the solutions for the other factors. Factors considered for the expanded service sire term included age of the bull at the time of mating, stud, inbreeding of the service sire, inbreeding of the mating (potential embryo), and an additive genetic effect. Both simulated and field data were used to study the objectives. In simulation, predictions were compared with true values, whereas with real data, predictions were compared with the bulls' average CR in set-aside data. Field data, using lactations 1 to 5, included 3,312,998 breedings of 737,626 Holstein cows in 1,419 herds distributed over 43 states and across 12 yr (1995 to 2006). The use of both multiple services and an expanded service sire term improved the accuracy of predictions. Multiple services contributed a 7 to 9% increase in accuracy, whereas the expanded service sire term improved accuracy by an estimated 12%. The amount of improvement in accuracy depends on the number of services available, but even for bulls with at least 500 matings, the combination of multiple services and an expanded service sire term can be expected to result in an overall increase in accuracy of at least 20%. Mean differences between predictions and bulls' average CR in set-aside data indicated that this improvement in accuracy can be brought about without introducing bias into the evaluations. Heritability estimates for artificial-insemination bull CR were essentially zero. Thus, use of an addi-

tive genetic effect for the service sire will not be of assistance in predicting bull fertility. All 4 of the other factors used in the expanded service sire term contributed to improved accuracy, although age of the bull at the time of mating was, by far, the major factor (correlation of 55.2% with future-year CR when included, 44.0% when not included). Allowing the stud effect to vary by year and using only the stud's most recent year solution in prediction were shown to be superior to using stud alone.

Key words: bull fertility, conception rate, prediction, artificial insemination

INTRODUCTION

In May 2006, Animal Improvement Programs Laboratory (AIPL) assumed responsibility for US evaluation of service sire fertility in dairy cattle. As an initial step, AIPL implemented Estimated Relative Conception Rate (ERCR) that had been previously developed by Dairy Records Management Systems and North Carolina State University (Raleigh, NC). The ERCR evaluations had been computed and published by Dairy Records Management Systems since 1986 and are described by Clay (2000) and Clay and McDaniel (2001). The general objective at AIPL is to explore possible methods to improve predictions through trait definition, editing, and statistical modeling. The 2 specific objectives of this research were to assess the use of 1) multiple services and 2) an expanded service sire term. It should be emphasized that, in contrast to bull evaluations for traits such as production and conformation, bull fertility evaluations are intended as phenotypic rather than genetic evaluations.

The trait used in ERCR is nonreturn rate at 70 d for first service only. Although nonreturn rate on first service is a well-defined trait and, as such, circumvents the problem of unknown outcomes for the last breedings on file, it also excludes a considerable amount of information by ignoring later breedings. Thus, one objective of this research was to determine the utility of using all breedings, rather than first service only.

Received October 1, 2007.

Accepted February 25, 2008.

¹Corresponding author: mkuhn@aipl.arsusda.gov

Most, if not all, current bull fertility evaluations, including ERCR, statistically model the bull's effect with a single random term, corresponding to the individual bull only. The basic model equation could be written as: $y = \text{nuisance variables} + \text{SSR} + e$, where y is the outcome of a breeding (success or failure); nuisance variables are factors such as herd, cow, and so on; and SSR is the effect of the service sire, fit as a random variable. A bull's evaluation, then, is simply his solution for the SSR term, with the additional step of converting to a probability in the cases in which threshold models are used. Furthermore, a diagonal variance-covariance matrix is used for the SSR term because the heritability of dairy bull fertility in AI is generally assumed to be zero. Thus, in contrast to genetic evaluations for performance and conformation traits, only a bull's own records contribute to his male fertility evaluation. However, factors such as service sire age and inbreeding have been found to affect conception rate (CR) in dairy cattle (Kuhn et al., 2006). Given the goal of phenotypic rather than genetic evaluation, modeling such factors as separate terms and then summing them along with the bull's own service sire solution may increase the accuracy of evaluation because more observations would contribute to estimation of the contributory factors. To illustrate, with 2 contributory factors (SSF_A , SSF_B) the basic model considered above would be expanded to $y = \text{nuisance variables} + \text{SSF}_A + \text{SSF}_B + \text{SSR} + e$, and a bull's predicted CR would be the sum of the solutions for SSF_A , SSF_B , and SSR. In a model with an expanded service sire effect, the SSR term accounts for variation caused by the bull and not accounted for by the other contributory factors, such that it can be interpreted as a residual service sire effect.

The literature suggests several factors to consider for an expanded service sire term. The long-standing view of inbreeding depression for reproductive traits has prompted consideration of the 2 inbreeding coefficients. Falconer (1989), for example, reviewed numerous early inbreeding studies and stated that "The most striking observed consequence of inbreeding is the reduction of the mean phenotypic value shown by characters connected with reproductive capacity . . ." Significant effects of service sire inbreeding and inbreeding of the mating in heifer breedings have also been found in previous research (Kuhn et al., 2006). Mating inbreeding is a characteristic of the bull to the extent that he is related to the population; as a bull's relationship to the population increases, the inbreeding of his embryos will tend to increase and, thus, may subsequently affect the viability of the conceptus as well.

Salisbury et al. (1978) presented several studies showing effects of bull age on fertility and, more recently, Kuhn et al. (2006) reported bull age effects on

CR in US Holstein heifers as well. Stud was an obvious consideration, given possible differences in bull management, semen processing procedures, postfreeze handling, and perhaps monitoring of semen performance in the field. Furthermore, previous studies (Gérard and Humblot, 1991; Kuhn et al., 2006) have shown differences among studs for CR.

As implied by the absence of a genetic term in the models of both Clay and McDaniel (2001) and Weigel (2004), heritability of AI bull fertility in the United States has either been assumed to be zero or ignored, given the objective of formulating a phenotypic rather than genetic predictor. Nonetheless, for the sake of completeness, heritability was also estimated because a nonzero heritability would improve the accuracy of the service sire term by allowing relative information to contribute to the bull's evaluation.

MATERIALS AND METHODS

General Aspects of Methodology

Both simulation and field data were used to assess the utility of multiple services and an expanded service sire term. In simulation, predictions of interest were computed and compared with the service sire's true effect. With real data, the approach was to 1) split the records into data for estimation and set-aside data, where set-aside data were records not used in estimation, 2) use only the estimation data to compute predictions of interest, and 3) compare predictions with service sires' arithmetic average CR in the set-aside data.

Two methods were used to divide records into estimation and set-aside data. The first was to split the data in half by herd, and the second was to set aside a future year for comparison; that is, to use all data up to a given year for estimation and then use data from the following year as set-aside data. In the split-data-by-herd approach, herds were randomly assigned to 1 of the 2 data sets. Another option would be to split the data by cow, randomly assigning individual cows rather than herds to 1 of 2 data sets. Simulation, however, showed no advantage to this approach over splitting data by herd, so data were split by herd to reduce the problem of small contemporary group sizes with real data.

The main statistic used to compare predictions with true values (simulation) or average CR in set-aside data (field data) was the correlation between the predictor and the true value or average CR (i.e., accuracy). Mean differences between predictors and true values or average CR (bias) were also computed. Standard deviations of differences (square roots of mean square errors) were also computed because mean square error serves as basically a composite of both bias (squared) and accu-

racy. The chi-square statistic, used by González-Recio et al. (2005), was also computed for comparisons with real data. The chi-square was calculated for each bull as

$$(NC - NC_{\text{expected}})^2 + (NF - NF_{\text{expected}})^2,$$

where NC was the number of observed conceptions in set-aside data, NC_{expected} was the number of expected conceptions, given the prediction, NF was the number of failed breedings for the bull in the set-aside data, and NF_{expected} was the number of expected failures, given the prediction. The overall chi-square statistic was then obtained by averaging across all bulls.

Preliminary comparisons showed that conclusions were the same when a minimum of 100, 500, or 1,000 records per bull were required in both the estimation and set-aside data for the split-herd approach. Thus, final comparisons for the split-herd approach used only bulls with at least 500 records in both data sets. For the future-year approach, however, this requirement resulted in too few bulls for comparison, so bulls were required to have 500 records in the estimation data but only a minimum of 100 records in the set-aside data.

Preliminary analyses using simulation indicated that linear models were as effective as threshold models in determining the usefulness of multiple services and expanded service sire terms (i.e., in addressing the particular objectives of this research). Thus, although the validity of a linear model for this binary trait warrants thorough investigation, linear models only were used for analyses of both simulated and field data to reduce the substantial amount of computing time required in this research and to limit the number of results to be presented.

Simulation also showed that when categorical contributory factors were fit as fixed effects, accuracy was higher than when using a single service sire term, but predictions were severely biased. However, when categorical contributory factors were fit as random, the advantage in accuracy was maintained with no bias. When contributory effects are treated as random, the variance components weight them appropriately, relative to the random residual service sire term; when fit as fixed effects, the solutions are unregressed and are simply too large relative to the random residual service sire component. Thus, further analyses considered (categorical) factors contributing to the service sire effect only as random effects.

Simulated Data

Description of Simulated Data. An underlying variable, y , was simulated according to the equation

$$y = \text{Herd} + \text{Year} + \text{Parity} + \text{Cow}_{\text{PE}} + \text{Cow}_{\text{BV}} \quad [1] \\ + \text{SSF}_A + \text{SSF}_B + \text{SSR} + e,$$

where Cow_{PE} and Cow_{BV} were permanent environmental and breeding value effects of the cow (mate of the service sire) and SSF_A , SSF_B , and SSR constituted the effect of the service sire, where SSF_A and SSF_B were simulated as contributory factors. Both SSF_A and SSF_B were categorical variables with 5 levels each, with relative frequencies of 0.30, 0.30, 0.20, 0.10, and 0.10 for SSF_A and 0.30, 0.25, 0.20, 0.15, and 0.10 for SSF_B . A total variance of 0.0078 (SD = 0.09) was attributed to service sire and was split equally (in half) between the SSR term and the contributory factors (SSF_A , SSF_B); the half attributed to the contributory factors was, in turn, divided equally between SSF_A and SSF_B . The total variance attributed to service sire effects (0.0078) corresponded to a proportion of variance attributable to service sires of approximately 0.7%. This rather low proportion is consistent with prior estimates of variance attributable to the service sire (Weigel, 2004). The variance-covariance matrix for SSR was of the form $I\sigma_{\text{SS}}^2$, where I was an identity matrix and the scalar σ_{SS}^2 was the variance of SSR .

An insemination was assigned the binary outcome of success (conception) if the underlying variable (y) exceeded a given threshold value, where thresholds were defined so as to give a desired CR. As with real data, CR was allowed to vary by service number, with CR decreasing as service number increased. The overall mean CR in the simulated data, across all services, was 35%. The simulation program generated data year-by-year for a total of 10 yr, maintaining 100,025 cows/yr and 250 service sires/yr. Across all 10 yr of simulation (1 replicate), there was an average of 2,472,572 breedings and 464,399 cows. The total number of service sires per replicate was 1,464.

Models for Analysis, Calculation of Predictions.

Four predictions were calculated with simulated data: 1) an expanded service sire term using all services for estimation, 2) an expanded service sire term using first services only, 3) service sire only (no expanded service sire term) using all services, and 4) service sire only using first service only. Equation [1] was the model used to obtain solutions for the case of predictions with an expanded service sire term. For predictions with no expanded term, solutions were obtained by dropping SSF_A and SSF_B from equation [1]. Predictions using the expanded service sire term were calculated as $100 \times (\text{population mean} + (\text{s}\hat{\text{s}}\text{r} + \text{S}\hat{\text{S}}\text{F}_A + \text{S}\hat{\text{S}}\text{F}_B))$, where $\text{s}\hat{\text{s}}\text{r}$, $\text{S}\hat{\text{S}}\text{F}_A$, and $\text{S}\hat{\text{S}}\text{F}_B$ were solutions from model [1]. Predictions not using an expanded service sire term were computed as $100 \times (\text{population mean} + \text{s}\hat{\text{s}}\text{r})$. The same

model equations were used to obtain solutions for the cases of using first service only, the only difference being that data were restricted to first breedings only.

Final results (e.g., correlations) were averages across 50 replicates. Correlations, for example, were calculated in each replicate for each predictor, with final estimates being the average correlation across 50 replicates. Estimated rather than true variances were used for prediction with simulated data because use of true variances could overestimate the potential of an expanded service sire term. Variances were estimated by REML (using AIreml developed by Tsuruta and Miztal) with 5 samples of approximately 160,000 cows each, where samples were taken from each of 5 independent replicates. The final variance components used for estimation were averaged across the 5 replicates.

Field Data

Description of the Data. Breeding records were being supplied by only 2 of the 4 major dairy record processing centers in the United States. Although data were predominantly from the eastern two-thirds of the United States, the region most heavily covered by the 2 supplying processing centers, there were 43 states with a minimum of 500 breedings in the data. Natural services are also reported to AIPL but were excluded from this research because current interest was in prediction of AI fertility. Lactations 1 through 5 were included, with a maximum of 7 breedings per lactation. Matings were also included only if the cow had a test day at least 70 d after the breeding so as to allow sufficient time for reporting of another reproductive event. Another edit applied to the field data was to eliminate matings on the same cow that occurred close in time. When 2 matings occurred less than 10 d apart, only the later breeding was kept. Presumably, repeat matings within short time periods are often the result of misdiagnosed heats on the first insemination, or perhaps the animal was bred on a timed AI program and was later observed in heat.

Breedings were restricted to Holstein cows with Holstein service sires, the latter restriction being made to avoid the need to model possible heterosis in the embryo and also because the number of crossbred matings currently available is minimal. Other minor edits included exclusion of cows known to be donor dams, exclusion of matings for which the age, stud, or inbreeding of the service sire were unknown, and exclusion of a lactation for which the cow's standardized milk yield was less than 4,536 kg. The edit on standardized milk yield was done to eliminate lactations in which the cow may have experienced health, injury, or other unrecorded problems. Mean milk yield was 11,440 kg with a SD of 2,046

kg. Thus, the 0.06% of lactations with yields less than 4,536 kg were more than 3 SD below the mean.

Three herd-year edits were applied: 1) at least 50% of the herd's lactating cows had to have at least 1 breeding reported, 2) herd average CR had to be from 10 to 90%, and 3) herds had to have at least 30 cows with a breeding for each year they were included in the data. The first edit was used to eliminate herds that report breedings of only a portion (often only a small portion) of their cows, whereas the second was to eliminate, among other possible anomalies, herds that report only successful breedings. The third edit was made to ensure reasonable herd-year sizes for analysis. The final edited data contained 3,312,998 breedings for 737,626 Holstein cows in 1,419 herds distributed over 43 states and across 12 yr (1995 to 2006), with an overall mean CR of 29.6%. There were a total of 17,549 service sires, many of which, however, had only a few matings and thus did not contribute to comparisons with set-aside data.

Model for Analysis, Calculation of Predictions.

The basic model equation used for analysis of field data was:

$$y = \text{HY} + \text{Year-State-Month} + \text{Parity} + \beta_1 \times \text{milk} \\ + \beta_2 \times \text{milk}^2 + \beta_3 \times \text{Cow_Age} + \beta_4 \\ \times \text{Cow_Age}^2 + \beta_5 \times \text{DIM} + \beta_6 \times \text{DIM}^2 + \text{Cow_PE} \\ + \text{A} + \text{SSR} + (\text{expanded_SSR}) + e, \quad [2]$$

where y was the result of the breeding (conception or failure) and HY and Year-State-Month were the fixed effects of herd year and year-state-month of breeding, respectively. Parity was also fit as a fixed effect, whereas lactational, standardized milk yield, cow age, and DIM at breeding were fit as linear and quadratic covariates. The terms Cow_PE and A were the random permanent environmental and additive genetic effects of the cow, respectively, and SSR was the random effect of service sire. The variance-covariance matrices for cow_PE and SSR were of the form $I\sigma^2$, where I was an identity matrix and the scalar σ^2 was the estimated variance for each effect. The variance-covariance matrix for cow breeding values was the usual animal model matrix of $A\sigma_a^2$, where A was a matrix of additive relationships and the scalar σ_a^2 was the estimated additive genetic variance. Model [2] was used in both the split-herd and future-year analyses.

The expanded_SSR term always contained the bull's own effect (SSR) and also various alternative combinations of the 4 candidate factors for an expanded service sire term. To serve as a basis for comparison, service sire alone was used and it was also used in conjunction

with each of the 4 factors separately, which allowed assessment of the contribution of each individual factor. All 4 factors were also included along with SSR to determine the total benefit of the expanded service sire term. Predictions were calculated as $100 \times (\text{population CR} + \text{solutions for terms in expanded_SSR})$.

Predictors Compared in the Split-Herd Approach. A total of 11 predictors were computed using the split-data-by-herd approach. Similar to simulation, the first 4 predictions were the expanded service sire term and service sire only predictors, each with both multiple services and first service only. The other 7 predictors for the split-herd approach involved fitting each of the 4 factors separately (multiple services only) to allow some assessment of the magnitude of contribution of each factor. A total of 10 replicates were conducted in the split-herd approach by simply resampling herds a total of 10 times. As in simulation, the final results (mean differences, correlations, etc.) were averages across all replicates; correlations, for example, were calculated in each replicate and then averaged across replicates to form the final estimates.

Predictors Compared in the Future-Year Approach. A primary difficulty with the future-year approach is that older, more popular bulls predominate in the comparisons simply because the older bulls are most likely to have an adequate number of matings for comparison in both the estimation data and the future-year (set-aside) data. This also implies that fewer bulls would be available for comparisons. In contrast, the split-data-by-herd approach includes a broader range of bulls, because each bull has his matings split essentially in half.

In this research, however, the split-herd approach also had some disadvantages, in particular, the use of contributory variables related to time. Assignment of service sire age, for example, was straightforward in estimation, but assignment of age for calculation of predictions was not straightforward in the split-herd approach because an individual bull's age could have varied considerably in the set-aside data. A bull has only one value to compare with in the set-aside data (his average CR), but when that value encompasses a span of years, which age solution to use to predict CR in set-aside data is unclear. Thus, in spite of its limitations, there were several reasons for using the future-year approach, in addition to the split-data-by-herd approach. First, it is a more direct test of the objective; the objective is not to predict bull CR for matings occurring simultaneously in time, but rather to predict future CR. Second, it can serve as a separate, essentially independent confirmation of at least general results from the split-herd approach. And, finally, use of the time-dependent contributory variables in the expanded service sire

term is much more straightforward in the future-year approach than with the split-data-by-herd approach. Only 1 replicate was performed with the future-year approach. Estimation data were restricted to matings between 2002 and 2004, whereas the year 2005 served as the set-aside year for comparison.

Given that AI companies may change procedures and personnel over time, another variable of interest was stud \times year, in contrast to only the main effect of stud, which assumes a constant effect over time. Stud \times year would be especially difficult to use in the split-herd approach because of an inability to assign year when calculating the prediction. In contrast, with the future-year approach, assignment of stud \times year in the predictor was straightforward; bulls simply received the most recent year solution for their stud.

A total of 6 predictors were compared using the future-year approach. The first 4 were the expanded service sire term and service sire only predictors, each with multiple services and first services only. The next 2 predictors both used stud \times year, in place of stud, in the model for estimation, but the first predictor also included stud \times year in the prediction, whereas the second did not. Some AI companies have argued, in effect, that stud effects should be removed from bull fertility evaluations. Inclusion of stud \times year in the model for estimation implies that stud effects were partitioned out of all contributory factors (e.g., the SSR term in model [2]) and thus did not contribute at all, either directly or indirectly, to the predictions formed by including stud \times year in the model but not in the predictor; the objective was to determine whether removing stud effects did, in fact, improve predictions. Service sire age in the future-year predictors was the bull's median age in the future-year data, calculated as age at the beginning of the future year plus 6 mo.

Covariates Versus Categorical Variables for Quantitative Factors Contributing to the Expanded Service Sire Term. Simulation indicated that if an effect (e.g., inbreeding) can be properly modeled with covariates, then 1) there is no bias associated with modeling quantitative factors as covariates, 2) there is a very slight advantage in accuracy (0.16%) with the use of covariates, and 3) within the realm of linear models (i.e., models linear in the parameters), if the factor cannot be properly modeled with linear or quadratic covariates or both, then categorization of the quantitative variables is generally preferred.

A fixed-effects model was used to determine the nature of the relationship (linear, quadratic, or other) of CR with service sire age, service sire inbreeding, and inbreeding of the mating; cow_PE, breeding value, and SSR were dropped from model [2] and service sire inbreeding and age, and mating inbreeding were all cate-

Table 1. Bulls used in calculation of comparative statistics in split-herd and future-year analyses: number of bulls and, average number of matings and first services per bull

Method	Bulls, n	Matings for estimation, ¹ n	Matings in set-aside data, ² n	First services for estimation, n
Split-herd ³	442	2,171	2,091	1,008
Future-year	262	3,051	1,342	1,511

¹Bulls were required to have a minimum of 500 matings to be included in the comparisons.

²Minimum of 500 matings in set-aside data required in the split-herd approach; only 100 required in the future analysis.

³For the split-herd approach, numbers in table are averages across 10 replicates.

gorized and added to the model, along with stud, and were fit as fixed effects. Categories were formed simply by rounding actual values, for each variable, to the nearest whole number. A maximum was also imposed for each variable; thus, the highest category contained all observations that were as large as or larger than the indicated category number. Plots of the solutions were used to ascertain relationships with CR.

Estimation of Heritability

Heritability was estimated from a sample of 40,953 cows, using calvings between January 2002 and July 2005. Herds were required to have a minimum of 20 breedings for each year and an average CR between 15 and 85% to be included. Bulls were required to have at least 1,000 matings, in the overall data, to be included in the sample. The model equation for estimation was the same as model [2] except that service sire was used in place of cow for the additive genetic term ("A"). Because threshold models model the underlying variable rather than the observed variable (as in a linear model), threshold model estimates of variance components are generally larger than those from a linear model. Thus, in contrast to the other analyses, heritability was estimated from both a linear model and a threshold model because estimates for heritability were expected to be low. To reduce computing time, to improve convergence, and for the sake of using a larger sample, a sire effect was used for the threshold model, rather than an animal effect, as in the linear model. The numerator of the heritability estimate for the threshold model was therefore 4 times the estimated sire variance.

RESULTS AND DISCUSSION

The number of bulls and average number of matings per bull in the split-herd and future-year analyses are given in Table 1. Results for simulation and real data, for all 4 statistics used to compare models, are shown in Table 2.

Simulation Results

In simulation, correlations of predictors with true values (Table 2) showed clear advantages in accuracy with the use of both multiple services and an expanded service sire term. Using both multiple services and an expanded service sire term maximized accuracy, which had a total advantage of approximately 16% over the predictor using first services only and service sire only in the predictor. Use of all services increased accuracy by 6% with the expanded service sire predictor and by approximately 12% when service sire only was used. Use of the expanded service sire term improved accuracy by approximately 4% when all services were used and by roughly 10% when using first services only for estimation.

Mean differences were essentially zero for the expanded service sire predictors. The mean differences for the service sire only predictions, however, were small but still different from zero. The reason for the slight bias in the service sire only predictor was unclear. Fortunately, however, the mean differences for the service sire only predictor were, at most, only of tangential interest in this research; the conclusion of primary interest from the mean differences in simulation was that the use of an expanded service sire term per se causes no bias in the predictor. The standard deviation of differences was also minimized by the use of an expanded service sire term with multiple services. Thus, all results from simulation supported the use of both multiple services and an expanded service sire term.

Quantitative Service Sire Components (Inbreeding, Age) as Covariates or Categorical Variables

A plot of the CR solutions for each quantitative factor in the expanded service sire term is given in Figure 1. Mating inbreeding (plots A and B) exhibited a clear linear relationship with CR, especially in plot B of only the first 12 categories, which excluded small group sizes. Thus, mating inbreeding was fit as a linear covariate only in all analyses. Solutions for service sire in-

Table 2. Comparisons of predictions to true service sire (SSR) conception rate (CR) in simulation and average CR in set-aside data (real data: split-herd, future-year approaches)

Method	Predictor	Correlation		Mean difference		SD of difference		Chi-square ¹	
		All services	1st service only	All services	1st service only	All services	1st service only	All services	1st service only
Simulation ²	Expanded	87.2	81.0	-0.01	-0.01	1.50	1.79	—	—
	SSR only	83.0	71.4	0.35	0.35	1.71	2.15	—	—
Split-herd ³	Expanded	56.1	45.5	-1.17	-0.42	2.84	3.25	9,991	12,610
	SSR only	44.0	37.5	0.21	0.80	3.03	3.12	7,649	11,001
Future-year	Expanded	32.1	29.4	0.92	0.97	3.53	3.56	5,116	5,972
	SSR only	29.3	23.6	1.14	0.64	3.58	3.63	6,027	5,177
	Expd SY	38.1	—	0.22	—	3.42	—	3,525	—
	Expd no SY	31.6	—	0.41	—	3.51	—	3,496	—
Split-herd	Age, Covar	46.2	—	0.14	—	2.99	—	6,922	—
	Age, Catgl	55.2	—	-1.23	—	2.86	—	10,196	—
	SSR F	45.2	—	0.17	—	3.01	—	7,436	—
	Mating F	45.1	—	0.20	—	3.01	—	7,516	—
	Stud	45.1	—	0.01	—	3.01	—	6,985	—

¹Average across bulls.

²Predicted CR computed as expanded: $ssf_A + ssf_B + ssr$; SSR only: ssr , where ssf_A , ssf_B , and ssr were solutions for the contributory factors (ssf_A , ssf_B) and for the service sire term, respectively.

³Predicted CR computed as expanded: $b_1 \times F_{Mtg} + b_2 \times F_{SSR} + age + stud + ssr$; SSR only: ssr ; Expd SY: $b_1 \times F_{Mtg} + b_2 \times F_{SSR} + age + stud \times year + ssr$ [same as expanded but with stud-year (SY) instead of stud]; Expd No SY: $b_1 \times F_{Mtg} + b_2 \times F_{SSR} + age + ssr$ (stud-year in model for estimation, but not in predictor); Age, Covar: $b_A \times AgeYear + ssr$; Age, Catgl: $age + ssr$; SSR F: $b_2 \times F_{SSR} + ssr$; Mating F: $b_1 \times F_{Mtg} + ssr$; Stud: $stud \times year + ssr$, where b_1 and b_2 were solutions for linear regressions of conception on mating inbreeding (F_{Mtg}) and SSR inbreeding (F_{SSR}); age, stud, stud \times year, and ssr were solutions for their respective (categorical factors), AgeYear was the bull's actual age in years.

breeding (plots C and D) were more erratic, but overall its relationship with CR appeared to be linear as well. Plot D of service sire inbreeding solutions, which used only categories 2 through 10 (the majority of matings and bulls), showed a clear linear trend. Thus, service sire inbreeding was fit only as a linear covariate as well. The relationship of CR with service sire age (plot E), however, was clearly not linear or quadratic. Service sire age was therefore considered as a categorical variable and, for comparison, as a covariate as well.

Field Data: Correlations Between Predictors and Average CR in Set-Aside Data

With field data, the correlations (Table 2) of predicted CR with average CR in the set-aside data were clearly highest with the use of multiple services and the expanded service sire term. In the split-herd approach, multiple services resulted in correlations that were 10.6% (expanded predictor) and 6.5% (service sire only) higher than with first service only; multiple services resulted in 3 and 6% higher correlations in the future-year approach. Considering the stud \times year predictor from the future-year analysis (Expd SY in Table 2), the expanded service sire predictions had an advantage of approximately 9% in correlation over the service sire only predictor and an advantage of approximately 12% in the split-herd approach. Overall, the split-herd analysis indicated an increase of nearly 20% in accuracy with the use of multiple services and an expanded ser-

vice sire term, relative to the case of service sire only using first services only, as in the evaluation currently used; the future-year analysis showed a total advantage in correlation of 15%.

As shown by the correlations from the future-year analyses (Expanded vs. Expd SY in Table 2), stud \times year is much preferred over stud alone, in spite of the fact that the estimation data covered only 3 yr. The expanded predictor using stud \times year was 6% more accurate than the predictor using only stud. Furthermore, when stud \times year was in the model for estimation but was not included in the predictor, accuracy was 6.5% lower than when stud \times year was included in the predictor. Thus, a complete removal of stud effects from predictions of service sire fertility would substantially reduce the accuracy of evaluation.

The individual contribution of each contributory factor can be assessed by comparing the split-herd correlations at the bottom of Table 2 to the SSR only split-herd correlation in Table 2. Correlations for the individual factors showed that service sire age at mating was, by far, the major factor contributing to improved prediction. The correlation for age fit as a categorical variable was 11% higher than the correlation for the service sire only predictor. Although predictors with only service sire inbreeding, mating inbreeding, or stud in the expanded term also improved accuracy, their correlations were only approximately 1% higher than the predictor using service sire only. Nonetheless, in spite of their much more modest contribution, their use in an ex-

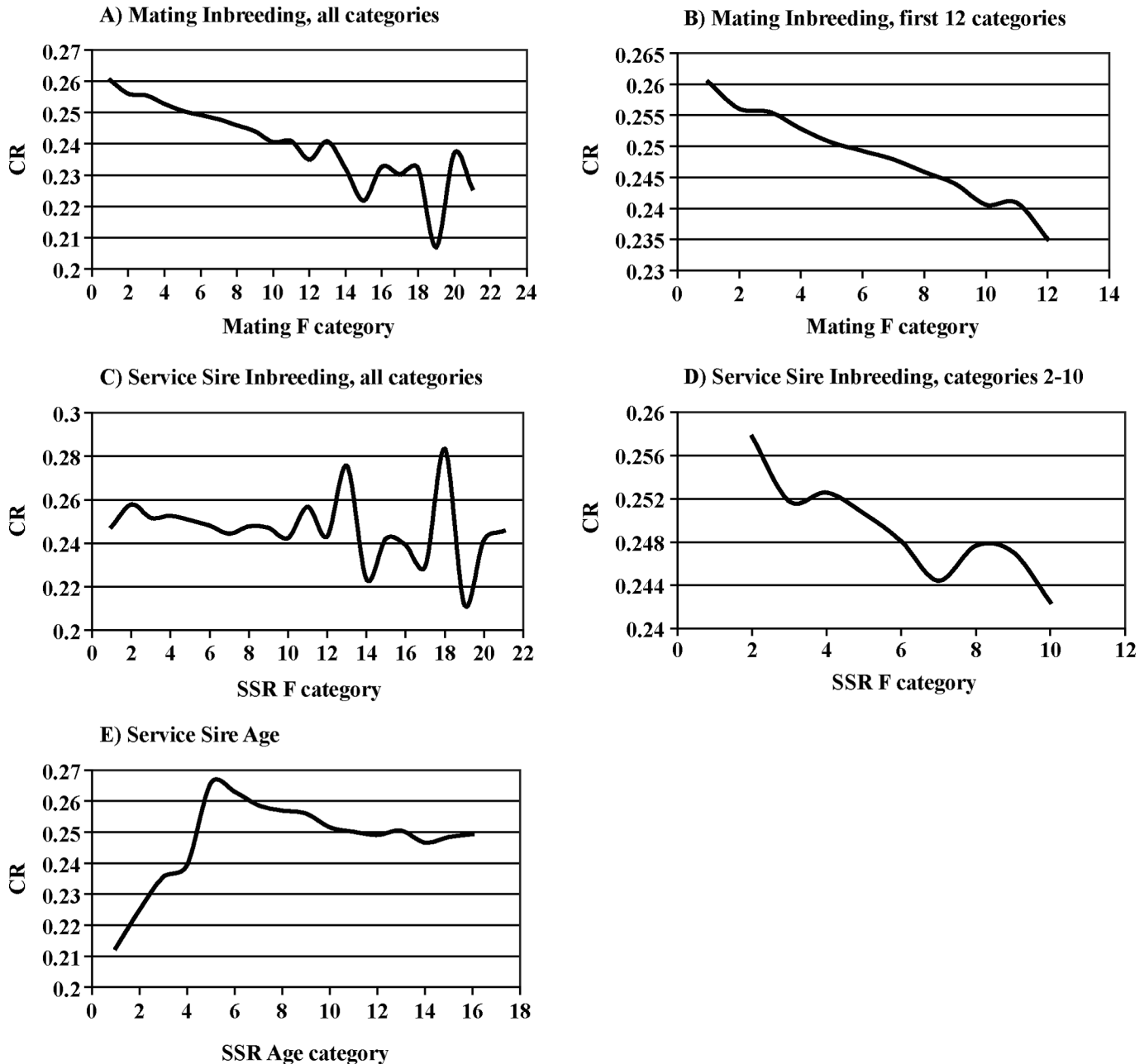


Figure 1. Plots of conception rates (CR) versus factors for consideration in an expanded service sire term: inbreeding (F) of the mating, inbreeding of the service sire (SSR), and SSR age at mating.

panded service sire term will still improve accuracy and should therefore be included in future predictors. Furthermore, the improvement attributed to stud is an underestimate because stud \times year is the preferred variable, as indicated by the future-year results. Results for the individual factors also showed a clear preference for fitting age at mating as a categorical variable rather than as a covariate. Correlation with average CR in the set-aside data was 9% higher when age was

fit as a categorical variable rather than as a covariate. The contribution of service sire age may also be somewhat underestimated in the split-herd results because a bull's average age in the estimation data was used to compute his prediction, although his age could have spanned several categories in the set-aside data, which had a 12-yr span.

There are several characteristics of the split-herd and future-year correlations in Table 2 which, although not

Table 3. Number and percentage of bulls in each service sire age group, for bulls included in the comparisons for the split-herd and future-year approaches

Service sire age group	Split-herd		Future-year	
	Bulls, n	Bulls, %	Bulls, n	Bulls, %
1	186	42.0	0	0.0
2	92	20.8	0	0.0
3	8	1.9	5	1.9
4	5	1.2	2	0.8
5	107	24.2	1	0.4
6	37	8.5	36	13.7
7	3	0.7	51	19.5
8	2	0.4	67	25.6
9	1	0.3	40	15.3
10	0	0.0	30	11.5
11	0	0.0	17	6.5
12	0	0.0	13	5.0

of direct relevance to the objectives of this research, warrant some explanation. Perhaps the most noticeable aspect is that, overall, the future-year correlations were just lower in magnitude than the split-herd correlations. This was because bulls were required to have only 100 matings in the set-aside data for the future-year comparisons (to increase the number of bulls included), whereas bulls were required to have 500 matings in the set-aside data for the split-herd approach. As can be seen in Table 1, bulls averaged approximately 700 fewer matings in the set-aside data for the future-year approach compared with the split-herd approach. When bulls were required to have only 100 matings in both data sets for both approaches, the correlations for the 2 different approaches were of more similar magnitude. Another apparent peculiarity was that the future-year approach generally showed less advantage than the split-herd approach. This was due to the inherent characteristic of the future-year approach to include predominantly older bulls, which is illustrated in Table 3. In the future-year analysis, there were no bulls in the first (youngest) 2 age groups and only 8 total bulls in the next 3 age groups and, as was indicated in Figure 1, fertility changes the most across the first 5 age groups. In contrast, in the split-herd analyses there was a much better distribution of bulls across the first 5 age groups. Thus, the advantages found in the future-year analysis were lower mainly because service sire age, the major factor contributing to improved accuracy, was unable to make as much of a contribution in the future-year analysis.

The future-year approach also showed less advantage for multiple services than the split-herd approach, which was due to the number of services available for estimation in the 2 different approaches. The amount of advantage will decrease as the number of services increases; a bull with 5,000 first services, for example,

will benefit less from the use of multiple services than a bull with 500 first services. As can be seen in Table 1, bulls used for analysis averaged approximately 1,500 first services in the future-year approach but only approximately 1,000 first services in the split-herd analyses. As a further illustration, 68% of the bulls had 1,000 or fewer first services in the split-herd analyses, whereas only 54% of the bulls had 1,000 or fewer services in the future-year analyses. Thus, there was simply less room for benefit of multiple services in the future-year approach than in the split-herd approach. In spite of the differences in magnitudes of correlations between the split-herd and future-year approaches, the future-year analyses served their purpose: 1) they confirmed the general results found in the split-herd analyses (both analyses showed higher correlations with an expanded service sire term and multiple services) and 2) they distinguished between the variables stud and stud \times year.

Field Data: Mean Differences, Standard Deviations of Differences, Chi-Square Statistics

Mean differences (Table 2) indicated general agreement between predictors and average CR in the set-aside data. Differences were generally close to zero (favorable), usually less than 1% and often less than 0.5%. The most noticeable mean difference was that for the expanded service sire predictor using multiple services in the split-herd approach, which was -1.17 . As indicated by the mean differences for the individual factors (split-herd approach), this "larger" (larger in magnitude) mean difference was due to service sire age at mating. The larger mean difference associated with service sire age occurred in the split-herd approach only and was due to the necessity of assigning bulls only 1 particular age in their predictor, although their average CR in the set-aside data could span up to 12 yr. For a bull that began his progeny test in 1995, was made active in, for example, 2000, and then remained in active service for 3 or 4 yr, the majority of his matings would generally be after he was made active (after 5 yr of age), where, as was shown in Figure 1, higher fertility exists. His average age, however, would put him at a midpoint age and he would therefore be under-predicted, relative to his average CR in the set-aside data. The mean difference of only 0.22 (essentially zero), for the expanded service sire predictor using stud \times year in the future-year analysis, confirms that service sire age does not, in fact, cause a bias and that the result from the split-herd analysis was indeed due to the inability to adequately assign age in the predictor for all bulls in the split-herd approach.

Table 4. Solutions for contributory factors from the future-year analysis

Inbreeding			Stud					
Bull	Mating	Year	A	B	C	D	E	F
-0.00123	-0.00152	2002	0.027	0.029	0.035	0.010	0.007	0.015
		2003	0.000	-0.003	0.003	-0.003	0.003	-0.013
		2004	-0.007	0.002	-0.001	-0.002	0.013	-0.020

Standard deviations of differences (square roots of mean square errors) and the chi-square statistics were also generally minimized with the expanded service sire term and multiple services. The larger chi-square statistic for the expanded service sire predictor using all services in the split-herd analysis, relative to the service sire only predictor (9,991 vs. 7,649), was due to the larger mean difference for the expanded predictor, discussed previously. In the future-year analysis, there was essentially no difference in chi-squares for the predictors with and without stud \times year included (3,525 vs. 3,496). Simulation indicated that this statistic appears to be much more sensitive to mean differences than to differences in correlation, and the mean differences for these 2 predictors were nearly the same.

Solutions for Contributory Factors

Solutions from the future-year analysis with stud \times year in the model are shown in Tables 4 and 5. Conception rate decreased about 0.1% for every 1% increase in both service sire inbreeding and mating inbreeding. The result for mating inbreeding in this research was nearly identical to that of Cassell et al. (2003). The stud \times year solutions are listed for only the 6 most widely used studs (5 US, 1 Canadian) in the US. The range in these stud solutions was 2.8, 1.7, and 3.3% for years 2002 to 2004, respectively. There was also considerable

year-to-year variation within stud, which reflects the superiority of the stud \times year variable over stud alone for prediction of bull CR. Although difficult to discern from Table 4, there is likely a correlation of stud effects across year. Future research could consider fitting a block diagonal variance-covariance matrix for stud \times year effects, allowing stud effects across year to be correlated. Such a model may further improve accuracy of the stud \times year solutions, and subsequently of bull CR predictions, by allowing additional years to contribute to the final year estimate.

Solutions for bull age effects (Table 5) were consistent with plot E in Figure 1. Fertility increased up to age 5, was maximum at 5 yr of age, and then decreased somewhat up to approximately age 9 or 10. These results also suggest that 9 categories may be sufficient for bull age. In fact, an expanded model was tested using only 9 categories for age, and results were nearly identical to those using 12 age groups; this is also consistent with Figure 1 (plot E), which showed that CR plateaued with increasing age around age group 9 or 10.

Several older studies on bull age were reviewed in Salisbury et al. (1978) and generally supported the pattern of results in Table 5. VanDemark et al. (1956) reported that volume of semen increased through the first 4 yr and that sperm concentration was maximized in the second and third year, with only slightly lower concentrations in the fourth year. Hahn et al. (1969) found that several semen characteristics were better in younger bulls than in older bulls. Tanabe and Salisbury (1946) reported peak AI bull fertility at 2 yr of age, whereas Bishop (1970) reported peak fertility at somewhat older ages of 3 to 4 yr. In general, the early research on bull age effects on fertility support the general pattern illustrated in Table 5 and Figure 1 (plot E), which is increasing fertility to an age of 3 to 5 yr, followed by some decline thereafter. Taylor et al. (1985) provided perhaps one of the more extensive studies on the effect of bull age at the time of collection. They also found the general pattern of increasing CR with increasing age, followed by a decline after reaching a maximum. They observed a maximum, however, at ages of approximately 19 to 20 mo; their sample sizes were quite small, however, for ages of 4 yr or less and thus estimates were probably subject to a nontrivial

Table 5. Solutions for service sire age group from the future-year analysis with number of bulls and matings used for estimation

Service sire age group	Solution	Bulls, ¹ n	Matings, n
1	-0.018	3,090	203,037
2	-0.016	4,419	262,266
3	-0.008	1,580	22,972
4	-0.002	681	6,094
5	0.018	1,026	139,765
6	0.013	1,275	320,901
7	0.009	1,139	258,752
8	0.006	901	158,099
9	0.001	626	98,657
10	0.002	426	41,129
11	-0.002	344	17,249
12	0.002	753	33,104

¹Bulls counted in each age group to which they contributed.

amount of sampling error. Results of Taylor et al. (1985) for bull ages 8 and greater were variable. There appeared to be some continued decline in fertility after 8 yr of age but at a substantially lower rate than the decline from ages 5 to 8 yr.

It should be emphasized that "bull age" in this research refers to age of the bull at the time of mating, not age when the semen, used for insemination, was collected; bull age at the time of mating is the only information currently available in the AIPL database. Thus, bull age at mating may have been 10 yr, but semen collected when the bull was 8 or 9 yr old, for example, may have been used in the mating. In spite of this caveat, results for bull age in Table 5 are generally consistent with previous research, and, more important, results in Table 2 clearly showed marked improvement in predictions when bull age at mating was used in the predictor. At the same time, however, it is worthwhile to bear in mind that more detailed data collection could possibly lead to yet further improvement in predictors. In particular, bull age at the time of collection may improve the accuracy of prediction even more than bull age at mating. It is conceivable that information on date of collection could also improve the accuracy of the stud \times year term; "year" is currently year of mating, but year of collection may provide a more accurate prediction. Routine acquisition of date of collection for the semen used in the mating is not in the foreseeable future for the United States. However, Select Sires, for example, has recently announced bar coding of semen straws (Dickrell, 2007) and such technology may enhance the possibility of acquiring collection dates in the future. In any event, routine collection of the more detailed information is encouraged while, in the meantime, bull age at mating will serve as a useful variable for improving the accuracy of predictions of bull CR.

Heritability

Heritability estimates were 0.02 and 0.013% (i.e., essentially zero) from the linear and threshold models, respectively. It is likely that screening of bulls on semen characteristics measured by lab assays, along with possible modification of sperm concentrations based on compensable semen traits, eliminates (from AI bull fertility) any genetic variation that may exist in the population. Sperm morphological characteristics, for example, could perhaps have at least a low or even moderate heritability, but elimination of defective bulls would eliminate most, if not all, genetic variance for the trait.

It may be important to compare estimates of heritability for AI bull fertility by using data from the same time period and population. If processing and screening

practices used by AI organizations do, in fact, affect heritability of AI bull fertility, then these practices may vary by location and over time as methods for screening or processing improve over time. Estimates of AI bull fertility in the literature, however, appear to be sparse, especially for the United States. Stålhammar et al. (1994) estimated young bull fertility by using nonreturn rates, primarily on first service, where data for their Swedish Red and Whites and Swedish Friesians spanned the years 1981 to 1985. Although somewhat higher than the estimates from this study, their estimates of heritability, ranging from 0.1 to 0.6%, were still quite low. Several older studies (e.g., Murray et al., 1977; Hansen, 1979) obtained somewhat higher estimates of "heritability," but it is unclear whether the additive genetic component was truly separated from the bull's environmental component in those studies.

Estimates of total service sire variance provide an upper limit on heritability because they will include both genetic and nongenetic factors. Weigel (2004) estimated service sire variance for US Holstein AI bulls at 0.3% using a threshold model. This recent estimate using US data is consistent with the very low heritabilities found in this research. Given estimates of essentially zero for heritability, future predictors will not attempt to use a genetic term for prediction of US AI bull fertility.

CONCLUSIONS

Use of both multiple services and an expanded service sire term that includes the effects of service sire age at mating, service sire inbreeding, mating inbreeding, and stud \times year will improve accuracy of prediction of US AI bull fertility. Mean differences between predictors and average CR in set-aside data indicate that these improvements can be made without introducing bias. Improvements in accuracy depend on the number of services available for the bull, but even for bulls with at least 500 breedings, a minimum overall increase in accuracy of approximately 20% can be expected. Heritability of AI bull fertility in the United States is nearly zero; therefore, use of an additive genetic effect for the service sire will not be of assistance in prediction of bull CR. Future research on prediction of bull fertility will include a comparison of linear and threshold models, methods to deal with services where the outcome is not known with certainty, edits and modeling of nuisance variables, and perhaps models that allow the service sire (residual) effect to vary across time.

ACKNOWLEDGMENTS

The authors thank Ignacy Misztal and Shogo Tsuruta for the use of their BLUPF90, AIreml, and cblup90reml

programs. Appreciation is also extended to Richard Saacke for reviewing the manuscript before submission. Dairy Records Management Systems (Raleigh, NC), Ag-Source (Verona, WI), and Minnesota DHIA (via Agri-Tech Analytics, Visalia, CA) are acknowledged for supplying breeding data, as are the American dairy farmers who pay for data collection in the United States.

REFERENCES

- Bishop, M. W. H. 1970. Aging and reproduction in the male. *J. Reprod. Fertil. Suppl.* 12:65–87.
- Cassell, B. G., V. Adamec, and R. E. Pearson. 2003. Maternal and fetal inbreeding depression for 70-day nonreturn and calving rate in Holsteins and Jerseys. *J. Dairy Sci.* 86:2977–2983.
- Clay, J. S. 2000. Computing estimates of fertility for mating bulls from non-return data. Pages 45–49 in *Proc. 18th Techn. Conf. on Artificial Insemination Reprod. Natl. Assn. Anim. Breeders*, Columbia, MO.
- Clay, J. S., and B. T. McDaniel. 2001. Computing mating bull fertility from DHI nonreturn data. *J. Dairy Sci.* 84:1238–1245.
- Dickrell, J. 2007. *Innovation 2007*. Dairy Today 22:22–23.
- Falconer, D. S. 1989. *Introduction to Quantitative Genetics*. 3rd ed. Longman Scientific and Technical, Essex, UK.
- Gérard, O., and P. Humblot. 1991. Influence of interactions between semen extender and number of spermatozoa on nonreturn rate estimates of fertility for individual Holstein bulls. *Theriogenology* 36:727–736.
- González-Recio, O., Y. M. Chang, D. Gianola, and K. A. Weigel. 2005. Number of inseminations to conception in Holstein cows using censored records and time-dependent covariates. *J. Dairy Sci.* 88:3655–3662.
- Hahn, J., R. H. Foote, and G. E. Seidel Jr. 1969. Quality and freezability of semen from growing and aged dairy bulls. *J. Dairy Sci.* 52:1843–1848.
- Hansen, M. 1979. Genetic investigations on male and female fertility in cattle. *Livest. Prod. Sci.* 6:325–334.
- Kuhn, M. T., J. L. Hutchison, and G. R. Wiggans. 2006. Characterization of Holstein heifer fertility in the United States. *J. Dairy Sci.* 89:4907–4920.
- Murray, B. B., L. R. Schaeffer, and E. B. Burnside. 1977. Estimation of breeding values for non-return rates of Canadian Holstein Friesian sires. *Dairy Ind. Res. Rep.* Guelph University, Ontario, Canada.
- Salisbury, G. W., N. L. VanDemark, and J. R. Lodge. 1978. *Physiology of Reproduction and Artificial Insemination of Cattle*. 2nd ed. W. H. Freeman and Company, San Francisco, CA.
- Stålhammar, E., L. Janson, and J. Philipsson. 1994. Genetic studies on fertility in AI bulls. II. Environmental and genetic effects on non-return rates of young bulls. *Anim. Reprod. Sci.* 34:193–207.
- Tanabe, T. Y., and G. W. Salisbury. 1946. Influence of age on breeding efficiency. *J. Dairy Sci.* 29:337–344.
- Taylor, J. F., R. W. Everett, and B. Bean. 1985. Systematic environmental, direct, and service sire effects on conception rate in artificially inseminated Holstein cows. *J. Dairy Sci.* 68:3004–3022.
- VanDemark, N. L., L. J. Boyd, and F. N. Baker. 1956. Potential services of a bull frequently ejaculated for consecutive years. *J. Dairy Sci.* 39:1071–1072.
- Weigel, K.A. 2004. Improving the reproductive efficiency of dairy cattle through genetic selection. *J. Dairy Sci.* 87(E Suppl.):E86–E92.