Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data

E. López de Maturana,*† A. Legarra,† L. Varona,‡ and E. Ugarte*
*Neiker, Basque Institute for Agricultural Research and Development, PO Box 46, 01080 Vitoria-Gasteiz, Spain
†Station d’amélioration génétique des animaux, Institut National de la Recherche Agronomique, Auzeville, BP 52627, 31326 Castanet Tolosan Cedex, France
‡Area de Producció Animal, Centro UdL-IRTA.25198, Lleida, Spain

ABSTRACT

A method based on the analysis of recursive multiple-trait models was used to 1) estimate genetic and phenotypic relationships of calving ease (CE) with fertility traits and 2) analyze whether dystocia negatively affects reproductive performance in the next reproductive cycle. Data were collected from 1995 through 2002, and contained 33,532 records of CE and reproductive data of 17,558 Holstein cows distributed across 560 herds in official milk recording from the Basque Country Autonomous Community (Spain). The following fertility traits were considered: days open (DO), days to first service, number of services per pregnancy (NINS), and outcome of first insemination (OFI). Four bivariate sire and sire-maternal grandsire models were used for the analyses. Censoring existed in DO (26.49% of the data) and NINS (12.22% of the data) because of cows having been sold or culled before reaching the next parturition. To avoid bias, a data augmentation technique was applied to censored data. Threshold models were used for CE and OFI. To consider that CE affects fertility and the genetic determination of CE and fertility traits, recursive models were applied, which simultaneously considered CE as a fixed effect on fertility performance and the existence of a genetic correlation between CE and fertility traits. The effects of CE score 3 (difficult birth) with respect to score 1 (no problem) for days to first service, DO, NINS, and OFI were 8 d, 31 d, 0.5 services, and −12% success at first insemination, respectively. These results showed poorer fertility after dystocia. Genetic correlations between genetic effects of fertility traits and CE were close to zero, except for the genetic correlations between direct effects of DO and CE, which were positive, moderate, and statistically different from 0 (0.47 ± 0.24), showing that genes associated with difficult births also reduce reproductive success.

Key words: recursive model, genetic parameter, fertility, dystocia

INTRODUCTION

The main objective in animal breeding is to obtain the desired products more efficiently in relation to the present generation by changing the genetic merit of animals in coming generations (Groen et al., 1997). In dairy cattle, the traditional breeding goal has been to obtain high-yielding cows to increase milk, fat, and protein yields. Presently, however, as a consequence of the milk quota system and decreased milk prices, economic efficiency can be further improved by adequate emphasis on other characteristics, called functional or secondary traits (Essl, 1998; Pryce et al., 2004). The term “functional trait” is used to summarize those characteristics of an animal that increase economic profit, not by higher output of products but by reduced costs of input (Groen et al., 1997). Examples of these traits are fertility, calving ease (CE), or milking speed.

Calving ease is one of the most economically significant secondary traits (Dekkers, 1994; Dematawewa and Berger, 1997), especially for first-calf heifers; it measures the presence or absence of dystocia and its intensity. Dystocia, defined as a prolonged or difficult parturition, affects the profitability of herds, animal welfare, and acceptability of the production system by the consumer (Carnier et al., 2000). Difficult births increase direct costs of the herd (veterinary fees, calf or cow death or both, and extra farmer labor), as well as indirect costs, such as an increase in the risk of subsequent unfavorable health events, an increase in culling rate, and a reduction in yield (Dekkers, 1994; Dematawewa and Berger, 1997). Moreover, dystocia can negatively affect reproductive traits, such as days open or number of services per pregnancy (Dematawewa and Berger, 1997).

In Holstein cattle, CE records have been collected systematically in the Basque Country Autonomous Community (Spain) by trained technicians since 1992.
Trained technicians are needed because calving performance is scored subjectively and might require a special recording routine (Philipson et al., 1994). A genetic evaluation of sires arose as a request by dairy farmers, with the aim of reducing the frequency of calving problems. Genetic evaluation of sires for CE has been conducted routinely since 1995 (Alday and Ugarte, 1997). To improve the evaluation system, data collection and the model for evaluation were modified; and a sire-maternal grandsire threshold model (Van Tassell et al., 2003) was implemented in 2003.

Female fertility is another important functional trait in dairy cattle because it affects direct reproduction costs and influences calving interval, calving season, and involuntary culling rate (Boichard et al., 1997), thereby decreasing the profitability of the herd. Moreover, fertility and yield have an unfavorable relationship in Holstein cattle (Pryce et al., 2004), and reproductive performance has been found to deteriorate as milk yield increases.

Effects of dystocia on fertility measures have not been thoroughly studied, but there seems to be a negative effect of dystocia on days open (DO) and number of services, which worsens fertility (Dematawewa and Berger, 1997, 1998). Moreover, a moderate and positive genetic relationship between CE and days open has been shown (Lee et al., 2003), which would imply that difficult calving leads to delay estrus and conception.

Neither of these studies considers the fact that a complex relationship exists between CE and fertility traits. Phenotypically, CE affects the cow’s fertility in the next reproductive cycle by a cause-and-effect relationship. For example, a difficult parturition might stress the cow and delay the appearance or the intensity of the heat. Retentions of placental membranes or delays in uterine involution are other consequences of a difficult calving (Rajala and Grohn, 1998). On the other hand, CE is not a phenomenon external to the cow, as is age of cow or herd. Calving ease is a trait influenced by both a calf direct genetic effect (the effect attributable to genes of the calf on its ease of birth) and maternal genetic effects (the effects of the genes of the calf's dam on CE through the environment provided by the dam). Obviously, fertility traits are also affected by environmental effects and genetic effects. These considerations complicate the analyses. Considering CE only as a fixed effect would ignore the fact that both CE and fertility are (partially) determined genetically and that they may be genetically correlated. On the other hand, considering only the genetic correlation would ignore the cause-and-effect relationships between CE and fertility traits. Gianola and Sorensen (2004) outlined a solution for this kind of biological system, describing the use of recursive multiple-trait models in a quantitative genetic context. These authors extended quantitative genetic theory to suit situations in which recursiveness or simultaneity exists between the phenotypes involved in a multivariate system, assuming an infinitesimal and additive model of inheritance.

A recursive multiple-trait model is a particular case of the more general “structural equation models” (Johnson and Wichern, 2002). Loosely speaking, these are models in which phenotypes interact among themselves; effects are observed with noise (and have to be inferred); and models produce different residual effects. According to Johnson and Wichern (2002, p. 524) “structural equation models are sets of linear equations used to specify phenomena in terms of their presumed cause and effect relationship variables. In their most general form, the models allow for variables that cannot be measured directly.” A typical example is a multivariate model that includes covariates. It is possible that the covariates themselves are not directly observed, but they can be predicted (to some degree of accuracy) from one or several other measured variables. Moreover, covariates included in the model may be related in some way (say, correlated). The same happens with the quantities of interest (the “true” phenotypes), which may have to be predicted from others (the “observed” phenotypes). A typical example would be the prediction of feed efficiency from consumption and growth. The recursive multiple-trait model is a particular case in which phenotypes (i.e., different traits) interact among themselves and, moreover, in which trait 1 affects trait 2 but not the opposite. The case in which trait 2 also affects trait 1 is called simultaneity. An example is the relationship between milk yield and SCC (Wiggans and Shook, 1987). High SCC, which are usually associated with subclinical mastitis, may inhibit milk yield; milk yield in turn has a dilution effect on SCC.

The relationship between CE (corresponding to first calving) and fertility is a recursive one, because CE might affect fertility in the following reproductive cycle, but obviously fertility cannot affect CE, which preceded it. The model for fertility traits, as we show later, includes a term for CE. The recursive multiple-trait methodology can handle all the information properly and provide estimates for genetic parameters and estimates of the effect of CE on fertility traits.

Another problem related to fertility traits is the existence of censored and missing data. Censoring in fertility (Schneider et al., 2005) comes from the fact that cows left the study (milk recording) with some fertility information, before reaching the next parturition, because of culling, death, or sale. As a consequence, we ignore the outcome of the reproductive cycle. This might cause bias if culling is due to poor fertility. If we plainly discard censored data, and dystocia causes a dramatic
loss of fertility leading to culling, its effect will be underestimated. This information can be analyzed by a survival analysis (Schneider et al., 2005) or by a data augmentation technique (Donoghue et al., 2004).

The objective of this work was to study the phenotypic and genetic relationships between CE and fertility traits by using recursive multiple-trait models (Gianola and Sorensen, 2004) and handling censored and categorical data. Estimates of the effect of the different levels of CE on fertility measures, and estimates of the genetic relationships between fertility and CE are shown.

MATERIALS AND METHODS

Definition of Traits

CE. Calving ease is a trait that measures the presence or absence of dystocia and its intensity. Dystocia can be defined as a prolonged or difficult parturition. The biological aspects of dystocia in dairy cattle were reviewed by Philipsson et al. (1979) and Meijering (1984). On a general basis, genetic factors affecting CE can be separated into maternal and fetal (or direct) components. Maternal calving ease includes the pure maternal effect (i.e., the uterine influence of the dam on her calf’s birth weight and the influence of her own genotype on the pelvic opening). Direct CE (DCE) corresponds to the effect of the calf’s own genes (e.g., size).

Female Fertility. Female fertility is a trait that can be defined in many different ways. For example, it can be described as the ability to produce a live offspring during an economically and physiologically approved period (Hyppänen and Juga, 1998). De Jong (1997) defined good cow fertility as an animal in lactation that shows her heat in time and gets pregnant after the first insemination. Both definitions show the subjectivity of this trait.

Fertility measures chosen for this study were DO (measured in days), days to first service (DFS, measured in days), number of inseminations per conception (NINS), and outcome of first insemination (OFI). Outcome of first insemination was coded as 1 when females calved after the first insemination, whereas cows that failed to calve were coded as 0. Calving interval was not used because of its similarity to DO.

Data Characteristics

In the Basque Country Autonomous Community (Spain), CE data collection has been ongoing since 1992. Data are collected monthly at the same time as milk recording (officially recognized Basque milk recording scheme). The recording technician asks the farmer how the births have occurred since the last visit. The technician then scores each calving with the criterion described in Table 1. Calving ease information was matched with the reproductive data collected in the AI recording system, which includes dates of all services, and also with data from the milk recording system. Calving records between 1995 and 2002 and data from the first 8 lactations were considered in the analysis. Moreover, every cow was required to have first lactation records to avoid culling bias and because dystocia mainly affects heifers. Calving ease records were edited as described in Alday and Ugarte (1997). In addition, multiple parities and parities showing a CE score of 5 were discarded; these performances are considered to lack a CE genetic component, because they are originated by pathologies such as abnormal presentation of the fetus and have little relationship to causes (pelvic opening, calf size) of the other CE score.

The Basque Country Autonomous Community has 3 provinces. Each one has its own technicians, which caused differences in data collection. To correct this, all records from province × year combinations with less than 20% recorded male calves or with more than 95% of records scored 1 were discarded (Ducrocq, 2000). Other edits included discarding data from heifers whose calving age was less than 18 mo or more than 40 mo. Data from multiparous cows whose calving age was less than 28 mo (too young) or more than 206 mo (too old) were discarded as well. Contemporary groups, formed as the interaction among herd, year, and the technician who scored the record, had to comprise at least 5 records.

The distribution of CE scores is summarized separately by parity in Table 2. The editing procedure that
Fertility records were allocated to herd-year-season (HYS) contemporary groups based on the month of calving. Two calving seasons were defined: November to April and May to October. Contemporary groups had to be formed by at least 5 records, and the cows had to have the first insemination recorded. An age within parity effect was also defined. For first-parity cows, the age groups were <26, 26 to 29, and >29 mo of age; for second parity, the age groups were 31 to 36, 37 to 40, and >40 mo of age; for third or greater parities, <54, 54 to 62, >62 mo defined the respective age groups.

Data Analysis

Recursive Multiple-Trait Model. Recursive multiple-trait models were described in the Introduction. Gianola and Sorensen (2004) described the implications of these models and methods used in quantitative genetics. They proposed a Bayesian strategy for implementation through Markov chain Monte Carlo methods. However, the method is neither straightforward nor obvious to implement in a general manner. The Appendix shows that only recursive models, but not simultaneous models, can be implemented using software usually applied by geneticists to address genetic analyses by fitting (in this case) CE as another fixed effect.

Censored Data. Censored data for DO and NINS were handled by data augmentation (Tanner and Wong, 1987; Sorensen and Gianola, 2002). The methodology applied to this specific problem (right censoring) was explained by Korsgaard et al. (2003) and consists in generating, within a Gibbs sampling scheme, samples from the predictive distributions of the censored records. The predictive distributions are truncated normal distributions, whose truncation point was the observed phenotype (largest DO or NINS corresponding to the observation period), the mean was the sum of (fixed and random) effects affecting that record, and the variance was the residual variance. This methodology has already been used by Donoghue et al. (2004) in the analysis of reproductive data.

Categorical Data. Outcome at first insemination and CE were analyzed by a threshold model. These traits have 2 and 3 ordered categories, respectively. The threshold model assumes that the observed category is determined by the value of an underlying continuous random variable, called liability (Wright, 1934). A data augmentation technique was used within a Gibbs sampling scheme, as described in detail by Albert and Chib (1993), Sorensen et al. (1995), and Sorensen and Gianola (2002). The first threshold for both traits was arbitrarily set to 0. The second threshold for CE was set to 1. For reasons of identifiability, the residual variance for OFI was set to 1, because the threshold and the residual variance cannot be estimated simultaneously for binary traits. Thus, we used the algorithm of Korsgaard et al. (2003) to sample residual variances with some diagonal elements equal to 1. The number of inseminations, although of a categorical nature, was considered as a continuous trait because it had a high

### Table 3. Type of data (censored, complete data, or missing values) available for each fertility trait

<table>
<thead>
<tr>
<th>Trait1</th>
<th>Censored data</th>
<th>Complete data</th>
<th>Missing values</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFS, %</td>
<td>--</td>
<td>54.3</td>
<td>45.70</td>
<td>33,532</td>
</tr>
<tr>
<td>DO, %</td>
<td>26.49</td>
<td>73.46</td>
<td>0.95</td>
<td>33,532</td>
</tr>
<tr>
<td>NINS, %</td>
<td>12.22</td>
<td>73.46</td>
<td>14.31</td>
<td>33,532</td>
</tr>
<tr>
<td>OFI, %</td>
<td>--</td>
<td>73.46</td>
<td>26.54</td>
<td>33,532</td>
</tr>
</tbody>
</table>

1Fertility measures: DFS = days to first service; DO = days open; NINS = number of services per pregnancy; OFI = outcome of first insemination.

### Table 4. Means and standard deviations for censored and complete fertility data

<table>
<thead>
<tr>
<th>Trait1</th>
<th>Censored data</th>
<th>Complete data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>DFS, d</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>DO, d</td>
<td>281.68</td>
<td>168.35</td>
</tr>
<tr>
<td>NINS, no.</td>
<td>2.36</td>
<td>1.42</td>
</tr>
<tr>
<td>OFI, %</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

1Fertility measures = DFS = days to first service; DO = days open; NINS = number of services per pregnancy; OFI = outcome of first insemination.
Table 5. Numbers of records and levels of effects of data used in the analyses

<table>
<thead>
<tr>
<th>Variable</th>
<th>CE</th>
<th>DFS</th>
<th>DO</th>
<th>NINS</th>
<th>OFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of records</td>
<td>33,532</td>
<td>33,532</td>
<td>33,532</td>
<td>33,532</td>
<td>33,532</td>
</tr>
<tr>
<td>Number of sires of cows in data</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
</tr>
<tr>
<td>Number of maternal grandsires</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
<td>1,046</td>
</tr>
<tr>
<td>Number of contemporary groups</td>
<td>3,501</td>
<td>4,600</td>
<td>4,600</td>
<td>4,600</td>
<td>4,600</td>
</tr>
<tr>
<td>Number of cows in data</td>
<td>17,558</td>
<td>17,558</td>
<td>17,558</td>
<td>17,558</td>
<td>17,558</td>
</tr>
<tr>
<td>Number of service sires</td>
<td>764</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

1CE = calving ease; DFS = days to first service; DO = days open; NINS = number of services per pregnancy; OFI = outcome of first insemination.

number of categories (5) that were distributed following a symmetric, centered distribution considered to be normal. Continuous and threshold univariate analyses were carried out to study this trait, and similar results were obtained (E. López de Maturana, unpublished data). Therefore, for simplicity we chose to consider NINS as a continuous trait.

Models. A sire model was used for fertility measures and a sire-maternal grandsire model was used for CE. Animal models including maternal effects were not adopted because reaching convergence of the Gibbs sampler for CE was difficult. This result was reported previously by Luo et al. (2001) and extensively described for a threshold animal model with maternal effects. Linear models were applied to DO, DFS, and NINS. As discussed, threshold models were adopted for OFI and CE.

Four bivariate analyses were carried out: 1) DO-CE, 2) DFS-CE, 3) NINS-CE, and 4) OFI-CE because of the computational requirements. Indeed, a multivariate analysis considering all traits was not feasible because the correlation between NINS and OFI tends to be 1: When NINS is equal to 1, OFI is also equal to 1 (both traits give the same information).

In matrix notation, the models were

\[
\begin{align*}
U_F &= X_F b_F + W_F p_F + Z_u U_u + e_F \\
U_{CE} &= X_{CE} b_{CE} + Z_{u_{CE}} u_{CE} + Z_{mgs_{CE}} u_{mgs_{CE}} + e_{CE}
\end{align*}
\]

where \( U_F \) is a vector of fertility traits (liabilities for OFI), \( U_{CE} \) is a vector of liabilities for CE, and \( b_F \) and \( b_{CE} \) are vectors of systematic effects. In all fertility models, \( b_F \) includes the “fixed” effect of CE, following the analysis of recursive models shown in the Appendix 1 (Gianola and Sorensen, 2004).

For DO and NINS, systematic effects included in \( b_F \) are the contemporary group (HYS, with 4,600 levels for DO and NINS), the month of successful service (12, plus 1 level for unknown month) and parity-age (9 levels). For DFS, systematic effects are the same as for DO and NINS, except that month of successful service is not included; in that case, the number of levels for HYS is also 4,600. For OFI, \( b_F \) includes HYS (4,600 levels), parity-age, the month of first service (12 plus 1 level for unknown month), and the technician who inseminates at first service (127 levels).

For CE, \( b_{CE} \) contains the effects included in the routine genetic evaluation: contemporary group (interaction between herd, year of calving, and technician who collects CE data, with 3,501 levels), the month of calving (12 levels), the interaction between number of parity and sex of calf (4 levels), and the breed of service sire (2 levels, Holstein and no Holstein). Table 6 gives the effects included in each model.

For 1CE = calving ease; DFS = days to first service; DO = days open; NINS = number of services per pregnancy; OFI = outcome of first insemination.

Table 6. Description of models for days to first service (DFS), days open (DO), number of inseminations (NINS), outcome of first insemination (OFI), and calving ease (CE)

<table>
<thead>
<tr>
<th>Trait</th>
<th>Model1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFS</td>
<td>hys + parity-age + pe + s + e</td>
</tr>
<tr>
<td>DO</td>
<td>hys + mss + parity-age + pe + s + e</td>
</tr>
<tr>
<td>NINS</td>
<td>hys + mss + parity-age + pe + s + e</td>
</tr>
<tr>
<td>OFI</td>
<td>hys + mfs + parity-age + s + e</td>
</tr>
<tr>
<td>CE</td>
<td>hytc + mc + np-sexcalf + breedss + sc + mgs + e</td>
</tr>
</tbody>
</table>

1hys = interaction between herd, year and season of calving; parity-age = interaction between parity and age of calving; pe = permanent effect; s = sire of the cow; e = residual effect; mss = month of successful service; mfs = month of first service; hytc = interaction between herd, year of calving, and technician that collects the CE data; mc = month of calving; np-sexcalf = interaction between parity and sex of calf; breedss = breed of service sire; sc = sire of calf; mgs = maternal grandsire of calf; e = residual effect.
chains (see Figures 1 and 2). The Markov chain did not converge to the posterior marginal distribution (Figure 1). We tried several models, but none of them converged. The complexity of the bivariate analysis between 2 threshold traits (CE and OFI) might make the cow permanent effect difficult to estimate.

\( \mathbf{u}_s \) and \( \mathbf{u}_{CE} \) are the vectors of sire genetic effects for fertility traits and CE. \( \mathbf{umgs}_{CE} \) is the vector of maternal grandsire genetic effects for CE (i.e., the maternal grandsires of the calves). \( \mathbf{u}_s \), \( \mathbf{u}_{CE} \), and \( \mathbf{umgs}_{CE} \) are considered to be distributed following a multivariate normal distribution with mean 0 and covariance matrix \( \mathbf{G}_0 \otimes \mathbf{A} \), where \( \mathbf{G}_0 \) is

\[
\begin{bmatrix}
\sigma_{sDO}^2 & \sigma_{sDOsDFS} & \sigma_{sDOsNINS} & \sigma_{sDOsOFI} & \sigma_{sDOsCE} & \sigma_{sDOmgsCE} \\
\sigma_{sDFS}^2 & \sigma_{sDFSsNINS} & \sigma_{sDFSsOFI} & \sigma_{sDFSsCE} & \sigma_{sDFSmgsCE} \\
\sigma_{sNINS}^2 & \sigma_{sNINSsOFI} & \sigma_{sNINSsCE} & \sigma_{sNINSmgsCE} & \sigma_{sNINS}^2 \\
\sigma_{soFI}^2 & \sigma_{soFIC} & \sigma_{soFI}^2 & \sigma_{soFI}^2 & \sigma_{soFI}^2 \\
\sigma_{sCE}^2 & \sigma_{sCE}^2 & \sigma_{sCE}^2 & \sigma_{sCE}^2 & \sigma_{sCE}^2 \\
\sigma_{mgsCE}^2 & \sigma_{mgsCE}^2 & \sigma_{mgsCE}^2 & \sigma_{mgsCE}^2 & \sigma_{mgsCE}^2 \\
\end{bmatrix}
\]

and \( \mathbf{A} \) is the numerator relationship matrix among sires.

Figure 1. Trace plot of the samples of genetic correlation between direct effects of the outcome of first insemination and calving ease traits (\( \rho_{OFIDIR-CEDIR} \)), considering a permanent effect in the model for the outcome of first insemination.

Figure 2. Trace plot of the samples of genetic correlation between direct effects of the outcome of first insemination and calving ease traits (\( \rho_{OFIDIR-CEDIR} \)), excluding the permanent effect in the model for the outcome of first insemination.
However, because bivariate analyses were run, the following matrix was estimated in each analysis:

\[
G_0 = \begin{bmatrix}
\sigma^2_f & \sigma_{sp}^{CE} & \sigma_{spmg}^{CE} \\
\sigma_{sp}^{CE} & \sigma^2_s & \sigma_{sCEmg}^{CE} \\
\sigma_{spmg}^{CE} & \sigma_{sCEmg}^{CE} & \sigma^2_{mg}^{CE}
\end{bmatrix}.
\]

The term \(\sigma^2_i\) is the genetic sire variance of the trait \(i\), and term \(\sigma_{mg}^{CE}\) corresponds to genetic variance of the maternal grandsire effect for CE. Covariances \(\sigma_{spmg}^{CE}\) and \(\sigma_{sp}^{CE}\) are the genetic covariances between sire effect for fertility traits and maternal grandsire effect for CE, between sire effect and the maternal grandsire effect for CE, and between sires effects for fertility traits and CE, respectively.

\(e_F\) and \(e_{CE}\) are the vectors of residual terms for the 2 traits. They are jointly distributed following a multivariate normal distribution of mean 0 and covariance \(R_0\), where

\[
R_0 = \begin{bmatrix}
\sigma^2_{eF} & \sigma_{eF,CE} \\
\sigma_{eF,CE} & \sigma^2_{CE}
\end{bmatrix}.
\]

\(\sigma^2_i\) corresponds to residual variance for trait \(i\), whereas \(\sigma_{eF,CE}\) corresponds to residual covariance between CE and fertility traits. \(X_F, X_{CE}, W_F, Z_{sCE}, Z_{sp}, Z_{mg}^{CE}\) are known incidence matrices with the appropriate dimensions for each trait. Fixed effects are those included in \(b_F\) and \(b_{CE}\); random effects are \(p_F, u_{sp}, u_{mg}^{CE}\); and the residual terms, represented by \(e_F\) and \(e_{CE}\).

**Implementation.** Estimation was carried out by Bayesian inference using Gibbs sampling. The reader is referred to the previous paragraphs for details related to censored data (for some records of NINS and DO), threshold models (for CE and OFI), and recursive models (considering that CE affects fertility traits). The rest of the Gibbs sampling scheme was similar to a multiple-trait, maternal-effects analysis (see, for example, Sorensen and Gianola, 2002). The missing values were also handled by data augmentation, as described by Sorensen and Gianola (2002), by generating a sample of the missing phenotype as the sum of effects, plus a random residual. The prior distributions for the fixed effects and (co)variance parameters were uniform, bounded between \(10^6\) and \(-10^6\). The full posterior conditional distributions for the location parameters (fixed and random effects) were univariate normal. The full posterior distributions for variance components were inverted Wishart distributions.

In each analysis, Gibbs sampling was carried out through a unique chain of 300,000 iterations, discarding the first 50,000 samples and retaining one every 50 samples. Thus, 5,000 samples were used to compute posterior means and standard deviations. The length of burn-in period was assessed by visual examination of trace plots of the estimates of genetic parameters. Features of the marginal posterior distributions were obtained using the Bayesian Output Analysis package (available at http://www.public-health.uiowa.edu/boa).

**Simpler Analyses.** For the sake of comparison, simpler analyses for fertility traits were carried out to estimate the effect of CE. These analyses included CE as a fixed effect, but they did not consider CE as a trait. Therefore, the existence of genetic correlations between CE on fertility traits was ignored.

**Posterior Analysis of the Effect of CE on Fertility Traits.** Estimates and confidence intervals of the difference between estimates of the effect of CE = 3 (dystocia) and CE = 2 (moderate difficulty) with respect to the effect of CE = 1 (normal parturition) were obtained from the marginal posterior distributions of these effects. The significance level was calculated as the percentage of samples higher than 0.

The effect of CE on OFI, as obtained in the Gibbs sampling, was estimated on the underlying scale and its meaning is not obvious. We transformed each sample in the underlying scale to the visible scale (i.e., percentage of success). This method involves transforming the estimates on the underlying scale to the probability scale. The method is similar to the one used for presentation of PTA for CE (Van Tassell et al., 2003) and can be summarized as follows:

\[
\text{Difference in the visible scale of (CE = 2) with respect to (CE = 1) = }
\Phi [\mu + \text{effect (CE = 2), } \sigma^2_e] - \Phi [\mu + \text{effect (CE = 1), } \sigma^2_e],
\]

where effect (CE = 2) is the effect estimated for CE = 2 on the underlying scale, \(\Phi(x, \sigma^2_e)\) is the cumulative normal distribution of mean \(x\) and residual variance \(\sigma^2_e\), and \(\mu\) is the average of the population. As mentioned, the residual variance is equal to 1. The value on the underlying scale corresponding to the mean on the observable scale is 0.0258. This value was obtained from tables of the normal distribution considering the population mean of OFI reported in Table 4.
Figure 3. Relationship between genetic effects of fertility traits and calving ease traits, where $S_F$ is the genetic effect of sire of cow for fertility traits; $D_F$ is the genetic effect of cow for fertility traits (direct effect); $M_{CE}$ is the genetic effect of maternal grandsire of calf for calving ease traits; $D_{CE}$ is the genetic effect of calf for calving ease traits (direct effect). For fertility data, $DO =$ days open, $DFS =$ days to first service, $NINS =$ no. of services per pregnancy, and $OFI =$ outcome of first insemination.

Genetic effects were transformed to the observed scale, effects and $P$-values were computed as described previously.

**Estimation of Genetic Correlations**

Because of the complexity of the models, Figure 3 displays the relationships between all genetic effects of fertility measures and CE. Because sire and sire-maternal grandsire models were used, the estimates of the additive genetic covariances among sire effects were transformed to an animal model using the following formulas (Kriese et al., 1991). Genetic covariances between direct effects of the $i$th fertility trait and CE were obtained from

$$\sigma_{DF_iDCE} = 4 \times \sigma_{SF_iSCE},$$

where $\sigma_{DF_iDCE}$ corresponds to the covariance between direct effects of the $i$th fertility trait and CE and $\sigma_{SF_iSCE}$ is the covariance between sire effects of the $i$th fertility trait and CE.

Thus, genetic correlations between direct effects of the $i$th fertility trait and CE were obtained from

$$\rho_{DF_iDCE} = \frac{\sigma_{DF_iDCE}}{\sqrt{\sigma_{DF_i}^2 \times \sigma_{DCE}^2}},$$

where

$$\sigma_{DF_i}^2 = 4 \times \sigma_{SF_i}^2$$

and

$$\sigma_{DCE}^2 = 4 \times \sigma_{SCE}^2,$$

where $\sigma_{SF_i}^2$ and $\sigma_{SCE}^2$ correspond to sire variances of the $i$th fertility trait and CE, respectively. $\sigma_{DF_i}^2$ and $\sigma_{DCE}^2$ correspond to direct genetic variances of the traits.

Estimates of maternal genetic variance were calculated following

$$\sigma_{MCE}^2 = 4 \sigma_{mg\sigma_{CE}}^2 - \frac{1}{4} \sigma_{DCE}^2 - \sigma_{DCE-MCE}^2,$$

and

$$\sigma_{DCE-MCE}^2 = 4 \sigma_{CE-mg\sigma_{CE}}^2 - \frac{1}{2} \sigma_{DCE}^2.$$

Genetic correlations between direct genetic effects of fertility traits and the maternal effect of CE were generated from these expressions.
Table 7. Posterior means and standard deviations (SD) for the effects of calving ease scores on fertility traits

<table>
<thead>
<tr>
<th>Trait</th>
<th>Calving ease score</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>DO, d</td>
<td>0.00</td>
<td>0.00</td>
<td>12.63</td>
<td>6.79</td>
</tr>
<tr>
<td>DFS, d</td>
<td>0.00</td>
<td>0.00</td>
<td>3.34†</td>
<td>2.49</td>
</tr>
<tr>
<td>NINS, no.</td>
<td>0.00</td>
<td>0.00</td>
<td>0.19*</td>
<td>0.09</td>
</tr>
<tr>
<td>OFI-liability</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.13*</td>
<td>0.06</td>
</tr>
<tr>
<td>OFI-observed scale, %</td>
<td>0.00</td>
<td>0.00</td>
<td>-5%*</td>
<td>2%</td>
</tr>
</tbody>
</table>

1DFS = days to first service; DO = days open; NINS = number of services per pregnancy; OFI-liability = estimate of conception rate on the underlying scale; OFI-observed scale = estimate of outcome at first insemination (expressed as percentage) on the observed scale.

2Calving ease scores: 1 = no problem; 2 = slight assistance; 3 = needed assistance or caesarean because of the calf’s size (dystocia). Estimates for unassisted parities were set to 0. Estimates for second and third calving ease scores are contrasts to the first score.

†P < 0.10; *P < 0.05; **P < 0.01.

\[
\sigma_{DF_i,MCE} = 4 \times \sigma_{sF_i,mgsCE} - \frac{1}{2} \times \sigma_{DF_i,DCE} = 4 \times \sigma_{sF_i,mgsCE} - 2 \times \sigma_{sF_i,sCE},
\]

where \( \sigma_{DF_i,MCE} \) corresponds to covariance between the direct genetic effect of the \( i \)th fertility trait and maternal effect of CE, \( \sigma_{sF_i,mgsCE} \) corresponds to covariance between sire effect of the \( i \)th fertility trait and maternal grandsire effect of CE, obtained from the analyses.

Thus, the genetic correlations of direct effects of fertility traits with the maternal effect of CE were calculated as

\[
\rho_{DF_i,MCE} = \frac{\sigma_{DF_i,MCE}}{\sqrt{\sigma_{DF_i}^2 \times \sigma_{MCE}^2}}.
\]

RESULTS AND DISCUSSION

Effect of CE Scores

Table 7 has the features of posterior distributions of CE scores on fertility traits obtained using recursive methodology, with estimates of first score set to 0. Because calvings with a CE of 3 (dystocia) cause the most problems for farmers, only estimates of the effect of this level on fertility measures is discussed here. Analyses of the features of posterior distributions (means, SD) of CE scores on fertility traits indicated statistical evidence proving that the effect of CE = 3 was significantly different from the effect of CE = 1. The level of significance was different in each case.

**DFS.** The estimate of effect of score-3 calvings was not significant. In spite of that, the changes in DFS showed a slight (\( P < 0.10 \)) trend toward an increase of DFS for the cow as CE scores increased.

**DO.** The estimate of the effect of score-3 calving was statistically significant (\( P < 0.05 \)), implying a delay in the pregnancy of the cow by 31 d. This result is consistent with those provided by Dematawewa and Berger (1997), who reported an increase of DO of 34 d, although they are not directly comparable because of the differences in data and applied models.

**NINS.** The effect of dystocia was statistically significant (\( P < 0.05 \)). This fact indicates that 0.5 more inseminations are needed to impregnate the cow in the next reproductive cycle after a difficult calving. These results are similar to those provided by Dematawewa and Berger (1997), who reported that 0.22 more services were necessary after a difficult calving.

**OFI.** The effect of dystocia was statistically significant (\( P < 0.01 \)). However, as explained, the value of the estimates for the effect of CE on OFI does not make biological sense, because estimates are expressed on the underlying (unobservable) scale. Therefore, we transformed the differences back to the observed scale. Thus, parities scored 3 for calving difficulty, in comparison with score-1 parities, impaired fertility, because they decreased the incidence of success at first insemination by 12% (\( P < 0.01 \)).

As for the simpler analysis, Table 8 shows the estimates of CE scores on fertility traits obtained without considering the recursive methodology, that is, by analyzing only fertility traits without considering that CE has a genetic background and may have a genetic correlation with the fertility traits. In general, the results indicated that the effects of dystocia on fertility measures were underestimated (except for OFI, because similar estimates have been obtained from both analyses), compared with the results obtained by applying
the recursive methodology. The recursive model gives a better description of the biological relationships, although the estimation might be difficult and a good data set might be required to estimate all the unknowns in the model.

**Genetic Parameters**

Table 9 provides features of the genetic parameters after applying the transformations explained in the Materials and Methods section. As has been observed in previous studies (e.g., Dematawewa and Berger, 1998),

Table 9. Posterior means, standard deviations, medians, and bounds of the high posterior density (HPD) intervals (95%) for heritabilities of fertility traits (ρ²) and calving ease; and genetic correlations (ρ) between direct genetic effects of fertility traits and direct and maternal genetic effects of calving ease.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>HPD 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>h²DFS</td>
<td>0.09</td>
<td>0.03</td>
<td>0.09</td>
<td>0.04 to 0.14</td>
</tr>
<tr>
<td>h²DO</td>
<td>0.06</td>
<td>0.02</td>
<td>0.06</td>
<td>0.03 to 0.10</td>
</tr>
<tr>
<td>h²NINS</td>
<td>0.04</td>
<td>0.01</td>
<td>0.04</td>
<td>0.02 to 0.06</td>
</tr>
<tr>
<td>h²OFI</td>
<td>0.12</td>
<td>0.03</td>
<td>0.12</td>
<td>0.06 to 0.18</td>
</tr>
<tr>
<td>h²CEDIR</td>
<td>0.07</td>
<td>0.02</td>
<td>0.07</td>
<td>0.02 to 0.11</td>
</tr>
<tr>
<td>h²CEMAT</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.01 to 0.09</td>
</tr>
<tr>
<td>ρDFS-CEDIR</td>
<td>0.44</td>
<td>0.26</td>
<td>0.46</td>
<td>−0.07 to 0.89</td>
</tr>
<tr>
<td>ρDO-CEDIR</td>
<td>0.47*</td>
<td>0.24</td>
<td>0.49</td>
<td>0.04 to 0.91</td>
</tr>
<tr>
<td>ρNINS-CEDIR</td>
<td>0.24</td>
<td>0.26</td>
<td>0.24</td>
<td>−0.25 to 0.78</td>
</tr>
<tr>
<td>ρOFI-CEDIR</td>
<td>−0.33</td>
<td>0.26</td>
<td>−0.36</td>
<td>−0.79 to 0.16</td>
</tr>
<tr>
<td>ρOFI-CEMAT</td>
<td>0.15</td>
<td>0.27</td>
<td>0.17</td>
<td>−0.38 to 0.66</td>
</tr>
<tr>
<td>ρDO-CEMAT</td>
<td>0.13</td>
<td>0.25</td>
<td>0.13</td>
<td>−0.36 to 0.60</td>
</tr>
<tr>
<td>ρNINS-CEMAT</td>
<td>0.12</td>
<td>0.28</td>
<td>0.12</td>
<td>−0.43 to 0.66</td>
</tr>
<tr>
<td>ρOFI-CEMAT</td>
<td>0.11</td>
<td>0.30</td>
<td>0.11</td>
<td>−0.45 to 0.71</td>
</tr>
<tr>
<td>ρCEMAT</td>
<td>−0.31</td>
<td>0.28</td>
<td>−0.35</td>
<td>−0.78 to 0.23</td>
</tr>
</tbody>
</table>

1DO = days open; DFS = days to first service; NINS = number of inseminations; OFI = outcome of first insemination; CE = calving ease; DIR = direct genetic effects; MAT = maternal genetic effects. Posterior means of heritabilities for DFS, DO, and NINS were low. Heritability for OFI (0.12) was larger than that found by González-Recio and Alenda (2005) using data originating from the same population. It is necessary to point out that the application of the recursive methodology in this study might have affected the results. Posterior means of heritabilities for DCE and maternal CE were low (0.07 and 0.05), although they are within the range of values reported in other studies (Djemali et al., 1987; Van Tassell et al., 2003). The estimated genetic correlation between direct and maternal components for CE was negative (−0.31 ± 0.28). This value is similar to that used in the genetic evaluation for CE (−0.41; E. López de Maturana, unpublished data). A negative association has also been reported by other authors (Djemali et al., 1987; Van Tassell et al., 2003) and reflects the genetic antagonism between direct and maternal effects.

Only the estimate of the genetic correlation between direct effects of DO and CE was significantly different from 0 (0.47 ± 0.24, P < 0.05). The moderate genetic association between direct effects of DO and CE indicates that what genetically provokes difficult births will provoke a genetic, although moderate, worsening of fertility. The estimate of a genetic correlation between direct effects of DFS and CE was positive and moderate, although it was not significantly different from zero (0.44 ± 0.26, P > 0.05). The estimate of a genetic correlation between DCE and a direct effect of OFI was moderately negative, but it was not significantly different from 0 (−0.33 ± 0.26, P > 0.05). The estimate of a genetic correlation between DCE and a direct effect of DFS is moderately positive, but it does not differ significantly from 0 (0.44 ± 0.26, P > 0.05). Estimates of genetic correlations between a maternal CE effect and direct effects of fertility traits were close to zero.
CONCLUSIONS

After a difficult calving, the reproductive performance of the cow is affected. Although difficult calvings do not significantly affect the interval of DFS, OFI is negatively affected. Therefore, the DO interval and NINS are also negatively affected by difficult calvings. Among these measures, the number of inseminations is accepted by farmers as a desirable goal, and it inherently has more association with physiological functions. All these results may be used to help quantify the cost of dystocia in production systems. Moreover, this study shows that the effect of CE might be included in fertility models.

The estimate of a genetic correlation between direct effects of DO and CE was the only correlation significantly different from 0. This result indicates that genes of the calf related to difficult births are moderately genetically correlated with reduced reproductive success of the cow.

The use of the recursive model methodology allowed us to distinguish between the effect of CE on fertility and the genetic relationship between CE and fertility traits. Estimates of effects and parameters are more realistic and give a better explanation of biological phenomena. Carrying out the analysis is not difficult, but a great number of parameters must be estimated.

ACKNOWLEDGMENTS

We thank EFRIFE (Federation of Holstein Associations from the Basque Country Autonomous region of Spain), who provided the data sets. This work was made possible by grants provided to E. López de Maturana by the “Cándido Iturriaga” foundation and the Department of Education of the Basque Government. A. Legarra was partly financed by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). We thank the referees for comments and for calling our attention to censored data.

REFERENCES


ten days and days to calving in beef cattle. J. Anim. Sci. 82:987–993.
APPENDIX

Equation [23] in Gianola and Sorensen (2004) describes the relationship between phenotypes for a given record by considering that trait 1 affects trait 2, and vice versa (a simultaneous model). Here, we consider 2 traits for simplicity, but the reasoning applies to any number of traits. The equation for the ith record is as follows:

$$Ay_i = Xb_i + u_i + e_i,$$

where \(b_i, u_i, \) and \(e_i\) are fixed, random, and residual terms for the ith record, and

$$\Lambda = \begin{pmatrix} 1 & -\lambda_{1,2} \\ -\lambda_{2,1} & 1 \end{pmatrix}$$

is a matrix containing ones in the diagonal and regression coefficients \(\lambda_{ij}\) in the off-diagonals, with the sign changed. These coefficients describe how trait \(i\) is affected by trait \(j\), that is, the regression coefficient of trait \(i\) on trait \(j\). However, an equivalent model can be described as

$$y_i = Xb_i + Wy_i + u_i + e_i,$$

where \(W = -(\Lambda - I)\) is a matrix that contains 0 in the diagonals and the \(\lambda\) coefficients. This can be further transformed in

$$y_i = Xb_i + Yb_w + u_i + e_i,$$

where

$$b_w = \begin{pmatrix} \lambda_{1,2} \\ \lambda_{2,1} \end{pmatrix}.$$

is a matrix of records for each trait, ordered conveniently. The last representation can be regarded as a model in which traits are “fixed effects” for other traits, and \(b_w\) is a vector of solutions for those “fixed effects.”

This form of presenting the simultaneous model is equivalent to that presented by Gianola and Sorensen (2004; equation [68]), although perhaps easier to recognize. However, the problem in solving for \(b_w\) or, equivalently, \(\Lambda\), is that, for statistical analysis, the determinant of \(\Lambda^{-1}\) is involved in the likelihood function (equation [37] in Gianola and Sorensen, 2004). Intuitively, the reason is that if trait 1 influences trait 2 and the opposite, certain conditions have to be fulfilled for the system to be possible. For a Gibbs sampling scheme, there is no closed form for the fully conditional posterior distribution; Gianola and Sorensen (2004) propose the use of a Metropolis-Hastings algorithm. However, as stated by Gianola and Sorensen (2004), for recursive models (where trait 2 does not influence trait 1), \(\Lambda\) will be a triangular matrix with the diagonal elements equal to 1, for example,

$$\Lambda = \begin{pmatrix} 1 & -\lambda_{1,2} \\ 0 & 1 \end{pmatrix}$$

whose determinant is always 1. Therefore, its elements can be obtained as the solutions of \(b_w\) in the system, which is implemented like any other fixed effect. This is the strategy we used in this work. However, it cannot be applied for simultaneous models, where there are pairs of traits in which trait 1 affects trait 2 and trait 2 affects trait 1.

Example

This example is inspired by Veerkamp and Thompson (1999). Suppose a cow’s record, for a given test day, is 30 kg for milk yield and 500 kg for live weight. A typical model to estimate the genetic correlation between milk yield and live weight would be (model for milk yield):

$$30 = m_{\text{milk yield}} + u_{\text{milk yield}} + e_{\text{milk yield}}$$

(model for live weight)

$$500 = m_{\text{live weight}} + u_{\text{live weight}} + e_{\text{live weight}}$$

in matrix form,
\[
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + u + e,
\]

where \( m, u, \) and \( e \) are general population means, additive genetic effects of the cow, and residuals for that record, respectively. The terms \( u \) and \( e \) follow bivariate normal distributions.

**Recursive Model.** If we consider that live weight influences milk yield (a recursive model), and following Gianola and Sorensen (2004), the models are in matrix algebraic form:

\[
A
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + u + e 
\]

where

\[
A = \begin{pmatrix}
1 - \lambda_{1,2} \\
0 & 1
\end{pmatrix}
\]

The term \( \lambda_{2,1} \) is set to zero because it is assumed that milk yield does not influence live weight. As discussed before, this is equivalent to:

\[
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + \begin{pmatrix}
\lambda_{1,2} \\
0
\end{pmatrix} \begin{pmatrix}
0 & 500 \\
30 & 0
\end{pmatrix} + u + e 
\]

or to

\[
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + b_w \begin{pmatrix}
0 & 500 \\
30 & 0
\end{pmatrix} + u + e.
\]

These models are equivalent to (model for milk yield):

\[
30 = m_{\text{milk yield}} + \lambda_{1,2} 500 + u_{\text{milk yield}} + e_{\text{milk yield}}
\]

(model for live weight):

\[
500 = m_{\text{live weight}} + u_{\text{live weight}} + e_{\text{live weight}}.
\]

This is straightforward to fit into any common genetic parameter estimation software (e.g., VCE or MTDFREML) just by including live weight as a covariate for milk yield and analyzing both traits simultaneously. As explained before, inferences are correct because likelihood functions for both representations ([1] and [2]) of the model are equivalent.

**Simultaneous Model.** If we consider that milk yield also influences live weight, an appropriate model would be

\[
A
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + u + e,
\]

where

\[
A = \begin{pmatrix}
1 - \lambda_{1,2} \\
-\lambda_{2,1} & 1
\end{pmatrix}.
\]

We may also write this model as:

\[
\begin{pmatrix}
30 \\
500
\end{pmatrix} = m + \begin{pmatrix}
\lambda_{1,2} \\
\lambda_{2,1}
\end{pmatrix} \begin{pmatrix}
0 & 500 \\
30 & 0
\end{pmatrix} + u + e.
\]

However, the likelihood function of model [4] is not the same as that of model [3] (and therefore not correct). The reason is that the determinant of \( A^{-1} \), which is involved in the likelihood function for the first model, is ignored in the second. An example would be an estimate of

\[
\begin{pmatrix}
\lambda_{1,2} \\
\lambda_{2,1}
\end{pmatrix} = \begin{pmatrix}
10 \\
0.1
\end{pmatrix}
\]

which could be legitimate under the second model. This estimate is not admissible in the first model because its likelihood is null (the determinant of \( A^{-1} \) is zero).