Genetic Evaluation of Dairy Cattle Using Test-Day Models

J. Jensen
Danish Institute of Agricultural Sciences
Department of Animal Breeding and Genetics
Research Centre Foulum
P.O. Box 50, DK-8830 Tjele
Denmark

ABSTRACT

Recently there has been considerable interest in modeling individual test-day records (TDR) for genetic evaluation of dairy cattle as a replacement for the traditional use of estimated accumulated 305-d yields. Some advantages of test-day models (TDM) include the ability to account for environmental effects of each test day, the ability to model the trajectory of the lactation for individual genotypes or groups of animals, and the possibility of genetic evaluations for persistency of production. Also, the use of test-day models avoids the necessity of extending short lactations on culled animals and animals with records in progress. The disadvantages of TDM include computational difficulties associated with analyzing much larger datasets and the need to estimate many more parameters than in a traditional 305-d lactation model. Several different models have been proposed to model the trajectory of the lactation, including so-called “biological functions,” various polynomials and character process models. At present, there is not universal agreement on which models to use in routine prediction of breeding values and better methods to compare models are desirable. Obtaining accurate estimates of the dispersion parameters to use in TDM remains a challenge. Methods used include a two-step procedure in which the dispersion parameters are estimated in a series of multivariate models followed by a reduction in order of fit using covariance functions, and a one-step procedure in which the parameters of TDM are estimated using restricted maximum likelihood or Bayesian methods in a random regression model. Further research should focus on including multiple lactation data and accounting for heterogeneity variance.

(Key words: genetic evaluation, test-day model, random regression, model choice)

INTRODUCTION

In recent years, the topic of genetic evaluation of dairy cattle using test-day models (TDM) has been investigated by several research groups, and some countries have already implemented routine genetic evaluation of large commercial dairy populations using a TDM. There are several advantages of using TDM compared with the traditional use of yields aggregated over, say, 305 d of lactation. Among these advantages are more precise adjustment for temporary environmental effects on the test day (TD), avoidance of the use of extended records for culled cows and for records in progress, and the possibility of genetic evaluation for persistency of lactation. Major disadvantages are that the volume of data to be analyzed is much larger and that models used will tend to have many more parameters that need to be estimated compared with the traditional models for aggregated lactation yields.

Several reviews on the use of TDM in genetic evaluation of dairy cattle have been presented (Swalve, 1995, 1998, 2000). Gengler (1996) reviewed methods for the estimation of persistency of lactation. Therefore, this paper is not a full review on the use of TDM, but instead presents topics that received less attention in those reviews and presents some new developments. These new developments are related to random regressions (RR) in several dimensions, choice of RR models, and estimation of dispersion parameters. Also, areas are identified in which additional research should be carried out. To put things into perspective, some general aspects of TDM will also be reviewed. Models considered are for single traits such as milk or protein, but in all cases straightforward extensions to the multivariate situation are available.
TEST-DAY MODELS

The analysis of TD records (TDR) entails the analysis of repeated records on an individual. A large amount of literature on this topic is found in the statistical journals. For reviews, see Lindsey (1993) and Vonesh and Chincilla (1997).

A summary of effects that can influence a TDR can be found in Swalve (1995). Many of those effects are the same as those that influence lactation yields (Swalve, 2000). However, special attention needs to be given to stage of lactation that often is modeled using some kind of submodel (Danell, 1990; Grossman and Koops, 1988; Guo and Swalve, 1995). Swalve (2000) divided TDM in the following classes:

Two-step TDM is a model in which the TDR are corrected for TD environmental effects, typically using a model with fixed effects only, and then the residuals from this model are combined into lactation measures that can be analyzed with traditional models for lactation yields. Such models have been presented by Jones and Goddard (1990) and are routinely used in Australia and New Zealand.

TDM with fixed regressions is a model in which the lactation curve is modeled as a fixed effect and the random component of the model are specified as a traditional repeatability model. A generalization of this model to multiple traits was presented by Reents et al. (1995a, 1995b), and such a model is currently used in Germany. A simplified scalar version of a TDM with fixed regressions for records in a single parity could be:

\[ y = \text{HTD} + \sum b_i x_i + a + p + e \]  

where HTD is the fixed herd TD effect, \( a \) is the random additive genetic effect, \( p \) is the random permanent environmental effect associated with each cow, and \( e \) is the random residual. The lactation curve is modeled using the regression parameters \( b_i \) and \( x_i \) are the corresponding covariates. The regressions are typically nested within classes of fixed effects such as age, season, and region. A further generalization could be to nest regressions within each herd-year to take variation in shape of the lactation curve between herds into account. Different feeding strategies can affect the shape of the lactation curve drastically. Considering the random effects, the model is a simple repeatability model that assumes constant additive genetic and permanent environmental variances throughout the lactation. Furthermore, it is assumed that the genetic and permanent environmental correlations between yield at different DIM are unity, regardless of the distance between the days chosen.

RR-TDM is an extension of the TDM with fixed regressions. Here, it is assumed that the shape of the lactation curve is also influenced by random genetic and permanent environmental effects. In this way, the genetic and permanent environmental correlation between yields at different DIM can be less than one. Furthermore, the model can accommodate heterogeneous additive genetic and permanent environmental variances during the lactation. The degree of heterogeneity may depend on the functions chosen to model the trajectory of lactation. The general concept of using RR in an animal breeding context was suggested by Henderson, Jr. (1982) and suggested for use in genetic evaluation of dairy cattle by Schaeffer and Dekkers (1994). In simplified scalar form, the model is:

\[ y = \text{HTD} + \sum b_i x_i + \sum a_k x_k + \sum p_k x_k + e \]  

where \( a_k \) is the additive genetic effect corresponding to regression coefficient \( j \), \( x_j \) are the corresponding covariates, and similarly for the permanent environmental effect subscripted by \( k \). The different subscripts indicate that the covariates in different parts of the model are not necessarily the same. When compared with equation [1], this corresponds to nesting the regressions within the additive genetic and the permanent environmental effects. The covariates \( x_i \) can, in principle, be any covariate but are usually relatively simple functions of DIM such as polynomials, orthogonal polynomials (e.g., Legendre polynomials) or the parameters of the Wilmink function (Wilmink, 1987), i.e.,

\[ x_1 = 1, \quad x_2 = \text{DIM}, \quad \text{and} \quad x_2 = \exp(-0.05\text{DIM}). \]

Many other functions can be used. As an example, Ptak and Schaeffer (1993) used various combinations of linear, quadratic, and logarithmic functions of DIM. In several cases, authors have attempted to develop functions that describe the lactation curves based on biological knowledge of the milk secretion process, e.g., Wood (1967) and Wilmink (1987). More recently, Pollott (2000) proposed a model of the lactation curve based on biological knowledge of the milk secretion process at the cellular level in the mammary gland. The proposed model has two parts. The first part models the increase in milk production mediated by cell proliferation and differentiation, and the second part models the decline in milk production due to programmed secretory cell death (apoptosis). Rekaya et al. (2000) used the nonlinear Wood curve in a hierarchical model to study the genetic variation in the shape of the lactation curve and the relationships between curve parameters and functions of the curve, such as total lactation yield and persistency. In most cases, however, the functions used in genetic evaluation of dairy cattle have been based on relatively simple polynomial functions that are known...
to fit the data reasonably well. Another advantage of such models is the relative ease of implementation compared with the use of models that are nonlinear in the parameters.

**Multiple-trait (MT) TDM** is a model in which records at each DIM, or records in classes of DIM defined in intervals along the lactation, are considered as separate traits. These traits are then analyzed using multivariate methodology (Meyer et al., 1991, Jensen et al., 1997). The advantage of the MT-TDM is that no structure is assumed for the (co)variances among records taken at different DIM. The disadvantage is that many “traits” must be defined, leading to difficulties in estimating the fixed effects and the dispersion parameters pertaining to the additive genetic and permanent environmental effects in the model. Also, because no structure is assumed on the development of (co)variances over time, the parameters can be “jumpy,” perhaps due to the relatively small datasets that are typically used for parameter estimation. Such behavior is in contrast to the expectation of a smooth development over the lactation. A special case was presented by Wiggins and Goddard (1997), in which many traits were reduced to a few by canonical transformation into “pseudo traits” corresponding to the largest eigenvalues of the covariance matrices for all the traits used in the model. Often, only relatively few such “pseudo traits” are necessary to represent almost all variation in the original data. A general description of the use canonical transformations in the estimation of covariance matrices or in genetic evaluation was presented by Jensen and Mao (1988). Extensions that yield approximate canonical variates in models with multiple random effects, such as models with a permanent environmental effect, were presented by Lin and Smith (1990). Furthermore, Duchêne and Besbes (1993) extended the use of canonical transformations to situations with missing values. This was accomplished by the use of an EM-algorithm that replaces missing values with their expectation in each round of the iteration process used to solve for model parameters.

**Covariance function models.** A covariance function (CF) is a function describing the (co)variance among records/traits that are measured at different DIM during the lactation. Such functions can be derived using a RR-TDM, or the covariance function can be estimated using methods such as generalized least squares applied to full fit estimates from MT models (Kirkpatrick et al., 1990). Alternatively, the coefficients of the CF can be estimated directly from the data using REML (Meyer, 1997). If the same functions are used, CF models are equivalent to RR models (Meyer and Hill, 1997; van der Werf et al., 1998). The covariance function can also be used to model the covariances between different traits such as milk, fat, and protein. The advantage of this, especially in the multivariate case, is that a reduced fit can be used, and this can lead to considerably fewer parameters that must be estimated. Such a reduction in the dimension of the model is currently used in the Finnish national evaluation system for dairy cattle (Lidauer and Mantysaari, 1999).

**Character process models (CPM).** An alternative covariance function not considered by Swalve (2000) is the CPM presented by Jaffrézic and Fletcher (2000) and Fletcher and Geyer (1999). The CPM does not attempt to model the random part of the trajectory of the lactation per se, but instead directly estimates the parameters in a model for the covariance structure. Using assumptions of correlation stationarity (correlation is only dependent on the distance in age between two measures), the CPM generally leads to fewer parameters to estimate. The assumption of correlation stationarity can be relaxed somewhat by applying an appropriate transformation of the time scale (DIM).

**Spline models.** White et al. (1999) used cubic splines to model the fixed and random effects on the lactation curve. The cubic spline model consists of a series of cubic polynomials, each defined in an interval on DIM. They are constructed such that they are continuous at the “knots” when moving from one interval to the next. Splines are often used for smoothing data, e.g., Wold (1974). The advantage of splines is that they offer greater flexibility than the functions typically used in RR models. However, they also increase computational demands. Currently, the number of intervals and their borders (knots) must be determined in advance. However, it should be possible to devise an algorithm that estimates the optimum number of knots for fixed and random effects.

From the summary above, it should be clear that many different models can be proposed for TDR, and currently no consensus exists for which model best fits lactation data. It is unlikely that a single “best model” can be found, since local circumstances might determine what effects should be included in the model. In principle, the model that maximizes genetic progress in the population should be chosen for genetic evaluation, even though this may be hard to verify for specific models.

The different models are clearly connected. The TDM with fixed regressions assumes a genetic correlation of unity between records obtained at different DIM. At the other extreme, the MT-TDM assumes no structure on these covariances. The CF model offers a compromise between these extremes. The covariance function is defined as $\Sigma = \Phi(DIM) \text{K} \Phi(DIM)'$, where $\text{K}$ is a matrix of coefficients and $\Phi(DIM)$ are functions of DIM, an
example is the Legendre polynomials used by Kirkpatrick et al. (1990) and Kirkpatrick et al. (1994). If the order of fit of the covariance function is equal to the number of traits in the multiple-trait approach, the two models are equivalent (Meyer, 1998). If the order of fit is reduced to one, the model becomes equivalent to a repeatability model. Intermediate choices of order of fit will tend to yield increasingly smooth functions that describe the variances and covariances of production at different DIM.

HETEROGENEITY OF (CO)VARIANCE

The total variance of a TDR is the sum of genetic, permanent environmental, and residual variance. Heterogeneity of variance through the lactation can occur in each of these components. Several studies have shown that this is the case in practice (e.g., Pander et al., 1992; Olori et al., 1999; Brotherstone et al., 2000; Kettunen et al., 2000; Rekaya et al., 2000; Jakobsen, 2000).

An example from Jakobsen (2000) is shown in Figure 1. The analysis was based on 8075 first-parity Holstein cows. The figure shows the heritability of a TDR in first parity as a function of DIM for different traits (milk, fat, and protein) and for different submodels for the lactation curve. Clearly the estimates of heritability of TDR were not constant throughout the lactation. For milk and protein production there was a tendency towards higher heritability estimates in midlactation, which is in accordance with many other similar investigations. For fat production, the heritability was more constant throughout the lactation.

The RR TDM also models the correlation structure among TDR at different DIM within lactation. As an example, results from Jakobsen (2000) are again used and shown in Figure 2. Genetic correlations between TD close together are close to unity, and the correlations gradually decline as the distance between TD increases. The figure is typical of several studies that modeled the lactation curve using RR models (Brotherstone et al., 2000, Kettunen et al., 2000).

The residual variance is generally also heterogeneous over the lactation and over parities. As an example, Figure 3 shows results from an analysis on Canadian Jerseys by Jensen et al. (2001), who modeled the residual variance in 14 different groups of DIM defined within each of the first three parities. The variance was highest in early lactation and then gradually declined. The variance generally increased with parity, although differences were small in late lactation. Clearly, residual variance increases by parity and is also heterogeneous over the lactation. Rekaya et al. (2000) used change point techniques to model the changing residual variance.
variability using fewer parameters, yielding smoother development of the residual variance as a function of DIM.

In summary, heterogeneity of variance seems to exist for all random effects in a RR TDM. Results indicate that TDM containing only fixed regressions will not fit the data well. The differences in heritability also would indicate that models based on the two-step TDM would need to consider the varying accuracy of each TDR when combining individual records into lactation measures.

Other sources of heterogeneity of variance may also exist. Veerkamp and Goddard (1998) investigated covariance functions that were dependent on herd production level and found clear evidence that genetic variance increased with herd production level. Proper methods to account for this kind of heterogeneity should be developed. Naive methods defining herd production levels as covariates in the model may not be optimal, because herd production level is a function of the data being analyzed, and this will complicate the likelihood function used to describe the sampling process.

MULTIDIMENSIONAL TDM

The models presented so far have all had DIM as the time scale. However, the time scale in a TDM can be any meta-meter. One possibility when analyzing records on several lactations is to include parity as an extra time scale. The DIM scale would model lactation trajectories within lactation and parity would model developments over parities. Such a model could have advantages over a model where each lactation is considered a separate trait, especially if one wants to include all available lactations in the analysis. The latter is often desired if results of the evaluation are to be used for management purposes in addition to genetic selection. Initial experience (Jensen et al., 2001) using a model including RR both within and across parities seems to be able to model TDR over several parities with fewer parameters than a full “one parity one trait” model. More work in this area seems necessary to find ways to incorporate records from all lactations in TDM.

An example of a multidimensional TDM was presented by Veerkamp and Goddard (1998). As mentioned above, they developed a covariance function that depended on DIM and on herd production level. Their analysis clearly showed that additive genetic variance depended on herd production level as well as DIM.

Another interesting use of multidimensional RR models was presented by Ravagnolo and Misztal (2000). In addition to a simple RR model for TDR, they included a RR on a heat index derived from weather information. This model was used to study the genetic variation in heat tolerance in dairy cows. They found clear indications of genetic variation in heat tolerance and also found a slightly negative genetic correlation between heat tolerance and production traits.

PERSISTENCY OF LACTATION

As mentioned earlier, the use of RR TDM can yield measures of lactation persistency. According to a review of Gengler (1996), many different definitions of persistency of lactation exist. A useful definition would be related to the flatness of the lactation curve and, for ease of interpretation, it should be independent of the level of production. The reasons for being interested in a flatter lactation curve at a given level of production include the possibility of using cheaper feed (i.e., roughage) around peak yields (Solkner and Fuchs, 1987) and a reduction in stress due to high peak production (Zimmerman and Sommer, 1973).

In the Canadian TDM, a relatively simple measure of persistency is computed. It is based on the decline in production from d 60 and until d 280 of lactation. (Schaeffer et al., 2000). The advantage of this measure is that it is simple and easily understood by the end users.

Jakobsen (2000) compared several definitions of persistency and concluded that a measure similar to the one used in Canada was optimal if all TD in the period from 60 to 280 DIM was used in the prediction of persistency. Her criteria were the accuracy (h²) of predicted breeding values for persistency and a low genetic correlation between persistency and total lactation yield.

Jakobsen (2000) also investigated the genetic relationship between persistency and disease resistance in dairy cattle. Based on records on production and disease, defined as total number of veterinary treatments for any disease, on about 8000 primiparous Danish Holsteins she found a genetic correlation between persistency and disease resistance in the range 0.20 to 0.50. Therefore, the hypothesis by Zimmerman and Sommer (1973) of an advantageous relationship between flat lactation curves and higher disease resistance was supported.

Dekkers et al. (1996) estimated the economic value of persistency in dairy cattle from the expected impact on feed cost and milk returns. They concluded that the economic value was approximately 5% relative to the economic value of production traits. If the positive relationship between persistency and disease resistance mentioned above can be verified in further studies, the economic importance of persistency may be higher than the estimates of Dekkers et al. (1996). Information on persistency would especially be useful in situations in which disease resistance cannot be measured directly, due to lack of large-scale recording programs, and there-
fore the actual economic value of persistency may well depend on local conditions and production systems.

**METHODS FOR PARAMETER ESTIMATION**

Models for TDR often require the estimation of many parameters. As an example, the model used in Canada includes milk, fat, protein, and SCC, and records in each of the first three parities are regarded as different traits (i.e., a total of 12 traits is included; Schaeffer et al., 2000). For each trait, a five-parameter polynomial is used so the genetic covariance matrix is of order $60 \times 60$, and 1830 (co)variances needed to be estimated for both the additive genetic and permanent environmental effects. This illustrates the considerable challenge in estimating all these parameters precisely.

Several options exist for estimation of parameters. In general, parameters are estimated using either REML (Patterson and Thompson, 1971) or Bayesian methods based on Markov Chain Monte Carlo (MCMC) methodology (Gilks et al., 1996). Applying these methods directly to RR models on large datasets can lead to extremely tedious computations and, in practice, parameter estimation for such models has relied on the analysis of relatively small subsets of the available data.

In REML, the algorithms currently in use are either based on the derivative-free (Graser et al., 1987), the EM, or the AI algorithm (Johnson and Thompson, 1995, Jensen et al., 1997). These algorithms use, respectively, no, first, or first and second derivatives of the restricted likelihood function. In general, use of more derivatives leads to higher numerical stability of the algorithm (Misztal, 1994). For high-dimensional models, the DF and EM algorithms can require many rounds of iteration, whereas the AI algorithm usually requires fewer than 10 rounds before convergence (Madsen et al., 1994). An efficient implementation of a multivariate AI algorithm can be found in Jensen and Madsen (1994). Some iterates in the AI algorithm can yield estimates that are outside the parameter space, in which case the EM algorithm or a combination of the EM and AI algorithm must be used (Jensen et al., 1997). The computational demand per round for the three algorithms is of similar magnitude, with a slight increase when more derivatives are used. The major part of the computational work in all three algorithms involves creating and factoring Henderson’s mixed model equations at each iteration. On most contemporary computers this would limit the number of equations to about 500,000 or fewer.

The problem of setting up and factoring the mixed model equations can be circumvented when using MCMC methods such as Gibbs sampling. Furthermore, these methods based on Bayesian statistics have the advantage of easy inclusion of prior knowledge into the estimates. The computational demands of one round of iteration in these methods is much less than for one round in the REML methods and is of similar magnitude to one round of iteration using the “iteration on data” algorithm for solving the mixed model equations (Schaeffer and Kennedy, 1986). Memory requirements for the algorithm are also of similar magnitude, leading to the possibility of analyzing very large datasets. However, the Gibbs sampler may require many rounds of iteration to reach the stationary posterior distribution and may require many rounds of iteration to ensure accurate estimates due to high autocorrelations between successive samples (Raftery and Lewis, 1992). Therefore, for situations in which both methods are possible to compute, the amount of computation using MCMC methods is often much larger using Gibbs sampling than when using AI-REML. However, for large datasets, the computations can at least be carried out using MCMC methods due to much smaller memory requirements. Such analyses can take several weeks, even on very fast computers.

An alternative to estimating parameters of the RR models directly is to estimate the corresponding covariance function based on estimates from simpler models. The strategy is to use the multiple-trait TDM on traits defined for different DIM along the lactation. Such a model will contain fewer parameters and can be used on a dataset containing many more animals than when applying the RR model directly. The resulting estimates can then be combined into a covariance function using generalized least squares (Kirkpatrick et al., 1990) or similar techniques. Misztal et al. (2000) described a “constructive approach” to estimating covariance functions based on estimates from a variety of sources. The approach consists of defining different functions for variances and covariances along the trajectory of the lactation and estimating the coefficients of these models using generalized least squares as in Kirkpatrick et al. (1990) or similar methodology. The advantage of this approach is that, when parameters at different points along the trajectory are obtained, many different covariance functions can be compared with a minimal effort. However, such a procedure will always contain an “ad hoc” element, and may not lead to estimates that are fully consistent with the models where they are to be applied in practice.

**MODEL CHOICE**

The problem choosing the optimum model has received much attention in the statistical literature (e.g., Burnham and Anderson, 1998; McQuarrie and Tsai, 1998; Raftery, 1996). This literature cannot be reviewed
here, but instead provides a few methods that could be of practical value and are easy to implement.

The International Bull Evaluation Service (INTERBULL) has a set of tests that must be fulfilled in order to have data from national evaluations accepted in international evaluations. The rules are based on work of Boichard et al. (1995). These tests focus on checking estimated genetic trends and age effects in the national models. The primary focus is to ensure fair international comparisons. These rules were developed for lactation models but can be generalized to be used for TDM as well.

If REML or MCMC methods are used for parameter estimation directly in the RR TDM, then more formal methods for model comparisons can be used. With REML, it is possible to make likelihood ratio comparisons and compare these with the proper chi-square statistic. Often such tests tend to favor complex models with many parameters. A more conservative comparison criteria is to use the Bayesian information criterion (BIC), Schwartz (1978) which, in a REML analysis, is:

\[
\text{BIC} = (\log\text{-likelihood}) - 1/2 
\times (\text{number of parameters in the model}) \times n^*,
\]

where \( n^* = n - p \) for \( n \) the number of observations, and \( p \) the number of estimable fixed effects in the model. The selected model is the one that maximizes this criterion.

If Bayesian estimation methods are employed using Gibbs sampling, it is possible to use Bayesian model selection techniques. Different models applied to the same data can be compared using marginal likelihoods (Newton and Raftery, 1994) and by comparing discrepancy measures between observed data and data from predictive distributions under different models (Gelfand, 1996). The marginal likelihood examines the global fit of the model, whereas the predictive distributions can be used to study the ability of the model to predict selected subsets of the data.

The marginal likelihood is the basis for computing Bayes factors (e.g., O’Hagan, 1994) and is:

\[
p(y|M_i) = \int p(y|\theta_i, M) p(\theta_i|M_i) d\theta_i
\]

An MCMC estimator of the marginal likelihood as proposed by Newton and Raftery (1994) is:

\[
\hat{p}(y|M_i) = \left[ \frac{1}{m} \sum_{j=1}^{m} p(y|\theta_i^{(j)}, M_i) \right]^{-1}
\]

where \( m \) is the length of the Gibbs chain applied and \( \theta_i^{(j)} \) is a sample from \( p(\theta_i|y) \).

An alternative to the computation of Bayes factors via the marginal likelihood is to compute the posterior probability of each competing model using reversible jump Markov chain Monte Carlo (Green, 1995). In this method, the model dimension is included as an extra parameter in the sampling procedure. However, efficient implementation of this algorithm in high-dimensional models such as TDM may be difficult.

The predictive distribution can be used to investigate how well the model is able to predict a selected subset of the data. Let \( y' = [y_r, z] \) where \( z \) is a selected subset of \( y \) that for this purpose is set to missing and \( y_r \) is the remaining data. The models can be investigated by computing discrepancy measures involving the observed data and samples from the predictive density \( p(z|y_r, M_i) \). Simulation of samples from this density can be easily implemented in a general multivariate Gibbs sampler for normally distributed traits that allow for missing observations.

A natural extension of model choice using Bayesian methods is Bayesian model averaging (Hoeting et al., 1999). If several models are likely, predictions averaged over these models provide better predictive ability than using results only from the model with the highest marginal likelihood only (Madigan and Raftery, 1994). The reason is that when predictions are based on a single model, the inaccuracy due to the model selection process is not taken into account (Hodges, 1987).

If we want to predict a breeding value for animal \( i \), \( \alpha_i \), say, the posterior distribution of \( \alpha_i \) is:

\[
p(\alpha_i|y) = \sum_{k=1}^{K} p(\alpha_i|M_k, y)p(M_k|y).
\]

This is the average of the posterior distributions of the breeding value for animal \( i \) in each of the \( K \) models considered, weighted by the posterior probability of each model.

The posterior probability for model \( M_k \) is:

\[
p(M_k|y) = \frac{p(M_k)p(y|M_k)}{\sum_{l=1}^{K} p(M_l)p(M_l|y)}
\]

where

\[
p(y|M_k) = \left[ \frac{1}{m} \sum_{j=1}^{m} p(y|\theta_k^{(j)}, M_k) \right]^{-1}
\]

which is the same quantity needed in the evaluation of Bayes factors in the previous section. The posterior mean of the breeding value for animal \( i \) is therefore:
where $\hat{\alpha}_{ik}$ is the posterior mean for animal $i$ by using model $k$. An overview of Bayesian model averaging is given by Hoeting et al. (1999), and more theoretical developments are given by Raftery (1993) and Draper (1995). A theoretical proof of the improved predictive ability by using predictions averaged over competing models as compared to using a single “best” model is given by Madigan and Raftery (1994).

Model averaging has not been applied to TDM, but from theoretical arguments an efficient implementation of such an algorithm should be beneficial, especially because of the many parameters that generally are needed in TDM. However, such an algorithm will further increase the already heavy computational burden encountered when applying RR models to large datasets. Results from Madigan and Raftery (1994) indicate that averaging over a few models with the highest posterior probability would be beneficial, and it may be possible to compute the model weights in advance, simultaneously with estimating the dispersion parameters, and then apply these weights in the routine runs.

CONCLUSIONS

Several different models and evaluation strategies have been proposed for the analysis of TDR from dairy cattle. No consensus currently exists on which model fits the data optimally. The range of models used to analyze TDR seems to be even bigger than the range of models used for the analysis of lactation records. Heterogeneity of variance exists in all sources contributing to the total phenotypic variance, and methods employed must take this into account. Carefully chosen and tested RR or covariance function models will be able to account for all the heterogeneous variances. Such models may well include different covariables in the fixed, the additive genetic, or the random environmental parts of the model.

If all lactations are to be included in the genetic evaluation, methods for this must be developed and evaluated. One possibility is to use multidimensional TDM where the dimensions are parity and DIM. Other effects such as herd production level have a considerable effect on genetic variance, and it seems worthwhile to study new methods to take this into account.

It may be possible to find functions that fit the lactation curve better than the models currently used. Use of spline functions or nonlinear functions might be useful alternatives.

Estimating the many population parameters needed in TDM remains a challenge. Optimum methods should be explored together with methods of selecting the best or the best few models to be used in routine evaluations. Due to the many parameters needed in most TDM it may be advantageous in terms of increased accuracy of evaluations to use some form of robust estimation as for example model averaging.

ACKNOWLEDGMENTS

This review was prepared during a sabbatical at University of Guelph, Ontario, Canada, and University of Wisconsin, Madison. Financial support from the Danish Research council, “Agronomfonden,” Denmark, Department of Animal and Poultry Science, University of Guelph, and Department of Animal/Dairy Science, University of Wisconsin, Madison is greatly appreciated.

REFERENCES


Journal of Dairy Science Vol. 84, No. 12, 2001