Issues in Defining a Genetic Evaluation Model

G. R. Wiggans

Animal Improvement Programs Laboratory, Agricultural Resarch Service,
United States Department of Agriculture, Beltsville, MD 20705-2350, USA
301-504-8334 (voice) 301-504-8092 (fax) wiggans@aipl.arsusda.gov http://aipl.arsusda.gov

Abstract

Genetic evaluation systems provide the basis for genetic improvement programs. National evaluations are combined into an international evaluation for use worldwide. An evaluation model partitions phenotypic records into genetic and environmental effects. Some influences on lactation records may be removed by adjustments before analysis. Some data may be excluded from the analysis because they are affected by factors that are not adequately accounted for by either adjustment before analysis or by the model. National evaluation systems have been tailored to the dairy industries in particular countries. Differences include type of model used, adjustments before analysis, effects included in the model, assumed parameters, solution methods, and reported results. The design goal is to avoid biases by accounting for effects that would influence the estimates of genetic merit. With the increasing importance of international use of evaluations, emphasis is placed on harmonization. Knowledge of practices worldwide can lead to adoption of the best practice for a specific situation as systems are modified to accommodate changes in milk recording and dairy husbandry and advances in computing power and algorithms.

Introduction

The goal of a genetic evaluation system is to generate accurate evaluations that are the basis for genetic improvement programs. The worldwide nature of dairy cattle breeding requires that national evaluations be appropriate for combination into an international evaluation. An evaluation model partitions phenotypic records into genetic and environmental effects and accounts for shared genes through the use of pedigree data. The model is the blueprint for which factors are assumed to affect the phenotypic records and how those effects are distributed. The model should account for environmental factors that might bias estimates of genetic merit.

Some influences on lactation records may be removed by adjustments before analysis based on the model. Some data may be excluded from the analysis because they are affected by factors that are not adequately accounted for by either preadjustment or the model. Those developing national evaluation systems have tailored their systems to their particular situations. Many variations exist on which effects are considered and how they are modeled. Changes in milk recording and dairy husbandry as well as advances in computing power and algorithms have led to updates of genetic evaluation systems. When evaluating possible changes, consideration should be given to harmonization among evaluation systems that contribute to international evaluations because similarity among contributing systems simplifies understanding the results.

Traditional model development

Increased computer power and development of computational techniques has made possible an increase in the complexity of models, thus reducing the likelihood that factors that are not included in the model will bias genetic estimates. The following effects typically are included in a model:

- **Genetic effects.** The random animal genetic effect is the genetic estimate that is the goal of the evaluation. In an animal model, a relationship matrix specifies the expected covariances between animals based on their ancestry. Unknown-parent groups represent ancestors when those are unknown or do not create ties.

- **Environmental effects.** The contemporary group is the primary environmental effect and represents the herd environment where the cow produces. Because yield is greatly affected by the environment, appropriate modeling of this effect is critical to the accuracy and successful use of evaluations. Environments change over time; therefore, the shorter the period that an effect spans, the greater its potential precision. The tradeoff is that the shorter the period, the fewer the number of observations and, therefore, the greater the sampling error of the estimate. A test-day model can define the environment as the effect of a specific test-day in a herd. Age-season effects account for the effect that calving age and season have on yield. In a test-day model, age at milking may replace age at calving. This effect may differ over time and region.

Differences among national systems

Differences in genetic evaluations systems reflect differences in milk recording system, population sizes, computing and statistical techniques that were available when the system was developed, preferences and experience of users, and resources available for making enhancements. Such differences may contribute to an interaction between genotype and environment. Correlations of <1 between evaluations in different countries may result partly from evaluation system differences. The same input data could give different rankings.

Data

Different traits are reported in different countries, and how those traits are reported also varies. Most systems include the milk components of fat and protein (as weights, not percentages). Reporting of somatic cell scores is quite common. In some countries, components are not recorded or analyzed because most milk is not used in manufacturing. Traditionally, lactation records have been analyzed, but the trend is toward analysis of test-day yields.

Countries also vary on how much information is included. Some analyses allow only first-lactation records to minimize computer requirements and to avoid a possible problem with selection bias. Many systems have an upper limit on age or parity to exclude records that are less representative of a cow's genetic merit. Other differences include the completeness of pedigree. Sire identification may be required, which reflects the historical focus on sire evaluation. Including all records may give better estimates of environmental effects if unknown parents can be represented adequately. In lactation models, a minimum lactation length may be imposed. A long minimum length may be set to ensure greater accuracy of the records, but this practice may introduce some culling bias because cows that are culled before their lactations reach that minimum length are eliminated from the analysis. To avoid a need to model heterosis, information from crossbred animals may be eliminated. However, if crossbreds are a significant portion of the population, this approach may not be acceptable. When herds contain several breeds, inclusion of all breeds in a single analysis may be desirable to maximize the size of contemporary groups.

Model

Recently, test-day models have been adopted by several countries. Test-day models better account for environmental effects and better model variation in testing schemes than do lactation models. Some test-day models allow for estimation of genetic differences in persistency. Lactation models are less computationally demanding and allow continued use of all historical data.

A different distinction between models is the animal versus sire model. The name refers to the genetic effect included in the model. A sire model was adopted first because of its lower computational requirements. With improvements in computing power and algorithms, animal models became practical. An animal model makes it possible to include information from all relatives in each animal's evaluation and provides solutions for cows and bulls simultaneously.

Another model difference is single trait versus multitrait. A multitrait model considers correlations among traits so that the information from one trait can affect the evaluation of another. This characteristic is particularly useful when information for one trait might be missing. In a model with each parity treated as a separate trait, a genetic correlation of <1 and a phenotypic correlation different from an overall repeatability can be assumed between parities. This assumption allows for genetic differences in maturity rate, which in turn can reduce changes in bull evaluations as daughters age and second-crop daughters come into production. A multitrait model can protect against culling bias when the culling has been on a correlated trait. A canonical transformation is able to create a set of uncorrelated traits from a set of correlated ones in many cases. This process makes possible multitrait analysis with little additional effort over single-trait analysis.

Observations often are adjusted prior to analysis to minimize computational requirements and to prevent problems with convergence from solving for effects for classes with few observations. The factors used for preadjustment can be smoothed so that values for nearby classes are similar. Over time, preadjustment factors may become dated. Effects included in the model necessarily always fit the data. Preadjustments can be multiplicative and thereby adjust the variance as well as the mean.

Parameters and assumptions

All evaluation systems require an assumption for the heritability of a trait. That estimate specifies what portion of differences are assumed to be genetic and is influenced by the milk recording system and model. The more of the variation accounted for by the fixed effects, the lower the variance assigned to the random effects, which leads to increased heritability.

Multiple parities may be assumed to be repeated measures of the same trait, or a multitrait model could be used to allow appropriate correlation among parities. The assumed correlation would vary with the evaluation system.

Genetic variation may be heterogeneous across herd. If genetic variance is proportional to production level, then genetic differences between the same animals would be less in low-production herds than in high ones. Data can be multiplied by factors to standardize variance, or variance differences can be included in the model.

Solution algorithms

The solution method should not affect the results; however, estimates of genetic trend have been found to be sensitive to the degree of convergence achieved. Even when changes between rounds of iteration are small, estimates of differences over time may continue to increase. Generally, the complexity of the model that is possible to use is determined by the solution methods available. Iteration on the data was critical to enabling the application of animal models to large data sets. Canonical transformation can make multitrait analysis practical in some situations. Faster processing can lead to faster genetic progress by delivering evaluations sooner and reducing generation interval.

Calculation of reliability

A measure of accuracy is desirable for any evaluation system. It provides a measure of the degree of risk associated with the evaluation. Because most evaluations are solved without inverting the coefficient matrix, no exact measure of accuracy is available. Various approximations are used. These usually fit the typical case well, but may be much less accurate for some unusual cases. The recent Interbull requirement to provide effective number of daughters for bull evaluations may lead to more harmonization in the calculation of reliability.

Model effects

When herd size is small, contemporary groups sometimes are extended beyond a single herd. Flexible assignment systems have been applied so that the time period covered by a group is extended to allow a minimum number of records. With test-day models, the contemporary group is those cows that are being milked on a specific test day. This group may be divided further by parity or a management factor within herd, such as milking frequency or milking string.

The reliability of data reporting affects which effects can be considered in the model. Stage of pregnancy has been shown to affect milk yield, and an advantage should not be given to a cow because she required extra time to conceive. However, if breeding dates are unreliable or selectively reported, accounting for this effect is complicated.

Accuracy of reporting also may be modeled. In a test-day model, a test day on which three milkings were weighed might receive more weight than if only one milking was recorded.

Lactation length also may be treated differently in various models. In test-day models, lactation length may be ignored except that a minimum number of test days may be imposed. In lactation models, short records may be extended to 305 days, and this extension may only be applied to records in progress.

Unknown-parent groups

Unknown-parent groups have been widely adopted with animal models to account more precisely for the origins of animals. With recent rapid genetic improvement, identification of ancestors by time period is important. Also, country of origin is important for populations that have been affected by large scale importations, such as black-and-white cattle in Europe. Evaluation systems differ in the number and size of ancestor groups formed, and whether their effects are considered to be fixed or random. The selection path (for example, sire to son) also may be considered. Because of small group sizes, sire-to-son, and dam-to-son paths often are combined. If group sizes are large enough, the risk of biasing genetic evaluations is less if groups are evaluated as fixed effects. If adequate goup size cannot be achieved, treating groups as random will avoid unreasonable group solutions.

Effect incidence

In some genetic evaluation systems, the same effect (for example, parity) occurs as a data preadjustment and in the model. Preadjustment may allow finer adjustment within larger groups that are estimated within the model.

The same effect may appear in more than one place in the model. For example, herd is part of the definition of contemporary group and also is used for herd-specific solutions for age or lactation stage.

Highly correlated effects also may be included in the model (for example, days open and calving interval).

Impact of model choices

In choosing among the various model options, a more complete model is expected to be more accurate. However, which option is best may not be obvious. Options can be tested by applying models to subsets of data. By excluding the most recent years, evaluations can be calculated and their accuracy in predicting results from the full data set determined. Parent averages from the data subset can be compared with evaluations from the full data set for recent animals.

Harmonization

Increased harmonization should increase international confidence in evaluations and simplify explanation of procedures. The goal of harmonization may act as an impetus for national organizations to update evaluation systems. However, an emphasis on harmonization should not be allowed to reduce experimentation, to delay innovation, or to be an undue imposition on national evaluation authority.

An alternative to harmonization of genetic evaluation systems is to combine records instead of evaluations and calculate estimated breeding values at a central site. A feasibility study for this approach is currently being undertaken by Interbull. Even with a single evaluation system, harmonization of reported data would be necessary.

Review of results

Careful review of evaluations after they have been calculated is necessary to detect problems. Typical problems are errors in the data, incomplete implementation of changes, and data falling outside the limits of the programs either in time or magnitude. A powerful test is to see how the results compare with those from the previous evaluation. This is the first thing that users notice, in particular if a bull in widespread use has his evaluation decrease substantially. A list of bulls of interest with the maximum changes should be reviewed and the reasons for those changes investigated if they are greater that what is expected based on the increase in data.

It is useful to process the evaluations from the distribution formats to insure that the last formatting step was successful. Maximum and minimum values should be checked to be sure that they are possible values. Most evaluation systems generate considerable documentation at the various steps. Often more than can be fully reviewed. Comparison of counts with those from the previous run is helpful in seeing if the changes are of the expected magnitude.

The users of the evaluations are usually the best source for identifying possible problems in evaluations. These concerns lead to research projects to detect possible biases and to develop ways to eliminate them. The Interbull center relies on trend validation to access if an evaluation system meets international standard. They also estimate sire genetic standard deviation each run which allows determination if there are changes in the national system. Such changes have been the basis for exclusion of current national evaluations from the Interbull across country evaluations.

Recommendations

National authorities should update evaluation models as computer capacity allows and adopt procedures similar to those used in other countries with similar conditions when appropriate. Data quality should be monitored to ensure that model requirements continue to be met. Periodic assessment of the success of the evaluation system is needed.

Research capability is a necessary part of a national evaluation system. Knowledge of practices worldwide can lead to adoption of the best practice as systems are modified to accommodate advances in computing power and algorithms as well as changes in milk recording, dairy husbandry, and reproductive techniques. Complex, specialized computer code may provide more accurate evaluations but can also make evaluation systems more difficult to maintain when conditions change. Increased sharing of computer programs could save time and reduce differences.

For breeders, timeliness of evaluations is more important than either evaluation accuracy or system harmonization. Phenotypic records should be processed into genetic evaluations quickly and on schedule to avoid delays in selection of animals of superior merit.