Test-Day Model for National Genetic Evaluation of Somatic Cell Count in New Zealand

B. L. Harris and A. M. Winkelman
Livestock Improvement Corporation Ltd, Private Bag 3016, Hamilton, New Zealand

Introduction

In New Zealand (NZ) herd testing is done for milk, fat, protein and somatic cell count (SCC). Currently, routine genetic evaluation is done for milk, fat and protein yields but not for SCC. A new system has been developed to calculate SCC breeding values (BVs) for the dairy industry. The current evaluation of milk, fat and protein involves combining herd test results to predict a 270-day yield deviation, corrected for test day effects. These yield deviations are then used as the records for genetic evaluation. SCC is measured and reported as a concentration (1000s of cells per ml of milk). Hence, expressing the SCC as an accumulated yield deviation would be inappropriate. Many countries have implemented, or are moving towards, random regression (RR) test-day (TD) models for genetic evaluation of herd test data (de Roos et al., 2003; Liu et al., 2003). The use of a TD model for genetic evaluation of SCC would not require the calculation of an accumulated or average yield as the record for evaluation. Additionally RR TD models have the benefits of improved accuracy of evaluation by simultaneously accounting for the fixed effects and genetic, permanent environmental (PE) and temporary environmental (TE) random effects that affect the trait across time. Because of the benefits of using a RR TD model, we chose to develop such a model for the genetic evaluation of SCC in NZ.

Genetic Parameter Estimation

Winkelman (2002) compared genetic and environmental parameters for first-lactation somatic cell score (SCS, log-transformed SCC) estimated using multiple trait (MT) models to those estimated using repeatability and RR TD models with constant or heterogeneous TE variances. For the RR models, Legendre polynomials of orders of fit (where order is degree + 1) from 2 to 5 were fitted to the sire and cow effects. Fitting the TE as a continuous function of time was accomplished using the link function approach described by Jaffrezic et al. (2000). The study found heterogeneity over the lactation of sire, cow and TE variances for SCS. Little benefit was found in fitting a Legendre polynomial of order greater than four to the sire and cow effects. However, the heterogeneity of the sire variance was lower than that of the cow and TE variances, suggesting a lower order Legendre polynomial would provide an adequate fit for this effect. Subsequent work, where MT RR models were used to estimate parameters for lactations 1 and 2, has been done. A subset of the data described by Winkelman (2002) was used for this analysis. The subset included first lactation records for cows in Sire Proving Scheme (SPS) herds in years 1992 to 1997, and their second and third lactation records. For computational ease, and to limit the amount of curvature, third-order Legendre polynomials were fitted to the sire and cow effects. Differences between the variances estimated using third- and fourth-order Legendre polynomials occurred mainly at the beginning and end of the lactation. The increased curvature at the extremes of the interval obtained with the higher order fit was likely a consequence of overparameterisation rather than an improved fit. Hence, the decision was made to use third-order Legendre polynomials to model heterogeneity of genetic and cow effects in the national evaluation. Attempts were made to use a MT RR model to estimate variance components for SCS in lactations 1, 2 and 3. Orders of fit of 2 and 3 both resulted in singularities in the information matrix. A MT repeatability model found that the estimated genetic correlations between lactations 1 and 2 was 0.95 and between 2 and 3 was 0.99. The work suggested that second and third lactation SCS are genetically almost the same trait. Therefore, the model chosen to estimate variance components for the national evaluation of SCS was a MT RR model fitting third-order Legendre polynomials to the sire
and cow effects in lactations 1 and 2. Heterogeneity of TE was estimated using the link approach as described earlier. Using this model, variance components were estimated using Livestock Improvement's SPS herds from years 1995 to 2002. A logarithmic transformation to the base 2 was used to transform SCC. ASREML (Gilmour et al., 2001) was used to estimate the variance components. The estimated variances for random effects by stage of lactation are shown in Figure 1 (lactation 1). The shapes of the curves were similar across lactations and also similar to those obtained from the analysis by Winkelman (2002). The genetic correlation between lactations 2 and 3, estimated using a MT repeatability model, was high (0.98).

Figure 1: Estimated variances for random effects

Statistical Model for National Genetic Evaluation

The model for the national genetic evaluation of SCS was a MT RR TD model where lactation 1 and lactations 2 and 3 are modelled as different traits. A TD SCS record, measured on day $d$ of lactation, was modelled as:

$$
y_{ijklmn} = h_d + a_{ijkl} + ind_{ijkl} + \sum_{i=1}^{6} h_d w_{ijkl} + \sum_{i=1}^{6} r_{ijkl} w_{ijkl} + \sum_{i=1}^{5} \phi_i(d) x_{ijkl} + \sum_{i=1}^{3} \phi_i(d) p_{ijkl} + e_{ijklmn}
$$

$Y_{ijklmn}$ is the TD record for SCS, $i$ denotes trait 1 or 2, $htd_{ij}$ is the $ith$ herd-year-season-TD fixed effect for trait $i$, with season referring to spring or autumn calving period, $age_{ijkl}$ is the $kth$ month class for the age at calving fixed effect for trait $i$ nested within the $lth$ breed class, with breed classes Holstein-Friesian, Jersey, Ayrshire, Holstein-Friesian x Jersey crossbred, and other.

$ind_{ijkl}$ is the $mth$ induced lactation fixed effect class for trait $i$ with induced classes representing induced or not induced calving,

$h_{it}$ is the $sth$ heterosis with covariate value $w_{ith}$ for animal $n$, for trait $i$,

$rc_{is}$ is the $sth$ recombination loss with covariate value $w_{irs}$ for animal $n$, for trait $i$,

$\phi_i(d)$ is the standardised Legendre polynomial of order $t$ calculated at days in milk $d$,

$s_{it}$ is the stage of lactation fixed effect for days in milk $d$ and order $t$ for trait $i$,

$d_{in}$ is the random additive genetic effect for animal $n$ for trait $i$,

$p_{in}$ is the random non-additive genetic and permanent environment effect for animal $n$ for trait $i$,

$e_{ijklmn}$ is the TE effect associated with record $Y_{ijklmn}$.

For the national SCS evaluation, genetic groups were assigned by breed, gender of missing parent, birth year and country of origin. Four breed classes were assigned genetic grouping, namely, Holstein-Friesian, Jersey, Ayrshire-Red, and other breeds. Genetic groups were assigned in 5 year intervals from 1960 to 1980 then yearly, with the first birth year group being prior to 1960. Country of origin was defined as NZ, North American and Other. Gender of missing parent was defined as female or male. If a genetic group had less than 200 animals per group birth years were clustered. No clustering occurred across breed, origin or gender genetic groups.

Heterogeneous Variance Adjustment

The method of Meuwissen et al. (1996) for joint estimation of BVs and heterogeneous variances was used. Essentially, the heterogeneous variances model is a multiplicative mixed model which scales the phenotypic record toward a common base phenotype variance. The estimation of the heterogeneous variances model parameters
occurs during the iterative procedure for estimating BVs. The method accounts for covariances among observations and allows the heterogeneity to be modelled in a flexible way. Pool and Meuwissen (2000) extended the earlier approaches to account for unequal variances due to stage of lactation within a TD framework. Their work was adapted by Lidauer and Mantysaari (2001) for milk production TD models. This approach was adopted for our national SCS evaluation. A multiplicative adjustment factor, $\lambda_{ijk}$, for a stratum $ijk$ within trait was calculated as:

$$\lambda_{ijk} = \exp^{0.5(\bar{d}_j + \beta_j + \beta_k)}$$

where the $\beta$ values are estimates for the heteroskedasticity in the TD model. The model for estimating heteroskedasticity within each trait was:

$$s_{ijk} = \mu + \beta_i + \beta_j + \beta_k + \xi_{ijk}$$

where $s_{ijk}$ is the pseudo-observation for each trait, $\mu$ is the overall mean, $\beta_i$ fixed effect for lactation by year-season subclass, $\beta_j = \sum_{t=2}^5 \phi_t(\bar{d}_j)s_t$ is a 5th order Legendre polynomial for the average stage of lactation ($\bar{d}_j$) in days for contemporary group $j$, $\beta_k$ is the random effect for herd-lactation-season and $\xi_{ijk}$ was the random residual. The repeatability for herd-lactation-season random effect was estimated using SPS data containing 11,241 herd-lactation-seasons and 43,292 TD contemporary groups.

National Evaluation Data

Herd test SCC data was extracted from the New Zealand national database. The extract was done in February 2004 and included all herd test SCC records from seasons 1992 onwards. On average cows had approximately 3.5 herd tests per lactation. Lactation averages of SCC and SCS showed increasing overall levels with increasing lactation number. Contemporary group was defined in the model as herd-year-season-TD within each of lactations 1, 2 and 3. There was some variation in the numbers of cows tested within a herd-year-season. The average number of cows within a herd-year-season was 43.5 (range 1 to 869), 35.3 (range 1 to 632) and 23.4 (range 1 to 632) for lactations 1, 2 and 3, respectively. The breed compositions were very stable across lactation with approximately 56% Friesian, 17% Jersey, 22% Friesian-Jersey crosses and 1.5% Ayrshires with the remainder being other breeds and other crosses.

National Evaluation Results

The distributions of BVs for sires enrolled in the national evaluation program are given in Figure 2. Table 1 provides the summary statistics for sire BVs within and across breed. The genetic trend for SCS by breed is given in Figure 3. The magnitude of heritability and range in SCS sire BVs indicate that selection for SCS would be effective. The positive trend genetic trend in SCS shown in Figure 3 is undesirable and provides further justification for including SCS in the national economic selection index.
Table 1: Summary statistics for estimated breeding values for sires enrolled in the national evaluation

<table>
<thead>
<tr>
<th>Breed</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across Breed</td>
<td>15556</td>
<td>0.116</td>
<td>0.360</td>
<td>2.388</td>
<td>-1.037</td>
</tr>
<tr>
<td>Holstein Friesian</td>
<td>7986</td>
<td>0.061</td>
<td>0.297</td>
<td>1.637</td>
<td>-0.978</td>
</tr>
<tr>
<td>Jersey</td>
<td>5919</td>
<td>0.029</td>
<td>0.284</td>
<td>1.544</td>
<td>-1.037</td>
</tr>
<tr>
<td>Ayrshire</td>
<td>1185</td>
<td>0.648</td>
<td>0.310</td>
<td>2.331</td>
<td>-0.346</td>
</tr>
<tr>
<td>Other*</td>
<td>466</td>
<td>0.823</td>
<td>0.425</td>
<td>2.388</td>
<td>-0.465</td>
</tr>
</tbody>
</table>

* Predominantly Shorthorn, Guernsey and Brown Swiss.

References


Figure 2: Sire breeding values across breed

Figure 3: Genetic trend for somatic cell score