Use of test-day records beyond three hundred five days for estimation of three hundred five–day breeding values for production traits and somatic cell score of Canadian Holsteins

J. Bohmanova,*† F. Miglior,‡ and J. Jamrozik*

*Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, Canada, N1G 2W1
†Dairy and Swine Research and Development Centre, Agriculture and Agri-Food Canada, Sherbrooke, Quebec, Canada, J1M 1Z3
‡Canadian Dairy Network, Guelph, Ontario, Canada, N1K 1E5

ABSTRACT

The Canadian Test-Day Model includes test-day (TD) records from 5 to 305 d in milk (DIM). Because 60% of Canadian Holstein cows have at least one lactation longer than 305 d, a significant number of TD records beyond 305 DIM could be included in the genetic evaluation. The aim of this study was to investigate whether TD records beyond 305 DIM could be useful for estimation of 305-d estimated breeding value (EBV) for milk, fat, and protein yields and somatic cell score. Data were 48,638,184 TD milk, fat, and protein yields and somatic cell scores from the first 3 lactations of 2,826,456 Canadian Holstein cows. All production traits were preadjusted for the effect of pregnancy. Subsets of data were created for variance-component estimation by random sampling of 50 herds. Variance components were estimated using Gibbs sampling. Full data sets were used for estimation of breeding values. Three multiple-trait, multiple-lactation random regression models with TD records up to 305 DIM (M305), 335 DIM (M335), and 365 DIM (M365) were fitted. Two additional models (M305a and M305b) used TD records up to 305 DIM and variance components previously estimated by M335 and M365, respectively. The effects common to all models were fixed effects of herd × test-date and DIM class, fixed regression on DIM nested within region × age × season class, and random regressions for additive genetic and permanent environmental effects. Legendre polynomials of order 6 and 4 were fitted for fixed and random regressions, respectively. Rapid increase of additive genetic and permanent environmental variances at extremes of lactations was observed with all 3 models. The increase of additive genetic and permanent environmental variances was at earlier DIM with M305, resulting in greater variances at 305 DIM with M305 than with M335 and M365. Model M305 had the best ability to predict TD yields from 5 through 305 DIM and less error of prediction of 305-d EBV than M335 and M365. Model M335 had smaller change of 305-d EBV of bulls over the period of 7 yr than did M305 and M365. Model M305a had the least error of prediction and change of 305-d EBV from all models. Therefore, the use of TD records of Holstein cows from 5 through 305 DIM and variance components estimated using records up to 335 DIM is recommended for the Canadian Test-Day Model.

Key words: random regression model, test-day record, variance-component estimation

INTRODUCTION

The test-day model has become a model of choice for national genetic evaluation of production traits in many countries (International Bull Evaluation Service, 2009). The vast majority of countries that use the test-day model include only test-day (TD) records up to 305 DIM. The threshold point of 305 DIM was adopted from lactation models in which yields were summed over a standard lactation length of 305 DIM. The number 305 originates from the seasonal calving practice for a dairy cow to have one calf per year (coinciding with fresh forage grazing), milk for 305 d, and rest for 60 d (days dry). Modern dairy cows are high-producing cows capable of maintaining production for a considerable period, and because drying off cows at high yield is a problem, rebreeding of these cows is often deliberately delayed. Moreover, high-producing cows are often in a negative energy balance at the peak of their lactation because they are not able to consume enough food to cover their energy requirements. Such animals have to mobilize their body reserve, which leads to decreased conception rate and, consequently, longer days open (Friggens et al., 2007).

Peak of lactation is generally considered the period of greatest risk for culling. Cows with extended lacta-
tion have, therefore, a lower proportion of days at risk per lifetime than cows with a standard lactation length (Knight, 2005). Cows with longer lactations are usually more profitable, mainly because they are less affected by pregnancy and, consequently, have lactations with greater persistency (Arbel et al., 2001; Bohmanova et al., 2009).

The increasing trend in milking cows for longer than 305 DIM (Cole et al., 2009) generated a noticeable number of TD records beyond 305 DIM. Six member countries of Interbull have decided to include TD records beyond 305 DIM in their national genetic evaluations (International Bull Evaluation Service, 2009). In Canada, only TD records from 5 through 305 DIM are used in the national genetic evaluation for production traits.

The 2 possible consequences of including TD records beyond 305 DIM are 1) increase of accuracy of 305-d EBV or 2) lack of fit because of inability of regression functions to properly fit the peak or end of the lactation curve when fitting longer lactations.

The Canadian Test-Day Model is a multiple-trait (milk, fat, protein, and SCS), multiple-lactation (first 3 lactations) random regression model (RRM) fitting TD records up to 305 DIM (Schaeffer et al., 2000). The objective of this study was to investigate whether TD records beyond 305 DIM are useful for estimation of 305-d EBV for milk, fat, and protein yield and SCS for Canadian Holsteins using the Canadian Test-Day Model.

### MATERIALS AND METHODS

#### Data

Data were TD milk, fat, and protein yields and SCS from the first 3 lactations of Holstein cows calved from 1988 to 2007. Milk, fat, and protein yields were preadjusted for the effect of pregnancy using adjustment factors as described by Bohmanova et al. (2008a). Three data sets were created for variance-component estimation by random sampling of 50 herds with at least 50 cows from the Canadian Holstein database. These data sets contained TD records with DIM from 5 to 305 (VC305), from 5 to 335 (VC335), and from 5 to 365 (VC365). Only TD records with all traits present were included in data sets for variance-component estimation.

Six additional data sets were extracted from the Canadian Dairy Network database for genetic evaluation purposes. Three data sets contained records up to April 2007 (D07) with DIM up to 305 (D07305), 335 (D07335), and 365 (D07365). Three additional data sets contained only TD records up to April 1999 (D99) with DIM up to 305 (D99305), 335 (D99335), and 365 (D99365). The data sets are described in Table 1.


### Table 1. Description of 3 data sets used for estimation of variance components (VC305, VC335, VC365) and 6 data sets used for genetic evaluation (DO7305, DO7335, DO7365, D999305, D999335, D999365) and description of models

<table>
<thead>
<tr>
<th>Item</th>
<th>VC305</th>
<th>VC335</th>
<th>VC365</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-day records, no.</td>
<td>89,272</td>
<td>93,746</td>
<td>96,593</td>
</tr>
<tr>
<td>Range of test-day records, DIM</td>
<td>5-305</td>
<td>5-335</td>
<td>5-365</td>
</tr>
<tr>
<td>Cows with a record, no.</td>
<td>6,220</td>
<td>6,220</td>
<td>6,220</td>
</tr>
<tr>
<td>Animals in pedigree, no.</td>
<td>18,161</td>
<td>18,161</td>
<td>18,161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data set used for estimation of variance components</th>
<th>VC305</th>
<th>VC335</th>
<th>VC365</th>
<th>VC335</th>
<th>VC365</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC305</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VC335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VC365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data set used for estimation of breeding values</th>
<th>M305</th>
<th>M335</th>
<th>M365</th>
<th>M305a</th>
<th>M305b</th>
</tr>
</thead>
<tbody>
<tr>
<td>M305</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Data sets VC305, VC335, and VC365 contained test-day records with DIM from 5 to 305, 5 to 335, and 5 to 365, respectively. Data sets DO7305, DO7335, and DO7365 contained records up to April 2007 with DIM up to 305, 335, and 365, respectively. Data sets D999305, D999335, and D999365 contained test-day records up to April 1999 with DIM up to 305, 335, and 365, respectively.

2M305 = model with data up to 305 DIM; M335 = model with data up to 335 DIM; M365 = model with data up to 365 DIM; M305a = model with data up to 305 DIM and variance components estimated by M335; M305b = model with data up to 305 DIM and variance components estimated by M365.
Models

Three multiple-trait, multiple-lactation RRM were used in fitting TD records up to 305 DIM (M305), 335 DIM (M335), and 365 DIM (M365). Models were defined as

\[
y_{ijklmnop} = HTD_{ijk} + \sum_{n=0}^{6} \alpha_{ijn} x_n(t) + \text{dim}_{-}c_{lm} + \sum_{n=0}^{4} \beta_{ijno} w_n(t) + \sum_{n=0}^{4} \gamma_{ijno} w_n(t) + c_{ijklmnop},
\]

where \( y_{ijklmnop} \) is the \( p \)th TD record of the \( o \)th cow for a trait \( i \) (TD milk, fat, protein yields or SCS) in lactation \( j \); \( HTD_{ijk} \) is the \( k \)th herd \( \times \) test-date effect for a trait \( i \) and lactation \( j \); \( \alpha_{ijn} \) is the \( n \)th fixed regression coefficient for a trait \( i \) and lactation \( j \) specific to the \( k \)th region \( \times \) age \( \times \) season class; \( n \) is the number of covariates; \( \text{dim}_{-}c_{lm} \) is the fixed effect of the \( m \)th DIM class \((m = 5, 6, 7, \ldots)\); \( \beta_{ijno} \) is the \( n \)th random regression coefficient for the additive genetic effect of cow \( o \) for trait \( i \) and lactation \( j \); \( \gamma_{ijno} \) is the \( n \)th random regression coefficient for the permanent environmental effect of cow \( o \) for trait \( i \) and lactation \( j \); \( x(t) \) is a vector of covariates of size 7 describing the shape of the lactation curve of fixed regressions evaluated at \( t \) DIM; \( w(t) \) is a vector of covariates of size 5 used for random regressions; and \( c_{ijklmnop} \) is the residual. Ten, 11, and 12 classes of residual variance with 30 d each were defined for each lactation for M305, M335, and M365, respectively. Legendre polynomials of order 6 and 4 were fitted for each lactation for M305, M335, and M365, respectively.

In matrix notation, each model can be described as

\[
y = Xb + Zu + Wp + e,
\]

where \( y \) is a vector of observations; \( b \) is a vector of the herd \( \times \) test-date effect, DIM classes, and regression coefficients for region \( \times \) age \( \times \) season class; \( u, p, \) and \( e \) are vectors of additive genetic, permanent environmental, and residual effects, respectively; and \( X, Z, \) and \( W \) are incidence matrices.

The data were assumed to follow

\[
y \mid b, u, p, R \sim MVN(Xb + Zu + Wp, R)
\]

and

\[
\begin{bmatrix}
u \\
p \\
e
\end{bmatrix} \mid \begin{bmatrix} G \otimes A & 0 & 0 \\ 0 & P \otimes I & 0 \\ 0 & 0 & R \end{bmatrix}
\]

where \( R \) is a block diagonal matrix with blocks of size \( 4 \times 4 \) specific to a particular DIM interval and lactation. Residuals were assumed to have the same variance within residual intervals and heterogeneous variance between intervals, and residuals for different DIM were assumed to be uncorrelated. Matrices \( G \) and \( P \) are the random regression covariance matrices \((60 \times 60)\) for the genetic and permanent environmental effects, respectively. The matrix \( A \) is the additive genetic relationship matrix, and \( I \) is an identity matrix; \( MVN \) denotes the multivariate normal distribution.

A Bayesian approach via Gibbs sampling was applied to estimate model parameters. A single, long chain of 100,000 samples was generated for each model. The first 20,000 samples were discarded as a burn-in, and the remaining samples were used to compute posterior means of model parameters. Convergence of Gibbs chains was monitored by visual inspections of plots of samples for selected parameters. The prior distributions of parameters of the model were

\[
b \propto \text{constant},
\]

\[
u_g \mid \Sigma_g \sim MVN(0, A \otimes \Sigma_g),
\]

\[
u_p \mid \Sigma_p \sim MVN(0, I \otimes \Sigma_p),
\]

\[
u_e \mid \Sigma_e \sim MVN(0, I \otimes \Sigma_e),
\]

\[
\Sigma_g \mid \nu_g, \Sigma_g^2 \sim IW(\nu_g, \nu_g^2 \Sigma_g^2),
\]

\[
\Sigma_p \mid \nu_p, \Sigma_p^2 \sim IW(\nu_p, \nu_p^2 \Sigma_p^2),
\]

\[
\Sigma_e \mid \nu_e, \Sigma_e^2 \sim IW(\nu_e, \nu_e^2 \Sigma_e^2),
\]

where \( \Sigma_g, \Sigma_p, \) and \( \Sigma_e \) are genetic, permanent environmental, and residual covariance matrices, respectively; \( MVN \) denotes the multivariate normal distribution; and \( IW \) denoted the inverse Wishart distribution with \( \nu \) degrees of belief and the scaling factor \( \Sigma \). Values adopted for the scaling factors \((\Sigma_g^2, \Sigma_p^2, \Sigma_e^2)\) were obtained by combining results from previous studies that considered TD with DIM \( \leq 305 \) and assumptions made on variances past 305 DIM. Conservative degrees of belief were chosen to represent the vague prior information.

Two additional models (M305a and M305b) were fitted. These models used data up to 305 DIM and variance components estimated by M335 and M365 that were transformed to a 5–305 DIM scale as in Togashi and Lin (2004). The (co)variance matrices of random
regression coefficients (5 × 5) of models M305a (G_a) and M305b (G_b) were converted from a (co)variance matrix of random regression coefficients (5 × 5) estimated by M335 (G_335) and M365 (G_365), respectively. As an example, the transformation is described in details for G_a only.

Let matrices K_a (301 × 301) and K_335 (301 × 301) be genetic (co)variance matrices of daily yields for DIM from 5 through 305 defined as

\[ K_a = \Phi G_a \Phi', \]  

and \[ K_{335} = \Phi G_{335} \Phi', \]  

where \( \Phi (301 \times 5) \) and \( \Phi' (301 \times 5) \) are matrices of Legendre polynomials of order 4 for DIM from 5 through 305 and 5–335 DIM scale, respectively.

Pre- and postmultiplying equation [1] by \( \Phi' (5 \times 301) \) and \( \Phi \) leads to

\[ \Phi' K_a \Phi = \Phi' G_a \Phi' \Phi. \]  

Rearranging equation [3] leads to

\[ G_a = (\Phi' \Phi)^{-1} \Phi' K_a \Phi (\Phi' \Phi)^{-1}. \]  

It was assumed that \( K_{335} = K_{305} \); therefore,

\[ G_a = (\Phi' \Phi)^{-1} \Phi' K_{335} \Phi (\Phi' \Phi)^{-1} \]

\[ = (\Phi' \Phi)^{-1} \Phi' G_{335} \Phi (\Phi' \Phi)^{-1}. \]  

Two genetic evaluations were carried out for each of the 5 models with variance components previously estimated. The first evaluation used D99 data, and the second one used D07 data. Mixed-model equations were solved by iteration on data with a preconditioned conjugate gradient algorithm using a block diagonal preconditioner (Lidauer et al., 1999). A description of the 5 models is summarized in Table 1.

Model Comparison

Predictive ability of models was investigated by computing correlations (RHO) between observed TD records in D07 and records predicted by models, the residual variance (RV), and percentage of squared bias (PSB). Ali and Schaeffer (1987) defined PSB as

\[ PSB_{jnr} = \frac{\sum_{r=1}^{o} (y_{jnr} - \hat{y}_{jnr})^2}{\sum_{r=1}^{o} (y_{jnr})^2}, \]

where \( y_{jnr} \) is the \( r \)th observed record of the \( j \)th trait and \( n \)th lactation, \( \hat{y}_{jnr} \) is the \( r \)th predicted record of the \( j \)th trait and \( n \)th lactation, and \( o \) is the number of records. All 3 statistics were computed for TD records for 5 through 305 DIM only.

Rank correlations were calculated between 305-d EBV and EBV for lactation persistency estimated by M305, M335, and M365 for 9,420 bulls with at least 25 daughters and for 3,817,767 cows. Lactation persistency EBV (PER) was calculated as a difference of 280-d and 60-d EBV:

\[ PER_{io} = \sum_{n=0}^{4} \beta_{ion} z_{n} (280) - \sum_{n=0}^{4} \beta_{ion} z_{n} (60), \]

where \( \beta_{ion} \) is the \( n \)th estimate of the random regression coefficients for the additive genetic effect of animal \( o \) and trait \( i \), and \( z_{n}(280) \) and \( z_{n}(60) \) are values of the \( n \)th Legendre polynomial for 280 and 60 DIM, respectively.

Models were further compared by the error of prediction (ERP) of young bulls’ EBV defined by Sullivan et al. (2005) as

\[ ERP = \sqrt{\frac{\sum_{i=1}^{n} (ebv07_i - pa99)^2}{n}}, \]

where \( ebv07 \) is the 305-d EBV estimated from D07, \( pa99 \) is the average of the animal’s parent EBV estimated from D99, and \( n \) is the number of bulls with no daughters in D99 and at least 25 daughters in D07. For milk, fat, and protein yield, \( ebv07 \) was defined as 305-d EBV, and for SCS, it was defined as average daily EBV. To account for differences in average of EBV between the 2 genetic evaluations, EBV from D07 were shifted by subtracting the average change in EBV from D99 to D07 for a set of bulls whose average EBV was not expected to change. The adjustment was based on 1,867 bulls with at least 25 daughters in D99, no new daughters, and no more than 10 new granddaughters in D07.

To compare stability of EBV, average change of 305-d EBV (AD) of 97 “newly proven bulls” (bulls that received their first official EBV in 1999; that is, had at least 20 daughters with TD records >120 DIM) was calculated as follows:

\[ AD = \frac{\sum_{i=1}^{97} (ebv99_i - ebv07_i)}{97}, \]
where ebv99 is the 305-d EBV estimated from D99. Similarly as in ERP, ebv07 was shifted by subtracting the average change in EBV from D99 to D07 for a set of bulls whose average EBV was not expected to change.

RESULTS

Figure 1 shows the distribution of first-parity TD records from 5 to 600 DIM of Canadian Holstein cows. The distribution is nearly uniform from 5 through 275 DIM, with approximately 71,000 records per day. In later stages of lactation, the number of records per day is notably smaller. At 305, 335, 365, and 400 DIM, the number of TD records per day was 30, 52, 67, and 79% smaller, respectively, than in the interval from 5 through 275 DIM.

The proportion of lactations longer than 305 DIM increased over time (Figure 2). The data set used in this study consisted of data from 1988 through 2005, but because there was a notably smaller number of recorded calvings in 1988 and 1989 (beginning of recording of TD records) compared with other years and because most cows that calved in 2006 or 2007 were still lactating past April 2007 (end of our data set), these 4 yr were not plotted in Figure 2 to avoid a bias in the trend.

In 1990, 45% of first, 34% of second, and 33% of third lactations were longer than 305 d. This percentage reached 50, 40, and 33% in 2005. A larger number of cows tends to lactate beyond 305 DIM in first lactation compared with later lactations.

In this study, 3 ranges of TD records were investigated: the currently used 5–305 DIM range and the 5–335 DIM and 5–365 DIM ranges. With the last 2 scenarios, 1 and 2 more TD records per cow per lactation were allowed to be included in the analyses (assuming a 30-d interval between subsequent tests). Considering only lactations shorter than 366 DIM, 45, 40, and 40% of first, second, and third lactations were ≥306 DIM, respectively. Not all cows lactated beyond 305 DIM; therefore, widening the range of TD records resulted in an average increase of only 0.3 and 0.4 TD records per cow and lactation in the 5–335 DIM and 5–365 DIM scenarios, respectively.

Variance Components

As shown in Table 2, very small changes in average daily heritability for milk, fat, and protein yields and SCS in the first 3 lactations were reported among models. Larger additive genetic (Figure 3) and permanent environmental variances (results not shown)
were observed at the beginning and end of lactation compared with mid lactation in all 3 original models (M305, M335, and M365). The increase of variances was observed 30 d before the last available TD record of a particular model. The increase occurred at earlier DIM with M305 compared with M335 and M365, resulting in larger genetic and permanent environmental variances at 305 DIM with M305.

Residual variance was the same in all 3 models until 275 DIM (Figure 4). In M305, residual variance was noticeably smaller in the interval from 276 through 305 DIM than in previous intervals. In M335 and M365, residual variances remained constant until 305 DIM and decreased in subsequent interval(s). In all models, residual variance was the smallest in the last interval.

Both additive genetic and permanent environmental variances were elevated at the beginning of lactation. Moreover, residual variance was the same in all 3 models from 5 through 275 DIM; therefore, no noticeable differences in daily heritability were reported from 5 through 275 DIM (Figure 5). The 3 models had slightly different trends of daily heritability from 275 DIM through 305, which was caused by differences among models in residual variance in this interval. However, the differences in heritability were very small (≤0.02). In general, daily heritability slightly increased 30 d before the last available TD record of a particular model and then decreased. Fat and protein yields and SCS and milk yield in later lactations showed a similar trend in additive genetic and permanent environmental variances and daily heritability as milk yield in first parity (results not shown).

As given in Figure 6, genetic correlations of 0.68, 0.79, and 0.81 were estimated between cumulative 305-d milk yield and 305-d TD yield in M305, M335, and M365, respectively, indicating that a TD record at 305 DIM was weighted differently in M305 than in M335 and M365. Test-day record at 5 DIM and TD record 30 d before the last available TD record in a particular model had the lowest correlations with cumulative 305-d yield.

![Table 2. Posterior mean estimates of average daily heritability of milk, fat, and protein yield and SCS for first 3 lactations](image)

<table>
<thead>
<tr>
<th>Model</th>
<th>Lactation 1</th>
<th>Lactation 2</th>
<th>Lactation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>Fat</td>
<td>Protein</td>
<td>SCS</td>
</tr>
<tr>
<td>M305</td>
<td>0.45</td>
<td>0.33</td>
<td>0.41</td>
</tr>
<tr>
<td>M335</td>
<td>0.45</td>
<td>0.33</td>
<td>0.42</td>
</tr>
<tr>
<td>M365</td>
<td>0.45</td>
<td>0.34</td>
<td>0.42</td>
</tr>
</tbody>
</table>

1M305 = model with data up to 305 DIM; M335 = model with data up to 335 DIM; M365 = model with data up to 365 DIM.

![Figure 2. Proportion of lactations longer than 305 DIM from 1990 to 2005.](image)
Genetic Evaluation

Small differences in ability of models to predict TD records from 5 through 305 DIM measured by PSB, RHO, and RV are reported (Table 3). Model M305 was the best model for SCS based on PSB, RHO, and RV. For milk, fat, and protein yields, M305 had the smallest RV, but M365 had the smallest PSB. Differences among models in RHO were very small. Model M305 had better predictive ability than both M305a and M305b. Considering only a part of lactation, M305 predicted TD records from 275 through 305 DIM better than M335 and M365, and M335 predicted TD records

Figure 3. Posterior mean estimates for additive genetic variance of daily milk yield in first lactation (M305 = model with data up to 305 DIM, M335 = model with data up to 335 DIM, M365 = model with data up to 365 DIM).

Figure 4. Posterior mean estimates for residual variance of daily milk yield in first lactation (M305 = model with data up to 305 DIM, M335 = model with data up to 335 DIM, M365 = model with data up to 365 DIM).
from 306 through 335 DIM better than M365 (results not shown). Similar to residual variance (Figure 4), RV was in all 3 models the smallest in the last 30 d of the lactation. This indicates that models tend to fit end of lactation better than the rest of lactation, and additional data beyond 305 DIM cause a lack of fit for TD records from 275 through 305 DIM.

Comparing the 3 original models (M305, M335, and M365), M335 had the smallest AD (Figure 7). Model M305a had the smallest AD when comparing all 5 models. The largest differences in AD between the worst and the best model were 25 kg for milk and 0.7 kg for fat, 0.3 kg for protein, and 0.005 for SCS. Error of prediction for milk was lower with M305 than with M335 and M365 (Figure 8). Modified models (M305a and M305b) had smaller ERP than original models (M305, M335, and M365); model M305a had the smallest ERP from all models.

Rank correlations between 305-d EBV and PER from different models for bulls with at least 25 daughters were all >0.99. Rank correlations for cows were >0.99 for 305-d EBV and >0.96 for PER. This may indicate that the range of TD records used for estimation of breeding values did not have an effect on ranking of bulls and cows but, in the long run, could improve stability of breeding values. No notable differences were found among models in terms of convergence rate in solving mixed-model equations.

**DISCUSSION**

The increase of additive genetic and permanent environmental variances and the decrease of residual variances did not occur in all models at the same DIM but approximately 30 d before the last available TD. This indicates that the rapid changes of variances at extremes of lactation are not due to data structure (fewer records) but are the artifacts of RRM. Kachman (2004) noted that variance functions are formed by even polynomials (the highest order of the polynomial is even). This results in a tendency of the variance to grow rapidly at extreme time points. Bohmanova et al. (2008b) reported that variance functions in RRM with linear splines are quadratic between knots and, similarly as in RRM with Legendre polynomials, grow rapidly at the beginning and end of lactation. However, the increase of variances was smaller in RRM using linear splines than in RRM using Legendre polynomials (Bohmanova et al., 2008b). It seems that regardless of the type of function used in RRM, this edge effect of variance functions will always be present.

The edge effect of variance functions was observed both at the beginning and at the end of the lactation. Models M305a and M305b used variance components estimated from M335 and M365, respectively, and TD records only up to 305 DIM. The different ranges of DIM used for variance-component estimation and genetic
evaluation resulted in elimination of the overestimation of variances at the end of lactation. However, the edge effect was still present in the first 30 d of lactation. This artifact could be removed by reducing the range of TD records used for genetic evaluation to 35 through 305 DIM and by estimating 305-d EBV as a sum of EBV from 35 through 305 DIM. An overestimation of genetic-variance functions causes an overestimation of accuracy of 305-d EBV. Therefore, accuracy of EBV estimated from a model with reduced overestimation of genetic variances will be less biased; however, changes in estimates of accuracy are likely to be small.

López-Romero et al. (2003) reported larger additive genetic and permanent environmental variances from 5 through 30 DIM and from 310 through 335 DIM than in the rest of lactation in Spanish Holsteins. Larger

Table 3. Percentage of squared bias (PSB), correlation between observed and predicted records (RHO), and residual variance (RV)

<table>
<thead>
<tr>
<th>Lactation</th>
<th>Trait</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSB</td>
<td>RHO</td>
<td>RV</td>
<td>PSB</td>
<td>RHO</td>
<td>RV</td>
</tr>
<tr>
<td>First Milk</td>
<td>4.7</td>
<td>0.79</td>
<td>16.1</td>
<td>4.5</td>
<td>0.79</td>
<td>16.2</td>
</tr>
<tr>
<td>Fat</td>
<td>5.8</td>
<td>0.75</td>
<td>0.029</td>
<td>5.7</td>
<td>0.74</td>
<td>0.03</td>
</tr>
<tr>
<td>Protein</td>
<td>5.3</td>
<td>0.7</td>
<td>0.018</td>
<td>5.2</td>
<td>0.69</td>
<td>0.018</td>
</tr>
<tr>
<td>SCS</td>
<td>17.5</td>
<td>0.8</td>
<td>1.3</td>
<td>17.8</td>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>Second Milk</td>
<td>5.6</td>
<td>0.88</td>
<td>18.2</td>
<td>5.5</td>
<td>0.88</td>
<td>18.1</td>
</tr>
<tr>
<td>Fat</td>
<td>6.9</td>
<td>0.84</td>
<td>0.036</td>
<td>6.8</td>
<td>0.84</td>
<td>0.037</td>
</tr>
<tr>
<td>Protein</td>
<td>6.2</td>
<td>0.83</td>
<td>0.02</td>
<td>6.2</td>
<td>0.83</td>
<td>0.02</td>
</tr>
<tr>
<td>SCS</td>
<td>16.1</td>
<td>0.83</td>
<td>1.3</td>
<td>16.5</td>
<td>0.83</td>
<td>1.3</td>
</tr>
<tr>
<td>Third Milk</td>
<td>6.3</td>
<td>0.89</td>
<td>19.8</td>
<td>6.2</td>
<td>0.9</td>
<td>19.7</td>
</tr>
<tr>
<td>Fat</td>
<td>7.5</td>
<td>0.86</td>
<td>0.041</td>
<td>7.4</td>
<td>0.85</td>
<td>0.041</td>
</tr>
<tr>
<td>Protein</td>
<td>6.9</td>
<td>0.85</td>
<td>0.022</td>
<td>6.8</td>
<td>0.85</td>
<td>0.022</td>
</tr>
<tr>
<td>SCS</td>
<td>14.7</td>
<td>0.84</td>
<td>1.3</td>
<td>15.1</td>
<td>0.84</td>
<td>1.4</td>
</tr>
</tbody>
</table>

1M305 = model with data up to 305 DIM; M335 = model with data up to 335 DIM; M365 = model with data up to 365 DIM; M305a = model with data up to 305 DIM and variance components estimated by M335; M305b = model with data up to 305 DIM and variance components estimated by M365.

Figure 6. Genetic correlations between cumulative 305-d milk yield and test-day milk yield from 5 through 365 DIM (M305 = model with data up to 305 DIM, M335 = model with data up to 335 DIM, M365 = model with data up to 365 DIM).
Figure 7. Average change of 305-d EBV for milk yield of bulls from their first official proof until 7 yr later (M305 = model with data up to 305 DIM, M335 = model with data up to 335 DIM, M365 = model with data up to 365 DIM, M305a = model with data up to 305 DIM and variance components estimated by M335, M305b = model with data up to 305 DIM and variance components estimated by M365).

Figure 8. Error of prediction (n = 1,867) for milk yield (M305 = model with data up to 305 DIM, M335 = model with data up to 335 DIM, M365 = model with data up to 365 DIM, M305a = model with data up to 305 DIM and variance components estimated by M335, M305b = model with data up to 305 DIM and variance components estimated by M365).
permanent environmental variances were observed from 4 through 60 DIM and 270 through 330 DIM in the Luxembourgian dairy population (Mayeres et al., 2004). de Roos et al. (2004) used RRM with and without herd-specific regression curves for estimation of variance components for Dutch Holsteins. Variances at borders of lactation estimated from the model with the herd-specific effect were smaller than variances estimated from the model without the effect. However, an increase of variances was reported at the beginning (from 5 through 35 DIM) and at the end of lactation (from 305 through 335 DIM) in both models. Results from these studies suggest that the optimal range of TD records for estimation of 305-d EBV seems to be population specific, and recommendations presented in this study cannot be generalized to other populations.

The shape of daily heritability for production traits varies significantly among studies (Misztal et al., 2000). In this study, daily heritability for milk yield was the smallest at the beginning of lactation followed by a steep increase in later stages of lactation. This was in agreement with de Roos et al. (2004). A U-shaped trend of heritability; that is, greatest values at the extremes of the lactation, was reported by López-Romero et al. (2003). Daily heritability estimated by Olori et al. (1999) and Strabel and Jamrozik (2006) increased continuously with stage of lactation.

Model M305 had better predictive ability of records from 5 through 305 DIM and smaller ERP than models M335 and M365, which was expected because both M335 and M365 had to fit much longer lactations and, consequently, had a lack of fit from 275 through 305 DIM. Vargas et al. (2000) compared the goodness of fit of functions for cows with extended lactation using a data set with TD records up to 305 DIM only and a data set with all available TD records. Better fit was obtained from models with TD record up to 305 DIM only. Both M305a and M305b were superior to M305 in terms of stability and ERP of EBV, probably because these 2 models used variance components that had fewer artifacts than variance components of M305. Additional TD records beyond 305 DIM did not improve estimation of 305-d EBV. This was caused by lack of fit when fitting longer lactations and because cumulative 305-d milk yield was only moderately genetically correlated with TD yield beyond 305 DIM.

**CONCLUSIONS**

Models that used TD records only up to 305 DIM were superior to models with a larger range of TD records in terms of predictive ability and ERP. Models M335 and M365 had fewer artifacts in additive genetic, permanent environmental, and residual variances from 5 through 305 DIM. The model that used TD records from 5 through 305 DIM only and variance components that were transformed to a 5–305 DIM scale from the estimates based on TD records from 5 through 335 DIM is recommended for estimation of 305-d EBV in the Canadian Test-Day Model. However, this suggestion is only applicable to Canadian Holsteins, and separate studies would be required for other dairy populations.

**ACKNOWLEDGMENTS**

This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: http://www.sharcnet.ca). Funding was provided by DairyGen Council of Canadian Dairy Network and the Natural Sciences and Engineering Research Council of Canada. Two anonymous reviewers are thanked for their helpful comments and suggestions.

**REFERENCES**


