Effects of Random Regression Test-Day Models on EBVs and Genetic Trends in Persistency

Marco H. Pool (M.H.Pool@id.wag-ur.nl) and Theo H.E. Meuwissen
Institute Animal Science & Health, Box 65, 8200 AB Lelystad, Netherlands

Introduction

Breeding values (**EBVs**) of dairy cows currently predicted in the Netherlands are from an animal lactation model based on information of realized or projected 305-d lactation yield (Wilmink, 1987). Today, the use of single test-day records instead of 305-d lactation yield is of more interest for the genetic evaluation of dairy cows as well as a tool for farm-management. The advantage of a TDM compared to a lactation model is that the former can correct more precise for effects as herd-test-dates (Swalve, 1995) simultaneously with the prediction of breeding values. Above that, EBVs are provided for each day along the trajectory and therefore TDM can describe differences in the shape of lactation curves between animals (Schaeffer and Dekkers, 1994), i.e. differences in the progression of milk production during the lactation period.

In this study, the random regression (**RR**) approach described by Schaeffer and Dekkers (1994) was applied. In the RR-TDM the lactation curve is modeled by a random function for each animal and the underlying covariance function structure is estimate directly from the data (Meyer, 1998). Not the lactation model but a re-peatability test-day model (**RPM**) was used to mimic the current situation. Because genetic variation is assumed constant during the whole lactation period, the 305-d EBV yield from the RPM model or for any part of the lactation period is just the average genetic daily prediction times the length of the (part)lactation. The genetic merit of an animal in the RR-TDM is, however, different for each DIM along the lactation trajectory which means that features of a RR-TDM are not fully expressed if EBVs are summarized into 305-d productions. It is better to express the EBVs from a RR-TDM as genetic merit lactation curves (Swalve, 2000) or by parameters that describe the persistency of a lactation. It is expected that RR-TDM leads to a reduced generation interval and to more accurate predictions of genetic merit (Swalve, 2000). In this study, EBVs were predicted with a TDM and compared to a RPM. We compared the TDM with a RPM instead of a lactation model because results from a RPM would be better comparable to a TDM, (i.e., comparable models and similar fixed effect corrections). Because, the RPM models only the level of production, as in a lactation model, it was assumed that selection for milk production was for the same trait, i.e. same type of lactation curve.

Materials and Methods

Data

First lactation test-day records from Holstein-Friesian dairy cows on 9680 herds (calved between April '90 and January '99) in the northern part of the Netherlands were used. Selection criteria for the first lactation records were: known birth date (i.e. for pedigree information), at least 50% Holstein-Friesian, age at calving 22 to 32 months, all test-day records realized on the same farm, and from herds with at least 25 first lactations in the data. The data contained 282,548 first lactations and 2,563,919 test-day records. Cows were offspring of 8323 sires and 197,226 dams. Pedigree had 577,769 entries (4,5% sires).

Model

Test-day records were modeled by Legendre polynomials (**LEG**) (Kirkpatrick et al., 1994) with a 0-th order for the RPM [LEG(0)], and a 4-th order for the TDM [LEG(4)]. Parameter estimates from Pool et al. (2000) were used. LEG(0) fits one parameter for the genetic component and one for the permanent environmental part (i.e., 2 RR coefficients per animal). LEG(4) fits five for both the genetic and permanent environmental component (i.e., 10 RR coefficients per cow). The general model of analyses was:

$$y_{ij} = \mathbf{x_{ij}}' \boldsymbol{\beta} + \boldsymbol{\phi_{ij(m)}}' \mathbf{k_{i(m)}} + \boldsymbol{\phi_{ij(m)}}' \mathbf{l_{i(m)}} + e_{ij} \qquad [1]$$

where, y_{ij} = test-day milk yield j for animal i; $\mathbf{x}_{ij}'$ = incidence row vector for fixed effects $\boldsymbol{\beta}$; $\boldsymbol{\beta}$ = [μ; ys; age; cDIM; HTD]', with μ = the overall mean, ys = year season of calving (3 monthly within year), age = age at calving (4 monthly), cDIM = weekly classes for days in milk (**DIM**) and models the shape of the average lactation curve (44 classes), and HTD = herd-test date effect (181,716 classes); $\boldsymbol{\phi_{ij(m)}}$' = $\mathbf{t}_{ij} \boldsymbol{\Lambda}_m$, which model the RR coefficients of test-day record j for animal i, with $\mathbf{t}_{ij}$ = (1 by m) row vector with standardized DIM (ranging from –1 to 1) to the power equal to the order of fit [0...(m-1)], with m = order of fit, and $\boldsymbol{\Lambda}_m$ = matrix of polynomial coefficients on the Legendre scale; $\mathbf{k}_{i(m)}$ and $\mathbf{l}_{i(m)}$ are (m by 1) vectors of covariance function coefficients from animal i for the additive genetic and permanent environmental effects; e_{ij} = residual term of test-day j of animal i and models measurement errors.

The genetic (co)variance function was: Var($\mathbf{u}$) = $\mathbf{A} \otimes \mathbf{G}$, where $\mathbf{A}$ = additive genetic relationship matrix, $\otimes$ = Kronecker product function, and $\mathbf{G}$ = var($\mathbf{k}_{(m)}$) = additive genetic covariance matrix of the RR coefficients. Similarly for permanent environment: Var($\mathbf{p}$) = $\mathbf{I} \otimes \mathbf{P}$, where $\mathbf{I}$ = identity matrix, and $\mathbf{P}$ = var($\mathbf{l}_{(m)}$) = permanent environmental covariance matrix of the RR coefficients. And the residual variance structure (σ_e^2) was assumed diagonal and constant over DIM.

Model Comparison

EBVs were estimated for a series of times twice per year (June and Dec.). Starting in 1991 with one year of test-day data, up to 1999 with nine years of data. Deleting all test-day records after a certain date created sub sets. Average EBVs were compared for different groups of animals. Young bulls were grouped by year of birth and compared based on different amounts of information (i.e. running and complete lactations for first versus second crop daughters, respectively). All EBVs were compared to the average EBV of cows born in the year 1989 (i.e., first cows with test-day records from complete lactations).

Result and Discussion

Correlations between 305-d EBVs from the RPM and the RR-TDM where in general high. For young bulls (grouped by year of birth) correlations between 305-d EBVs, ranged from 0.85 to 0.99 when compared for pedigree and second crop daughter information, respectively. Schaeffer et al. (2000) presented correlations from a multi-trait RR-TDM which were slightly lower as in this study. Standard deviations for 305-d EBVs were in the RR-TDM as high as for the RPM, and indicate that the RR-TDM models more variability between animals. In the December 1999 run, the trend in 305-d EBVs for young bulls (Figure 1) was 106 kg of milk per year with the RPM and 94 kg of milk per year with the RR-TDM. The shape of the genetic lactation curve (based on the RR-TDM, Figure 2) was similar for different groups of young bulls, indicating that selection in the past had been for an increase of the level of production, and not for more or less persistent animals. The slightly lower trend in level of 305-d EBVs from the RR-TDM compared to the RPM could be affected by differences in fixed effect corrections but is more likely the result of biased estimates from the lactation model. Comparisons presented so far are based on a 305-d lactation yield and do not include all information that can be provided by the RR-TDM. Therefore they should be expressed as curves with the genetic merit plotted over the whole lactation trajectory (Figure 3). Curves with genetic merit from the RPM (dotted lines) are flat and express only the average level of production. Conversely, the genetic merit lactation curves of the RR-TDM show a variable additive genetic daily EBV across DIM. Lactation curves for different groups of animal were similar in shape and seemed to vary only in level of production. However, curves of individual animals do differ for both the level and progression of production (Figure 4). For example, the daily EBVs increased continuously for sires 3 and declined for sire 1. With RPM (dotted lines) EBVs for sires 2 and 4 were comparable (Δ 0.6 kg per day). In the RR-TDM the genetic merit lactation curve for sire 4 is more persistent than for sire 2. The 305-d EBV for sire 2 was equal in both models but for sire 4 the EBV from the RR-TDM was 90 kg higher than in the RPM, indicating downwards biased EBVs for the more persistent animals from RPM.
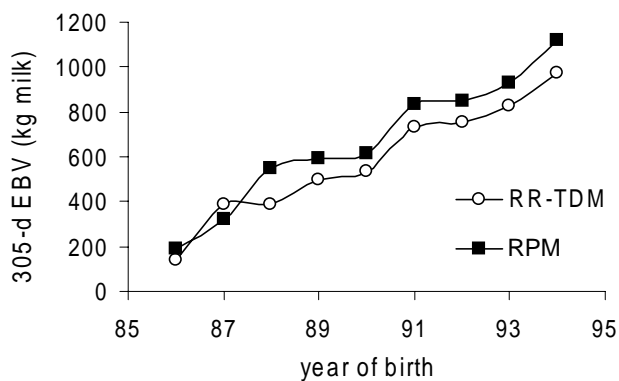
Figure 1.   Trend in 305-d EBVs for youngbulls from a random regression test-day model (RR-TDM) and a repeatability test-day model (RPM). Youngbulls are grouped by birth year and EBVs are standardized to cows born in 1989.
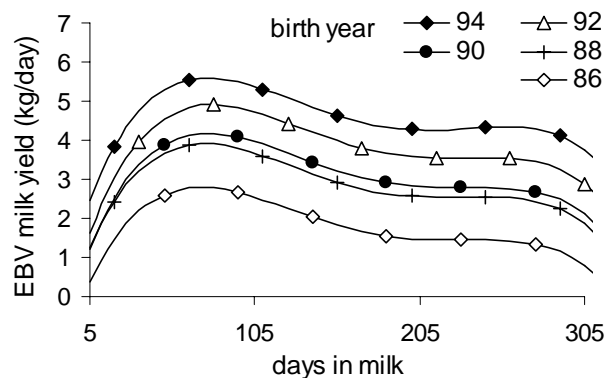
Figure 2.   Genetic merit lactation curves for groups of progeny tested bulls (grouped by birth year) from a random regression test-day model (RR-TDM). EBVs shown were from the December 1999 run.
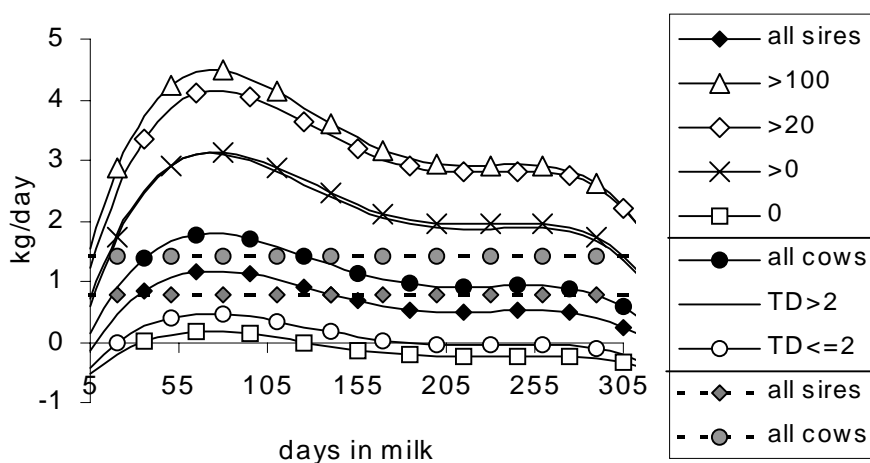

Figure 3.   Genetic merit lactation curves for different groups of animals from a random regression test-day model (solid lines) and a repeatability test-day model (dotted lines). EBVs shown were from the December 1999 run (i.e. the complete data set with 2.56 million test-day records from 282,548 first lactations).
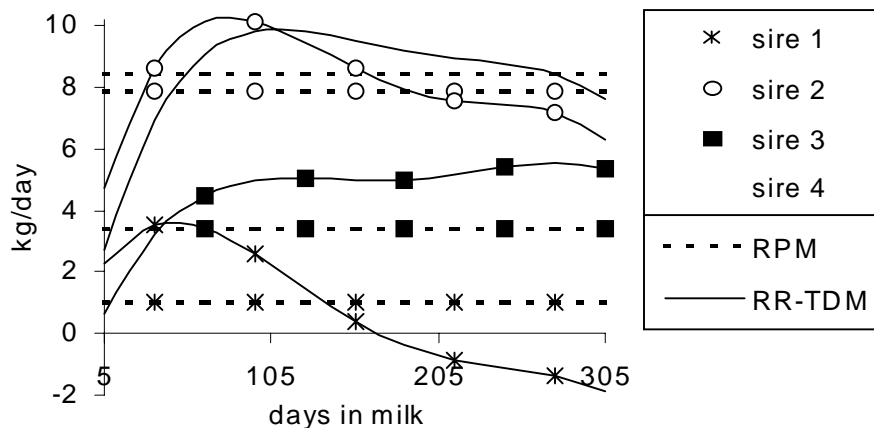
Figure 4.   Genetic merit lactation curves for four sires for both the random regression test-day model (solid lines) and the repeatability test-day model (dotted lines). EBVs shown were from the December 1999 run for sires born in 1992.

Table 1.   Differences in EBV for groups of youngbulls (EBV$_{(first\ crop)}$ minus EBV$_{(2nd\ crop)}$)

| Birth-year | overall | | more persistent | | less persistent | |
	RPM	RR-TDM	RPM	RR-TDM	RPM	RR-TDM
1992	-36.6	3.8	-111.7	-15.9	41.8	24.2
1990	-31.7	4.0	-67.9	-9.3	2.5	16.6
1988	-28.8	1.95	-100.9	-40.9	46.7	46.8
average	-23.5	14.3	-66.1	1.7	20.2	27.3

In Table 1 values for the stability of EBVs are calculated to express differences between EBV of young bulls when they are on average 4.5 and 6 years old (i.e., EBVs based on running and complete lactations from first versus second crop daughters, respectively). EBVs of young bulls (grouped by birth year) predicted by the RPM were at the age of 4.5 years in general lower than at the age of 6 years old, where the RR-TDM showed only slightly lower EBVs at the age of 6 years old. However, differences were relatively small. To compare differences in persistency between the RPM and RR-TDM, young bulls were grouped for persistency and again compared. The genetic production on day 280 minus day 60, as suggested by Jamrozik et al. (1998), was used as measure for persistency (p280-p60). When average 305-d EBVs of young bulls are compared based on their persistency, differences were larger. A sire was assumed to be more persistent when the EBV$_{p280}$ minus EBV$_{p60}$ was smaller than on average and the other way around when less persistent. For the more persistent young bulls (Table 1, col. 4 and 5) the RPM showed a larger increase in average 305-d EBVs compared to the RR-TDM. For the less persistent young bulls differences were comparable. This shows that EBVs of more persistent bulls are predicted more precise with the RR-TDM.

References

Kirkpatrick, M., Hill, W.G. & Thompson, R. 1994. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Gen. Res. Camb. 64,* 57-69.

Meyer K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *GSE 30,* 221-240.

Pool, M.H., Janss, L.L.G. & Meuwissen, T.H.E. 2000. Genetic parameters of Legendre polynomials for first parity lactation curves. *J. Dairy Sci. 83,* 2640-2649.

Pool, M.H. 2000. Test-day models, breeding value estimation based on individual test-day records. *Thesis,* Wageningen-UR.

Schaeffer, L.R. & Dekkers, J.C.M. 1994. Random regressions in animal models for test-day production in dairy cattle. *Proc 5th WCGALP. Guelph, Canada 18,* 443-446.

Swalve, H.H. 1995. Test day models in the analysis of dairy production data - a review. *Archiv für Tierzucht 38,* 11-31.

Swalve, H.H. 2000. Theoretical basis and computational methods for different test-day genetic evaluation models. *J. Dairy Sci. 83,* 1115-1124.

Wilmink, J.B.M. 1987. Comparison of different methods of pre-dicting 305-day milk yield using means calculated from within-herd lactation curves. *Livest. Prod. Sci. 17,* 1-17.