Experiences with a Genetic Evaluation Using Test Day Data

R. Reents, L. Dopp and F. Reinhardt
Vereinigte Informationssysteme Tierhaltung w.V. (VIT), Heideweg 1, 27283 Verden, Germany
Internet: www.vit.de

Abstract

A multiple lactation test day model was applied to predict genetic merit for dairy production traits in German Holstein cattle. The model for test day genetic evaluation included a fixed herd-test-date effect, fixed regressions on functions of days in milk, random permanent environmental effects within lactation, random animal genetic effects, and residual effects. Test day evaluations for cows were compared to evaluations obtained with a lactation model as used for official evaluation in Germany. Correlations between EBV from test-day model and EBV from lactation model were in a range of .87 to .92 for older cows with complete test day information, but were as low as .83 to .85 for young cows with few test day records. Adjustment for heterogeneous variances across herds is one of the most important next steps to apply a test day model for official routine evaluation in Germany.

Introduction

Analysis of test-day records in genetic evaluation procedures instead of lactation observations has received attention. Removing residual variance connected with individual herd-test-day effects was a first step to make better use of individual test day records on milk production traits (Jones and Goddard, 1990; Van Tassel et al., 1992; Everett et al., 1994). In this approach a correction on a herd-test-day basis is done prior to aggregation to a lactation yield. Ptak and Schaeffer (1993) described a model that defines all environmental and genetic effects on a test day basis. Beside the consideration of a herd-test day effect in the model this test-day model gives more flexibility, since:

- Terminated lactations have not to be extended to a 305 day yield (no selection bias due to culling for production in early lactation)
- Records in progress can be analysed without extending them to 305 day yield
- Different recording schemes (in terms of interval between test day results and of accuracy of individual test day observations, e.g. am/pm schemes) across herds can be accomodated, since the amount of information for each individual test day result can be considered in the evaluation

The model of Ptak and Schaeffer (1993), which is a first lactation, single trait model was extended by Reents et al. (1995 a,b) to multiple lactations, in which test day records are considered as different traits across lactations. The multiple lactation test-day model (MLTDM) included both random animal additive genetic and permanent environmental effects as different traits for each parity.

This model so far is used for official genetic evaluation for Somatic Cell Count data in Canada (Reents et al., 1995b) and Germany (Reents, 1996). Starting in December 1995 a modification of this procedure is used twice a year for producing intermediate genetic evaluation of milk production traits of Holstein cattle in Germany, from which only evaluations for bulls are released to the AI studs for early information. Results on sire evaluations were presented at the Interbull meeting in Veldhoven, The Netherlands, 1996. In the meantime characteristics of the cow evaluations, which were obtained simultanously with the sire evaluations in the animal model, were analysed.

Objectives

The objectives of this study were

a) to describe the application of MLTDM for genetic evaluation for the German conditions

b) to analyse EBVs for cows estimated with the MLTDM

Materials and methods

Data

Data consisted of test day records from the database maintained at Vereinigte Informationssysteme Tierhaltung w.V. (VIT), Verden, Germany, which contains about 75% of all Holstein cows in Germany. Different from the official evaluation, where lactation records from 1979 to present are used, test day records in MLTDM are only used from 1990 to present, because not in all regions of Germany test day results are available from before 1990. The second reason is that a population wide analysis of protein content was started in the former GDR in 1990. Edits were on:

- age of calving in months (20 to 40, 30 to 56, and 44 to 75, for lactations 1 to 3, respectively)
- day in milk of the sample between 5 and 325
- and interval between consecutive tests from 7 to 90 days.

Following edits in the June 1997 evaluation 70,836,504 records from lactation 1 to 3 on 4,623,277 cows remained. Pedigree was completed for cows with identification of dam and maternal grandsire from the national pedigree file. Pedigrees for bulls with daughter records or granddaughter records were completed for several generations. Unknown parents were assigned to phantom parent groups, grouped by birth year of offspring (5 years per interval).

Model

For genetic evaluation of test-day observations, a multiple trait test day model (MLTDM) with repeated observations within each lactation was used.

The statistical model for analysis of test-day records was:

$$y_{ijkmn} = HTD_{im} + A_{jm} + P_{jm} + RASC_{km} + b_{km1}(D/c) + b_{km2}(D/c)^2 + b_{km3}\ln(c/D) + b_{km4}[\ln(c/D)]^2 + e_{ijkmn}$$

where

y_{ijkmn} is the n^{th} test day observation of the j^{th} cow in parity m
HTD_{im} is a fixed herd-test-date effect
A_{jm} is an animal additive genetic effect (random)
P_{jm} is a within lactation permanent environmental effect to account for common environmental effects associated with all test-day records of the j^{th} cow in parity m (random)
$RASC_{km}$ is a region x age_of_calving x season_of_calving x calving_interval subclass mean effect in parity m
b_{km1} and b_{km2} are regression coefficients on the linear and quadratic effects of D/c, where D is days in milk and c=381
b_{km3} and b_{km4} are regression coefficients on the linear and quadratic effects of ln(c/D)
e_{ijkmn} is a random residual effect

Regression coefficients were estimated within 945 lactation x RASC groups, resulting from: 3 lactations, 3 regions, 5 age_of_calving groups, 3 season_of_calving groups (Jan-Mar; Apr-Aug; Sep-Dec), 7 groups for calving_interval (< 320 days, 321-350 days, ..., > 470 days). Contemporary groups for second and third lactation records from a specific herd-test-date were combined into a common herd-test-day class to increase the size of subcells.

The statistical model for the official genetic evaluation with the multi lactation model (MLM) is:

$$y_{ijm} = HYS_{im} + a_{jm} + e_{ijm}$$

where

y_{ijm} is the yield of cow j in part-lactation m (three 100 day parts from first lactation, the second lactation and the third lactation are considered as 5 genetically different traits)
HYS_{im} is a fixed herd-year-season effect
a_{jm} is a random additive genetic effect of animal j, and
e_{ijm} is a random residual effect

Variance components used for MLTDM were as in Reents et al. (1995a) and for MLM as used in the official evaluation in Germany.

Results and discussion

EBVs from MLTDM are on a per day production scale. Currently these evaluations are standardised to a scale with mean of zero and standard deviation of 1 in order to avoid a confusion with the official evaluations on a 305 day basis. If EBVs from MLTDM are required on a 305 day equivalent the per day production proof can be multiplied by 305 as proposed by Ptak and Schaeffer (1993). Table 1 displays standard deviations of estimated breeding values for cows from different birth years. Cows from birth year 1990 had the opportunity to finish first 3 lactations from which test day results are included. Cows from birth year 1994 are youngest animals which had only incomplete data for first lactation. Standard deviations for milk yield and protein yield were consistently higher for MLTDM EBVs compared to the lactation proofs from MLM.

Table 2 displays correlations for the different groups of cows between MLTDM and MLM. For older cows with completed lactations correlations between both methods were in a range of .87 to .92. For younger cows from birth year 1994 correlations between EBVs from both models were about .81 to .85.

Beside statistics about all cows in the population also TOP cows from both models were of interest for selection of bull dams. The top 1000 cows ranked by the German production index RZM (mean=100, SD = 12) from both models were compared. Since for the test day model a slightly higher genetic trend was observed TOP 1000 cows were on average one year younger from MLTDM compared to the TOP 1000 cows from MLM. From the TOP 1000 cows from MLM 680 were also present in the TOP 1000 list from MLTDM. From the TOP 1000 cows from MLM 476 differed only by " 2.5 points from the EBV in the MLTDM. However, for representation of cows from individual herds a significant reduction or an increase could be observed (Table 3). For the two herds G and M the number of TOP cows in the national TOP list reduced significantly in the MLTDM ranking because these were small herds with few cows with extremly high phenotypic yields and therefore also high within herd standard deviation. It has always been argued that this type of herds is overrepresented in the TOP list because for these small herds the management effect in MLM is defined by herdclass-region-year-season effect instead of a herd-year-season effect of larger herds. The consequent use of an herd-test day classification in MLTDM reduces the number of cows from these herds in the German TOP list. Herds L and G2 are herds with resonable size and resonable within herd variance, thus nearly the same number of cows is selected by both models. Herds D and N are characterised by a large number of cows and an extremely high within herd variance. The model MLTDM is in favor for such herds, thus it has been concluded that for ranking of individual TOP cows a correction for heterogeneous within herd variance is also necessary for the MLTDM. A new procedure for correction for within herd variance in a test day model is under development at VIT and pilot runs will be done in the fall of 1997.

Since December 1995 four evaluation runs were carried out with the MLTDM. Stability of proofs in consecutive runs was slightly higher than for the MLM model when only new test day data accumulates. For some releases not only new data from formerly included regions but also complete test day data from regions not included in the VIT database had to be added to the dataset. Then for these runs stability of proofs from one release to the next was lower.

The MLTDM incorporates very early information on daughter performance, i.e. when there are results from the daughter=s first test day. It was of interest if evaluations of very young bulls change consistently when more data (i.e. 6 months of test day data) was added for the next run. Therefore young bulls were ranked for milk yield EBVs based on results from the December 1996 MLTDM. About 900 bulls were devided in 11 groups of about 80 bulls in each group. EBVs of the same bulls were averaged for the next evaluation with updated test day data. Table 4 displays these averages for the MLTDM. It can be concluded that early proofs give an unbiased estimate of the EBV. However, it has to be noted that proofs from individual bulls can change when new data is added. For comparison the same statistic was added for the MLM from two consecutive runs.

Conclusions

Application of a test day model for dairy production traits is feasible even on a large scale national dataset. Ranking from bull EBVs compared to a lactation model change substantially, i.e. correlations between MLTDM results and MLM are in a range of .88-95 (Reents and Dopp, 1996). Correlations of EBVs from MLTDM with other countries lactation model EBVs are in a similar range as for MLM (not reported here).

Correlations between cow EBVs from the different models were slightly lower than for bull EBVs. Analysis of TOP cows showed that an adjustment for heterogeneous variances across herds is necessary. This is not an easy task since the consideration of test day results from different stage of lactation in one management group effect (i.e. herd-test-day) has to be accounted for as well. A procedure for incorporation of heterogeneous variances of test day data is under development at VIT.

References

Everett, R.W., Schmitz, F. and Wadell, L.H. 1994. A test-day model for monitoring management and genetics in dairy cattle. *J. Dairy Sci. 77, (Suppl.1),* 267.

Jones, L.P. and Goddard, M.E. 1990. Five years experience with the animal model for dairy evaluations in Australia. *4th Wld. Congr. Gen. Appl. Livest. Prod.,* Edinburgh, UK, *Vol XIII,* 382.

Ptak, E. and Schaeffer, L.R. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci. 34,* 23.

Reents, R., Jamrozik, J., Schaeffer, L.R. and Dekkers, J.C.M. 1995a. Estimation of genetic parameters for test day records of somatic cell score. *J. Dairy Sci. 78,* 2847.

Reents, R., Dekkers, J.C.M. and Schaeffer, L.R. 1995b. Genetic evaluation for somatic cell score with a test day model for multiple lactations. *J. Dairy Sci. 78,* 2858.

Reents, R. 1996. The use of test day models in genetic evaluation. *INTERBULL Bulletin No. 12,* 1996.

Reents, R. and Dopp, L. 1996. Genetic Evaluation for Dairy Production Traits with a Test Day Model. *INTERBULL Bulletin No. 14,* 1996.

Van Tassel, C.P., Quaas, R.L. and Everett, R.W. 1992. Parameter estimates for 305-day ME records and 305-day test-day residual records. *J. Dairy Sci. 75, (Suppl.1),* 234.

Table 1. Standard deviations of estimated breeding values (EBV) for cows from test-day model (MLTDM) and EBV from multi lactation model (MLM).

			Estimated breeding values				
Model	birth year	N	Milk yield	Fat %	Fat yield	Protein %	Protein yield
MLM	$ 87	4,146,396	510	.23	18.7	.11	13.8
MLTDM	$ 87	3,320,315	547	.21	19.3	.09	14.9
MLM	= 90	615,471	518	.24	19.1	.12	13.5
MLTDM	= 90	496,196	540	.22	18.9	.10	14.3
MLM	= 94	141.855	477	.21	17.4	.10	13.3
MLTDM	= 94	148.842	539	.21	19.9	.09	14.3

Table 2. Correlations between estimated breeding values (EBV) for cows from test-day model (MLTDM) and EBV from multi lactation model (MLM).

birth years	N	trait				
		Milk yield	Fat %	Fat yield	Protein %	Protein yield
$ 87	2.877.241	.86	.89	.83	.87	.84
= 90	463,005	.91	.92	.87	.91	.88
= 94	102.317	.83	.85	.83	.84	.81

Table 3. Number of cows from specific herds in the TOP 1000 German RZM list from test-day model (MLTDM) and EBV from multi lactation model (MLM).

Herd	Number of TOP cows in MLM	Number of TOP cows in MLTDM	explanation of the effect (see also text)
G	32	2	herds in MLM in herdclasses
M	10	3	
L	7	7	no effect because most cows in HYS classification and resonable within herd SD
G2	7	6	
D	6	24	extremely high within herd SD, up to now no correction for heterogeneous variance in MLTDM.
N	5	13	

Table 4. Mean EBVs of young bulls (born 1990-1991) on consecutive evaluation releases for milk production. Ranking was based on EBVs from Dec 1996 for MLTDM and Feb 97 for MLM.

Group of bulls	Number of bulls	MLTDM		Number of bulls	MLM	
		Dec 96	June 97		Feb 97	Aug 97
1	80	1867	1848	79	1685	1659
2	79	1462	1441	80	1322	1310
3	80	1247	1206	80	1154	1150
4	81	1071	1044	80	1004	989
5	80	930	915	81	882	870
6	80	791	752	80	768	763
7	81	662	649	81	638	610
8	80	535	496	80	511	487
9	80	369	355	80	371	375
10	80	184	202	81	203	181
11	89	-264	-265	88	-181	-194