Alternative Modeling of Fixed Effects in Test Day Models to Increase Their Usefulness for Management Decisions

P. Mayeres [1] , J. Stoll [2] , R. Reents [3] and N. Gengler [1,4]

[1]Animal Science Unit, Gembloux Agricultural University, B-5030 Gembloux, Belgium
[2]Luxembourgish Herdbook Federation, Ettelbruck, Luxembourg
[3]United Datasystems for Animal Production (VIT) Verden, Germany
[4]National Fund for Scientific Research, B-1000 Brussels, Belgium

Abstract

Current genetic evaluation models and especially fixed effects are setup in order to minimize bias in genetic evaluations but not for optimal usefulness for management. This is especially true for the definition of contemporary group effects. Very moderate alternative modeling of this effect as the sum of fixed herd-year of test and herd-month of test over a time period and a random herd test day deviation increased their usefulness for management nearly without producing ranking changes in animals (rank correlations > 0.99) for milk, fat and protein yields. Additionally this contemporary group definition protects against too small herd test day groups regressing this effect towards the expected value for a given herd, year of test and month of test. Results showed that correlation of contemporary group effects defined as the sum of the two new fixed effects were highly correlated (> 0.91) and absolute differences could essentially be explained by the existence of very small herd test day classes.

Introduction

Contrary to 305-d lactation models, which consider that all effects (e.g. contemporary group) are the same during the entire lactation, test day models (**TDM**) allow for variation during lactation by accounting for the effects at the day of test.

Numerous advantages of this approach were given in the literature (Swalve *et al.,* 2000). However one advantage often neglected and seldom developed is the possibility to use test day model results for herd management purposes. Results from milk recording transmitted to breeders are still limited to simple reports of the performed yields (milk, fat, protein and eventually somatic cell score and/or urea) of the last test together with cumulative records. Additionally estimated breeding values (**EBV**) are also reported. With test day data, a lot of additional information is available. However, geneticist are mostly thinking of reporting EBV for new traits as persistency, longevity or rate of maturity. However interest of milk producers is broader. Interest in breeding values, especially on the cow side, may be rather low, however other results are computed but have then to be communicated to the farmer. Furthermore recording organizations must justify the costs for milk recording. An interesting information therefore is the potential ability of TDM to predict future test day yields. The predicted values can then be used in different ways. First this values can be compared with the real values measured at the farm by using advanced decision making theory (e.g., Kalman Filter, Van Bebber *et al.,* 1999). Then out of the prediction for a given day of a given cow, a prediction for a larger time period could be developed that would permit management decisions on an individual and/or herd level. Historically this purpose is somewhat linked to the estimation of lactation yields. A certain number of methods were proposed in this context: Test Interval Method (**TIM**) and Centering Date Method (**CDM**) which are computing cumulative production adjusted for intervals among test. Recently several more sophisticated methods were proposed as e.g. Multiple Trait Prediction (**MTP**) (Schaeffer *et al.,* 1996) and Best Prediction Method (**BP**) (VanRaden, 1997). TDM should be a clear improvement over those old methods as they should model optimally the mean and variance structures among known test days, e.g. accounting for cow specific lactation curves. Predictions would then be obtained directly from the solutions. As long as we are interested into deviations of animals from mean yields this should work well but fixed effects and especially contemporary group definition may be more problematic. Since early days of test day model research (Ptak & Schaeffer, 1993) use of a herd-test date fixed effect (**HTD**) was considered optimal as it allows theoretically unbiased comparison of animals because of the theoretically unbiased estimation of contemporary groups yield levels. Therefore most TDM contemporary groups are based on this effects. One might however identify and especially in the context of prediction of future yields, different shortcomings. First future HTD can not simply be predicted because this effect does not model any time or seasonal trends. Secondly, current definition of HTD is not very robust against low number of animals in contemporary groups. Especially with seasonal calvings, even large herds may have from time to time very few animals that just freshened or were not yet dry.

The aim of this paper was therefore primarily to study the evolution of this HTD effect with the data of the Luxembourgish dairy population and to see how we can first predict its value at the next test date. Secondly we propose a simple remodeling of this effect in order to use it in a management purpose.

Materials and Methods

Data

Data were extracted by the VIT (Vereinnigte Informationssysteme Tierhaltung Germany) who manages Luxembourgish data and contained 499524 first lactation test day records (between days in milk 4 and 330) for milk, fat and protein yields from 58881 Holstein and Red and White cows. The pedigree file contained 106896 animals.

Precorrection

The official Luxembourgish dairy population data (herdbook, conformation, performances, ...) are processed by VIT and included in the German data bases. Given the small size of the Luxembourgish dairy population it was decided to precorrect milk, fat and protein test day yields for age-season-stage of lactation, lactation curves coming from the current VIT test day model:

$$y_c = y - KASRZ_{km} + b_{km1}(D/c) + b_{km2}(D/c)^2 + b_{km3}\ln(c/D) + b_{km4}[\ln(c/D)]^2$$

where y_c is the corrected test day yield, y is the measured test day yield, $KASRZ_{km}$ is the constant effect during lactation (age at calving * season interval * interval between calving * breed * region * class of year of calving), D is the day in milk and c a constant term equal to 380.

Model 1

The precorrected data set was then introduced in the following random regression model (**RRM**):

$$\mathbf{y_c = Xb + Q(Wh + Za + p) + e}$$

where **b** vector of HTD fixed effects, **h** vector of common herd period of calving environmental random regression coefficients, **a** vector of genetic random regression coefficients, **p** vector of permanent environmental random regression coefficients, **e** vector of residual effects, **X**, **W**, **Z** are incidence matrices, **Q** is the covariate matrix for the second order Legendre polynomials. This model is very close to the classical RRM used currently except for the common herd environmental effect. This effect was introduced as preliminary research showed that its introduction improved consistency of heritabilities and genetic correlations during the lactation as parts of the formerly genetic (co)variances were considered environmental (Rabier, 2002). The common herd effect was defined inside 5 periods of calving of 2 years (<1992/04/01, 1992/04/01-1994/03/31, 1994/04/01-1996/03/31, 1996/04/01-1998/03/31, >1998/03/31).

Study of fixed HTD solutions

HTD solutions obtained from Model 1 were studied. At population level means were computed for every month across all the herds and years.

Model 2

If predictability of future tests is a major issue, replacing the fixed HTD effect by alternative fixed effects spanning over several test days with an additional HTD random effect could be a promising modification. The prediction of a given herd mean at a given test day would then simply be the sum of solutions of the new fixed effects. Therefore an alternative model was defined:

$$\mathbf{y}_c = \mathbf{Um} + \mathbf{Tt} + \mathbf{Xb} + \mathbf{Q}(\mathbf{Wh} + \mathbf{Za} + \mathbf{p}) + \mathbf{e}$$

where **m** vector of herd test month period (**HTMp**) fixed effects, **t** vector of herd test year (**HTY**) fixed effect, **b** vector of herd test day random effects (**HTDr**), **U, T, X**, are incidence matrix, all the other symbols stay the same.

In order to allow stable predictions for current test years, the HTY effect spanned the two last years. The HTMp effect was defined inside time periods of four, respectively five years for newer years allowing for major changes inside a herd over the years.

Comparison of solutions from Model 1 and Model 2

Possible bias in rankings due to the use of Model 2 were studied by comparing breeding values for lactation yields obtained by integration of the Legendre polynomials for each animal from 0 to 305 days in milk. Comparison of fixed effect solutions from Model 2 (HTY + HYMp) with fixed HTD solutions from Model 1 allowed to access the potential for predictability of herd test day mean effects.

Variance components

The variance components used in this study were based on those computed by Rabier (Rabier, 2002) using subsamples of the same data, Model 1 EM-REML. For Model 1 and Model 2 the same variance components were used, only for Model 2, error variance was artificially subdivided into two parts, a part considered linked to HTDr and a part considered being the reduced error variance. The 1/1 ratio was chosen for this preliminary study.

Results and Discussion

Analysis of the HTD effects from Model 1

HTD solution showed a large dispersion. This underlines the large differences among herds. Figure 1 gives the evolution of monthly HTD solutions for milk.

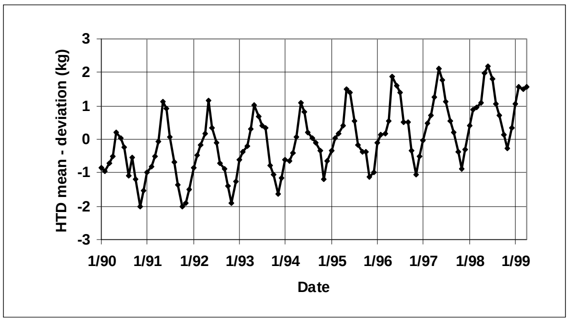

Figure 1. Evolution of the monthly HTD milk solutions across time.

Two trends are visible in these HTD solutions. First overall production is progressing over time. Secondly, over time monthly HTD means show very similar patterns within year, where the maximum corresponds generally to pastern release and the minimum to the month of November.

Comparison of solutions from Model 1 and Model 2

The comparisons of EBVs from both models showed that the new modeling proposition lead to few rerankings. Table 1 gives the Spearman rank correlation for all the cows in the pedigree file. Values were consistently over 0.99. Table 1 gives also the distribution parameters of the absolute difference and this for milk, fat and protein yields.

Table 1. Comparison of 305 day lactation EBVs from both models.

Trait	Rank Correlation	Absolute difference (kg) Mean	Std	Max
Sires				
Milk	0.993	22.456	25.444	404.06
Fat	0.993	0.7506	0.7904	10.394
Protein	0.991	0.5916	0.6579	11.942
Cows				
Milk	0.991	31.526	32.644	903.44
Fat	0.993	1.0767	1.0877	38.722
Protein	0.991	0.8209	0.8254	22.749

Figure 2 and 3 show the frequency distribution of the absolute difference for milk for the sires and for the cows (cows with records and female ancestors) in the pedigree file.

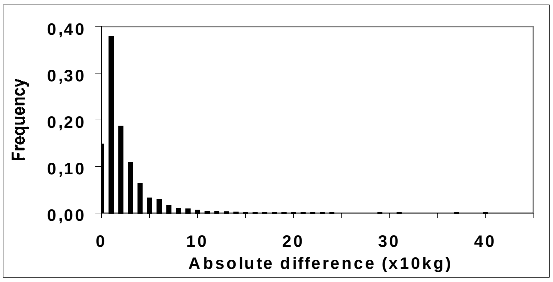

Figure 2. Frequency distribution of the absolute EBV differences for sires.

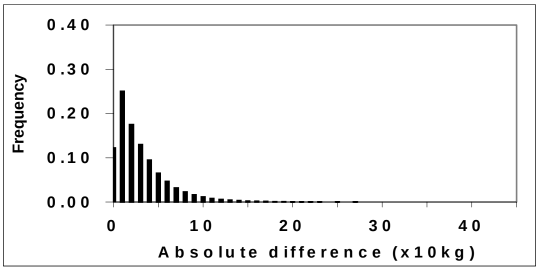

Figure 3. Frequency distribution of the absolute EBV differences for cows.

Rerankings were surprisingly rare showing that potential bias in rankings is limited.

Table 2. Comparison of HTY + HTMp from Model 2 with HTD from Model 1.

		Absolute difference		
Trait	Correlation	Mean	Std	Max
Milk	0.918	1.00	0.91	17.6
Fat	0.919	0.046	0.042	0.87
Protein	0.919	0.037	0.032	0.49

Table 2 gives details about the comparison of the absolute difference between the contemporary group solution defined as (HTY + HTMp) and HTD. Absolute differences were generally very small with some exceptions.

The highest difference, 17.6, occurred for a test date in which only one test day yield was recorded (Figure 4). In such a situation the fixed HTD solution from Model 1 makes anyway not much sense.

Table 2 gives also the Pearson correlations comparing solutions from HTY + HTMp (Model 2) with those from HTD (Model 1). Results here showed that prediction of HTD results from the sum of HTY + HTMp solutions would have been possible. As shown in Figure 4 most larger differences that were observed are simply due to the fact that original HTD classes were small.

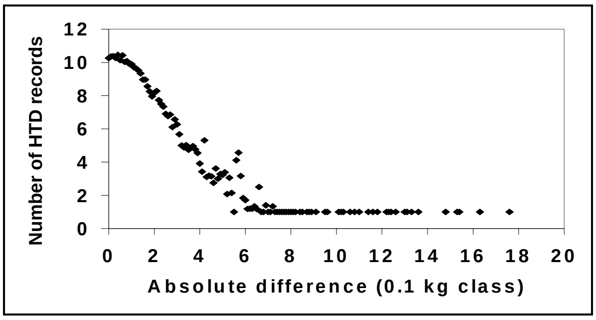

Figure 4. Dispersion diagram plotting the mean number of HTD records within 0.1 kg classes of absolute difference for milk yields.

Conclusions

Additional research will be needed as we reported only for first lactation yields and we did not reestimated (co)variance components. However the present results are highly encouraging. The alternative modeling of fixed effects to increase their usefulness for daily herd management can be done inexpensively and give very interesting results. The expected bias in the EBVs was extremely small and nearly no reranking occurred. Differences between fixed HTD effects and new HTY + HTMp effects were generally small and were mostly due to the small size of some HTD classes. Obviously one might doubt on the usefulness of very small HTD classes.

The introduction of the new model allows a better predicting of the herd mean performances over time. This is somewhat unusual thinking, because genetic evaluation systems are setup actually in a way to try to predict only unbiased EBVs. However in the future alternative use of results from genetic evaluation systems will become an important issue. Especially the possibility to model the data in a correct manner and to obtain jointly estimates of environmental and genetic effects having the desired (BLUE and BLUP) properties. Progress in this field is very important in order to justify the rising costs of present and future milk recording schemes.

Acknowledgments

Nicolas Gengler who is Chercheur qualifié of the FNRS (Brussels, Belgium) acknowledges his financial support. The authors gratefully acknowledge the financial support of the Luxembourgish Herdbook Federation, a breeders cooperative.

References

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

Rabier, F. 2002. Personal communication.

Schaeffer, L.R. & Jamrozik, J. 1996. Multiple-Trait Prediction of lactation yields for dairy cows . *J. Dairy Sci. 79*, 2044-2055.

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

Van Bebber, J., Reinsch, N., Junge, W. & Kalm E. 1999. Monitoring daily milk yields with a recursive test day repeatability model (Kalman Filter). *J. Dairy Sci. 82*, 2421-2429.

VanRaden, P. 1997. Lactation yields and accuracies Computed from test day yields and (co)variances by best prediction . *J. Dairy Sci. 80*, 3015-3022.