Comparison of Random Regression Test-Day Models for Production Traits of Dairy Cattle in Poland

***T. Strabel*[1], *J. Szyda*[2], *E. Ptak*[3], *J. Jamrozik*[4]**
[1]*Agricultural University of Poznań, Department of Genetics and Animal Breeding, Wołyńska 33, 61-627 Poznań, Poland; e-mail: strabel@jay.au.poznan.pl*
[2]*Agricultural University of Wrocław, Department of Genetics and Animal Breeding, ul. Kożuchowska 7, 51-631 Wrocław, Poland*
[3]*Agricultural University of Kraków, Department of Genetics and Animal Breeding, Al. Mickiewicza 24/28, 30-059 Kraków, Poland*
[4]*Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, ON, N1G 2W1, Canada*

Introduction

Several countries have already introduced test-day models for genetic evaluations of dairy cattle. Poland is among those, which are to implement it in the near future.

Many alternative models have been proposed for the analysis of test-day yields. A random regression model (RRM) becomes the model of choice for genetic evaluation. The selection of the best model requires decisions concerning effects included, especially submodels to define trajectories of group and individual lactation curves.

The aim of this study was to compare various RRM for genetic evaluation of dairy cattle in Poland for first lactation milk yield.

Material and Methods

Data consisted of test-day milk records of 6319 first lactations collected in 55 randomly selected herds. Structure of the data set is presented in Table 1. Cows were assigned to one of 10 subclasses for age-season of calving. The following overall model was used:

$$Y_{ijklm} = HTD_i + \Sigma b_{jl}z_{mnl} + \Sigma c_{kl}z_{mnl} + \Sigma a_{ml}z_{mnl} + \Sigma p_{ml}z_{mnl} + e_{ijklm},$$

where

Y_{ijklm} is the l-th milk yield of cow m within herd-test day effect i, belonging to herd k and j-th class of age-season of calving, HTD_i is herd-test day effect, b_{jl} and c_{kl} are fixed regression coefficients within age-season subclass j and herd-year (HY) k, a_{ml} and p_{ml} are random regression coefficients for animal (AG) and permanent environmental (PE) effects, respectively, e_{ijklm} is residual effect for each observation, z_{mnl} are corresponding covariates.

Several alternative submodels were used to describe lactation curves: Ali and Schaeffer (A&S) (1987), Wilmink (1987) subsequently denoted as W, second (L2), third (L3) and fourth (L4) order Legendre polynomials (Kirkpatrick *et al.,* 1990) and the combination of L2 and exp(-0.05DIM) (WL) were used as covariates. Compared models are presented in Table 2.

HTD effect was alternatively used as a fixed or random effect with variance $\mathbf{I}\sigma^2_h$. The covariance structure of remaining random effects in the models was defined as:

$$\mathrm{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{p}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\otimes\mathbf{A} & 0 & 0\\ 0 & \mathbf{P}\otimes\mathbf{I} & 0\\ 0 & 0 & \mathbf{I}\sigma^2_e\end{bmatrix},$$

where **A** is a matrix of additive genetic relationships among animals, $\otimes$ is a Kronecker product function, **G** and **P** are covariance matrices of the random regression coefficients for AG and PE effects, **I** is an identity matrix, and σ^2_e is a residual variance.

The (co)variance components were estimated with REML algorithm using the computer package of Misztal *et al.* (2002).

Bayesian information criterion (BIC) (Schwarz, 1978) and percentage of squared bias (PSB) (Ali and Schaeffer, 1987) were employed for models comparison. Analysis of accuracy of evaluations was also carried out.

Results

Lactation curves for the youngest group of cows calved in summer season estimated using different order of Legendre polynomials are presented in Figure 1. The typical shape of milk yield lactation curves was achieved only with L4 polynomial. Remaining submodels were unable to recover the peak of lactation. Lactation curves estimated with Wilmink function (not shown) were either extremely high or low at the beginning of the trajectory, depending on age-season class.

Table 2 shows statistics used for comparison of models. Models with A&S functions are not presented due to problems in reaching convergence criteria. Solutions for these models included many extreme values for herd-year and HTD effects, which could not find biological explanation. Similar problem of unreasonably extreme solutions were found for some levels of fixed HTD effect and L4 coefficients used for herd-year curves.

Comparison of models based on the BIC criteria confirmed the better fit of models with higher order Legendre polynomials for fixed regressions. Among best ranking models those with L3 for AG and PE curves had lower values of BIC. In general, models with lower order polynomials for PE effect than for AG were of advantage. Models with fixed HTD effect were better ranked by this criterion.

The values of PSB statistics ranked models differently than BIC. The largest PSB value of this criterion was obtained for model with extremely high solutions for some levels of fixed HTD. PSB, similarly to BIC, favored models with L3 for AG and PE effects rather than models other order polynomials. Among models with L3 for AG effect those with more complex submodels (L4) for fixed regressions (age-season and HY) showed slightly lower bias. PSB slightly preferred fixed over random HTD effect.

Correlation between true and predicted breeding values favored models with smaller degree of polynomials for PE than for AG effect, especially models with L3 for AG and L2 for PE effect. Less accurate prediction of genetic merit was obtained for models with L4 used for AG and PE effects. Differences between models with various submodels for fixed regression and the same submodels for random regressions were small; model with fixed HTD effect gave slightly more accurate results than models with random HTD.

Daily heritabilities obtained for selected models showed that different combinations of functions for AG and PE effect clearly influenced estimates. Shapes of heritability curves for models with the same functions for AG and PE regressions were similar - typical examples are presented in Figure 2. Relatively smooth curves without high values at periphery of trajectories were obtained with L3 or L4 for AG and PE effects. Different order of polynomials for those effects (L3 and L4, L4 and L3 etc. for AG and PE effects, respectively) resulted in highly oscillatory patterns with high estimates at the beginning, the end of first trimester and the end of lactation. WL submodel gave similar pattern of heritability changes with the highest values in the middle of lactation.

Discussion

The problem of selecting the best RRM based on the comparison criteria, that ranked models in different way, is not trivial and has been recently discussed also by Druet *et al.* (2003) and Lopez-Romero and Carabano (2003). It seems that some pre-selection decision can be made despite of minor differences between models.

Extreme values of solutions for fixed HTD effect, tend to disqualify models with fixed HTD. Similarly, models with W and the lowest order of Legendre polynomials did not satisfy in describing the overall age-season lactation curves and should be skipped from further analysis. Convergence problems and poor performance of models with A&S exclude also this function from consideration.

Criteria used for models comparison suggested, in general, that Legendre polynomials of order three should be used for individual animal genetic lactation curves. Various combinations of functions used for other effects, especially different order of polynomials for fixed regressions, did not allow for simple conclusions. On the other hand, minor differences between models leave space for the choice of the best fitting model to be made after further studies including additional lactations records and considering practical issues of implementation.

Conclusions

Criteria used for comparison ranked models in different way. However, differences between the best ranking models were rather negligible. Random rather than fixed HTD effect should be used for genetic evaluation of dairy cattle in Poland. The average lactation curve should be described by Legendre polynomials of order four. The best results in describing individual lactation curves were obtained with third order Legendre polynomials. For permanent environmental effect lower order polynomials fitted better. Models with different order polynomials for AG and PE effects yielded undesired shapes of heritabilities curves.

References

Ali, T.E. & Schaeffer, L.R. 1987. *Can. J. Anim. Sci.* 67, 637-644.

Druet, T., Jaffrezic, F., Boichard, D. & Ducrocq, V. 2003. *J. Dairy Sci*. 86, 2480–2490.

Kirkpatrick, Lofsvold, D. & Bulmer, M. 1990. *Genetics 124*, 979-993.

Lopez-Romero, P. & Carabano, M.J. 2003. *Livest. Prod. Sci. 82,* 81-96.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T. & Lee, D.H. 2002. *Proc. of the 7th World Congr. on Gen. Appl. to Livest. Prod.,* Montpellier, France.

Schwarz, G. 1978. *Ann. Stat. 6,* 461-464.

Wilmink, J.B.M. 1987. *Livest. Prod. Sci. 16,* 335-348.

Acknowledgement

The Polish State Committee for Scientific Research, Project KBN No 6 P06D 019 20 supported this work.

Table 1. Description of the data set.

Number of herds	55
Number of TD records	51,365
Number of lactations	6319
Number of HTD classes	3487
Minimum number of TD records per lactation	5
Minimum DIM of lactation	150
Minimum number of lactations per herd	30
Average number of daughters per bull	5.18
Average number of TD records per HTD	14.73

Table 2. Bayesian information criteria (BIC), percentage of squared bias (PSB) and correlation between true and predicted breeding values (ACC) for compared models (ranking position in brackets).

Model					Goodness of fit				ACC	
HTD*	Age-Season	HY	AG	PE	BIC		PSB			
F	L4	L4	L3	L2	361743	(1)	1641.090	(15)	0.3829	(13)
F	L2	L2	L2	L2	365288	(3)	0.990	(13)	0.3583	(8)
F	L3	L3	L3	L3	363425	(2)	0.866	(3)	0.3540	(6)
R	L2	L2	L2	L2	387497	(13)	1.008	(14)	0.3564	(7)
R	L2	L2	L4	L4	387329	(12)	0.923	(12)	0.3353	(1)
R	W	W	L3	L3	388543	(14)	0.870	(4)	0.3694	(12)
R	L3	L3	L3	L3	385587	(10)	0.879	(7)	0.3504	(3)
R	L4	L4	L4	L4	383675	(7)	0.887	(9)	0.3434	(2)
R	W	W	WL	WL	389017	(15)	0.887	(10)	0.3609	(10)
R	L4	L4	L3	L3	383492	(4)	0.860	(1)	0.3504	(4)
R	L3	L4	L3	L3	383493	(5)	0.860	(2)	0.3505	(5)
R	L3	L3	L3	L2	385340	(9)	0.896	(11)	0.3865	(15)
R	L4	L4	L3	L2	383523	(6)	0.878	(6)	0.3837	(14)
R	L3	L3	L3	L4	385594	(11)	0.883	(8)	0.3620	(11)
R	L4	L3	L3	L4	385273	(8)	0.877	(5)	0.3594	(9)

* F – fixed, R – random HTD effect.

Figure 1. Lactation curves for the youngest group of cows calved in summer season estimated with various submodels: × - L2, Δ - L3, ■ - L4.

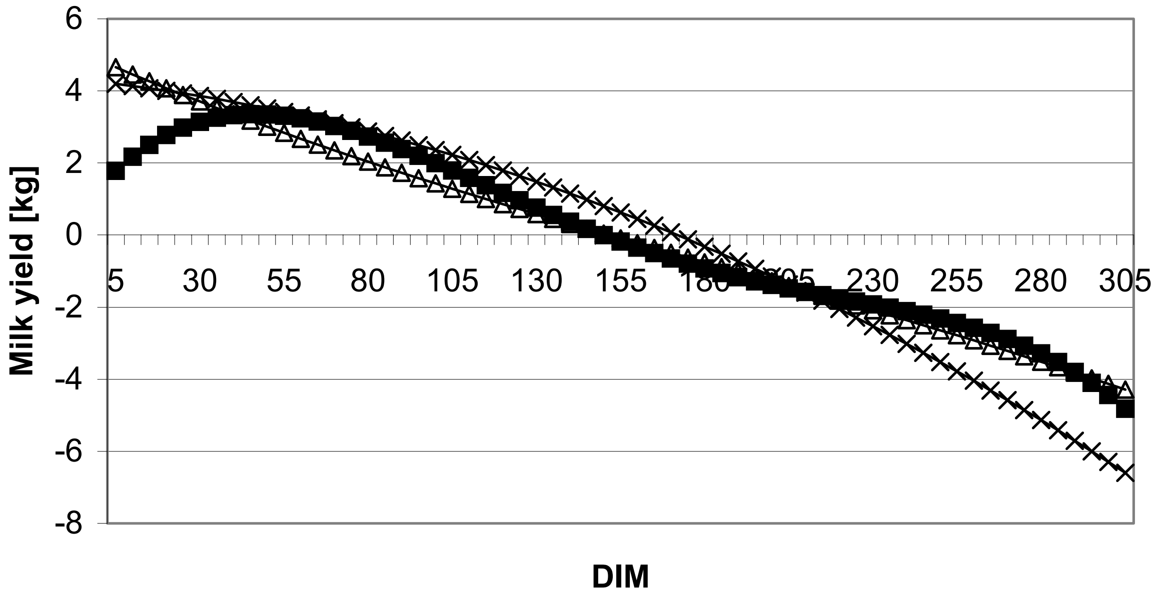

Figure 2. Daily heritability estimates for selected models with different functions for AG/PE effects: ○ - WL/WL, Δ - L2/L2, × - L3/L3, - - L4/L4, ◆ - L3/L2, ■ - L3/L4.