Comparison of Random Regression Test-Day Models for Polish Black and White Cattle

T. Strabel,¹ J. Szyda,² E. Ptak,³ and J. Jamrozik⁴

¹Agricultural University of Poznań, Department of Genetics and Animal Breeding, Wołyńska 33, 61-627 Poznań, Poland

²Agricultural University of Wrocław, Department of Genetics and Animal Breeding, ul. Koźuchowska 7, 51-631 Wrocław, Poland

³Agricultural University of Kraków, Department of Genetics and Animal Breeding, Al. Mickiewicza 24/28, 30-059 Kraków, Poland

⁴Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, ON, N1G 2W1, Canada

ABSTRACT

Test-day milk yields of first-lactation Black and White cows were used to select the model for routine genetic evaluation of dairy cattle in Poland. The population of Polish Black and White cows is characterized by small herd size, low level of production, and relatively early peak of lactation. Several random regression models for first-lactation milk yield were initially compared using the “percentage of squared bias” criterion and the correlations between true and predicted breeding values. Models with random herd-test-date effects, fixed age-season and herd-year curves, and random additive genetic and permanent environmental curves (Legendre polynomials of different orders were used for all regressions) were chosen for further studies. Additional comparisons included analyses of the residuals and shapes of variance curves in days in milk. The low production level and early peak of lactation of the breed required the use of Legendre polynomials of order 5 to describe age-season lactation curves. For the other curves, Legendre polynomials of order 3 satisfactorily described daily milk yield variation. Fitting third-order polynomials for the permanent environmental effect made it possible to adequately account for heterogeneous residual variance at different stages of lactation. (**Key words:** dairy cattle, test-day yield, random regression model)

Abbreviation key: ACC = accuracy of breeding value, AG = additive genetic, AS = age and season of calving, BIC = Bayesian information criterion, HTD = herd-test-date, HY = herd-year, PE = permanent environmental, PSB = percentage of squared bias, RRM =

random regression model, RV = residual variance, TDM = test-day model.

INTRODUCTION

A variety of models with different levels of sophistication are used for genetic evaluation of production traits in dairy cattle (Mark, 2004). The use of daily production data with test-day models (TDM) has several advantages over the traditional procedure of evaluating lactational records, such as the ability to account for environmental effects on each test day and to model individual lactation curves. Although TDM require analysis of much larger data sets and estimation of many more parameters than for the lactation model, they have already been implemented, or are being developed, in many countries.

The main nongenetic effects in TDM are the herd-test-date (HTD) effect and fixed regressions to account for the average shape of lactation curves. It has been shown that using HTD effect increases the accuracy of evaluation (Beard, 1983; Strabel and Szwaczkowski, 1995). Average lactation curves may be modeled in many alternative ways; for example, by lactation curve functions (Guo and Swalve, 1997; Jamrozik et al., 1997), fixed classes (Stanton et al., 1992; Strabel, 2004), or with the use of splines (Druet et al., 2003). Due to the differences between populations, various functions are preferred for genetic evaluation models in different countries (Liu et al., 2001; Reinhardt et al., 2002; Mrode et al., 2003).

The variance of daily production also changes during the course of lactation, and the correlation between daily yields decreases with increasing time between tests. For these reasons, random regression models (RRM) were proposed (Schaeffer and Dekkers, 1994). They assume that additive genetic (AG) and permanent environmental (PE) effects change the average shape of the lactation curve. Although various functions have

Received September 17, 2004.

Accepted June 15, 2005.

Corresponding author: Tomasz Strabel; e-mail: strabel@man.poznan.pl.

been used to model random curves (Jamrozik and Schaeffer, 1997; Ptak and Żarnecki, 1998; López-Romero and Carabaño, 2003), Legendre orthogonal polynomials are becoming the function of choice for this part of the model, as they sufficiently describe variation and help to avoid overestimation of genetic variances and heritabilities at the extremes of lactation (Rekaya et al., 1999; Brotherstone et al., 2000; López-Romero and Carabaño, 2003).

It has been found in several studies that residual variance (**RV**) is not constant throughout lactation (Jamrozik et al., 1997; Olori et al., 1999a,b; Brotherstone et al., 2000; Rekaya et al., 2000). The heteroskedastic model can be postulated by defining arbitrary classes for DIM, within which RV is assumed constant (López-Romero et al., 2003), or by using the change point technique (López-Romero et al., 2004) or the function of time to model RV (Jaffrezic et al., 2000).

The number of possible combinations of functions and effects included in the RRM makes comparison of all of them infeasible. Thus, the model selection process is usually restricted to analysis of preselected candidates. The choice of a function to describe fixed lactation curves is often made before the full model is applied to the data (Druet et al., 2003). Models can be compared based on several measures, including the size of the residual variance and average accuracy (**ACC**) of prediction for certain groups of animals. The likelihood ratio may indicate better models; however, this method tends to favor more complex models (Ducrocq, 2000; López-Romero and Carabaño, 2003). To avoid this propensity, a correction for the number of parameters used in the model can be applied, such as the Bayesian information criterion (**BIC**) (Schwarz, 1978). Additional methods, such as analysis of residuals, are often used, as the results of comparison of models based on a formal criterion are not always consistent for a given set of models.

Although the number of cows under milk recording has increased in recent years, the average herd size in Poland is still small. In 2002, there were over 464,000 cows in almost 20,000 herds, which gave on average 23.2 cows per herd (KCHZ, 2003). This resulted in a large proportion of cows without or with only a few contemporaries. Observations from such classes have limited value for genetic evaluation, and they reduce the reliability of estimates. Small herds usually lead to small HTD classes, and the problem of small contemporary groups remains. In such a case, increasing the size of the contemporary group or treating the herd-time effect as random are among the methods that can improve the quality of genetic evaluation by reducing the variances of residuals and prediction error (Ugarte et al., 1992; Visscher and Goddard, 1993; Oikawa and

Sato, 1997). Strabel and Szwaczkowski (1999) compared models with HTD formed by different strategies of clustering with a model containing fixed herd and random HTD effects and found the latter to be superior. The model with a random HTD gave the smallest ratio of error to total variance, and was the best in predicting future records. Similarly, Lidauer et al. (2003) modeled the herd effect for the Finnish dairy population by fixed herd-year and random HTD components. Mayers et al. (2002) found that the combination of random HTD and fixed herd-years gave protection against the negative consequences of HTD groups that were too small.

The overall lactation curve in Poland is characterized by a relatively early peak (Strabel, 2004). Rapid changes in the beginning of lactation may require more complex functions or higher-order polynomials to describe variation in this period. A similar problem was noted for dairy breeds in Switzerland (Kistemaker and Schnyder, 2004). The low peak yield of Polish Black and White cows is associated with relatively low production level when compared with other populations (Strabel and Misztal, 1999).

The objective of this study was to develop the random regression test-day model that would be applicable for routine genetic evaluation of Polish Black and White cattle.

MATERIALS AND METHODS

All first-lactation test-day records on Polish Black and White cattle calving from 1995 to 2001 (16.4 million test-day records) were available for this project. Edits included restriction of the range of age at calving to between 18 and 49 mo, daily milk yield between 0.1 and 90 kg, a minimum of 5 records per lactation, and a minimum of 150 DIM. After the edits, more than 4.6 million first-lactation test-day records remained, which gave on average: 38.2 records per herd-year of calving (**HY**), 8.1 cows per HY, and 5.2 cows (records) per HTD. The proportions of small herds, HY and HTD classes with first-lactation records are presented in Table 1. Almost 35% of the HTD classes contained single records, and slightly over 23% of the HY classes contained no more than 5 records. Cows were assigned to 1 of 10 subclasses for age-season (**AS**) of calving. Two seasons of calving (April–September and October–March) and 5 classes for age at calving (below 25 mo, 25 to 26, 27 to 28, 29 to 30, and over 30 mo) were defined.

The model selection process was split into 2 steps. In the initial preselection of models, several combinations of functions for fixed and random regressions were considered. This part of the study was conducted based on one selected data set (data set A). Based on the results

Table 1. Proportion of herds, herd-year classes (HY), and herd-test-date classes (HTD) with small numbers of records/cows in the complete data set.

Size	% of herds (records)	% of HY (records)	% of HTD (cows/records)	% of herds (cows)	% of HY (cows)
1	0.6	5.3	34.7	12.3	21.9
≤2	1.4	9.7	56.9	21.4	40.2
≤3	2.5	13.7	69.8	28.9	54.2
≤5	7.5	23.1	81.4	40.4	71.8
≤10	17.5	42.8	89.0	60.6	86.6

of this analysis, 3 selected models were further studied using 2 separate data sets (data sets B and C).

Data set A consisted of test-day milk yields of 6319 cows in 55 randomly selected herds that were required to have a minimum of 30 cows. The detailed structure of data set A is presented in Table 2. A slightly stricter criterion on herd size (minimum of 50 cows) was applied for selection of herds for the second part of the analysis (data sets B and C). Two distinct data sets were created and used to reduce the strength of the association between the data and the model. The records used in data set A could have been selected again in data set B or C. To avoid HY levels with fewer than 5 records, some classes of this effect were clustered: records from a single HY were moved to HY from neighboring years. The structure of data sets B and C is presented in Table 2. Both data sets were almost equal in size, but in data set B there was a smaller number of herds, higher average production, bigger HTD classes, and more cows per sire.

The following general model was applied in both parts of the study:

$$Y_{ijklm} = htd_i + \sum_{n=1}^{n_b} b_{jn} z_{mnl} + \sum_{n=1}^{n_c} c_{kn} z_{mnl} + \sum_{n=1}^{n_a} a_{mn} z_{mnl} + \sum_{n=1}^{n_p} p_{mn} z_{mnl} + e_{ijklm} \quad [1]$$

where Y_{ijklm} is the yield of milk on test-day l of cow m within herd-test day effect i , belonging to herd-year

class k and the j th class of AS, htd_i is the herd-test-day effect, b_{jn} are fixed regression coefficients specific to age-season subclass j , c_{kn} are fixed regression coefficients specific to herd-year k , a_{mn} are random regression coefficients specific to the AG effect of cow m , p_{mn} is the random PE effect, and e_{ijklm} is the residual effect for each observation. The z_{mnl} are Legendre polynomials on DIM (Kirkpatrick et al., 1990), and n_b , n_c , n_a , and n_p represent the order of fit for AS, HY, AG, and PE curves, respectively.

The covariance structure for models with random HTD effect was defined as:

$$\text{var} \begin{bmatrix} \text{htd} \\ \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{I}\sigma_h^2 & 0 & 0 & 0 \\ 0 & \mathbf{G} \otimes \mathbf{A} & 0 & 0 \\ 0 & 0 & \mathbf{P} \otimes \mathbf{I} & 0 \\ 0 & 0 & 0 & \mathbf{R} \end{bmatrix}, \quad [2]$$

where $\mathbf{I}$ is the identity matrix, σ_h^2 is the variance of the random HTD effect, $\mathbf{A}$ is the matrix of additive genetic relationships among animals, $\otimes$ is the Kronecker product, $\mathbf{G}$ and $\mathbf{P}$ are covariance matrices of the random regression coefficients for AG and PE effects, $\mathbf{R}$ is the diagonal matrix of the form $\mathbf{I}\sigma_e^2$, and σ_e^2 is residual variance. Models with fixed HTD effects all had the same (except for the HTD component) covariance structure.

All specific models were fitted by the REML method using the BLUPF90 software package (Misztal et al., 2002).

Table 2. Description of data.

	Data set		
	A	B	C
Mean daily yield, kg	16.0	15.7	17.6
SD, kg	5.9	6.2	6.2
Minimum number of lactations per herd	30	50	50
Total number of test-day records	51,365	51,981	52,157
Number of animals with records	6319	6207	6239
Number of herds	55	37	31
Number of daughters per bull	5.2	5.1	6.2
Number of test-day records per HTD	14.7	20.6	23.9
Cows with test-day records after 250 DIM, %	76	78	80
Average number of test-day records per cow	8.5	8.4	8.4

Preselection of Models

Several alternative submodels were used to describe the shape of the lactation curves. They were created by selecting second- (L2), third- (L3), and fourth- (L4) order Legendre polynomials for different parts of the model. The HTD effect was alternatively used as a fixed or random effect. The models are presented in Table 3. Model comparison criteria were the percentage of squared bias (**PSB**; Ali and Schaeffer, 1987), and the analysis of ACC of evaluations. The PSB was computed as

$$PSB = \frac{\sum_{l=1}^n (y_{ijklm} - \hat{y}_{ijklm})^2}{\sum_{l=1}^n (y_{ijklm})^2} \times 100 \quad [3]$$

where n is the number of records, y_{ijklm} is the observed record, and $\hat{y}_{ijklm}$ is the record predicted by the particular model in question, considering all the estimates of effects associated with this observation.

The accuracy of evaluation for the total yield in lactation was calculated as the average correlation between the true and predicted breeding values for all animals. The inverse elements of the mixed model equations for the diagonal block corresponding to random regression coefficients for animal genetic effects were used to make this calculation. They were combined to obtain prediction error variance for the total yield per lactation (for details see Jamrozik et al., 2000).

Further Model Comparison

Additional comparisons of the selected models were made using the same general TDM [1] as in the preselection step. Only models with random HTD were considered. Third-order Legendre polynomials (L3) were used to describe the HY, AG, and PE curves. For the AS effect, fourth- (L4) and fifth- (L5) order Legendre polynomials were also used, and the models with these effects were denoted AS4-PE3 and AS5-E3, respectively. Additionally, a model with L5 for the PE effect and L5 for the AS effect was tested. This model was denoted AS5-PE5. All models were analyzed alternatively with homogeneous or heterogeneous RV. All remaining parts for all compared models were the same as in the preselection step.

Matrix **R** for the models with homogeneous RV was a diagonal matrix containing the constant variance of residuals. For the models with heterogeneous RV, **R** was the diagonal matrix of the residual variances that depended on DIM. A second-order Legendre polynomial, following Druet et al. (2003), was applied:

$$R = \text{diag}[\sigma_{e_z}^2], \quad [4]$$

where $\sigma_{e_z}^2 = \exp\left(\sum_{n=0}^2 r_n z_n\right)$ to ensure positive values of variances; r_n were parameters that described the residual variance over DIM, and z_n were the Legendre polynomial covariates. No other sources of heterogeneity of variance were assumed.

The models were compared by BIC values and a detailed analysis of residuals was performed. The BIC was calculated as:

$$BIC = -2 \ln L + k \ln(\lambda), \quad [5]$$

where $\ln L$ is the maximum of the restricted log likelihood function, k is the number of variance components estimated, and λ is the difference between the number of observations and the rank of the fixed effects incidence matrix. The BIC criteria could only be used to compare models with equal fixed parts, as L was the restricted likelihood (REML method).

Residuals were calculated for each DIM between 5 and 305 as a difference between y_{ijklm} and $\hat{y}_{ijklm}$. Mean values and variances of these residuals (over all TD records) were estimated and plotted for selected DIM. Additionally, averages of mean daily residuals and their variances (over 305 DIM) were estimated for each model and they were later applied as a comparison criterion. Average residuals can be used to determine the accuracy of the model, and variances of residuals can evaluate precision of the compared models (Jamrozik et al., 1997).

RESULTS

Preselection of Models

Table 3 shows the statistics used for the model comparisons in the preselection step. In general, the lowest value of PSB was obtained for the group of models with L3 for the AG effect. The differences between models within this group were small. Legendre polynomials of order 3 for PE were preferred over other orders of polynomials for this effect. The fixed part of the model had a minor effect on the values of PSB. However, the solutions for the model with fixed HTD and L4 for HY curves included extreme values for herd-year and HTD effects. That led to a value of PSB much larger than in a similar model containing HTD effects defined as random (Row 1 vs. Row 2).

The evaluation of models based on the correlation between true and predicted breeding values favored models with unequal degrees of polynomials for the AG

Table 3. Percentage of squared bias (PSB) and correlations between true and predicted breeding values (ACC) for the compared models (ranking position in parentheses). Legendre polynomials of order 2 (L2) to order 4 (L4) were used, and the herd-test-date (HTD) effect was alternatively treated as fixed (F) or random (R).

HTD	Model				PSB (%)	ACC
	Age-Season	Herd-Year	AG ¹	PE ²		
F	L4	L4	L3	L2	(13) 1,641,090	(3) 0.38
F	L2	L2	L2	L2	(11) 0.990	(6) 0.35
F	L3	L3	L3	L3	(3) 0.866	(8) 0.35
R	L2	L2	L2	L2	(12) 0.990	(7) 0.35
R	L2	L2	L4	L4	(10) 0.923	(13) 0.33
R	L3	L3	L3	L3	(6) 0.879	(11) 0.35
R	L4	L4	L4	L4	(8) 0.887	(12) 0.34
R	L4	L4	L3	L3	(2) 0.860	(10) 0.35
R	L3	L4	L3	L3	(1) 0.860	(9) 0.35
R	L3	L3	L3	L2	(9) 0.896	(1) 0.38
R	L4	L4	L3	L2	(5) 0.878	(2) 0.38
R	L3	L3	L3	L4	(7) 0.883	(4) 0.36
R	L4	L3	L3	L4	(4) 0.877	(5) 0.35

¹AG = Additive genetic.

²PE = Permanent environmental.

and PE effects, especially the models with L3 for the AG and L2 for the PE effect. Less accurate prediction of genetic merit was obtained from models with L4 used for the AG and PE effects. The differences between models with various submodels for fixed curves and the same models for random curves were small. The differences between models with different ways of defining the HTD effect were minor, and the order of fit for fixed submodels was less important than the other differences between models.

Lactation curves for a selected AS class with different orders of Legendre polynomials are presented in Figure 1. The typical shape of the milk yield lactation curve was achieved only with the L4 polynomial. The other submodels were not able to describe peak yield accurately.

Figure 1. Age-season lactation curves with Legendre polynomials of different orders: L2 (.....), L3 (-----), and L4 (——). The model included fixed herd-test-day effect and third-order Legendre polynomials as regressions for herd-year, additive genetic, and permanent environmental effects.

Based on these results, 3 models with random HTD and L3 for fixed HY curves and random AG curves were selected for further comparisons: AS4-PE3, AS5-PE3, and AS5-PE5.

Further Model Comparison

Table 4 presents BIC and residuals statistics for the 3 compared models with homogeneous and heterogeneous RV. All values were similar for both data sets (B and C). In general, BIC slightly favored more complex models with a higher number of parameters for PE and heterogeneous RV over simpler models with homogeneous RV. However, model AS5-PE5 with homogeneous RV for data set B was indicated as the best. It should be emphasized that the differences in BIC values were small.

Increasing the order of fit for the fixed AS curves resulted in almost 50% reduction of the standard deviation of mean residuals for both data sets (Table 4), regardless of whether RV was fitted as homogeneous or heterogeneous. This suggested that undesirably large fluctuations of the mean residuals over lactation could be observed for a model with a lower order of fit for fixed curves. This can be confirmed by studying the plots of mean residuals for data set B (Figure 2). A reduced number of parameters for the AS effect resulted in more extreme values at the peripheries, in particular at the beginning, of lactation. The same consequences were observed for data set A (results not shown), which meant that employing at least fifth-order Legendre polynomials for the overall lactation curves was crucial. Neither increasing the number of parameters for the

Table 4. Bayesian information criterion (BIC), average standard deviation of mean residuals (across 305 DIM), average variances of residuals, and their standard deviations for the compared models.

Residual variance	Data set			
	B		C	
	Homogeneous	Heterogeneous	Homogeneous	Heterogeneous
Model ¹	BIC ²			
AS5-PE3	220	81	465	277
AS5-PE5	119	0	0	108
	Standard deviations of mean residuals ³			
AS4-PE3	1.30	1.35	1.20	1.24
AS5-PE3	0.77	0.79	0.72	0.74
AS5-PE5	0.70	0.75	0.65	0.68
	Variance of residuals ⁴			
AS4-PE3	2.93	3.04	2.78	2.94
AS5-PE3	2.93	3.04	2.78	2.93
AS5-PE5	2.56	2.77	2.34	2.55
	Standard deviations of variance of residuals ⁵			
AS4-PE3	0.77	0.81	0.70	0.71
AS5-PE3	0.77	0.79	0.71	0.71
AS5-PE5	0.61	0.65	0.56	0.52

¹ASi-PEj denotes the model with Legendre polynomials for age-season and permanent environmental curves of orders i and j, respectively.

²Relative to the best model for which BIC = 0.

³Variability of mean residuals across d 5 to 305.

⁴Average variance of residuals across d 5 to 305.

⁵Variability of variance of residuals across d 5 to 305.

PE effect nor fitting the heterogeneous RV had any significant effect on the standard deviation of mean residuals, suggesting that increasing model complexity in this part of the model was unnecessary.

Using Legendre polynomials to model RV had an influence on the variance of residuals (Table 4). Comparison of models based on this criterion showed that a higher order of fit for the PE effect (AS5-PE5 vs. AS4-PE3 and AS5-PE3 models) reduced the variance of re-

siduals across DIM. It also decreased the standard deviations of the variance of residuals over lactation. The values of this statistic increased for almost all the models when changing from homogeneous to heterogeneous RV. The plots of the residual variance across lactations (Figure 3) showed that the variability of residuals in-

Figure 2. Mean residuals for selected days in milk for models with varying orders of Legendre polynomials (3 to 5) for age-season (AS) and permanent environmental (PE) curves on data set C: $\triangle$ = AS4-PE3, $\diamond$ = AS5-PE3, $\square$ = AS5-PE5. Homogeneous (empty symbols) and heterogeneous (filled symbols) residual variances were assumed.

Figure 3. Variance of residuals for selected days in milk for models with varying orders of Legendre polynomials (3 to 5) for age-season (AS) and permanent environmental (PE) curves on data set C: $\triangle$ = AS4-PE3, $\diamond$ = AS5-PE3, $\square$ = AS5-PE5. Homogeneous (empty symbols) and heterogeneous (filled symbols) residual variances were assumed.

Figure 4. Additive genetic (—), residual (—), permanent environmental (---), and herd-test-date (.....) variances ($\times 100$) for analyzed models based on data set C. Legendre polynomials of varying orders (3 to 5) for age-season (AS) and permanent environmental (PE) curves were applied, and homogeneous or heterogeneous (HETE) error variances were assumed.

creased at the periphery of trajectories, mainly in the beginning of lactation.

Figure 4 shows AG, PE, HTD, and the residual variances for models applied to data set C. The general pattern for all the variances was as follows: lower values at the middle part of the lactation and higher values at the extremes. The estimates of HTD variances were similar for all models. The largest differences between models were observed for the PE effect. Models with 4 parameters for the PE effect and homogeneous RV had the highest values of variation at the peripheries of lactation. Fitting heterogeneous RV for these models reduced PE variation at the beginning and end of lactation. Increasing the order of fit for the PE effect changed the pattern of variation at the extremes and also led to further reduction in PE variance at the very beginning and very end of lactation. No apparent effect of different modeling on genetic variance was noticed. The

smallest values of error variance were along the trajectory, except for the very end of lactation for the most complex model (AS5-PE5). The differences between models, however, were small. Applying the same models for data set B resulted in a similar pattern of changes in the variation (results not presented).

Heritabilities of daily milk yields are shown in Figure 5. The differences in the overall level of heritabilities were high between data sets. For data set B, heritability oscillated around 0.15 for most of the middle of lactation, while for data set C this parameter was on average 0.05 higher. This discrepancy reflects a difference in the approximate standard errors of the heritabilities, which were lower for data set B (on average 0.0049 across DIM) than for data set C (on average 0.0055). Changes in the trajectories of PE and RV for different models had a limited impact on the pattern of heritabilities. Setting PE curves to order 5, which reduced PE

Figure 5. Daily milk yield heritabilities for data sets B (lower set of curves) and C (upper set of curves) for models with Legendre polynomials of varying orders (3 to 5) for age-season (AS) and permanent environmental (PE) curves. Homogeneous or heterogeneous (HETE) error variances were assumed: $\square$ = AS4-PE3, $+$ = AS4-PE3-HETE, $\circ$ = AS5-PE3, $*$ = AS5-PE3-HETE, $\times$ = AS5-PE5, $\triangle$ = AS5-PE5-HETE.

variance at the very beginning and very end of lactation, resulted in increases in heritabilities and their standard errors, especially at the end of lactation. This increase was particularly high for data set B, where heritability reached the range of 0.3 to 0.4.

Figure 6 presents an example of the influence of herd-specific effects on lactation curves in a selected herd. From 1995 to 1997, a negative effect can be observed.

The size of this effect increased with the number of DIM. Milk production in that period was not very high (average daily yield was 15.5 kg), and lactations were not persistent in that herd. In more recent years (1999 to 2001), the positive values for this effect might be explained by the possible improvement of environmental conditions (daily yield increased to 20.3 kg). Solutions for the HY effect in the last year (2001) had higher

Figure 6. Herd-year curves for one selected herd: $*$ = 1995, $\blacksquare$ = 1996, $+$ = 1997, $\times$ = 1998, $\blacklozenge$ = 1999, $\blacktriangle$ = 2000, and $---$ = 2001.

values after the lactation peak than for the beginning of lactation, which indicated that additional milk production was achieved by increasing yield persistency.

DISCUSSION

The problem of selecting the best RRM using comparison criteria that rank models in different ways is not trivial, and has recently been discussed by Druet et al. (2003), López-Romero and Carabaño (2003), and Ødegård et al. (2003). In the study of Ødegård et al. (2003) the choice of the model required simultaneous inspection of the estimated variance components, the likelihood-ratio test statistic, the likelihood-based criterion, residual variance, and the mean square error of prediction. The differences in the ranking of models by different criteria were usually quite small, consistent with the results of this study.

A comparison based only on the PSB and ACC criteria performed in the preselection step did not clearly indicate the best model. The selection of a model for further analysis was done in a stepwise manner. For the AG effect, Legendre polynomials of order 3 were chosen because such models were superior based on the PSB, and because models with other orders for this effect and equal remaining parts of the model ranked the worst. Presumably, the PE effect should be modeled with the same function as AG, because it is difficult to find a biological reason that the functions should differ. Such a model may be easier to implement. Finally, the PSB and ACC criteria did not unanimously indicate which submodel, AG or PE, should be described by the simpler function. Models with L2 for these effects could hardly be indicated as an alternative. Both PSB and ACC evaluated those models differently. Low complexity for the PE effect curve might not be enough to account for systematic changes at the periphery of the trajectories. Although models with fixed and random HTD effects ranked similarly with respect to the accuracy of evaluation and PSB, fixed HTD gave extreme solutions for some levels of this effect. Hence, random HTD was chosen for further analysis. Considering the order of fit for the fixed part of the model, it should be stressed that the low orders of Legendre polynomials were not satisfactory in describing the overall AS lactation curve. Hence, L4 was selected for this effect for further consideration. A similar submodel could be used for HY curves, but due to the problem of small herd size, functions with a lower number of parameters (requiring fewer observations to be estimated) were preferred. Consequently, a third-order Legendre polynomial was chosen.

Further in the comparison, fixed AS curve analysis showed that fifth-order Legendre polynomials were the

most suitable for the Polish Black and White population, as they clearly improved the distribution of residuals compared with fourth-order polynomials. The high order of polynomials required to describe lactations of Polish Black and White cows may be the result of their relatively low production. The average milk daily yield of all the data sets used in this study (16.4 kg) was equal to the average yield at the end of lactation in a study based on Canadian cows (Jamrozik and Schaeffer, 1997), and much lower than the average daily yield (22.9 kg) for the cow population in the Netherlands (de Roos et al., 2004) and the 20.3 kg for Finland (Kettunen et al., 2000). The shape of the lactation curve, its lower peak in particular, is also related to the level of production. Strabel (2004) found that the peak appears relatively early, in the second part of the first month of lactation. This part of the lactation is the most complex, and thus particularly sensitive to differences in modeling. In addition, modeling is difficult due to the limited number of records available for this stage of lactation. An overestimated yield at the very beginning of lactation and an underestimated yield at the peak were also found by Jamrozik et al. (1997), who used the Wilmink and the Ali and Schaeffer (1987) functions for the fixed part of the model. The same stage of lactation caused similar problems for the German population (Liu et al., 2001).

Random regression models for the national genetic evaluation also differ in the functions used to describe random curves. Although Legendre polynomials have become a standard for this part of the model, there are differences in their order between different implementations: the third order is used in Germany (Liu et al., 2001), the fourth in Canada (Kistemaker, 2003), and the fifth in the United Kingdom (Mrode et al., 2003). Although higher-order polynomials usually improve model plausibility, there are several problems associated with them. The AG variance follows more oscillatory patterns, which leads to extreme values at the peripheries of lactation and a negative correlation for the extremes of lactation (Pool et al., 2000; López-Romero and Carabaño, 2003; Strabel et al., 2003). Moreover, analysis of the eigenvalues of AG covariance matrices shows the diminishing importance of adding further parameters. Finally, the more parameters are used, the less accurately they are estimated, because fewer records are available for each estimate. Unreasonably high estimates of genetic variance for some parts of the lactation may lead to overestimation of the average genetic variance across the whole lactation. Consequently, the accuracy of genetic evaluation may be overestimated. Thus, the rankings of random regression models based on this criterion should be studied with care.

Szyda and Liu (1999) found heterogeneous variance of yield traits in test-day records of Polish Black and White cows throughout lactation. The results of this study confirmed the heterogeneous residual variance in test-day milk yields. However, the need to incorporate heterogeneous error variance into the model was not apparent, based on the statistical tests performed. In general, using heterogeneous RV increased the variability of residuals for almost all of the models compared. Although as the error variance increased at the peripheries of the lactation, reduction of PE variance in the same regions was noted. A similar effect was reported by López-Romero et al. (2004). A small effect of fitting heterogeneous error variance on the estimation of AG and PE variance was found by Olori et al. (1999b). Similarly, López-Romero et al. (2003) found no differences in AG variance with different numbers of classes for residual variance. López-Romero et al. (2003) observed a slight impact of ignoring heterogeneous RV when the order of fit for PE curves was higher than 3. Similarly, Ødegård et al. (2003) found that the PE effect absorbed most of the heterogeneity of residual variance, particularly for more complex models. A higher order of polynomials for PE (fourth) than for AG (third) was postulated by Pool et al. (2000). López-Romero and Carabaño (2003) also concluded that lower orders of polynomials for AG than for PE might be sufficient. Although the range of values of PSB obtained in this study was within the range of results presented by López-Romero and Carabaño (2003), no association between PSB and the order of fit for the PE effect in the preselection step was observed. The positive effect of increasing the number of parameters for the PE effect was confirmed only in terms of a reduction in the variance of residuals. Setting fifth-order polynomials for PE instead of third-order had no other positive effect on the compared models. On the other hand, the model with fifth-order Legendre polynomials for PE and third-order for AG could not very accurately describe PE and AG variance at the end of lactation, which resulted in unreasonably high estimates of heritabilities. Hence, it seems appropriate to model milk test-day yields with homogeneous residual variance, because such a model is more parsimonious.

Random regression models lead to various results for the level and pattern of daily milk yield heritability; this was discussed by Misztal et al. (2000), and confirmed in more recent studies. The shape of the heritability curve with lower values at the beginning and end of lactation, as obtained by Druet et al. (2003) for example, seems most desirable, as these periods are strongly influenced by nongenetic effects cumulated before calving and associated with the farmer's decision about drying off. Similar shapes of heritability curves were obtained by

several authors using multitrait analysis (Liu et al., 2000; Druet et al., 2003) and random regression models (Pool et al., 2000; Jakobsen et al., 2002; Auvray and Gengler, 2002). Opposite shapes, with large values estimated at the peripheries and low values in the middle part of lactation were also found (Jamrozik and Schaeffer, 1997; Strabel and Misztal, 1999; Kettunen et al., 2000; Samoré et al., 2002). Using the covariance function approach and Legendre polynomials of order 3, Szyda (2001) also obtained higher heritabilities in the middle part of lactation than at the peripheries for yield traits and the first lactation. The range for heritabilities obtained in that study was similar to that obtained by Strabel and Misztal (1999) for the same population, and comparable with the results obtained in this study.

The different values of the genetic parameters obtained for the 2 data sets (B and C) imply that variance component estimation employing RRM is particularly sensitive to data selection. This problem was also mentioned by Druet et al. (2003), who found large variation among genetic parameters obtained from 10 distinct data samples. These authors suggested pooling to obtain more reliable estimates with lower standard deviations. López-Romero and Carabaño (2003) found that the values of daily variances and the associated heritabilities were not the same across data sets from different regions of Spain. The importance of the data used for analysis was also raised by Pool and Meuwissen (2000). They found that the goodness of fit is improved when only records of complete lactations were used for variance component estimation. In this study, lactations with test-day records covering at least the first 150 DIM were used. The distributions of records for both analyzed data sets showed that the frequency of the records in the third trimester of lactation decreased, reducing the number of records for the last days of the 305-d period to 50% of records in the first 2 trimesters. The differences in heritability estimates for data sets B and C, although small, were not easy to investigate. A possible explanation for the higher estimates for data set C might be differences in the structure of the data sets.

A random HTD effect required the inclusion of other herd-specific fixed effects. The herd-year effect was added to the model in this study. This made it possible to analyze the changes in environmental conditions that influenced the shape of the lactation curves. A solution for this effect may help in making management decisions, indicating in which part of the lactation an increase in production can be achieved. Although this additional benefit appeared because of adjusting the model to a population with small herd size, this effect needs to be analyzed with caution. Seasonal calvings, if they exist in a herd, may interfere with changing

environmental conditions. In small herds, due to the small number of records, herd (instead of herd-year) lactation curves might have to be estimated.

CONCLUSIONS

The PSB and ACC criteria ranked the models in different ways. Analysis of fixed curves, the distribution of residuals, the shape of the variance curves for different DIM, and investigation of the extreme solutions for some levels of effects, made it possible to select the most plausible model for routine genetic evaluation of dairy cattle in Poland. A random rather than fixed HTD effect should be used. Relatively low production and early peak of lactation in Poland require age-season lactation curves to be described by fifth-order Legendre polynomials. Other curves included in the model should be defined with third-order Legendre polynomials, because for less parsimonious models, it may be more difficult to fit the data from small herds. Fitting third-order polynomials for the permanent environmental effect made it possible to adequately account for the heterogeneous variance at different stages of lactation. Due to the relatively large influence of the data set on variance component estimation, it is recommended to use all the available data when estimating genetic parameters for the routine genetic evaluation model.

ACKNOWLEDGMENTS

The Polish State Committee for Scientific Research, Project KBN No 6 P06D 019 20, supported this work. Thanks go to the anonymous reviewers for their useful comments on the paper.

REFERENCES

- Ali, T. E., and L. R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637–644.
- Auvray, B., and N. Gengler. 2002. Feasibility of a Wallon test-day model and study of its potential as tool for selection and management. *Interbull Bull.* 29:123–127.
- Beard, K. 1983. Prediction of total lactation yield in dairy cows. M.S. Thesis, University of Melbourne, Victoria, Australia.
- Brotherstone, S., I. M. S. White, and K. Meyer. 2000. Genetic modeling of daily milk yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70:407–415.
- de Roos, A. P. W., A. G. F. Harbers, and G. de Jong. 2004. Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in the Netherlands. *J. Dairy Sci.* 87:2693–2701.
- Druet, T., F. Jaffrezic, D. Boichard, and V. Ducrocq. 2003. Modeling lactation curves and estimation of genetic parameters for first-lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86:2480–2490.
- Ducrocq, V. 2000. Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects. *Interbull Bull.* 25:123–130.
- Guo, Z., and H. H. Swalve. 1997. Comparison of different lactation curve sub-models in test day models. *Proc. Annu. Interbull Meeting*. Vienna, Austria. *Interbull Bull.* 16:75–79.
- Jaffrezic, F., I. M. S. White, R. Thompson, and W. G. Hill. 2000. A link function approach to model heterogeneity of residual variances over time in lactation curve analyses. *J. Dairy Sci.* 83:1089–1093.
- Jakobsen, J. H., P. Madsen, J. Jensen, J. Pedersen, L. G. Christensen, and D. A. Sorensen. 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *J. Dairy Sci.* 85:1607–1616.
- Jamrozik, J., G. J. Kistemaker, J. C. M. Dekkers, and L. R. Schaeffer. 1997. Comparison of possible covariates for use in a random regression model for analyses of test day yields. *J. Dairy Sci.* 80:2550–2556.
- Jamrozik, J., and L. R. Schaeffer. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first-lactation Holsteins. *J. Dairy Sci.* 80:762–770.
- Jamrozik, J., L. R. Schaeffer, and G. B. Jansen. 2000. Approximate accuracies of prediction from random regression models. *Livest. Prod. Sci.* 66:85–92.
- KCHZ. 2003. Milk performance of cows under milk recording and evaluation and selection of AI bulls in 2002. Central Animal Breeding Office, Warsaw, Poland. [In Polish]
- Kettunen, A., E. Mäntysaari, and J. Pösö. 2000. Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test-day model. *Livest. Prod. Sci.* 66:251–261.
- Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics* 124:979–993.
- Kistemaker, G. J. 2003. The Canadian test-day model using Legendre polynomials. *Interbull Bull.* 31:202–204.
- Kistemaker, G. J., and U. Schnyder. 2004. Modification of the Canadian test-day model for dairy breeds in Switzerland. *Interbull Open Meeting*, Sousse, Tunisia.
- Lidauer, M., E. A. Mäntysaari, and I. Strandén. 2003. Comparison of test-day models for genetic evaluation of production traits in dairy cattle. *Livest. Prod. Sci.* 79:73–86.
- Liu, Z., F. Reinhardt, A. Bünger, L. Dopp, and R. Reents. 2001. Application of random regression model to genetic evaluations of test day yields and somatic cell scores in dairy cattle. *Interbull Bull.* 27:159–166.
- Liu, Z., F. Reinhardt, and R. Reents. 2000. Estimating parameters of a random regression test day model for first three lactation milk production traits using the covariance function approach. *Interbull Bull.* 25:74–80.
- López-Romero, P., and M. J. Carabaño. 2003. Comparing alternative random regression models to analyse first-lactation daily milk yield data in Holstein-Friesian cattle. *Livest. Prod. Sci.* 82:81–96.
- López-Romero, P., R. Rekaya, and M. J. Carabaño. 2003. Assessment of homogeneity vs. heterogeneity of residual variance in random regression test-day models in a Bayesian analysis. *J. Dairy Sci.* 86:3374–3385.
- López-Romero, P., R. Rekaya, and M. J. Carabaño. 2004. Bayesian comparison of test-day models under different assumptions of heterogeneity for the residual variance: The change point technique versus arbitrary intervals. *J. Anim. Breed. Genet.* 121:14–25.
- Mark, T. 2004. Applied genetic evaluations for production and functional traits in dairy cattle. *J. Dairy Sci.* 87:2641–2652.
- Mayers, P., J. Stoll, R. Reents, and N. Gengler. 2002. Alternative modeling of fixed effects in test-day models to increase their usefulness for management decisions. *Interbull Bull.* 29:128–132.
- Misztal, I., T. Strabel, J. Jamrozik, E. A. Mäntysaari, and T. H. E. Meuwissen. 2000. Strategies for estimating the parameters needed for different test-day models. *J. Dairy Sci.* 83:1125–1134.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). Page 28 in *Proc. 7th World Cong. Genet. Appl. Livest. Prod.*, Montpellier, France.
- Mrode, R. A., G. J. T. Swanson, and M. F. Paget. 2003. Implementation of the test day model for production traits in the UK. *Interbull Bull.* 31:193–196.
- Ødegård, J., J. Jensen, G. Klemetsdal, P. Madsen, and B. Heringstad. 2003. Genetic analysis of somatic cell score in Norwegian cattle

- using random regression test-day models. *J. Dairy Sci.* 86:4103–4114.
- Oikawa, T., and K. Sato. 1997. Treating small herds as fixed or random in animal model. *J. Anim. Breed. Genet.* 114:177–183.
- Olori, V. E., W. G. Hill, and S. Brotherstone. 1999a. The structure of the residual error variance of test day milk yield in random regression models. Workshop on Computational Breeding, Finland.
- Olori, V. E., W. G. Hill, B. J. McGuirk, and S. Brotherstone. 1999b. Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. *Livest. Prod. Sci.* 61:53–63.
- Pool, M. H., L. L. G. Janss, and T. H. E. Meuwissen. 2000. Genetic parameters of Legendre polynomials for first parity lactation curves. *J. Dairy Sci.* 83:2640–2649.
- Pool, M. H., and T. H. E. Meuwissen. 2000. Reduction of the number of parameters needed for a polynomial random regression model. *Livest. Prod. Sci.* 64:133–145.
- Ptak, E., and A. Żarnecki. 1998. Estimation of breeding values of Polish Black and White cattle using test day yields. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia XXIII:335–338.
- Reinhardt, F., Z. Liu, A. Bünger, L. Dopp, and R. Reents. 2002. Impact of application of a random regression test day model to production trait genetic evaluations in dairy cattle. *Interbull Bull.* 29:103–107.
- Rekaya, R., M. J. Carabaño, and M. A. Toro. 1999. Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle. *Livest. Prod. Sci.* 57:203–217.
- Rekaya, R., M. J. Carabaño, and M. A. Toro. 2000. Assessment of heterogeneity of residual variances using changepoint techniques. *Genet. Sel. Evol.* 32:383–394.
- Samoré, A. B., P. Boettcher, J. Jamrozik, A. Bagnato, and A. F. Groen. 2002. Genetic parameters for production traits and somatic cell scores estimated with a multiple trait random regression model in Italian Holsteins. Proc. 7th World Congr. Genet. Appl. Livest. Prod., Montpellier, France. CD-ROM Commun. no. 01–07.
- Schaeffer, L. R., and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada 18:443–446.
- Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Stat.* 6:461–464.
- Stanton, T. L., L. R. Jones, R. W. Everett, and S. D. Kachman. 1992. Estimating milk, fat, and protein lactation curves with a test day model. *J. Dairy Sci.* 75:1691–1700.
- Strabel, T. 2004. Age-season lactation curves for production traits of cows calving in Wielkopolska region of Poland. *J. Anim. Feed Sci.* 13:405–416.
- Strabel, T., and I. Misztal. 1999. Genetic parameters for first and second lactation milk yields of Polish Black and White cattle with random regression. *J. Dairy Sci.* 82:2805–2810.
- Strabel, T., and T. Szwaczkowski. 1995. Certain nongenetic effects on test-day milk yields in dairy cows. *Anim. Sci. Pap. Rep.* 13:55–64.
- Strabel, T., and T. Szwaczkowski. 1999. The use of test day models with small size of contemporary groups. *J. Anim. Breed. Genet.* 116:379–386.
- Strabel, T., J. Szyda, E. Ptak, and J. Jamrozik. 2003. Comparison of random regression test-day models for production traits of dairy cattle in Poland. *Interbull Bull.* 31:197–200.
- Szyda, J. 2001. Application of the covariance function approach with an iterative two-stage algorithm to the estimation of parameters of a random regression test day model for dairy production traits. *J. Appl. Genet.* 42:177–191.
- Szyda, J., and Z. Liu. 1999. Modeling test day data from dairy cattle. *J. Appl. Genet.* 40:103–116.
- Ugarte, E., R. Alenda, and M. J. Carabaño. 1992. Fixed or random contemporary groups in genetic evaluations. *J. Dairy Sci.* 75:269–278.
- Visscher, P. M., and M. E. Goddard. 1993. Fixed and random contemporary groups. *J. Dairy Sci.* 76:1444–1454.