



## Modeling lactation curves of Polish Holstein-Friesian cows. Part II: Prediction of 305-d lactation milk, fat and protein yields

A. Otwinowska-Mindur<sup>1</sup>, E. Ptak, W. Jagusiak and A. Satola

University of Agriculture in Krakow, Department of Genetics and Animal Breeding  
Al. Mickiewicza 24/28, 30-059 Krakow, Poland

**KEY WORDS:** dairy cow, lactation curve, multiple-trait prediction, milk yield

Received: 27 November 2012  
Revised: 4 November 2013  
Accepted: 11 March 2014

<sup>1</sup>Corresponding author:  
e-mail: rzmindur@cyf-kr.edu.pl

**ABSTRACT.** The objective of the study was to compare five lactation curve models, Ali and Schaeffer, Guo, Wilmink, and third- and fourth-order normalized Legendre polynomials, for their ability to reliably predict 305-d lactation yields from test-day records. The analysis covered 27 589 422 test-day yields from 3 350 638 first three lactations of 1 621 796 Polish Holstein-Friesian cows. All of the functions were fitted to the test-day yields of milk, fat and protein using a multiple-trait prediction (MTP) method. Based on the fitted curves, 305-d yields were calculated and the results were compared with the official lactation yields of cows assumed as true yields. Lactation curve models differed significantly ( $P < 0.01$ ) when 305-d yields of milk traits were estimated and compared. Legendre polynomials showed the smallest bias and the highest accuracy among five compared models for all milk traits, whereas the Guo model for milk and the Wilmink model for fat and protein had the highest bias and the lowest accuracy. All of the functions produced similar estimates of milk, fat and protein yields at peak day; the estimated day of peak depended on the model.

### Introduction

In the past, total 305-d yields of milk formed the basis for genetic evaluation of dairy cattle. This parameter is still useful for producers making their management and breeding decisions (Druet et al., 2003). In the last two decades, genetic evaluations of dairy bulls and cows based on test-day (TD) yields employing TD models have been widely studied (Strabel et al., 2005; Bohmanova et al., 2008). Many different mathematical functions have been applied to model lactation curves. Appropriate and accurate modeling of lactation curves for individual animals and herds may help breeders make the right breeding decisions regarding, for example, selection or culling. According to Grzesiak et al. (2006), cows

with a very high peak yield are unable to take the necessary amount of nutrition during the first stage of lactation, and this may cause a negative energy balance, reduced reproduction rate, and increased susceptibility to diseases. To lower the cost of milk recording, many breeders have abandoned the A4 method and chosen a cheaper method (AT4 or A8). In Poland 67% of cows are tested using AT4 and 5% of cows with A8 (PFHBiPM, 2013). At the same time, breeders expect predictions of cumulative yield with minimum error. Some researches maintain that well-fitted lactation curves may accomplish that (Olori et al., 1999; Quinn et al., 2005; Silvestre et al., 2006).

Various models have been examined for their ability to properly describe the pattern of milk yields

and to accurately predict 305-d cumulative milk yield from partial records. Among the most popular functions describing the course of lactation are those of Wood (1967), Ali and Schaeffer (1987), and Wilmink (1987). The functions differ mainly in the type of regression, the number of parameters, and the degree of relationship with the main characteristics of a typical lactation curve, such as peak yield, time at peak and persistency (Macciotta et al., 2005). With time, these functions have been replaced by orthogonal polynomials that model the lactation curve properly and tend to converge faster because of a lack of correlation among the coefficients. Recently, splines have been proposed as a good alternative to Legendre polynomials (Bohmanova et al., 2008). A variety of different mathematical models have also been examined for modeling lactation curves of the Polish population of cows (Ptak and Frącz, 2002, 2003; Frącz and Ptak, 2003; Ptak et al., 2004; Strabel et al., 2005; Grzesiak et al., 2006; Otwinowska-Mindur et al., 2013).

In Poland, a test interval method (TIM) is used as a standard method for calculating 305-d lactation yield from test-day (TD) yields. The method consists in interpolating mean yields over intervals between monthly tests and adding up the resulting values, with special adjustments made to the first and last TD yields (Stolzman, 1982). With TIM, yields of milk, fat and protein are processed separately, ignoring information from the other two traits (Schaeffer and Jamrozik, 1996). Other methods developed to calculate 305-d yields include best prediction (VanRaden, 1997) and a multiple-trait procedure (Schaeffer and Jamrozik, 1996). Best prediction (BP) uses observed TD yields to predict daily yields that were not observed. In this method, for each day of lactation the most probable yield is computed using selection index methodology, and next the total 305-d yield is taken as the sum of the predicted daily yields (VanRaden, 1997; Cole et al., 2009). In the multiple-trait procedure (MTP) by Schaeffer and Jamrozik (1996), lactation curve parameters for a single cow are estimated using information on standard curves for groups of cows with similar production characteristics and incorporating phenotypic correlations among yields of milk, fat and protein. This method can handle long intervals between TD yields (as in the A8 method) and also test days with milk only recorded, and can make 305-d predictions on the basis of just one TD record per cow. The MTP method lends itself to calculation of many parameters important to herd owners, such as peak yields, day of peak yield, expected daily yields and herd lactation curves. The MTP method

is not significantly better than TIM when data are available for all equally spaced tests, but if fat and protein information is missing for some TD, MTP allows prediction of 305-d yields for those traits with relatively high accuracy (Schaeffer and Jamrozik, 1996).

The objective of this study was to compare five mathematical functions used as lactation curve models for their suitability to estimate the 305-d lactation milk, fat and protein yields of Polish Holstein-Friesian cows. The best model could be applied in practice as an alternative to the currently used test interval method (TIM). The true 305-d yields of milk traits were not available so we assumed 305-d yields calculated by TIM as a basis for comparison.

## Material and methods

The data were records on 27 589 422 test-day (TD) yields of milk, fat and protein from 3 350 638 first three lactations of 1 621 796 Polish Holstein-Friesian cows (Table 1). The records were gathered in the SYMLEK national recording system, and were made available by the Polish Federation of Cattle Breeders and Dairy Farmers.

Cows calved in 1995–2009 at ages 18–48, 29–65 and 41–75 months, respectively, for the first, sec-

**Table 1.** Number of lactations and test-day records by parity, calving age, and season of calving

Lactation No.	Calving age, months	Number of lactations	Number of test-day records
1	18–24	358 050	2 934 134
	25–26	379 327	3 228 104
	27–28	283 423	2 424 341
	29–30	184 905	1 581 990
	31–48	271 422	2 290 253
Total		1 477 127	12 458 822
2	29–38	366 555	2 908 541
	39–41	276 278	2 275 468
	42–44	198 280	1 647 385
	45–65	261 320	2 159 625
Total		1 102 433	8 991 019
3	41–51	260 541	2 029 427
	52–55	209 195	1 685 066
	56–75	301 342	2 425 088
Total		771 078	6 139 581
Lactation No.	Season of calving	Number of lactations	Number of test-day records
1	April–September	701 726	5 885 777
	October–March	775 401	6 573 045
2	April–September	539 883	4 399 758
	October–March	562 550	4 591 261
3	April–September	378 837	3 008 543
	October–March	392 241	3 131 038

ond and third calving. TD records, 1 to 10 per lactation per cow, and TD yields between 5 and 305 days in milk (DIM) were considered. Data were collected with three methods: A4 (52% of records), AT4 (31% of records) and A8 (17% of records). There were 8 TD records per cow per lactation on average. In case of the A4 or AT4 methods, the average number of TD records was 9–10, and when the A8 method was used, the average number of records was 5–6, with lower means in third parity. Most of the lactations terminated naturally (drying or next calving: 70%–75%), much fewer terminated with culling (15%–20%) and the least were lactations in progress (9%). About a third of the naturally terminated lactations lasted less than 305 days. Daily yields exceeding 85 kg milk and 8.5% fat or protein were rejected. According to age at calving, the data were divided into five, four and three groups within the first, second and third lactations respectively (Table 1). Two seasons of calving were assumed (October–March and April–September).

A multiple-trait prediction (MTP) method was used to model the lactation curve and to estimate 305-d lactation yields from individual milk, fat and protein TD yields (Schaeffer and Jamrozik, 1996). The parameters of standard lactation curves were estimated within 24 subclasses by lactation, age at calving, and season of calving. To estimate the matrix of (co)variances among the curve parameters, only cows with first TD before 50 DIM and a minimum of 9 TD records per lactation were used. The total milk, fat and protein yields of 305-d lactation were calculated by adding up yields from 5 to 305 DIM. All of the 305-d yields obtained by MTP were compared with the yields from the SYMLEK system, that is, those predicted by the test interval method (TIM). In Poland the TIM method is used for lactations that last at least 200 days. Because of that condition, the 305-d yields calculated by TIM were available for 65%, 60% and 57% of the all data analysed by MTP for the first, second and third lactations, respectively.

The functions chosen to model the lactation curve were as follows: Ali and Schaeffer (1987), Guo (Guo and Swalve, 1995), Wilmink (1987), and third- and fourth-order Legendre polynomials (Kirkpartick et al., 1990). The form of the functions together with a comparison of their goodness of fit were given in the first part of the study under the general title ‘Modeling lactation curves of Polish Holstein-Friesian cows’ (Otwinowska-Mindur et al., 2013).

For comparison of models, the mean and standard deviation of the difference between estimated

and true 305-d yields were used as measures of bias and accuracy, respectively. The estimated yields were those calculated by MTP and a chosen lactation curve model, whereas yields determined by TIM were assumed as the true yields. The effects of such factors as lactation number (1, 2, 3), lactation curve model (5 functions), and method of testing (A4, AT4, A8) were investigated by analysis of variance through the GLM procedure of SAS (2004).

## Results

The means and standard deviations for peak day and peak yields predicted using the five models of the lactation curve are given in Table 2. Although all models gave very similar estimates of peak yields, the peak days were determined differently by each model. The earliest days for peak were generated by the Guo (GUO) and the Wilmink (WIL) functions in each of the three lactations. Peak day occurred somewhat later when estimated by the Ali and Schaeffer function (ALI) and third-order Legendre polynomials (LEG3), and at least one month later when fourth-order Legendre polynomials (LEG4) were used. Within each lactation, variation among estimates of peak milk yield obtained by the five models was comparable, but the standard deviation (SD) was highest for LEG4 and lowest for both three-parameter functions (WIL and GUO). Fat and protein yields at peak differed very little between the models. The SD of fat within each parity was slightly higher when ALI was used. For protein the standard deviations were at similar levels regardless of the lactation curve model used.

Only the ALI function sometimes modelled lactation curves with two peaks. The second production peak occurred on average at 202 DIM in first parity and about 195 DIM in later parities, with lower yields than in the first peak. In many cases, the maximal milk yield occurred at the beginning of lactation (5 DIM) when lactation curves were fitted using the GUO (24%–34%), WIL (42%–60%), LEG3 (44%–60%) and LEG4 (35%–48%) functions. This happened only in 1% of the lactations when the ALI function was used.

The means and standard deviations for 305-d milk, fat and protein yields calculated using the five lactation curve models and the TIM method are presented in Table 2. Total (305-d) milk yields averaged 5 852–5 882 kg (TIM: 5 880 kg), 6 519–6 559 kg (TIM: 6 547 kg) and 6 716–6 752 kg (TIM: 6 740 kg) in the first, second and third lactations, respectively. The Legendre polynomials (LEG3, LEG4) gave the estimates of 305-d milk yield that least differed from

**Table 2.** Means and standard deviations (SD) of peak day of milk yield, milk, fat and protein yields at peak, and 305-d lactation yields estimated using five lactation curve models and multiple-trait prediction method (MTP) or test interval method (TIM)

Lactation No.	Lactation model <sup>1</sup>	Day of peak milk yield <sup>2</sup>		Peak yields <sup>2</sup>						305-d yields <sup>3</sup>					
				milk, kg		fat, kg		protein, kg		milk, kg		fat, kg		protein, kg	
		mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
1	TIM	—	—	—	—	—	—	—	—	5 880	1 732	241	72	190	59
	ALI	42.26	37.50	23.89	6.99	0.97	0.40	0.72	0.24	5 854	1 734	241	74	191	60
	GUO	33.14	47.48	23.09	6.55	0.97	0.32	0.70	0.21	5 852	1 723	241	72	190	59
	LEG3	62.32	76.09	22.85	6.73	0.94	0.32	0.70	0.22	5 875	1 735	241	72	190	59
	LEG4	86.28	85.77	22.93	7.13	0.93	0.35	0.71	0.23	5 882	1 740	241	73	190	59
	WIL	31.98	46.72	23.19	6.53	0.99	0.35	0.71	0.22	5 862	1 722	241	72	190	59
2	TIM	—	—	—	—	—	—	—	—	6 547	1 998	272	85	216	67
	ALI	33.40	28.55	28.51	8.99	1.17	0.53	0.88	0.30	6 519	2 001	272	87	217	68
	GUO	19.36	24.97	28.28	8.71	1.20	0.44	0.88	0.28	6 530	1 998	273	86	216	67
	LEG3	42.04	66.48	27.45	8.86	1.14	0.44	0.86	0.28	6 550	2 003	273	86	217	67
	LEG4	61.83	77.10	27.47	9.41	1.12	0.47	0.86	0.29	6 559	2 011	273	86	216	67
	WIL	17.76	25.16	28.64	8.81	1.24	0.49	0.90	0.29	6 543	1 997	273	85	216	67
3	TIM	—	—	—	—	—	—	—	—	6 740	1 956	282	86	220	65
	ALI	33.87	28.17	29.66	9.04	1.23	0.57	0.91	0.31	6 716	1 961	283	89	221	66
	GUO	19.82	22.60	29.23	8.65	1.26	0.45	0.90	0.27	6 718	1 956	283	87	220	65
	LEG3	39.03	61.39	28.59	8.81	1.21	0.46	0.88	0.28	6 741	1 961	283	86	221	65
	LEG4	58.46	73.70	28.73	9.47	1.19	0.50	0.88	0.30	6 752	1 969	283	87	221	66
	WIL	18.31	23.32	29.59	8.72	1.30	0.51	0.92	0.28	6 733	1 956	283	86	221	65

<sup>1</sup> TIM – test interval method, ALI – Ali and Schaeffer function, GUO – Guo function, LEG3 – third-order Legendre polynomials, LEG4 – fourth-order Legendre polynomials, WIL – Wilmink function;

<sup>2</sup> day of peak and peak yields were calculated using records of 1 477 127 first, 1 102 433 second and 777 078 third lactations

<sup>3</sup> 305-d yields were calculated using records of 953 674 first, 652 189 second and 438 348 third lactations

**Table 3.** Means and standard deviations (SD) of absolute differences (DIFF-A) and percentage differences (DIFF-P) between predicted and true 305-d milk, fat and protein yields by different lactation and lactation curve models

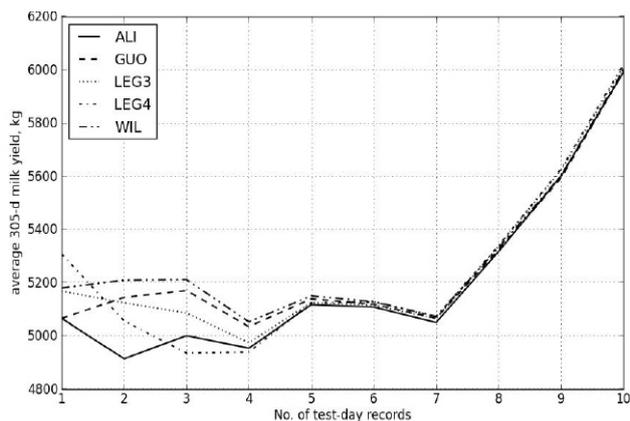
Lactation No.	Lactation model <sup>1</sup>	Milk, kg				Fat, kg				Protein, kg			
		DIFF-A		DIFF-P		DIFF-A		DIFF-P		DIFF-A		DIFF-P	
		mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
1	ALI	58.92	65.10	1.06	1.17	3.36	9.74	1.41	3.75	2.25	7.10	1.15	2.65
	GUO	67.08	67.28	1.16	1.11	3.22	3.81	1.36	1.47	2.46	5.58	1.25	1.30
	LEG3	41.67	52.40	0.73	0.88	2.57	3.51	1.08	1.32	1.70	5.43	0.84	1.07
	LEG4	35.89	54.62	0.63	0.92	2.42	4.19	1.02	1.62	1.57	5.56	0.77	1.25
	WIL	66.25	66.93	1.15	1.11	3.49	4.25	1.48	1.63	2.54	5.64	1.29	1.36
2	ALI	69.15	80.16	1.12	1.30	4.06	11.74	1.51	4.00	2.63	7.66	1.20	2.68
	GUO	71.06	74.85	1.12	1.13	3.93	4.78	1.46	1.59	2.56	5.76	1.15	1.27
	LEG3	53.24	66.45	0.84	1.00	3.27	4.48	1.21	1.48	2.06	5.64	0.92	1.14
	LEG4	46.34	68.95	0.73	1.03	3.15	5.14	1.17	1.72	1.96	5.80	0.87	1.30
	WIL	70.36	73.88	1.11	1.13	4.36	5.36	1.63	1.77	2.71	5.84	1.23	1.34
3	ALI	70.61	84.63	1.10	1.31	4.39	12.66	1.57	4.15	2.80	7.67	1.25	2.71
	GUO	74.10	77.79	1.13	1.14	4.33	5.23	1.55	1.68	2.62	5.74	1.16	1.30
	LEG3	55.50	69.32	0.85	1.01	3.52	4.88	1.26	1.54	2.13	5.62	0.93	1.17
	LEG4	48.30	72.59	0.74	1.05	3.43	5.67	1.22	1.83	2.02	5.80	0.88	1.35
	WIL	72.18	76.42	1.10	1.13	4.85	5.95	1.74	1.89	2.79	5.83	1.25	1.38

<sup>1</sup> see Table 2

the TIM estimates. Three other models (ALI, GUO, WIL) slightly underestimated total milk yields but the differences in yields between the lactation curve models and TIM constituted less than 0.5% of TIM estimates. Mean total yields of fat and protein were the same regardless of the method used (MTP or TIM). Standard deviations of yields were greater

in later lactations than in the first one for each trait (milk, fat and protein). It is worth mentioning that the averages were calculated using data restricted to lactations that lasted at least 200 days.

When 305-d yields were predicted based on the data from all lactations, also those lasting several months, the total yields were often underestimated



**Figure 1.** Means of 305-d milk yields estimated using five lactation curve models (ALI – Ali and Schaeffer function, GUO – Guo function, LEG3 – third-order Legendre polynomials, LEG4 – fourth-order Legendre polynomials, and WIL – Wilmlink function) by different numbers of test-day records

by the MTP method. Figure 1 shows how average 305-d milk yields changed when data accumulated. The differences in milk yields between the lactation curve models and TIM were greater for later lactations than for the first one. The differences in 305-d milk predictions between the various lactation curve models depended on the number of TD yields used in calculations and amounted to less than 5% in the first lactation and less than 7.2% in the third lactation (Figure 1). When there were more than 3–4 TD yields available per cow, the estimated yields differed by less than 1%. Lactation yields estimated based on a few TD records (1–3) mostly differed between models. A clear upward trend in predicting the total yields was observed when more TD data were used (Figure 1). The same tendency of changes was seen for 305-d fat and protein yields (not presented).

There were statistically significant ( $P < 0.01$ ) differences in means among the five compared models, mainly between ALI and WIL in each lactation for 305-d milk and fat yields, and in later lactations for protein yield. For total milk yield there were significant differences ( $P < 0.01$ ) also between ALI and Legendre polynomials (LEG3 and LEG4). The 305-d protein yield in the first and third lactations differed significantly between ALI and LEG4 ( $P < 0.01$ ). There were no significant differences between ALI and LEG3 or LEG4 for total fat yield.

Because almost half of the data (48%) were from herds in which AT4 (31%) or A8 (17%) testing schemes were used, the influence of this factor on prediction of 305-d yields or peak yields was examined. As expected, both 305-d yields and peak yields of milk, fat and protein were smaller when A8 was the method of testing. The A4 and AT4 testing schemes provided similar predictions of peak yields irrespective of the model fitted to the data. Milk yield

at peak differed between the models at about 0.6 kg (1.5%–1.7%) in the first and 1.0 kg (3%) in the third lactation. The day of maximal milk yield depended on the testing scheme, especially when Legendre polynomials were used as the lactation curve model. For example, in the first lactation modeled by LEG3, the peak day occurred at 51 DIM for the A8 method and at 61 or 68 DIM when A4 or AT4 were used. In the third lactation those differences were smaller (A8: 33 DIM, A4: 38 DIM, AT4: 43 DIM).

The effect of testing method on the 305-d yields of all 3 traits was statistically significant ( $P < 0.01$ ), and there were significant differences between all pairs of testing methods (A4 vs A8, A4 vs AT4, and AT4 vs A8). In the future, this factor should be included when standard lactation curves in MTP are determined.

Table 3 compares the five lactation curve models in terms of their suitability for calculation of 305-d yields. Total yields determined by TIM were assumed as the true yields. The bias was expressed as the mean absolute difference between estimated and true yields and as the mean percentage of those differences in relation to TIM yields. In both cases, the bias was smallest for Legendre polynomials and largest for GUO when total milk yield was compared. Looking at the bias in fat and protein 305-d yields, the WIL model produced the greatest differences from the true 305-d yields, and LEG4 and LEG3 again the smallest. Generally, those differences were larger in later lactations than in the first one. The standard deviations (SD) of absolute and percentage differences between estimated and true yields showed similar accuracy for LEG3 and LEG4, slightly lower for WIL and GUO, and in many cases, the worst for ALI. Much higher SD was observed particularly for the differences in 305-d yields of fat and protein estimated by ALI. On a percentage scale, they were at least twice as high as for yields calculated with LEG3 or LEG4.

In most cases, the linear correlations between the total yields of milk, fat and protein estimated by TIM and those predicted by the five mathematical functions were positive and almost perfect ( $> 0.99$ ), with only one exception: the correlations between fat yields calculated by ALI and TIM (0.977–0.981). Such high values are attributable to two factors: first, the huge number of records used for calculation of these correlations ( $> 2$  millions), second, the coefficients of correlation were calculated only for relatively long-lasting lactations, that is, those that lasted at least 200 DIM. In such cases, the lactation curves are fitted more precisely than when only 1 or 2 test-day yields are available. In turn, one pos-

sible explanation of the lower correlations for fat yield is the relatively large number of lactations for which the 305-d fat yield prediction from ALI differed from that calculated by TIM by more than 30% (about 2 000 lactations).

## Discussion

Peak milk yields were lowest in the first lactation, and were reached later than in the second and third parities, which is consistent with results from Strabel (2004) and Sawa et al. (2007). The latter authors found that cows achieved maximum milk yield (23.5 kg) on DIM 62 during the first lactation, and slightly higher yields (27.3–28.7 kg) in earlier DIM (47–50) during later lactations. Strabel (2004) showed lower milk yields at peak (16–18 kg), which were obtained in the second part of the first month of lactation (21–35 DIM). He concluded that this relatively early peak of lactation might be explained in part by the fact that the first month after calving is the most complex physiological stage of lactation, thus, very sensitive to differences in modeling. Additionally, the limited number of records usually available for that part of lactation makes accurate modeling difficult, causing overestimation of yields at the very beginning and underestimation of peak yields (Strabel et al., 2005). Cole et al. (2009) reported that first-parity Holstein cows produced, on average, less milk at peak (33.40 kg) and exhibited higher average DIM (82.51) until peak yield, as compared to later lactations (milk yield: 42.88 kg, DIM: 58.02). Similar findings were described by Dematawewa et al. (2007): mature cows achieved higher and earlier peaks than maturing heifers. They modeled lactation curves using the WIL function and observed that first-parity cows reached maximum milk yield (33.55 kg) at 93 DIM, whereas cows in third and later parities achieved higher peak milk yield (44.42 kg) earlier (54 DIM). Tekerli et al. (2000) noted that the milk secretory tissue of cows in the first lactation reaches its peak activity later than in subsequent lactations.

Comparing five models of the lactation curve, Olori et al. (1999) found the WIL model to be best in predicting both the time (week 7) and yield (32.0 kg) at peak. They also examined the effect of the week of peak production on the goodness of fit of the models, and came to the conclusion that lactations peaking between 6 and 9 weeks were predicted more accurately than those peaking earlier or later. Dematawewa et al. (2007) examined nine different models and recommended the WIL function as the most accurate and simplest model for estimating

the peak of lactation. Quinn et al. (2005) used several models for analysis, including ALI, GUO and WIL, and estimated the peak at about week eight. It should be added that Olori et al. (1999) and Quinn et al. (2005) considered weekly yields; what is more, Olori et al. (1999) used only data on heifers from one herd, kept under uniform feeding and management conditions. Macciotta et al. (2005) suggested that three-parameter models such as the WIL function ensured a superior fit to data with a single curvature. According to them, five-parameter functions like ALI and LEG4 are more flexible and permit description of a larger number of lactation curve shapes; that is why they were often used in random regression analyses of TD data to model individual deviations from an average fixed curve. For the Polish population, Strabel et al. (2005) recommended fifth-order Legendre polynomials for modeling age-season lactation curves. They pointed out that the necessity to use the required polynomials to describe the course of lactations of Polish cows might be the result of their relatively low production.

The results presented in Table 2 confirm that the very high standard deviation of peak day, especially for Legendre polynomials (LEG3 and LEG4), is related to the mathematical rather than biological nature of the different shapes of lactation curves for milk yield, especially when modeled by polynomials. Polynomials are very flexible functions and can represent not only the typical form of the lactation curve (with an ascending phase to a peak and a descending phase thereafter), but also several other shapes. When the first TD yield was at the end of the first month of lactation, the yields in the first month were extrapolated, and that might be why LEG3 and LEG4 very often modeled lactation curves without a peak, that is, as continuously decreasing curves. Macciotta et al. (2005) confirmed that such atypical lactation curves were mainly a result of lack of records in the first days after calving. They also noticed that the occurrence of the curves without a peak arose from the peculiar combination of TD yields and their distribution along the entire lactation. Silvestre et al. (2006) confirmed that the goodness of fit of lactation models depends on data availability and structure. They stated that some models such as WIL or ALI were particularly affected by the interval between tests, mostly between calving and the first TD. This confirmed the need for further research on the effect of data structure on the accuracy of modeling lactation curve and indirectly of 305-d yields using fitted models. Many authors have pointed out that the availability of TD records before peak yield is crucial for correct estimation of

lactation curve shape (Macciotta et al., 2005; Strabel et al., 2005; Silvestre et al., 2006).

The 305-d yields of milk, fat and protein calculated using lactation curve models should be compared with true yields, meaning those calculated as a sum of daily yields. We had no access to such data, so the 305-d yields predicted based on the mathematical models were compared with those calculated by the TIM method treated as the true 305-d yields. The total yields of milk, fat and protein estimated using the various models of the lactation curve were in the range of those calculated by the TIM method (Table 2). This is connected mainly with the fact that the TIM method was used only for parities longer than 200 days, the ones with more TD yields. In such cases, the lactation curves were fitted more accurately than when only one or two TD yields were available.

The results given in Table 2 generally agreed with those obtained by Ptak and Frącz (2003) and Ptak et al. (2004) for the Polish HF population. In these studies, however, Legendre polynomials (i.e. second-order, LEG3 and LEG4) were investigated, and the mean 305-d yields predicted by all the lactation curve models were higher than those computed by TIM. According to Quist et al. (2007), MTP overestimated 305-d yields to a considerable extent in early lactation and more accurately estimated the yield in the later part of the lactation. Our results also showed that the accuracy of MTP depends on the number of TD records available but the opposite trend was observed in early lactation (Figure 1). When only a few TD records were available, the 305-d yields of milk were underestimated by MTP by about 500–700 kg in comparison with TIM estimates. One explanation might be that the MTP method uses so-called ‘standard lactation curves’ for calculation of total 305-d yields. The parameters of those standard curves represent the average milk production of groups of cows with the same production characteristics and play an important role mainly in early lactations when there is not much data available for a particular cow. Underestimation of a cow’s 305-d yield could also be connected in part with the low average milk production described by the standard curve or inaccurate prediction of parameters of the standard lactation curve for the group to which this cow belongs. Perhaps it would be reasonable to fit standard lactation curve parameters to better-described, more homogeneous groups of cows.

Quinn et al. (2005) compared various lactation curve models using Irish data and found that the ALI (five-parameter) model overestimated the ac-

tual total milk yield by almost 4%, while the WIL and GUO (three-parameter) models overestimated it only by 1.9% and 2.5%, respectively. ALI, however, showed a much better fit to the milk yield of Irish cows than WIL or GUO. The same dependencies among these models were observed in our study: ALI predicted 305-d yields worse than WIL or GUO, which showed similar accuracy of prediction. On the other hand, the ALI model ensured a superior fit to the data (Otwinowska-Mindur et al., 2013). Schaeffer and Jamrozik (1996) found that the two methods of predicting 305-d lactation yield they compared (MTP and TIM) gave similar results when two assumptions were fulfilled: tests were regularly spaced through a lactation, and each test included yields for all three traits: milk, fat and protein. This conclusion was also confirmed in our research.

In all cases, the coefficients of correlation between the 305-d yields predicted by lactation curve models and those estimated by TIM exceeded 0.97. This is in agreement with Ptak and Frącz’s (2002) finding that the 305-d milk, fat and protein yields predicted using WIL were almost perfectly positively correlated with those calculated by TIM ( $r \geq 0.98$ ). Frącz and Ptak (2003) also obtained very high correlations between the 305-d values of all three traits predicted by ALI and those produced by TIM ( $r \geq 0.98$ ). Ptak et al. (2004) found that the correlations between the yields predicted by Legendre polynomials (LEG3 and LEG4) and the yields computed by TIM exceeded 0.98, with the single exception of 305-d protein yields predicted using LEG3 ( $r = 0.68$ ). Modeling the lactation curve by LEG4 and comparing the results with the TIM estimates in the first three lactations, Ptak and Frącz (2003) obtained slightly lower correlations for 305-d milk (0.85–0.92) and fat (0.61–0.71) yields, and higher (over 0.94) for 305-d protein yield. Schaeffer and Jamrozik (1996) reported high correlations (0.991–0.997) between the yields predicted using a lactation curve model (MTP method) and the TIM estimates. They also observed that the correlation of MTP and TIM with true 305-d yields became closer as the number of tests increased. In this case the MTP method is more accurate than when there are only a few TD yields from the first few tests available.

The performance of the models in describing fat and protein yields was similar to that for milk yield. This is in agreement with results from Dematawewa et al. (2007), but Guo and Swalve (1995) proposed using different models for fat and protein traits than for milk yield.

## Conclusions

All of the functions compared in this study – Ali and Schaeffer (ALI), Guo, Wilmink, and third-order (LEG3) and fourth-order (LEG4) normalized Legendre polynomials – gave similar estimates for the amount of milk, fat and protein at peak of lactation, but the estimated day of peak milk yield depended on the model; lactation curves modeled using the Ali and Schaeffer, Guo, or Wilmink functions gave earlier peak days than those obtained from Legendre polynomials.

The present study showed that there were statistically significant differences among the five compared models in predicting 305-d yields of milk, fat and protein. In practice this means that breeders interested in predicting 305-d yields by the multiple-trait prediction method should use Legendre polynomials (LEG3 or LEG4); if breeders want to obtain more information about the course of a cow's lactation, they should consider the function that best fits the data (i.e. LEG4 or ALI).

## Acknowledgements

The use of Dr. Janusz Jamrozik's programmes is gratefully acknowledged.

## References

- Ali T.E., Schaeffer L.R., 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67, 637–644
- Bohmanova J., Miglior F., Jamrozik J., Misztal I., Sullivan P.G., 2008. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. *J. Dairy Sci.* 91, 3627–3638
- Cole J.B., Null D.J., VanRaden P.M., 2009. Best prediction of yields for long lactations. *J. Dairy Sci.* 92, 1796–1810
- Dematawewa C.M.B., Pearson R.E., VanRaden P.M., 2007. Modeling extended lactations of Holsteins. *J. Dairy Sci.* 90, 3924–3936
- Druet T., Jaffrezic F., Boichard D., Ducrocq V., 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
- Fraçz A., Ptak E., 2003. Prediction of 305-day milk yield of Black and White cows using Ali and Schaeffer lactation curve model (in Polish). *Rocz. Nauk. Zoot., Supl.* 17, 349–352
- Grzesiak W., Błaszczyk P., Lacroix R., 2006. Methods of predicting milk field in dairy cows – Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Comput. Electron. Agr.* 54, 69–83
- Guo Z., Swalve H.H., 1995. Modeling of lactation curve as a sub-model in the evaluation of test day records. *Interbull Meeting, Prague (Czech Republic) Interbull Bull.* 11, 4–7
- Kirkpatrick M., Lofsvold D., Bulmer M., 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics* 124, 979–993
- Macciotta N.P.P., Vicario D., Cappio-Borlino A., 2005. Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. *J. Dairy Sci.* 88, 1178–1191
- Olori V.E., Brotherstone S., Hill W.G., McGuirk B.J., 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.* 58, 55–63
- Otwinowska-Mindur A., Ptak E., Jagusiak W., Satola A., 2013. Modeling lactation curves of Polish Holstein-Friesian cows. Part I: The accuracy of five lactation curve models. *J. Anim. Feed Sci.* 22, 19–25
- PFHBiPM (Polish Federation of Cattle Breeders and Dairy Farmers), 2013. Evaluation and Dairy Cattle Breeding. Date from 2012 year (in Polish), pp. 70
- Ptak E., Fraçz A., 2002. Utilization of Wilmink lactation curve for the prediction of 305-day milk yield in the first lactation of Black-and-White cows (in Polish). *Rocz. Nauk. Zoot., Supl.* 15, 31–36
- Ptak E., Fraçz A., 2003. Modeling lactation curves using orthogonal polynomials (in Polish). *Zesz. Nauk. Prz. Hod.* 68, 191–197
- Ptak E., Satola A., Czaja H., 2004. Prediction of 305-d lactation milk, fat and protein yields using Legendre polynomials and test-day yields from different parts of lactation. *Anim. Sci. Pap. Rep.* 22, 173–183
- Quinn N., Killen L., Buckley F., 2005. Empirical algebraic modelling of lactation curves using Irish data. *Irish J. Agr. Food Res.* 44, 1–13
- Quist M.A., LeBlanc S.J., Hand K.J., Lazenby D., Miglior F., Keltom D.F., 2007. Agreement of predicted 305-day milk yields relative to actual 305-day milk weight yields. *J. Dairy Sci.* 90, 4684–4692
- SAS, 2004. SAS © 9.1.2 Qualification Tools User's Guide. SAS Institute Inc. Cary, NC
- Sawa A., Neja W., Bogucki M., Repuszewska D., 2007. Course of lactation in first calvers and older cows as related to milk yield (in Polish). *Rocz. Nauk. PTZ* 3, 205–216
- Schaeffer L.R., Jamrozik J., 1996. Multiple-trait prediction of lactation yields for dairy cows. *J. Dairy Sci.* 79, 2044–2055
- Silvestre A.M., Petim-Batista F., Colaco J., 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. *J. Dairy Sci.* 89, 1813–1821
- Stolzman M. (Editor), 1982. Guide for Cattle Breeders (in Polish). PWRiL, Warszawa, pp. 26–30
- 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., Szyda J., Ptak E., Jamrozik J., 2005. Comparison of random regression test-day models for Polish Black and White cattle. *J. Dairy Sci.* 88, 3688–3699
- Tekerli M., Akinci Z., Dogan I., Akcan A., 2000. Factors affecting the shape of lactation curves of Holstein cow from the Balikesir province of Turkey. *J. Dairy Sci.* 83, 1381–1386
- VanRaden P.M., 1997. Lactation yields and accuracies computed from test day yields and (co)variances by best prediction. *J. Dairy Sci.* 80, 3015–3022
- Wilmink J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* 16, 335–348
- Wood P.D.P., 1967. Algebraic model of the lactation curve in cattle. *Nature* 216, 164–165