Best prediction of yields for long lactations

J. B. Cole,[1] D. J. Null, and P. M. VanRaden
Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

ABSTRACT

Lactation records of any reasonable length now can be processed with the selection index method known as best prediction (BP). Previous prediction programs were limited to the 305-d standard used since 1935. Best prediction was implemented in 1998 to calculate lactation records in USDA genetic evaluations, replacing the test interval method used since 1969 to calculate lactation records. Best prediction is more complex but also more accurate, particularly when testing is less frequent. Programs were reorganized to output better graphics, give users simpler access to options, and provide additional output, such as BP of daily yields. Test-day data for 6 breeds were extracted from the national dairy database, and lactation lengths were required to be $\geq$500 d (Ayrshire, Milking Shorthorn) or $\geq$800 d (all others). Average yield and SD at any day in milk (DIM) were estimated by fitting 3-parameter Wood's curves (milk, fat, protein) and 4-parameter exponential functions (somatic cell score) to means and SD of 15- ($\leq$300 DIM) and 30-d ($>$300 DIM) intervals. Correlations among TD yields were estimated using an autoregressive matrix to account for biological changes and an identity matrix to model daily measurement error. Autoregressive parameters (r) were estimated separately for first (r = 0.998) and later parities (r = 0.995). These r values were slightly larger than previous estimates due to the inclusion of the identity matrix. Correlations between traits were modified so that correlations between somatic cell score and other traits may be nonzero. The new lactation curves and correlation functions were validated by extracting TD data from the national database, estimating 305-d yields using the original and new programs, and correlating those results. Daily BP of yield were validated using daily milk weights from on-farm meters in university research herds. Correlations ranged from 0.900 to 0.988 for 305-d milk yield. High correlations ranged from 0.844 to 0.988 for daily yields, although correlations were as low as 0.015 on d 1 of lactation, which may be due to calving-related disorders that are not accounted for by BP. Correlations between 305-d yield calculated using 50-d intervals from 50 to 250 DIM and 305-yield calculated using all TD to 500 DIM increased as TD data accumulated. Many cows can profitably produce for $>$305 DIM, and the revised program provides a flexible tool to model these records.

Key words: best prediction, milk yield, long lactation

INTRODUCTION

There is increasing interest in extended lactations for dairy cows (Knight, 2005), particularly as high-producing cows in intensively managed herds appear to suffer more from health and fertility problems than do lower producing cows (Hansen, 2000; Lucy, 2001; Windig et al., 2006). Cows with longer lactations may be able to produce the same amount of lifetime milk while incurring risks associated with difficult calvings and postpartum metabolic diseases less often, and it may be profitable to breed some cows later in lactation than is needed to maintain a 365-d calving interval (De Vries, 2006). The ability to model lactations longer than 305 d is needed to determine the optimum calving interval. Cows with long lactations also are less affected by pregnancy and its associated physiological effects.

Best prediction (**BP**; VanRaden, 1997; Cole and VanRaden, 2006) is a computational method derived from selection index that allows test-day (**TD**) data to be condensed into total lactation yields and persistencies. The use of the BP by the Animal Improvement Programs Laboratory (**AIPL**; USDA, Beltsville, MD) for calculation of lactation records for milk (**M**), fat (**F**), and protein (**P**) yields and SCS began in November 1999. The BP approach was applied for all cows calving in 1997 and later, replacing the use of the test interval method (**TIM**; Sargent et al., 1968) and projection factors (Shook et al., 1980a,b). Norman et al. (1999a) showed that the advantages of BP are small compared with the TIM for most 305-d lactations, but are larger for lactations with infrequent testing or missing component samples. Some biases in TIM were eliminated by the introduction of Shook factors (Shook et al., 1980b), but BP proved to be more useful when dealing with

many alternative testing plans. Best prediction permits more precise estimation of records for SCS in particular, because TD are adjusted for stage of lactation. Yield records calculated using BP have slightly lower standard deviations (**SD**) than those from TIM because BP regresses estimates toward the herd average.

Best prediction calculates actual daily and lactation yields from TD data. Daily BP of yields and SCS may be useful for cow management, particularly if combined with other data available in on-farm computer systems to more accurately predict actual daily yields. For example, daily estimates of yield could be adjusted for known effects of health conditions such as clinical mastitis (e.g., Rajala-Schultz et al., 1999) or breeding status (De Vries, 2006) to obtain more accurate estimates of the cow's actual daily yield. Predicted daily yields for the remainder of current lactations are also important for breeding and culling decisions (De Vries, 2006). It is possible to account for environmental effects such as regional climate differences, which could be of particular value to producers in the southeastern United States. Many factors known to affect yield, such as pregnancy effects, are not yet accounted for in BP, but can readily be accommodated with the new software. A single lactation curve shape was previously used for all breeds and has been replaced with breed-specific curves. Crossbred cows can be processed using the curves for the sire breed.

The International Committee for Animal Recording (ICAR, 2006) approved BP for lactation record calculation in 2002. It requires only phenotypic covariances and assumes that herd means and variances are known. Reverse prediction may be used to obtain daily yields from lactation yields and persistencies. Both single- and multiple-trait prediction are supported. The Canadian system for estimating 305-d records, multiple trait prediction (**MTP**; Schaeffer and Jamrozik, 1996), is similar to BP in many respects.

Dematawewa et al. (2007) used data from long lactations of US Holsteins (**HO**) to fit several different functions to lactation curves by using TD data for M, F, and P, concluding that Wood's curves (Wood, 1967) most parsimoniously described yield for both 305-d and 999-d lactations. The goal of that study was the accurate description of lactation curves for the whole population, rather than the prediction of daily or lactation yields for individual cows. However, curves describing the SD of M, F, and P were not examined, nor were curves calculated for SCS.

There are few reports in the literature of lactation curves for SCS. Rodriguez-Zas et al. (2000a) compared several nonlinear models for SCS and concluded that a function of Morant and Gnanasakthy (1989) best fit the data for 305-d lactations. In a related study, they

used nonlinear mixed effects models that included cow-specific lactation curves to study SCS (Rodriguez-Zas et al., 2000b). In the current study, several functions were fitted to TD SCS data to identify the most suitable model for those data.

The original BP program, AIPLDCR, used a mathematical function to calculate correlations among TD within and between traits to avoid repeated computation of large correlation matrices (Norman et al., 1999b). However, the functions used differed by parity in the number of terms used, making for sometimes-confusing explanations, and there were difficulties ensuring the positive-definiteness of the resulting correlation matrices. Simplified functions with more desirable computation properties were desired for the new BP program, BESTPRED.

The objectives of the current research were to develop curves for the SD of M, F, and P, as well as the mean and SD of SCS, for lactations longer than 305-d in each of the 6 major US dairy breeds, improve the modeling of TD correlations between and within traits, and validate BP of actual daily and 305-d lactation yields.

MATERIALS AND METHODS

Curves for Modeling Long Lactations

Milk, Fat, and Protein Yields. Breed- and lactation-specific curves were fitted for Ayrshire (**AY**), Brown Swiss (**BS**), Guernsey (**GU**), HO, Jersey (**JE**), and Milking Shorthorn (**MS**) cattle using an approach similar to that of Dematawewa et al. (2007). Test-day data of AY, BS, GU, JE, and MS cows calving between 1997 and 2003 were extracted from the national dairy database (**NDDB**) at AIPL. First through fifth lactations were used, the first TD was required to be ≤60 DIM, and cows were required to have at least 500 DIM in a lactation. The same data set as Dematawewa et al. (2007) was used for HO, and included TD data of 152,734 cows with calvings between 1997 and 2003 obtained from the NDDB. Summary statistics of M, F, and P yields and SCS standardized to 2 milkings per day, 305-d mature-equivalent basis for first and later parities of at least 60, 300, 500, and 800 DIM were calculated from that data set and are presented in Table 1.

The TD data were grouped into either 15-d (1 to 300 DIM) or 30-d (>300 DIM) intervals based on TD DIM. Means and SD of M, F, and P yields of the groups (Table 1) were modeled using Wood's formula (Wood, 1967) with the NLIN procedure of SAS/STAT software (SAS Institute, 2007):

$$Y_t = at^b e^{-ct},$$

where Y_t is the average daily yield or SD of M, F, or P (kg/d) in 15- or 30-d interval t, a is a scale parameter, and b and c are shape parameters. Separate curves were fitted for first-and later-parity cows. The data were regressed toward the mean as a result of grouping TD into intervals, but the shape parameters of the curves were the results of greatest interest, not the scale parameters.

SCS. Lactation curves for means and SD of SCS, which is $\log_2$-transformed SCC (Schutz et al., 1995), were also needed for BESTPRED. Test-day data for SCS were extracted from the NDDB subject to similar edits as M, F, and P, although cows were required to have least 800 DIM in a lactation. The TD data were grouped into 15- or 30-d intervals as described above for M, F, and P. Several functions were fitted to the data (Nelder, 1966; Wood, 1967; Wilmink, 1987; Morant and Gnanasakthy, 1989; Rook et al., 1993; Dijkstra et al., 1997), and a 4-parameter function (**C4**) described by Morant and Gnanasakthy (1989) provided the best fit as assessed by R^2:

$$Y_t = a - bt + \frac{ct^2}{2} - \frac{d}{t},$$

where Y_t is the average daily SCS or SD of SCS in 15- or 30-d interval t, and a, b, c, and d are curve parameters.

Modeling Correlations Among and Within Traits

Test-day data for HO cows were extracted from the NDDB and included lactations longer than 305 d. Records from first through fifth parities were used if lactation lengths were at least 250 d, records were made in a single herd, at least 5 tests were reported, and only twice-daily milking was reported. After edits, 171,970 first-parity and 176,153 later parity records were available. Correlations were calculated for 10-d intervals from 1 to 999 DIM by using an approach similar to that described by Norman et al. (1999b). Parameters obtained using HO data were assumed to be representative of other breeds as well and separate parameters were not estimated for each breed.

Correlations among TD yields were estimated using a simplified model that included an identity matrix (I) to model daily measurement error and an autoregressive matrix (E) to account for biological change. The matrix E is defined mathematically as $E_{ij} = r^{|i-j|}$, where i and j are TD DIM and $0 < r < 1$. The matrix of correlations within traits (B) was calculated as $B = b_1 I + b_2 E$, where the b_i are regression coefficients; separate functions were used to model the yield traits and SCS. Re-

Table 1. Number of observations (n), means, and standard deviations (SD) of 305-d mature-equivalent milk, fat, and protein yield and SCS of first- and later-parity Holstein cows with lactations of at least 60, 300, 500, and 800 DIM

Trait	Parity	DIM 60 n	60 Mean	60 SD	300 n	300 Mean	300 SD	500 n	500 Mean	500 SD	800 n	800 Mean	800 SD
Milk (kg)	1	303,577	11,434	2,017	286,412	11,493	2,009	30,264	12,023	1,920	746	12,147	2,008
	2+	188,900	10,833	2,093	174,452	10,912	2,084	14,086	11,500	2068	279	11,617	2,369
Fat (kg)	1	297,189	416	76	280,905	418	75	29,712	434	75	727	436	78
	2+	183,159	398	80	169,501	400	80	13,729	419	81	275	413	83
Protein (kg)	1	279,962	343	58	264,510	344	58	27,941	357	56	686	359	58
	2+	175,027	324	60	161,870	326	60	13,223	342	60	265	345	67
SCS	1	294,074	3.10	1.39	278,115	3.07	1.30	29,500	3.10	1.40	717	3.03	1.40
	2+	179,641	3.14	1.25	166,369	3.10	1.25	13,490	2.99	1.27	274	2.97	1.26

Table 2. Regression coefficients used to weight identity and autoregressive matrices for calculating correlations among test-days for milk, fat, and protein (MFP) and SCS for first and later parities

Parity	Trait	Regression coefficient	
		b_1	b_2
1	MFP	0.214	0.786
	SCS	0.199	0.801
2+	MFP	0.132	0.868
	SCS	0.199	0.801

Table 3. Correlations (matrix T) among yield traits and SCS

Trait	Milk	Fat	Protein	SCS
Milk	1.00	0.67	0.85	−0.08
Fat		1.00	0.77	−0.14
Protein			1.00	−0.10
SCS				1.00

gression coefficients were estimated separately for each trait and parity, and are presented in Table 2.

Suppose that T is a 4×4 matrix of phenotypic correlations among traits partitioned as

$$T = \begin{bmatrix} t_{MM} & t_{MF} & t_{MP} & t_{MS} \\ t_{FM} & t_{FF} & t_{FP} & t_{FS} \\ t_{PM} & t_{PF} & t_{PP} & t_{PS} \\ t_{SM} & t_{SF} & t_{SP} & t_{SS} \end{bmatrix} = \begin{bmatrix} T_m & T_{ms} \\ T_{sm} & T_{ss} \end{bmatrix},$$

where t_{ij} is the phenotypic correlation of trait i with trait j; the same correlations are used in BESTPRED for all parities (Table 3).The complete correlation matrix (C) then can be obtained as

$$C = \begin{bmatrix} M \otimes T_m & M^{\frac{1}{2}} \cdot S^{\frac{1}{2}} \otimes T_{ms} \\ S^{\frac{1}{2}} \cdot M^{\frac{1}{2}} \otimes T_{sm} & S \otimes T_s \end{bmatrix},$$

where $C_{ik,jl} =$ correlation of trait i at DIM k with trait j at DIM l, $\otimes$ is the Kronecker (direct) product operator, and $\cdot$ denotes element-wise matrix multiplication. The matrices M and S denote the functions used to calculate correlations among M, F, and P and SCS, respectively. The matrices T and C may be calculated separately for first and later parities, but currently only C differs by lactation.

Validation

Lactation Yields. Test-day records for each of the major US dairy breeds were extracted from the NDDB subject to the restrictions that each record took place in a single herd and included 20 or fewer TD. More rigorous edits were not applied because the goal of this analysis was to validate the BESTPRED programs against all lactation data in the NDDB, not just a highly selected subset of those data. All lactations meeting those criteria were used for the AY, BS, GU, and MS breeds, but

random samples of 1,000,000 records were taken from the HO and JE breeds because of the large number of available records. Cow identification numbers were sampled at random without replacement and all lactations for a cow meeting above restrictions were used. Product-moment correlations for 305-d yields computed by AIPLDCR and BESTPRED using multiple-trait methods were calculated within breed and parity group and used to compare the 2 programs.

Daily Milk Yields. Daily milk weight data (7-d averages from electronic meters) of Holstein cows were obtained from the University of Florida (Gainesville), Pennsylvania State University (University Park), and Virginia Polytechnic Institute and State University (Blacksburg) for validating the estimates of daily yield provided by BP. The on-farm data were used to calculate 305-d yields by summing the individual daily milk yields. Corresponding TD data were extracted from the NDDB and processed with BESTPRED to obtain BP of actual 305-d yields and yields for each day of lactation. Product-moment correlations were calculated for the 305-d actual and individual daily yields.

Accuracy of Prediction. Accuracy was assessed by comparing predictions of 305-d yields and SCS using TD data from various points in lactation with those calculated using all available TD data. Subsets of the Holstein data set used for fitting lactation curves were created that contained only TD data occurring on or before 50, 100, 150, 200, and 250 DIM and calculating 305-d yields and SCS from those data. Correlations were calculated between those estimates and 305-d yield calculated using all available TD up to 500 DIM for first and later lactations.

RESULTS

Curves for Modeling Long Lactations

Summary statistics including peak yield and DIM at peak are presented in Table 4. Lactations from all breeds were originally required to be at least 500 DIM for M, F, and P yields and 800 DIM for SCS to be used for modeling lactation curves. Dematawewa et al. (2007) required that lactations be at least 800 DIM for M, F, and P yield, but their interest was specifically in modeling lactation curves for cows with very long

lactations. The lactation curves used for BP need to accommodate cows with long lactations in a reasonable manner, but it is more important that the first 365 d of lactation be modeled with great precision than, say, d 600 to 700. The summary statistics in Table 1 show that cows with 500- and 800-d lactations have similar means and SD of mature-equivalent yields for both parity groups, but the number of records differs dramatically when moving from 500- to 800-d lactations. To overcome concerns about properly modeling lactation curves whose shapes are representative of the

majority of the data, the lactation curves for M, F, and P were refitted using a maximum lactation length of 500 DIM. Using lactations with a maximum of 500 DIM represents a reasonable compromise between cows milking only to 305 d and those milking longer; across breeds only about 5% of cows have lactations longer than 500 DIM.

Standard lactation curves for mean M yield and SCS for lactations up to 500 DIM are shown in Figures 1 and 2, and the parameters of those curves are presented in Table 5. Lactation curves for F and P yields (data

Table 4. Number of records, average lactation length, average number of test-day (TD) observations, average peak milk yields, and peak DIM for milk, fat, and protein yield and SCS for first and later parities[1]

Breed[2]	Trait	Parity	Lactation records	Lactation length (d)	TD records	Peak yield	Peak DIM (d)
AY	Milk (kg)	1	234,260	351	10	25.79	66.40
		2+	510,374	339	10	32.20	51.07
	Fat (kg)	1	234,260	351	10	0.98	71.18
		2+	510,374	339	10	1.23	52.07
	Protein (kg)	1	234,260	351	10	0.74	71.41
		2+	510,374	339	10	0.90	50.00
	SCS	1	24,206	579	16	2.96	107.41
		2+	28,851	585	16	3.54	129.42
BS	Milk (kg)	1	579,033	359	11	26.88	77.50
		2+	1,041,330	357	11	34.70	55.65
	Fat (kg)	1	579,033	359	11	1.07	80.47
		2+	1,041,330	357	11	1.40	58.40
	Protein (kg)	1	579,033	359	11	0.81	86.32
		2+	1,041,330	357	11	1.03	61.00
	SCS	1	109,566	614	17	3.64	131.53
		2+	184,749	616	17	4.20	139.51
GU	Milk (kg)	1	355,453	357	11	24.70	63.14
		2+	589,856	358	11	30.13	48.69
	Fat (kg)	1	355,453	357	11	1.01	78.72
		2+	589,856	358	11	1.26	58.12
	Protein (kg)	1	355,453	357	11	0.72	74.12
		2+	589,856	358	11	0.88	52.57
	SCS	1	65,382	620	18	4.07	128.10
		2+	105,797	621	18	4.16	143.10
HO	Milk (kg)	1	7,173,385	353	11	33.40	82.51
		2+	3,830,540	355	11	42.88	58.02
	Fat (kg)	1	7,173,385	353	11	1.20	75.93
		2+	3,830,540	355	11	1.54	55.55
	Protein (kg)	1	7,173,385	353	11	0.86	80.96
		2+	3,830,540	355	11	1.11	55.62
	SCS	1	1,356,070	612	18	3.83	125.08
		2+	551,445	600	17	4.11	129.20
JE	Milk (kg)	1	4,053,165	338	10	23.50	74.73
		2+	7,360,015	336	10	29.94	56.29
	Fat (kg)	1	4,053,165	338	10	1.04	81.26
		2+	7,360,015	336	10	1.33	63.02
	Protein (kg)	1	4,053,165	338	10	0.74	80.88
		2+	7,360,015	336	10	0.96	57.65
	SCS	1	435,855	610	18	4.03	123.67
		2+	602,592	602	17	4.19	134.99
MS	Milk (kg)	1	107,666	338	10	26.23	70.13
		2+	159,827	328	10	33.84	49.83
	Fat (kg)	1	107,666	338	10	0.94	68.58
		2+	159,827	328	10	1.22	48.77
	Protein (kg)	1	107,666	338	10	0.54	51.61
		2+	159,827	328	10	0.80	41.70
	SCS	1	8,534	611	18	3.74	139.82
		2+	5,835	592	17	3.68	139.62

[1]Data from 500-d lactations were used for milk, fat, and protein yields; data from 800-d lactations were used for SCS.

[2]AY = Ayrshire; BS = Brown Swiss; GU = Guernsey; HO = Holstein; JE = Jersey; MS = Milking Shorthorn.
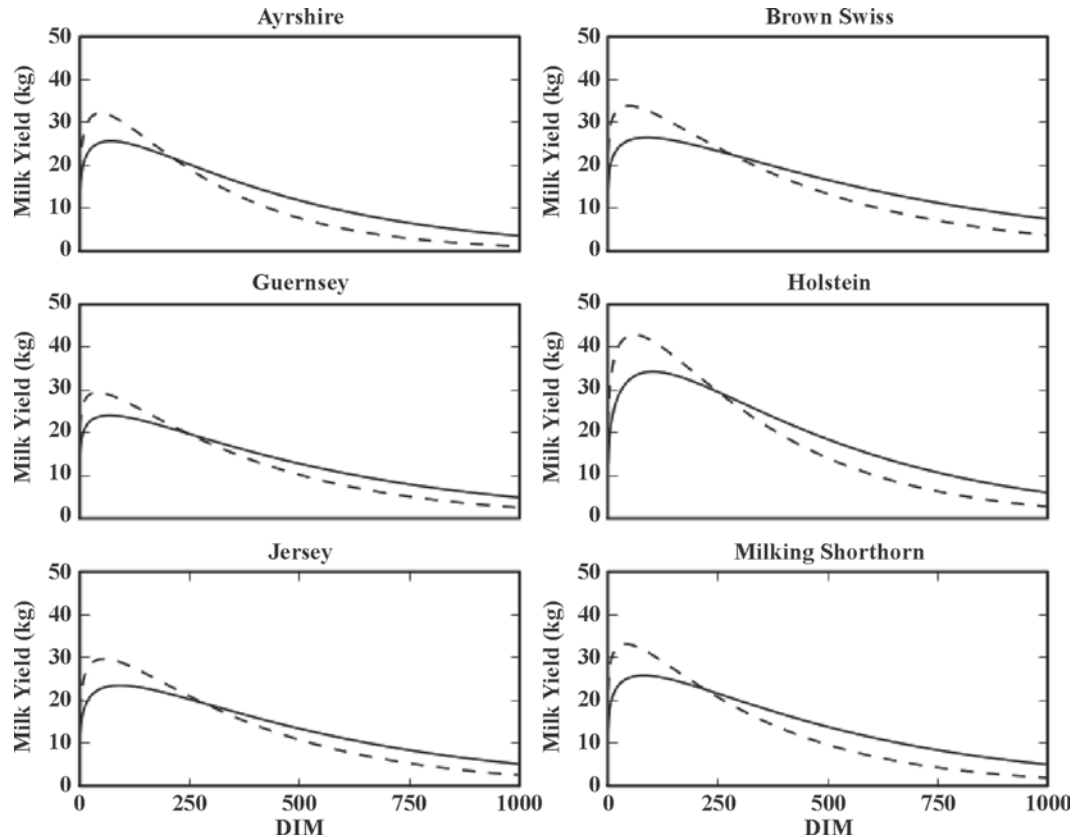
Figure 1. Lactation curves for mean milk yield (kg) of Ayrshire, Brown Swiss, Guernsey, Jersey, Holstein, and Milking Shorthorn dairy cattle for first (solid line) and later (broken line) parity groups.

not shown) were similar to those for M yield. In all breeds, first-parity cows have lower peak yields and flatter lactation curves on average than later parity cows, and are more persistent. Later parity cows consistently outperform first-parity cows for M, F, and P on a daily yield basis until late in lactation.

As in the paper of Rodriguez-Zas et al. (2000b), the Morant and Gnanasakthy (1989) model overestimated mean SCS early and late in lactation and underestimated it in the middle of lactation. However, the other models (Nelder, 1966; Wood, 1967; Wilmink, 1987; Rook et al., 1993; Dijkstra et al., 1997) provided much poorer fits to the data, with plots of the resulting curves asymptoting to positive infinity as DIM increased. Similar behavior was also observed with the Morant and Gnanasakthy (1989) model when 500-d lactations were used. Somatic cell scores for cows with 800-d lactations reached a plateau later in lactation and produced curves that were more biologically meaningful, although the curves may not represent how average cows would perform if milked to 800 DIM. If only 305-d lactations were being modeled, Wood's curves would probably provide reasonable fits,

and it might be desirable to use the same function for modeling all lactation curves. However, when Wood's formula was fit to SCS data for longer lactations, the resulting curves did not fit the data, increasing linearly beyond 305 d rather than reaching a plateau, unlike the Morant and Gnanasakthy function.

First-parity cows had consistently lower SCS than older cows for all breeds except MS, where the shape of the mature cow curve is clearly affected by a few extreme observations, although the differences were fairly small. Heroic efforts to construct a more reasonable curve for mature MS cows were deemed unnecessary because fewer than 3% of them have lactations longer than 500 d, whereas 90% have lactations less than 400 DIM, where the standard curves appear to be quite reasonable. The curves for HO and JE cows were similar to those published by Schutz et al. (1995), although our curves for first-parity cows did not cross the mature cow curves near 305 d as did theirs. Reents et al. (1995) found large differences between first- and later parity lactation curves for Canadian Holsteins, as was the case with US Holsteins in this study, but they found that
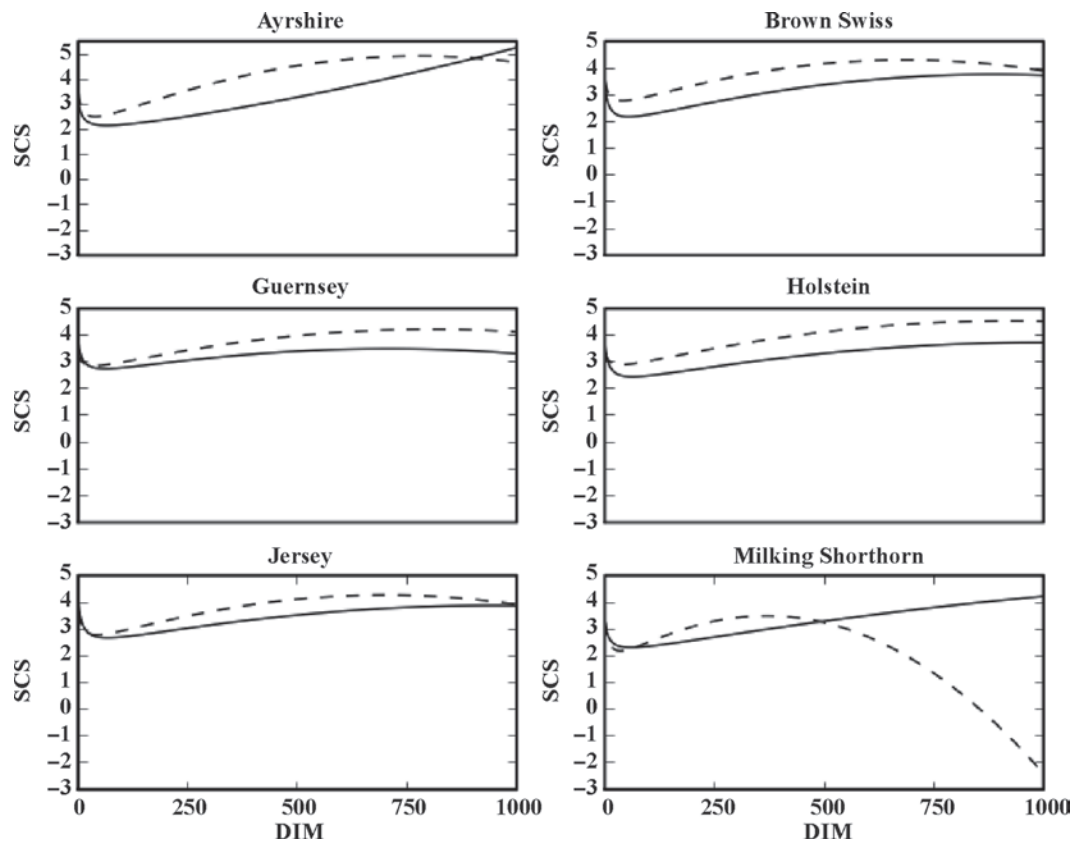
Figure 2. Lactation curves for mean SCS of Ayrshire, Brown Swiss, Guernsey, Jersey, Holstein, and Milking Shorthorn dairy cattle for first (solid line) and later (broken line) parity groups.

the older cows had lower SCS than first-parity cows until about 100 DIM. These differences underscore the challenge of correctly modeling SCS curves.

Lactation curves for the SD of M yield and SCS are shown in Figures 3 and 4, and the parameters of those curves are presented in Table 6. Lactation curves for the SD of F and P yields (data not shown) were similar to those for the SD of M yield. The SD curves are used in the calculation of covariances among TD, covariances among TD and lactation yields, and persistency of yield. Standard deviations for M, F, and P were larger for mature cows than for first-lactation cows at the beginning of lactation, generally converged in the middle of lactation (~500 DIM), and diverged beginning in the middle of lactation. Milking Shorthorns were an exception to that pattern, with very similar curves for SD of M and F, but widely divergent curves for P. These atypical results may be due to the relatively low number of records available for curve-fitting. The SD for SCS were higher for mature cows than for first-lactation cows at the beginning of lactation, generally converged in the middle of lactation (~500 DIM), and diverged

to varying degrees at the end of lactation. This is in agreement with the results of Gengler et al. (2004) and DeGroot et al. (2007), who found that phenotypic variances for HO M, F, and P were largest at the start and end of 305-d lactations for first through third parities. The latter study reported the same results for SCS.

Figure 5 shows mean yields, mean yields ± 1 SD, and mean yields ± 2 SD for M, F, and P yields and SCS in first-parity Holstein cows. The SD do not show large increases near the beginning and end of lactation as is sometimes the case in random regression TD models (Kachman, 2004), although in some cases, most notably SCS, the mean − 2 SD curves for SCS in both heifers and cows (data not shown) are below 0 for most of the lactation. Negative values make no biological sense for M, F, and P, and represent possible but implausible values for SCS, although they reflect the variability present in the 15- and 30-d groups used in fitting the lactation curves to the SD. A more rigorous set of data quality edits might result in groups having lower variation at the risk of introducing bias into the mean and SD curves by estimating them from a group of animals

Table 5. Parameter estimates of the Wood's curve for mean milk, fat, and protein yield, and of Morant and Gnanasakthy's curve C4 for mean SCS[1]

Breed[2]	Parity	Trait	Curve[3,4]	Parameter[5]			
				a	b	c	d
AY	1	Milk (kg)	W	13.796	0.19040	0.003	—
		Fat (kg)	W	0.650	0.122	0.002	—
		Protein (kg)	W	0.513	0.124	0.002	—
		SCS	M&G	1.791	−0.002	0.000	14.205
	2+	Milk (kg)	W	13.796	0.190	0.003	—
		Fat (kg)	W	0.650	0.122	0.002	—
		Protein (kg)	W	0.513	0.124	0.002	—
		SCS	M&G	1.772	−0.008	0.000	17.830
BS	1	Milk (kg)	W	15.062	0.162	0.002	—
		Fat (kg)	W	0.691	0.117	0.001	—
		Protein (kg)	W	0.485	0.152	0.001	—
		SCS	M&G	1.613	−0.005	0.000	17.813
	2+	Milk (kg)	W	23.382	0.130	0.003	—
		Fat (kg)	W	1.189	0.056	0.002	—
		Protein (kg)	W	0.822	0.096	0.002	—
		SCS	M&G	2.223	−0.006	0.000	13.581
GU	1	Milk (kg)	W	15.037	0.146	0.002	—
		Fat (kg)	W	0.603	0.141	0.001	—
		Protein (kg)	W	0.497	0.120	0.001	—
		SCS	M&G	2.330	−0.003	0.000	14.803
	2+	Milk (kg)	W	21.281	0.120	0.003	—
		Fat (kg)	W	0.962	0.091	0.002	—
		Protein (kg)	W	0.788	0.068	0.002	—
		SCS	M&G	2.471	−0.004	0.000	8.808
HO	1	Milk (kg)	W	13.010	0.267	0.003	—
		Fat (kg)	W	0.784	0.120	0.001	—
		Protein (kg)	W	0.463	0.203	0.002	—
		SCS	M&G	1.980	−0.003	0.000	16.883
	2+	Milk (kg)	W	22.009	0.216	0.004	—
		Fat (kg)	W	1.287	0.073	0.002	—
		Protein (kg)	W	0.854	0.132	0.002	—
		SCS	M&G	2.507	−0.004	0.000	8.980
JE	1	Milk (kg)	W	11.534	0.202	0.002	—
		Fat (kg)	W	0.503	0.195	0.002	—
		Protein (kg)	W	0.410	0.180	0.002	—
		SCS	M&G	2.159	−0.004	0.000	19.653
	2+	Milk (kg)	W	17.399	0.177	0.003	—
		Fat (kg)	W	0.790	0.163	0.003	—
		Protein (kg)	W	0.715	0.123	0.002	—
		SCS	M&G	2.213	−0.005	0.000	14.925
MS	1	Milk (kg)	W	13.625	0.188	0.002	—
		Fat (kg)	W	0.688	0.081	0.001	—
		Protein (kg)	W	0.539	0.091	0.001	—
		SCS	M&G	1.899	−0.003	0.000	14.152
	2+	Milk (kg)	W	22.425	0.143	0.004	—
		Fat (kg)	W	1.213	0.015	0.002	—
		Protein (kg)	W	0.863	0.055	0.002	—
		SCS	M&G	1.238	−0.012	0.000	20.772

[1]Holstein curve parameters were reported in Dematawewa et al. (2007).

[2]AY = Ayrshire, BS = Brown Swiss, GU = Guernsey, HO = Holstein, JE = Jersey, MS = Milking Shorthorn.

[3]W = Wood's (1967) curve: $Y_t = at^b e^{-ct}$, where Y_t is the average daily SD of milk, fat, or protein (kg/d) in 15- or 30-d interval t, and a is a scale parameter, and b and c are shape parameters.

[4]M&G = Morant and Gnanasakthy's (1989) curve C4: $Y_t = a - bt + \dfrac{ct^2}{2} - \dfrac{d}{t}$, where Y_t is the average daily SD of SCS in 15- or 30-d interval t, and a, b, c, and d are curve parameters.

[5]Wood's curve is a 3-parameter function, whereas Morant and Gnanasakthy's curve C4 is a 4-parameter function.

that is not representative of the average cow in the milking population. The key point illustrated in Figure 5 is that variation remains fairly constant across the lactation. Variation decreased slightly late in lactation for mature cows, but they were slightly more variable than heifers early in lactation, as expected.
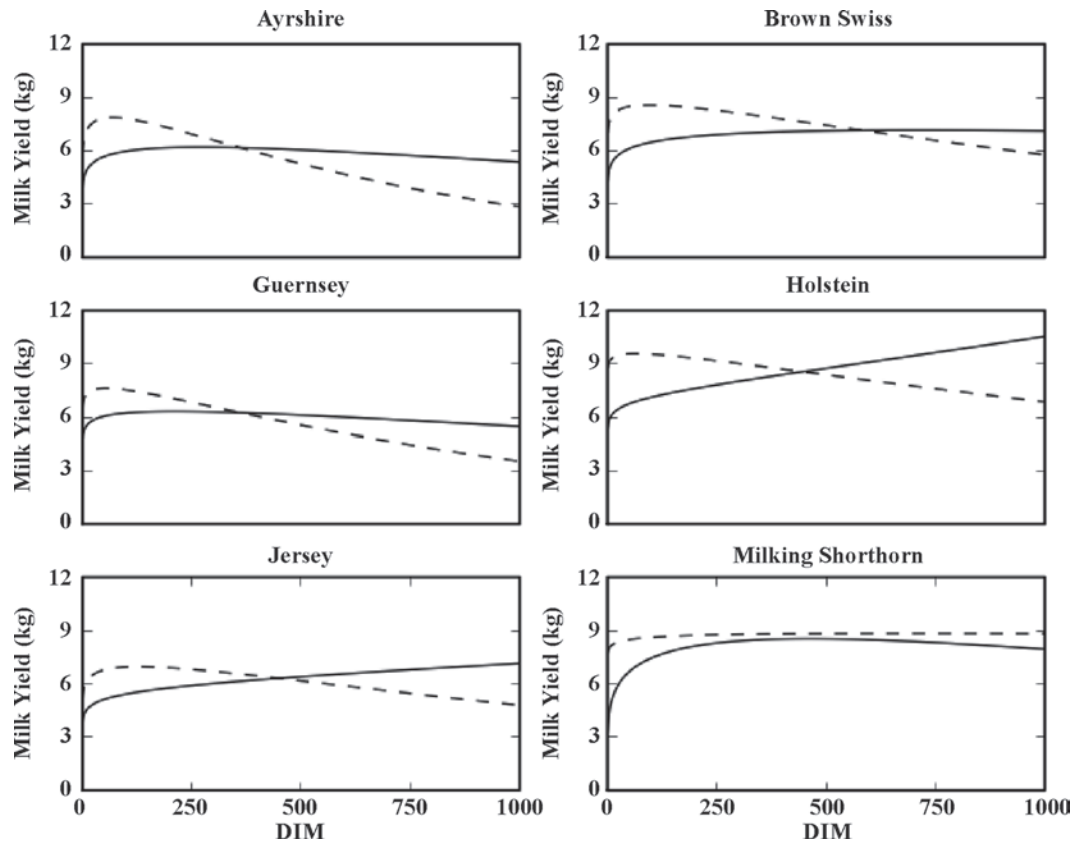
Figure 3. Lactation curves for the standard deviation of milk yield (kg) of Ayrshire, Brown Swiss, Guernsey, Jersey, Holstein, and Milking Shorthorn dairy cattle for first (solid line) and later (broken line) parity groups.

Modeling Correlations Among and Within Traits

The autoregressive coefficients, r, for M, F, P, and SCS were estimated separately for first and later parities. Values of r were 0.998 and 0.997 for M, F, and P in first and later parities, respectively, and 0.998 for SCS in both parity groups. The values for M, F, and P were slightly larger than previous estimates (Norman et al., 1999b) due to the inclusion of the identity matrix. Parameters were not previously calculated separately for SCS. Yields of M, F, and P on subsequent DIM are more similar in first than in later parities, but the differences are small. Somatic cell scores on adjacent DIM are identical across parities. Data files were checked to ensure no errors were made, and although parameter estimates were identical to 3 decimal places, the correct data were used in each analysis. Twelve to 20% of the variation in correlations between TD was due to daily measurement error and 73 to 80% was due to biological differences among animals.

The advantages of the models used to calculate correlations between traits (Table 2) over those previously used are those of parsimony and improved modeling of correlations of M, F, and P with SCS (Table 3). The model of the correlation matrix C presented by Norman et al. (1999b) included 4 terms for first and later parities, and factors differed between parities. The current model includes only an autoregressive matrix that accounts for biological changes over the course of lactation and an identity matrix that accounts for daily measurement error. Their model also assumed zero correlations between SCS and the other yield traits, which was a deliberate decision to allow daily SCS to fluctuate independently from the other traits. Goodness-of-fit for the original and new equations as assessed by the R^2 statistic were very similar (data not shown).

Validation

Lactation Yields. Product-moment correlations among 305-d yields for M, F, and P yields and SCS calculated using the AIPLDCR and BESTPRED programs are presented in Table 7 by breed and parity. Correlations for M, F, and P were greater for first than later parities in most cases, and correlations for SCS were larger for first parity. The lower correlations for
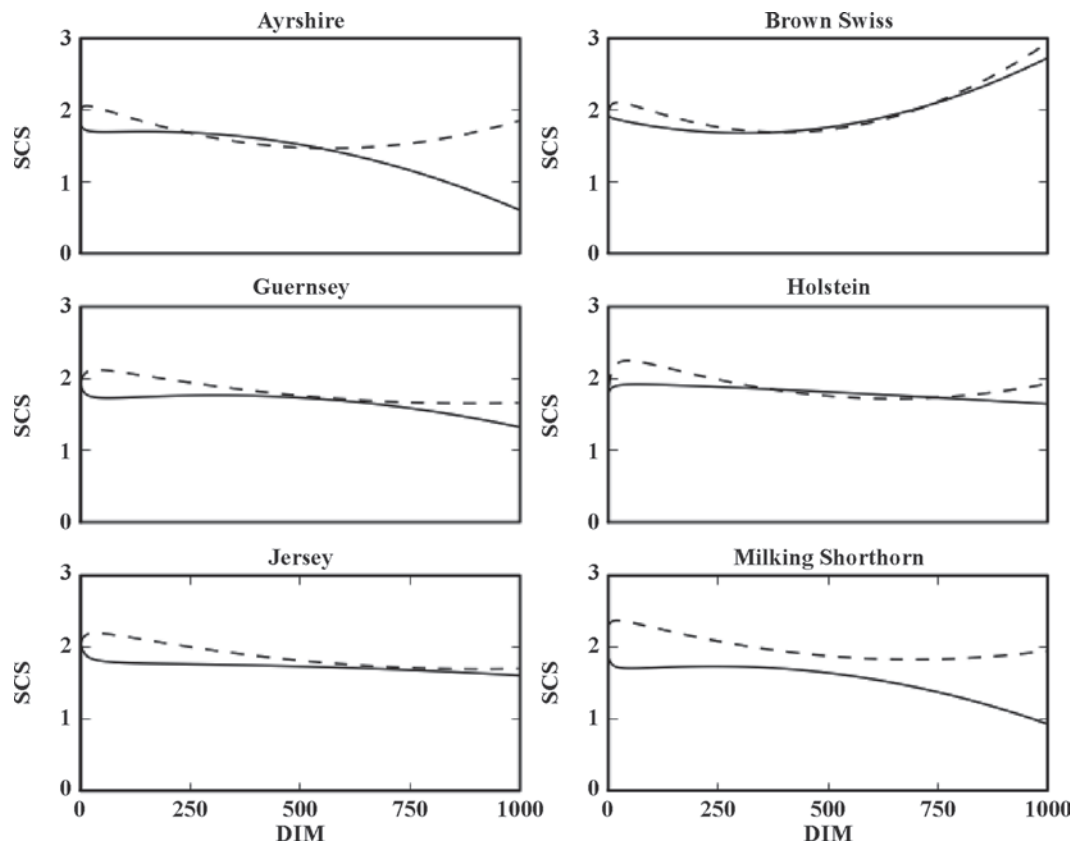
Figure 4. Lactation curves for the standard deviation of SCS of Ayrshire, Brown Swiss, Guernsey, Jersey, Holstein, and Milking Shorthorn dairy cattle for first (solid line) and later (broken line) parity groups.

later parity records may reflect heterogeneity among lactation curves for older animals that is not well modeled by grouping animals into first and later parities. Dematawewa et al. (2007) fitted curves for first, second, and third-and-later parity Holstein cows and found that the second-parity curves had shapes intermediate to the first and third-and-later groups, concluding that the third-and-later parity curve could be used for second-parity cows as well. That approach was also adopted for the other 5 breeds in this study; gains in accuracy from modeling later lactations separately are probably very small, and some of the breeds have small enough populations that it could be difficult to accurately calculate lactation curves for later parities individually.

Somatic cell score was the trait with the lowest correlation between AIPLDCR and BESTPRED results, which corresponds to the changes made to the SCS calculation in BESTPRED, most notably the calculation of correlations among TD. Records showing the largest difference between the 2 programs for each trait were examined (data not shown) and no systematic source of variation could be identified. Correlations were also calculated using linear-interpolated standard curves, rather than the smooth Wood's and Morant and

Gnanasakthy's curves, to determine how much of the differences could be explained by the new curves, and similar results to those in Table 7 were obtained (data not shown). Those results suggest that the updated autoregressive parameters and changes to modeling of correlations among TD within lactations had substantial effects on the lactation yields.

Daily Milk Yields. Correlations of 305-d milk yields and the best and poorest correlations of individual daily data using records from the university herds are presented in Table 8. Correlations ranged from 0.968 to 0.997 for first and 0.952 to 0.993 for later parities. These results are quite encouraging given that the daily weights were actually 7-d averages rather than true daily yields, and that the daily yields were affected by factors such as disease status (e.g., presence or absence of mastitis), climate, and other factors that are accounted for in the standard lactation curves only to the degree that they affect large numbers of cows whose records are stored in the NDDB. It is also possible that research being conducted in those herds affected daily yields in a way that is not representative of commercial herds. Finally, it is also worth considering that it can be extremely difficult to account for effects with durations

Table 6. Parameter estimates of the Wood's curve for standard deviations (SD) of milk, fat, and protein yield, and of Morant and Gnanasakthy's C4 curve for SD of SCS

Breed[1]	Parity	Trait	Curve[2,3]	Parameter[4]			
				a	b	c	d
AY	1	Milk (kg)	W	3.874	0.103	0.000	—
		Fat (kg)	W	0.226	0.034	0.000	—
		Protein (kg)	W	0.123	0.088	0.000	—
		SCS	M&G	1.644	−0.001	0.000	1.476
	2+	Milk (kg)	W	5.825	0.094	0.001	—
		Fat (kg)	W	0.369	0.012	0.001	—
		Protein (kg)	W	0.223	0.022	0.000	—
		SCS	M&G	2.174	0.002	0.000	−1.392
BS	1	Milk (kg)	W	4.331	0.091	0.000	—
		Fat (kg)	W	0.299	0.003	0.000	—
		Protein (kg)	W	0.131	0.098	0.000	—
		SCS	M&G	1.903	0.001	0.000	0.261
	2+	Milk (kg)	W	7.029	0.056	0.001	—
		Fat (kg)	W	0.495	−0.025	0.001	—
		Protein (kg)	W	0.242	0.028	0.000	—
		SCS	M&G	2.281	0.003	0.000	−2.707
GU	1	Milk (kg)	W	4.757	0.065	0.000	—
		Fat (kg)	W	0.327	−0.055	−0.001	—
		Protein (kg)	W	0.157	0.035	0.000	—
		SCS	M&G	1.647	−0.001	0.000	2.678
	2+	Milk (kg)	W	6.511	0.053	0.000	—
		Fat (kg)	W	0.402	−0.022	0.000	—
		Protein (kg)	W	0.250	−0.013	0.000	—
		SCS	M&G	2.243	0.001	0.000	−3.015
HO	1	Milk (kg)	W	5.381	0.054	0.000	—
		Fat (kg)	W	0.380	−0.061	−0.001	—
		Protein (kg)	W	0.180	0.017	−0.001	—
		SCS	M&G	1.955	0.000	−0.000	−1.515
	2+	Milk (kg)	W	8.755	0.028	0.000	—
		Fat (kg)	W	0.554	−0.053	0.000	—
		Protein (kg)	W	0.327	−0.036	0.000	—
		SCS	M&G	2.485	0.002	0.000	−6.391
JE	1	Milk (kg)	W	3.802	0.074	0.000	—
		Fat (kg)	W	0.225	0.045	0.000	—
		Protein (kg)	W	0.126	0.076	0.000	—
		SCS	M&G	1.752	0.000	0.000	3.132
	2+	Milk (kg)	W	5.114	0.080	0.001	—
		Fat (kg)	W	0.291	0.069	0.001	—
		Protein (kg)	W	0.203	0.040	0.000	—
		SCS	M&G	2.313	0.001	0.000	−2.561
MS	1	Milk (kg)	W	3.326	0.183	0.000	—
		Fat (kg)	W	0.318	−0.045	−0.001	—
		Protein (kg)	W	0.141	0.058	0.000	—
		SCS	M&G	1.628	−0.001	−0.000	2.224
	2+	Milk (kg)	W	7.733	0.024	0.000	—
		Fat (kg)	W	0.562	−0.124	−0.001	—
		Protein (kg)	W	0.349	−0.130	−0.001	—
		SCS	M&G	2.480	0.002	0.000	−1.593

[1]AY = Ayrshire, BS = Brown Swiss, GU = Guernsey, HO = Holstein, JE = Jersey, MS = Milking Shorthorn.

[2]W = Wood's (1967) curve: $Y_t = at^b e^{-ct}$, where Y_t is the average daily SD of milk, fat, or protein (kg/d) in 15- or 30-d interval t, and a is a scale parameter, and b and c are shape parameters.

[3]M&G = Morant and Gnanasakthy's (1989) curve C4: $Y_t = a - bt + \frac{ct^2}{2} - \frac{d}{t}$, where Y_t is the average daily SD of SCS in 15- or 30-d interval t, and a, b, c, and d are curve parameters.

[4]Wood's curve is a 3-parameter function, whereas Morant and Gnanasakthy's curve C4 is a 4-parameter function.

shorter than the interval between TD (Bohmanova et al., 2005), resulting in lower correlations between calculations based on TD data and daily data.

The largest and smallest individual daily correlations and corresponding DIM are also presented in Table 8.

In the Florida and Virginia data the highest correlations between actual daily yields and BP of yields were found in later lactation (increasing DIM), whereas the highest correlations were found near peak yield in the Pennsylvania data. Correlations were lowest at the beginning of

Table 7. Product-moment correlations between 305-d yields computed using the AIPLDCR and BESTPRED programs

Breed[1]	Parity	n	Trait			
			Milk	Fat	Protein	SCS
AY	1	55,936	0.996	0.997	0.995	0.991
	2+	137,817	0.993	0.992	0.991	0.989
BS	1	120,396	0.997	0.996	0.995	0.990
	2+	255,867	0.994	0.988	0.990	0.986
GU	1	94,483	0.996	0.996	0.996	0.992
	2+	184,055	0.993	0.992	0.993	0.991
HO	1	351,031	0.995	0.996	0.993	0.992
	2+	648,810	0.993	0.986	0.985	0.990
JE	1	356,258	0.997	0.993	0.994	0.993
	2+	763,461	0.992	0.986	0.989	0.992
MS	1	24,316	0.995	0.997	0.987	0.989
	2+	47,469	0.994	0.995	0.982	0.983

[1]AY = Ayrshire; BS = Brown Swiss; GU = Guernsey; HO = Holstein; JE = Jersey; MS = Milking Shorthorn.

lactation in all cases, which is expected as cows transition from the dry period into lactation. The minimum correlations increased considerably when data before 5 DIM were excluded, ranging from 0.28 to 0.66 (data not shown). Health data were not obtained from the universities, and it also is possible that the extremely low correlations seen at the beginning of the lactation are due to the presence of animals that experienced dystocia or metabolic problems (e.g., hypocalcemia) at calving, and such effects are not accounted for by BP.
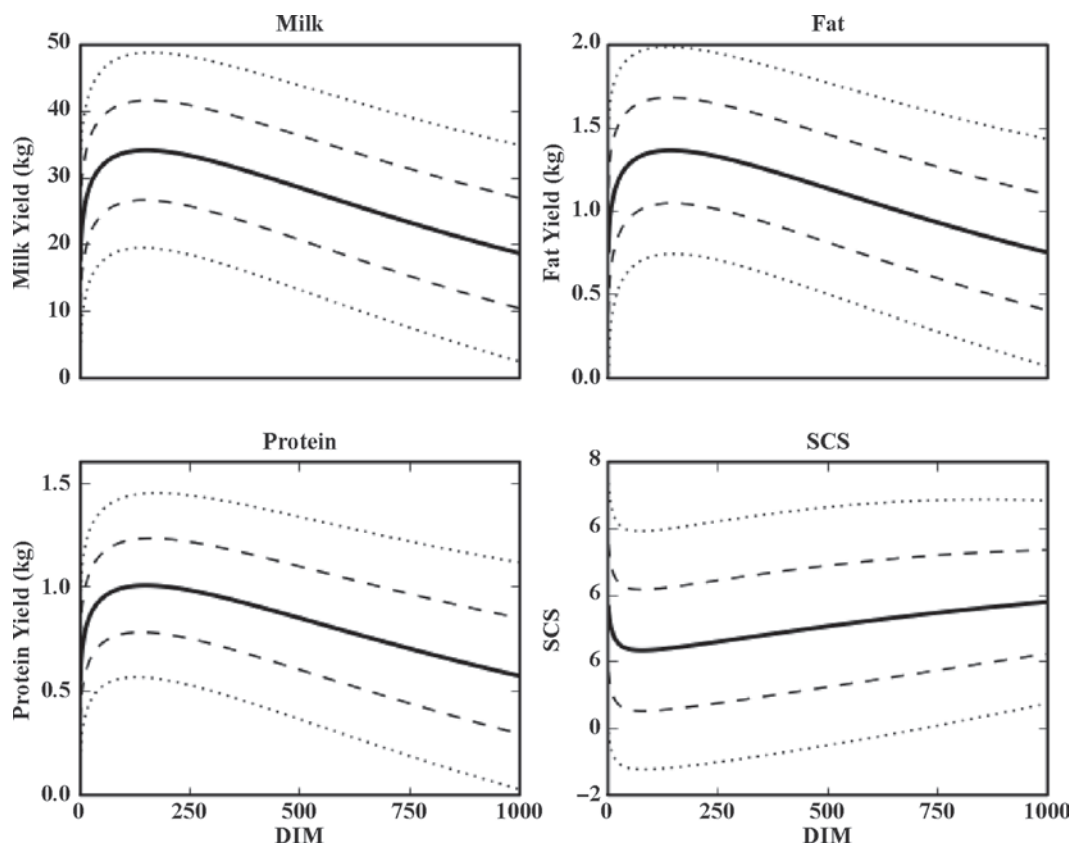

Figure 5. Lactation curves for milk, fat, and protein yield and SCS in first-parity Holstein cows showing the mean (solid line), the mean ± 1 standard deviation (SD) (broken lines), and the mean ± 2 SD (dotted lines).

Table 8. Product-moment correlations between 305-d yields (r_{305}) and the highest and lowest correlations (r_{high} and r_{low}) and corresponding DIM between best predictions of daily yields and actual daily yields recorded on-farm

Data set	Parity	305-d yield		Individual daily yields					
					Highest			Lowest	
		n	r_{305}	n	DIM	r_{high}	DIM	r_{low}
Florida	1	143	0.900	423	282	0.844	1	0.149
	2+	218	0.952	490	271	0.862	1	0.165
Pennsylvania	1	59	0.955	212	58	0.909	1	0.176
	2+	117	0.932	416	222	0.887	1	0.049
Virginia	1	33	0.988	208	249	0.948	1	0.015
	2+	68	0.976	391	218	0.954	2	−0.117

It is also possible that the Wood's curves are not fitting the data well very early in lactation.

In a recent validation study, Quist et al. (2007) compared estimates of actual daily and 305-d yields from the Canadian MTP system, as well as estimates of 305-d yield with the 305-d sum of daily yields, using data collected from on-farm milk meters. They found that concordance correlations increase with DIM, and that MTP overestimated lactation yields in every parity, although estimates of yield improved as TD data accumulated. They reported the best agreement between MTP estimates and actual daily yields in late lactation, which is similar to results from the Florida and Virginia data. Best prediction and MTP both use all available TD data, so it is not surprising that both methods provide better estimates in later versus early lactation.

Accuracy of Prediction. Correlations among 305-d yields and SCS calculated using different subsets of HO data are presented in Table 9. As expected, correlations of the earlier subsets with yield calculated using all TD increased as DIM and the number of available TD increased. Results for SCS suggest that the curves do not provide a good fit in early lactation, but do a much better job by 100 DIM when many cows will have had 2 or 3 tests.

DISCUSSION

The genetic merit of individual cows is not currently accounted for in BP. This is a deliberate decision made so that the BP routines can be used by other stakeholders in the US dairy industry such as the dairy records processing centers that do not have access to all of the data stored in the NDDB that are needed to compute genetic evaluations. The inclusion of genetic merit in BP might improve the predictions provided by BP, but the validation component of this study (Table 8) suggests that the magnitude of any improvements would be small and would probably not offset the additional complexity and lack of portability of the resulting programs.

Lactation curves were calculated using data from cows that milked at least 500 d, which may have produced curves biased in favor of animals with long lactations that have different shapes than those of 305-d lactations. Restricted maximum likelihood procedures (Patterson and Thompson, 1971) can account for bias such as that associated with selection when the data used to make selection decision are included in the analysis, and it is possible that a REML-type curve fitting procedure could produce curves unaffected by selection bias. Calculation of curves using only 305-d lactations may result in the opposite case, in which curves poorly describe longer lactations. It is also possible that a data sampling strategy can be devised that will provide the optimum balance between cows with typical lactations and those with long lactations. For purposes of genetic evaluation, any lactation curve biases will probably be accounted for by herd-year-season effects and will not adversely affect bull proofs. However, biases could have detrimental effects on calculations such as optimum calving interval that depend on estimates of daily yield. The same data were used in the current study and by Dematawewa et al. (2007), who found only small differences between curve parameters when fitting Wood's curves to 305-d versus 999-d lactations. This suggests that significant bias was not introduced by using curves fitted to lactations longer than 305 d.

To obtain lactation curves for SCS that were believed to be biologically plausible, data from lactations of at least 800 d were used for curve-fitting, but there is concern that those data may not be representative of the average cow in a typical lactation. It is unlikely that cows with chronically poor udder health, characterized by persistently high SCS, would remain in the herd that long. This suggests that improved functions for modeling SCS can be identified, perhaps through the use of splines (White et al., 1999) rather than parametric functions (e.g., Morant and Gnanasakthy, 1989). However, validation of BESTPRED results against AIPLDCR using different lactation curves showed that

Table 9. Correlations of 305-d milk, fat, and protein yields and SCS calculated using all available test-days (TD) and 305-d yield and SCS calculated using only TD occurring on or before the indicated DIM

| Trait | Parity | DIM | | | | |
		50	100	150	200	250
Milk (kg)	1	0.885	0.935	0.967	0.984	0.995
	2+	0.913	0.942	0.966	0.983	0.995
Fat (kg)	1	0.890	0.930	0.961	0.981	0.993
	2+	0.892	0.927	0.957	0.980	0.994
Protein (kg)	1	0.821	0.881	0.932	0.965	0.988
	2+	0.824	0.875	0.922	0.960	0.987
SCS	1	0.577	0.792	0.889	0.946	0.979
	2+	0.634	0.797	0.890	0.947	0.978

results were relatively insensitive to small changes in curve parameters (data not shown).

Norman et al. (1999b) used daily milk weights for Canadian cows and monthly test records of US cows to estimate phenotypic correlations between TD within herd-year. Their objective was to develop a mathematical function for estimating correlations that would eliminate the need to repeatedly calculate 365-by-365 correlation matrices. They found that correlations between daily yields for a designated interval between TD generally were largest for mid-lactation and smallest for early and late lactation. Several possible correlation structures were defined, including identity, first-order autoregressive matrices, and persistency matrices. The resulting models for first parity included 2 terms (autoregression and early DIM), and the model for later parities included 3 sources of variation (autoregression, early DIM, and middle DIM). The addition of more terms to the model generally improved fit, but gains were often very small, and sometimes there was difficulty ensuring the positive-definiteness of the resulting correlation matrix. The model presented in this paper is much simpler to explain than previous models, does not have problems with the positive definiteness of the resulting matrix, and produces results that are very highly correlated with the results of Norman et al. (1999b).

Correlations among M, F, and P were not used in the prediction of daily or lactation yields of SCS. Canavesi et al. (2007) reported that bull proofs were virtually identical when calculated using 3-trait (M, F, and P) and 4-trait (M, F, P, and SCS) random-regression TD models. Those results suggest that there is no advantage to be gained by estimating SCS simultaneously with the yield traits. In addition, Norman et al. (1999b) noted that correlation matrices were not guaranteed to be positive definite when correlations were allowed to vary for pairwise combinations of traits, so a correlation pattern that allowed SCS to differ from the other traits was selected. This allows daily fluctuations of SCS to be larger and independent of the other traits because off-diagonals between yield and SCS are zero, and the model is allowed to account for greater heterogeneity among daily SCS than daily yield.

Lactation yields initiated on or after January 1, 1997, were recalculated using the BESTPRED program. Milk, F, and P yields and SCS data were then extracted and genetic evaluations calculated using the yields from the new program. The resulting predicted transmitting abilities were sent to the International Bull Evaluation Service (Interbull, Uppsala, Sweden) for a test run. No significant problems were identified in the test run, and the updated version of BP described in this paper are now currently used in the United States. Several dairy records processing centers have also expressed interest in replacing their TIM calculations with BP.

The BESTPRED program includes several improvements over AIPLDCR: lactations of any reasonable length can now be modeled; lactation-to-date, 305-d, 365-d, and projected yields are provided; and BP of individual daily yields, TD yields, and standard curves are now output (Cole and VanRaden, 2007). It is public domain software and may be downloaded, with documentation, from the AIPL Web site at: http://www.aipl.arsusda.gov/software/bestpred/. Many cows can profitably produce for more than 305 d, and BEST-PRED provides a flexible tool to model these records.

There is increasing interest in milking cows for longer than 305 d, which has been the standard used in the United States since 1938. Several changes were made to the software used to process TD records into lactation yields to accommodate longer lactations. Correlations among TD are now modeled using simpler functions that account for both biological and day-to-day sampling differences. Curves for 999-d lactations have been developed for means and SD of yield that account for breed and parity differences among animals. Results from 2 validation approaches demonstrate that the lactation and daily yields are highly accurate. These revisions to best prediction provide a flexible tool to accurately model milk, fat, protein, and SCS yield in lactations of any reasonable length.

REFERENCES

Bohmanova, J., I. Misztal, S. Tsuruta, H. D. Norman, and T. J. Lawlor. 2005. National genetic evaluation of milk yield for heat tolerance of United States Holsteins. Interbull Bull. 33:160–162.

Canavesi, F., S. Biffani, and E. L. Nicolazzi. 2007. Three versus four traits random regression TD model genetic evaluation for the Holstein breed in Italy. Page 64 in Proc. European Assoc. Anim. Prod. 58th Annu. Mtg., Dublin, Ireland.

Cole, J. B., and P. M. VanRaden. 2006. Genetic evaluation and best prediction of lactation persistency. J. Dairy Sci. 89:2722–2728.

Cole, J. B., and P. M. VanRaden. 2007. A Manual for Use of BESTPRED: A Program for Estimation of Lactation Yield and Persistency Using Best Prediction. http://www.aipl.arsusda.gov/software/bestpred/. Accessed Dec.11, 2007.

De Vries, A. 2006. Ranking dairy cows for optimal breeding decisions. Pages 49–66 in Proc. 43rd Florida Dairy Prod. Conf., Gainesville.

DeGroot, B. J., J. F. Keown, L. D. Van Vleck, and S. D. Kachman. 2007. Estimates of genetic parameters for Holstein cows for test-day yield traits with a random regression cubic spline model. Genet. Mol. Res. 6:434–444.

Dematawewa, C. M. B., R. E. Pearson, and P. M. VanRaden. 2007. Modeling extended lactations of Holsteins. J. Dairy Sci. 90:3924–3936.

Dijkstra, J., J. France, M. S. Dhanoa, J. A. Maas, M. D. Hanigan, A. J. Rook, and D. E. Beever. 1997. A model to describe growth patterns of the mammary gland during pregnancy and lactation. J. Dairy Sci. 80:2340–2354.

Gengler, N., G. R. Wiggans, and A. Gillon. 2004. Estimated heterogeneity of phenotypic variance of test-day yield with a structural variance model. J. Dairy Sci. 87:1908–1916.

Hansen, L. B. 2000. Consequences of selection for milk yield from a geneticist's viewpoint. J. Dairy Sci. 83:1145–1150.

ICAR. 2006. International Agreement of Recording Practices. International Committee for Animal Recording, Rome, Italy.

Kachman, S. D. 2004. Relationship between the choice of a random regression model and the possible shapes of the resulting variance functions. J. Dairy Sci. 87(Suppl. 1):243. (Abstr.)

Knight, C. H. 2005. Extended lactation: Turning theory into reality. Adv. Dairy Technol. 17:113–124.

Lucy, M. C. 2001. Reproductive loss in high-producing dairy cattle: Where will it end? J. Dairy Sci. 84:1277–1293.

Morant, S. V., and A. Gnanasakthy. 1989. A new approach to the mathematical formulation of lactation curves. Anim. Prod. 49:151–162.

Nelder, J. A. 1966. Inverse polynomials, a useful group of multifactor response functions. Biometrics 22:128–141.

Norman, H. D., P. M. VanRaden, J. R. Wright, and J. S. Clay. 1999a. Comparison of test interval and best prediction methods for estimation of lactation yield from monthly, a.m.-p.m., and tri-monthly testing. J. Dairy Sci. 82:438–444.

Norman, H. D., P. M. VanRaden, J. R. Wright, and L. A. Smith. 1999b. Mathematical representations of correlations among yield traits and somatic cell score on test day. J. Dairy Sci. 82:2205–2211.

Patterson, H. D., and R. Thompson. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554.

Quist, M. A., S. J. LeBlanc, K. J. Hand, D. Lazenby, F. Miglior, and D. F. Kelton. 2007. Agreement of predicted 305-day milk yields relative to actual 305-day milk weight yields. J. Dairy Sci. 90:4684–4692.

Rajala-Schultz, P. J., Y. T. Gröhn, C. E. McCulloch, and C. L. Guard. 1999. Effects of clinical mastitis on milk yield in dairy cows. J. Dairy Sci. 82:1213–1220.

Reents, R., J. Jamrozik, L. R. Schaeffer, and J. C. M. Dekkers. 1995. Estimation of genetic parameters for TD records of somatic cell score. J. Dairy Sci. 78:2847–2857.

Rodriguez-Zas, S. L., D. Gianola, and G. E. Shook. 2000a. Evaluation of models for somatic cell score lactation patterns in Holsteins. Livest. Prod. Sci. 67:19–30.

Rodriguez-Zas, S. L., D. Gianola, and G. E. Shook. 2000b. An approximate Bayesian analysis of somatic cell score curves in Holsteins. Acta Agric. Scand. A 50:291–299.

Rook, A. J., J. France, and M. S. Dhanoa. 1993. On the mathematical description of lactation curves. J. Agric. Sci. 121:97–102.

Sargent, F. D., V. H. Lytton, and O. G. Wall Jr. 1968. Test interval method of calculating dairy herd improvement association records. J. Dairy Sci. 51:170–179.

SAS Institute. 2007. SAS OnlineDoc 9.1.3. SAS Institute Inc., Cary, NC. http://support.sas.com/onlinedoc/913/docMainpage.jsp. Accessed Dec. 11, 2007.

Schaeffer, L. R., and J. Jamrozik. 1996. Multiple-trait prediction of lactation yields for dairy cows. J. Dairy Sci. 79:2044–2055.

Schutz, M. M., P. M. VanRaden, G. R. Wiggans, and H. D. Norman. 1995. Standardization of lactation means of somatic cell scores for calculation of genetic evaluations. J. Dairy Sci. 78:1843–1854.

Shook, G. E., E. L. Jensen, and F. N. Dickinson. 1980a. Factors for estimating sample-day yield in am-pm plans. DHI Letters 56:25–30.

Shook, G. E., L. P. Johnson, and F. N. Dickinson. 1980b. Factors for improving accuracy of estimates of test-interval yield. DHI Letters 56:9–24.

VanRaden, P. M. 1997. Lactation yields and accuracies computed from test-day yields and (co)variances by best prediction. J. Dairy Sci. 80:3015–3030.

White, I. M. S., R. Thompson, and S. Brotherstone. 1999. Genetic and environmental smoothing of lactation curves with cubic splines. J. Dairy Sci. 82:632–638.

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.

Windig, J. J., M. P. L. Calus, B. Beerda, and R. F. Veerkamp. 2006. Genetic correlations between milk production and health and fertility depending on herd environment. J. Dairy Sci. 89:1765–1775.

Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. Nature 216:164–165.