Animal (2012), **6:10**, pp 1572–1582 © The Animal Consortium 2012
doi:10.1017/S1751731112000766

Analysis of lactation shapes in extended lactations

R. Steri[1†], C. Dimauro[1], F. Canavesi[2], E. L. Nicolazzi[2,3] and N. P. P. Macciotta[1]

[1]*Dipartimento di Scienze Zootecniche, Università di Sassari, Via Enrico De Nicola 9, 07100 Sassari, Italy;* [2]*Associazione Nazionale Allevatori Frisona Italiana, Via Bergamo 292, 26100 Cremona, Italy;* [3]*Istituto di Zootecnica, Università Cattolica del Sacro Cuore, Via Emilia Parmense 84, 29122 Piacenza, Italy*

(Received 7 April 2011; Accepted 30 January 2012; First published online 3 April 2012)

In order to describe the temporal evolution of milk yield (MY) and composition in extended lactations, 21 658 lactations of Italian Holstein cows were analyzed. Six empirical mathematical models currently used to fit 305 standard lactations (Wood, Wilmink, Legendre, Ali and Schaeffer, quadratic and cubic splines) and one function developed specifically for extended lactations (a modification of the Dijkstra model) were tested to identify a suitable function for describing patterns until 1000 days in milk (DIM). Comparison was performed on individual patterns and on average curves grouped according to parity (primiparous and multiparous) and lactation length (standard ≤305 days, and extended from 600 to 1000 days). For average patterns, polynomial models showed better fitting performances when compared with the three or four parameters models. However, LEG and spline regression, showed poor prediction ability at the extremes of the lactation trajectory. The Ali and Schaeffer polynomial and Dijkstra function were effective in modelling average curves for MY and protein percentage, whereas a reduced fitting ability was observed for fat percentage and somatic cell score. When individual patterns were fitted, polynomial models outperformed nonlinear functions. No detectable differences were observed between standard and extended patterns in the initial phase of lactation, with similar values of peak production and time at peak. A considerable difference in persistency was observed between 200 and 305 DIM. Such a difference resulted in an estimated difference between standard and extended cycle of about 7 and 9 kg/day for daily yield at 305 DIM and of 463 and 677 kg of cumulated milk production at 305 DIM for the first- and second-parity groups, respectively. For first and later lactation animals, peak yield estimates were nearly 31 and 38 kg, respectively, and occurred at around 65 and 40 days. The asymptotic level of production was around 9 kg for multiparous cows, whereas the estimate was negative for first parity.

Keywords: dairy cattle, extended lactation, mathematical model

Implications

The average lactation length of high-producing cows has increased markedly in recent years mainly because of negative relationships between level of production and reproductive efficiency. Nevertheless, a voluntary conception delay is an alternative of interest for many farmers.

The knowledge of temporal evolution of milk yield and milk components in extended lactations is a key element for accurate management strategies. For this purpose the use of different models may offer complementary information and allow calculation of two traits, the inflection point and the asymptotic level of production, which may have some relevance for management and genetic purposes.

Introduction

The occurrence of cows having lactations that exceed the standard length of 305 days has markedly increased in recent years. For example, more than 55% of US Holsteins have lactations longer than 305 days (VanRaden *et al.*, 2006) and about 25% of cows in Costa Rica are dried off after 330 days of lactation (Vargas *et al.*, 2000). An overall increase of lactation length of 30 days for the last decade has been estimated by Gonzalez-Recio *et al.*, (2004).

The main reason for extended lactations is conception failure in early lactation. Traditionally, calving interval is fixed at about 1 year and the cow is inseminated around 2 months after parturition, that is, approximately when the lactation peak is expected to occur. Several authors report a negative relationship between level of production and reproductive efficiency at this lactation stage, in intensively managed herds (Lucy, 2001; Dobson *et al.*, 2007). The negative energy balance experienced by the animal leads to massive mobilization of body reserves with a subsequent marked decrease in conception rate (Veerkamp *et al.*, 1995; Bertilsson *et al.*, 1997). Moreover, the intense selection for increasing yield has resulted in a high risk of health problems around the period between calving and lactation peak. Erb *et al.* (1984)

† E-mail: rsteri@uniss.it

estimated that around 60% of health disorders occur during the first 40 days of lactation.

Conversely, a voluntary waiting period for insemination in order to extend calving interval from 18 to 19 months represents a management strategy for cows able to maintain high production levels for long time (Knight, 2005). Actually, the possibility of increasing lactation length has been considered as an alternative to the common strategy of maximizing lactation peaks and minimizing calving intervals (Grossman and Koops, 2003; Auldist *et al*., 2007; Butler *et al*., 2010). Cows with extended lactations have a reduction of calving risks and of *postpartum* metabolic diseases (Cole and Null, 2009) with a corresponding decrease in insemination costs and in number of days dry. Intensively managed dairy cows typically achieve three lactations in their entire productive life, which exposes them to three 'peak risk' periods. The immediate benefit of extended lactations would be a reduction of this exposure to two cycles. Moreover, Knight (2005) pointed out that milk yielded in 3 years with two (extended lactation) or three (conventional lactation) cycles, is the same if persistency is improved by 1% in the extended lactations.

Knowledge of main features of long lactations could be a key element for management and breeding decisions. Mathematical modelling of lactation curve with regular functions of time represents the most frequently used approach for studying main features of lactation patterns. Till the early 2000s, most literature dealing with lactation curve modelling was focused on the standard length and models have been basically conceived to fit 305-day lactation patterns. Vargas *et al.* (2000) fitted some of the most popular functions to extended lactations, obtaining the best performances for the diphasic model. Grossman and Koops (2003) further developed the multiphasic approach. A more recent comparison performed by Dematawewa *et al.* (2007) on US Holsteins concluded that three- or four-parameter models, as the Wood (Wood, 1967) or the Rook (Rook *et al*., 1993) could be more suitable for describing long lactations in routine use, whereas more complex mechanistic models with a higher number of parameters should be used mainly for research purposes.

Some authors have argued that models conceived for fitting 305-day patterns may not adequately describe extended lactations and that specific functions should be developed (Grossman and Koops, 2003). Such a consideration may be correct for simple models characterized by a scarce flexibility but it may be questionable if flexible and more general functions are used. On the other hand, the increase of the number of parameters often leads to computational problems.

VanRaden *et al.* (2006), suggested a modification of the mechanistic model of Dijkstra *et al.* (1997). The model was effective in modelling the average extended lactations curves of US Holstein. Moreover, it was able to identify a phase with production nearly constant beyond 600 days in milk (DIM) (asymptotic phase) that may be used as a decision criteria for keep milking a cow.

The aim of this study is to assess the main features of lactation curves for milk yield (MY) and its components for different length classes. For this purpose, six empirical mathematical models currently used to fit 305 standard lactations and the modified Dijkstra function (DF) were tested to evaluate their efficiency in describing average and individual extended lactations of Italian Holsteins.

Material and methods

Data

Data were 241 354 test-day records for MY, fat (FP) and protein (PP) percentage and somatic cell score (SCS = $\log_2$(somatic cell count/100)−3), from 21 658 lactations of Italian Holstein cows. Data were recorded in the period from 2002 to 2006 by the Italian Breeders Association in herds located in Northern Italy. Lactations were grouped according to parity (1 = first; 2 = second and third) and to lactation length (standard ⩽305 days, and extended from 600 to 1000 days), whereas lactations with lengths between 306 and 599 days were excluded from this study. Lactations were discarded if the first test day occurred after 70 days from parturition, whereas records after 1000 days from calving were not considered. Summary statistics across class of lactation length are reported in Table 1.

Lactation curve models

Seven mathematical models available in literature were selected. Considered functions are characterized by an increasing number of parameters (three to seven), in order to

Table 1 *Distribution of lactations and TD available, means (and standard deviation) of lactation length and test days for lactation, according to parity and length classes*

	Lactation length			
	Standard		Extended	
Class of parity	Primiparous	Multiparous	Primiparous	Multiparous
TD records	92 061	1 20 982	17 263	11 048
No. of lactations	10 776	14 114	872	564
Mean ± s.d. of TD by lactation	8.5 (±0.62)	8.6 (±0.64)	19.78 (±2.46)	19.59 (±2.36)
Mean ± s.d. of lactation length	276.3 (±18.96)	275.2 (±19.81)	684.5 (±79.45)	680.9 (±78.99)

TD = test days.

test the effect of different degrees of flexibility on the ability to describe extended lactation patterns. The models were:

(1) the incomplete gamma function of Wood (1967) (WD):

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

that is the most popular and commonly used model of lactation curve, recommended also for fitting extended lactations (Dematawewa *et al*., 2007);

(2) the combined exponential and linear model of Wilmink (1987) (WIL):

$$Y_t = a + be^{-kt} + ct$$

used in random regression test day models (Schaeffer *et al*., 2000; Reinhardt *et al*., 2002) to model fixed regression effects;

(3) the modified version of DF proposed by VanRaden *et al.* (2006):

$$Y_t = b_0 + b_1 \exp\left[\frac{b_2(1-\exp(-b_3 t))}{b_3} - b_4 t\right]$$

for $t \rightarrow \infty$ the DF has an horizontal asymptote the value of which is expressed by b_0, the asymptote can be interpreted as the value of MY in the final phase of lactation;

(4) the five-parameter polynomial regression of Ali and Schaeffer (1987) (AS):

$$Y_t = a_0 + a_1 x + a_2 x^2 + a_3 \log(1/x) + a_4 (\log(1/x))^2$$

where $x = t$/lactation length, also this model has been used to fit individual curves in early versions of random regression test day models;

(5) a fourth-order Legendre orthogonal polynomial (LEG):

$$Y_t = a_0 P_0 + a_1 P_1 + a_2 P_2 + a_3 P_3 + a_4 P_4$$

where P_i is the function of time calculated using values published by Schaeffer (2004), in this form LEG is used as a sub-model in the random regression test day models for fixed and random effects, both in Italy (Muir *et al*., 2007) and Canada (Muir *et al*., 2004);

(6) a quadratic spline function (QSPL) with three knots:

$$Y_t = a + b_1 t + b_2 t^2 + \sum c_j (t - N_j)^2$$

(7) a cubic spline function (CSPL) with three knots:

$$Y_t = a + b_1 t + b_2 t^2 + b_3 t^3 + \sum c_j (t - N_j)^3$$

where N_j is the position of the knots, placed at 50, 200 and 600 days (the latter only for the longest class; Macciotta *et al*., 2010). In all models, Y_t represents the test day (MY, PP, FP or SCS) recorded at time t (days); a_i, b_i, c_i and k are parameters to be estimated.

The seven models were used to fit both individual and average lactation patterns for all the traits considered. Average curves were calculated as means of all records available for each DIM, for each parity group and within each lactation length class. Only DIM with at least five test-day observations were considered. The goodness of fit was assessed by both using adjusted coefficient of determination (R^2_{adj}) and the root means square error (RMSE). Moreover, the Durbin–Watson statistic (DW) was used to detect the presence of autocorrelation in the residuals from the regression analysis. In fact, the presence of autocorrelated residuals suggests that the function may be inappropriate for the data. A linear regression was used to fit the AS, LEG, QSPL and CSPL models, whereas the WD, WIL and DF were fitted by a nonlinear regression, based on the Marquardt compromise and using the values published by VanRaden *et al.* (2006) and Dematawewa *et al.* (2007) as initial parameter values.

Lactation curve traits, that is, time at which peak yield occurs (Tp), peak yield (Yp), cumulated milk production at 305 days (Y_{305}) and 1000 days (Y_{1000}) were calculated for average patterns. Individual curves were classified according to five levels of R^2_{adj} (<0.30, from 0.30 to 0.50, from 0.50 to 0.70, from 0.70 to 0.90, >0.90).

Results and discussion

Average curves

Goodness-of-fit statistics for the seven functions fitted to average standard curves are reported in Table 2. All models fitted MY well, with R^2_{adj} ranging from 0.971 to 0.996 and RMSE from 0.68 to 0.36 kg. Moderate values of autocorrelation between residuals were found for all models except for spline functions. For LEG, the value of DW was particularly low ($d_L = 1.71$ for $P < 0.01$), especially for the second-parity group. Goodness of fit for milk components was slightly worse for FP and SCS, whereas for PP it was similar to MY. In general WIL, DF and polynomials provided the best fit, whereas WD and LEG gave systematically poorer results, with DW values particularly low for PP and SCS. An exception was the behaviour of WD on FP. For all models the fit was better for the second-parity group than for the first.

For extended lactations (Table 3) a general reduction of goodness of fit compared with standard lactations was found, with lower R^2_{adj} and higher RMSE. For example, average R^2_{adj} and RMSE for MY in the first-parity group were 0.98 and 0.45 in standard and 0.91 and 1.52 in extended lactations, respectively. Such a reduction was markedly higher for milk composition traits, especially FP and SCS. On the other hand, a substantial absence of autocorrelations among residuals (DW ~ 2.00, $n = 713$) was observed for extended lactations. Possible reasons for the reduction in goodness of fit in this class can be found in the high variability of test-day data beyond 600 DIM. Differences observed between traits could also be ascribed to their intrinsic variability, which is larger for SCS and FP than for MY and PP.

As far as model comparison is concerned, no important differences were found for MY, whereas for milk components WD, WIL and LEG models performed worse than the other functions. Moreover, no significant increase in goodness of

Table 2 *Goodness of fit for average standard curves for MY, PP, FP and SCS according to parity class, for WD, WIL, modified DF, LEG, AS, QSPL and CSPL models*

			Model						
Trait	Parity	Statistics	WD	WIL	DF	LEG	AS	QSPL	CSPL
MY	1	R^2_{adj}	0.974	0.972	0.972	0.971	0.979	0.985	0.983
		RMSE	0.48	0.49	0.49	0.51	0.43	0.36	0.39
		DW	0.96	0.91	0.90	0.86	1.20	1.68	1.46
	2	R^2_{adj}	0.989	0.993	0.993	0.987	0.994	0.996	0.996
		RMSE	0.62	0.51	0.51	0.68	0.46	0.40	0.39
		DW	0.64	0.97	0.95	0.55	1.21	1.55	1.69
PP	1	R^2_{adj}	0.931	0.989	0.989	0.935	0.988	0.980	0.990
		RMSE	0.05	0.02	0.02	0.05	0.02	0.03	0.02
		DW	0.21	1.24	1.24	0.22	1.16	0.66	1.31
	2	R^2_{adj}	0.932	0.979	0.979	0.907	0.987	0.968	0.988
		RMSE	0.05	0.03	0.03	0.06	0.02	0.04	0.02
		DW	0.40	0.72	0.72	0.19	1.20	0.52	1.29
FP	1	R^2_{adj}	0.962	0.934	0.933	0.906	0.963	0.962	0.968
		RMSE	0.05	0.06	0.06	0.07	0.04	0.04	0.04
		DW	1.43	0.82	0.82	0.58	1.48	1.43	1.67
	2	R^2_{adj}	0.944	0.933	0.932	0.879	0.954	0.956	0.962
		RMSE	0.05	0.05	0.05	0.07	0.04	0.04	0.04
		DW	1.29	1.08	1.07	0.62	1.60	1.65	1.94
SCS	1	R^2_{adj}	0.810	0.894	0.893	0.823	0.894	0.886	0.892
		RMSE	0.14	0.10	0.10	0.13	0.10	0.11	0.10
		DW	1.19	1.99	2.00	1.20	1.96	1.86	1.95
	2	R^2_{adj}	0.926	0.958	0.958	0.936	0.958	0.955	0.959
		RMSE	0.13	0.10	0.10	0.13	0.10	0.11	0.10
		DW	1.23	2.01	2.02	1.33	2.04	1.88	2.10

MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score; WD = Wood; WIL = Wilmink; DF = Dijkstra function; LEG = Legendre polynomials; AS = Ali and Schaeffer; QSPL = quadratic splines; CSPL = cubic splines; R^2_{adj} = adjusted coefficient of determination; RMSE = root means square error; DW = Durbin–Watson.

Table 3 *Goodness of fit for average extended curves for MY, PP, FP and SCS according to parity class, for WD, WIL, modified DF, LEG, AS, QSPL and CSPL models*

			Model						
Trait	Parity	Statistics	WD	WIL	DF	LEG	AS	QSPL	CSPL
MY	1	R^2_{adj}	0.907	0.906	0.909	0.900	0.910	0.916	0.913
		RMSE	1.54	1.55	1.52	1.59	1.51	1.46	1.48
		DW	1.75	1.74	1.80	1.65	1.81	1.94	1.88
	2	R^2_{adj}	0.933	0.913	0.945	0.934	0.944	0.945	0.945
		RMSE	1.91	2.19	1.74	1.91	1.76	1.74	1.74
		DW	1.59	1.20	1.89	1.56	1.84	1.89	1.89
PP	1	R^2_{adj}	0.813	0.814	0.914	0.890	0.920	0.917	0.920
		RMSE	0.14	0.14	0.09	0.11	0.09	0.09	0.09
		DW	0.88	0.85	1.84	1.43	1.97	1.91	1.98
	2	R^2_{adj}	0.765	0.795	0.890	0.861	0.894	0.892	0.896
		RMSE	0.15	0.14	0.11	0.12	0.10	0.10	0.10
		DW	0.97	0.99	1.84	1.46	1.88	1.85	1.92
FP	1	R^2_{adj}	0.536	0.639	0.700	0.640	0.716	0.720	0.721
		RMSE	0.23	0.20	0.18	0.20	0.18	0.18	0.18
		DW	1.22	1.52	1.87	1.53	1.93	1.97	1.99
	2	R^2_{adj}	0.435	0.537	0.570	0.519	0.588	0.585	0.588
		RMSE	0.24	0.22	0.21	0.23	0.21	0.21	0.21
		DW	1.51	1.82	1.97	1.76	2.06	2.04	2.06
SCS	1	R^2_{adj}	0.531	0.562	0.574	0.538	0.580	0.578	0.581
		RMSE	0.40	0.39	0.38	0.40	0.38	0.38	0.38
		DW	1.88	1.99	2.04	1.88	2.07	2.06	2.08
	2	R^2_{adj}	0.459	0.451	0.506	0.483	0.508	0.508	0.513
		RMSE	0.50	0.50	0.48	0.49	0.48	0.48	0.47
		DW	1.80	1.76	1.95	1.87	1.96	1.96	1.98

MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score; WD = Wood; WIL = Wilmink; DF = Dijkstra function; LEG = Legendre polynomials; AS = Ali and Schaeffer; QSPL = quadratic splines; CSPL = cubic splines; R^2_{adj} = adjusted coefficient of determination; RMSE = root means square error; DW = Durbin–Watson.

fit was found moving from models with five parameters (DF and AS) to six (QSPL) and seven (CSPL).

Tables 4 and 5 report lactation curve traits for standard and extended average patterns, respectively. Comparison between predicted and observed values shows, in general, a greater accuracy for AS and DF. Spline functions performed better for extended than for standard lactations. WD and LEG prediction were the less accurate, particularly for PP and SCS in standard and for MY and PP in extended lactations, respectively.

As expected, first-parity cows had a lower peak yield (about 31 kg) and later occurrences (about 60 to 65 DIM), compared with higher parities (about 38 kg and 40 DIM, respectively). No practical differences between standard and extended cycles were observed for Yp and Tp predicted by AS, DF and spline functions, whereas for WD, WIL and LEG a large difference was found.

Prediction of total MY until 305 and 1000 days for standard and extended lactations are reported in Table 6. Estimates across models were rather similar for Y_{305} with means of 8343 ± 9 and 8811 ± 22 kg for first parity and 9196 ± 14 and 9864 ± 49 kg for second parity in standard and extended lactation, respectively. Differences were higher for Y_{1000}, ranging from 21 270 to 23 566 kg and from 20 221 to 23 711 kg predicted by WIL and CSPL for the first- and second- parity groups, respectively. The differences between parities tended to decrease passing from standard to extended cycle because of greater persistency for young cows.

Estimated curves for standard and extended patterns (Figures 1 and 2), further clarify above reported results (only plots for MY and SCS for the second-parity group are reported for brevity). All models were able to adequately describe the trajectory of MY and SCS for standard lactation length, except WD for SCS and LEG for both traits. For extended lactations, models with fewer parameters tended to underestimate peak yield and were unable to describe the change of slope that occurred in the second part of lactation. This resulted in an overestimation (underestimation for milk contents) in the middle of lactation and an underestimation (overestimation for milk contents) in the final phase of production. These results partially agree with those reported by other authors (Grossman and Koops, 2003; Dematawewa *et al*., 2006; VanRaden *et al*., 2006), that found underestimations of actual yield at later stage of extended lactations with standard models.

Moreover, WD was unable to reconstruct the peak phase correctly for PP (result not shown) and SCS. Spline

Table 4 *Tp, Yp and observed value (actual), for average standard lactations for MY, PP, FP and SCS according to parity class, predicted by WD, WIL, modified DF, LEG, AS, QSPL and CSPL models*

			Model							
Trait	Parity	Statistics*	WD	WIL	DF	LEG	AS	QSPL	CSPL	Actual
MY	1	Tp	63	65	65	70	62	47	46	59
		Yp	31.0	31.1	31.1	31.0	30.8	31.2	30.9	31.9
	2	Tp	43	43	43	50	37	42	36	36
		Yp	38.1	38.3	38.3	37.6	37.9	38.5	38.5	38.9
PP	1	Tp	56	38	38	60	39	45	38	39
		Yp	3.10	3.03	3.03	3.05	3.04	2.99	3.00	3.02
	2	Tp	63	39	39	64	38	45	38	38
		Yp	3.09	3.01	3.01	3.05	3.04	2.99	2.99	3.00
FP	1	Tp	92	79	79	89	91	74	94	96
		Yp	3.38	3.35	3.35	3.35	3.39	3.4	3.41	3.35
	2	Tp	91	69	69	85	82	49	72	78
		Yp	3.43	3.39	3.39	3.39	3.43	3.42	3.45	3.38
SCS	1	Tp	71	45	45	67	48	46	41	55
		Yp	6.32	6.22	6.22	6.21	6.24	6.13	6.18	6.04
	2	Tp	37	31	31	49	31	42	35	33
		Yp	6.55	6.4	6.39	6.47	6.39	6.35	6.34	6.26

Tp = time to the peak; Yp = production at peak; MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score; WD = Wood; WIL = Wilmink; DF = Dijkstra function; LEG = Legendre polynomials; AS = Ali and Schaeffer; QSPL = quadratic splines; CSPL = cubic splines.
*Yp = maximum value for MY and minimum value for protein, fat and somatic cell score.

Table 5 *Tp, Yp and observed value (actual), for average extended lactations for MY, PP, FP and SCS according to parity class, predicted by WD, WIL, modified DF, LEG, AS, QSPL and CSPL models*

			Model							
Trait	Parity	Statistics*	WD	WIL	DF	LEG	AS	QSPL	CSPL	Actual
MY	1	Tp	71	57	65	76	70	47	45	64
		Yp	30.7	31.2	31.5	30.1	31	31.6	31.1	32.5
	2	Tp	28	29	46	1	38	43	39	37
		Yp	37.0	36.4	38.3	37.5	38.1	38.9	38.3	41.3
PP	1	Tp	1	24	40	1	42	44	39	44
		Yp	2.56	3.11	2.90	3.01	2.95	2.90	2.92	2.89
	2	Tp	1	28	46	55	45	46	43	48
		Yp	2.74	3.10	2.91	3.07	2.95	2.92	2.94	2.90
FP	1	Tp	87	63	75	131	85	50	94	83
		Yp	3.53	3.41	3.26	3.41	3.33	3.33	3.32	3.07
	2	Tp	104	59	69	129	78	57	92	83
		Yp	3.58	3.47	3.34	3.45	3.40	3.40	3.38	3.03
SCS	1	Tp	10	31	40	3	43	45	42	79
		Yp	6.59	6.53	6.39	6.55	6.37	6.25	6.34	5.91
	2	Tp	1	19	38	1	32	43	37	63
		Yp	6.03	6.95	6.6	6.59	6.57	6.48	6.54	5.62

Tp = time to the peak; Yp = production to the peak; MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score; WD = Wood; WIL = Wilmink; DF = Dijkstra function; LEG = Legendre; AS = Ali and Schaeffer; QSPL = quadratic splines; CSPL = cubic splines.
*Yp = maximum value for MY and minimum value for protein, fat and somatic cell score.

functions were able to describe both the first and the middle phase of lactation accurately, whereas beyond ~600 DIM, they yielded a poor fit. A decrease of prediction ability for spline regression at the extremes of the lactation trajectory, especially if few records are available, has been previously reported (Macciotta *et al.*, 2010). LEG showed problems in fitting both extremes of the lactation trajectory for all traits.

An explanation of the difficulty for WD to adequately model the peak yield can be found in the high degree of correlations between its parameters. This is not the case for LEG because of the orthogonality property of polynomials.

Table 6 *Cumulative production until 305 day (Y_{305}) and cumulative production until the dry-off (Y_{1000}), for MY, according to parity and length classes, predicted by WD, WIL, modified DF, LEG, AS, QSPL and CSPL models*

Parity	Length class	Trait	WD	WIL	DF	LEG	AS	QSPL	CSPL
			Model						
1	Standard	Y_{305}	8335	8348	8348	8353	8328	8347	8343
	Extended	Y_{305}	8812	8774	8811	8848	8816	8806	8811
		Y_{1000}	21 516	21 270	21 561	22 272	21 827	23 031	23 566
2	Standard	Y_{305}	9189	9199	9200	9214	9168	9204	9197
	Extended	Y_{305}	9887	9769	9877	9928	9842	9876	9871
		Y_{1000}	21 207	20 221	21 804	20 696	22 057	22 494	23 711

MY = milk yield; WD = Wood; WIL = Wilmink; DF = Dijkstra function; LEG = Legendre polynomials; AS = Ali and Schaeffer; QSPL = quadratic splines; CSPL = cubic splines.

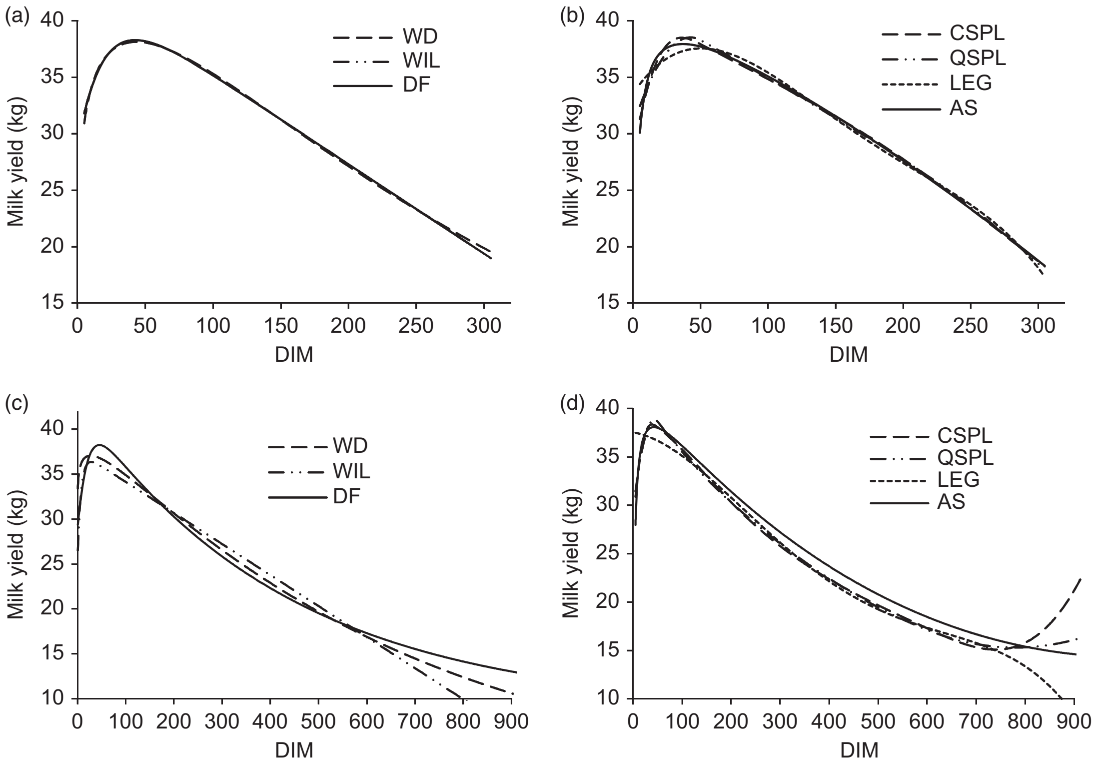

Figure 1 Estimated milk yield (MY) lactation curves for the second-parity group, by Wood (WD), Wilmink (WIL), modified Dijkstra function (DF), Legendre polynomials (LEG), Ali and Schaeffer (AS), quadratic splines (QSPL) and cubic splines (CSPL) models, for standard (a, b) and extended (c, d) classes. DIM = days in milk.

On the basis of the results on goodness of fit and prediction ability for average curves, DF and AS were able to reconstruct the regular patterns for both standard and extended average curves with good accuracies.

Main differences in shape between the two lactation length classes can be inferred from curves estimated by DF for the second-parity group in standard and extended lactations (Figure 3; similar results were obtained with the AS model). In agreement with previous studies (VanRaden *et al.*, 2006; Dematawewa *et al.*, 2007) no important differences were detected between standard and extended patterns in the initial phase of lactation, both for the magnitude of values and for the occurrences of their maximum (or minimum for milk components). However, a marked difference in persistency can be observed between 200 and 305 DIM, especially for MY. Such a difference, reported by several authors (Bertilsson *et al.*, 1997; Vargas *et al.*, 2000; Dematawewa *et al.*, 2007) may be explained by the effect of pregnancy on milk production, that results in a reduction of yield and in a change of milk composition around the 6th month of gestation (Bertilsson *et al.*, 1997; Brotherstone *et al.*, 2004).

The variation of slope between standard and extended patterns resulted in an estimated difference of about 7 and 9 kg/day for daily yield at 305 DIM in the first- and second-parity groups, respectively. Extended lactations had a higher total MY at 305 DIM (difference of 463 and 677 kg for the

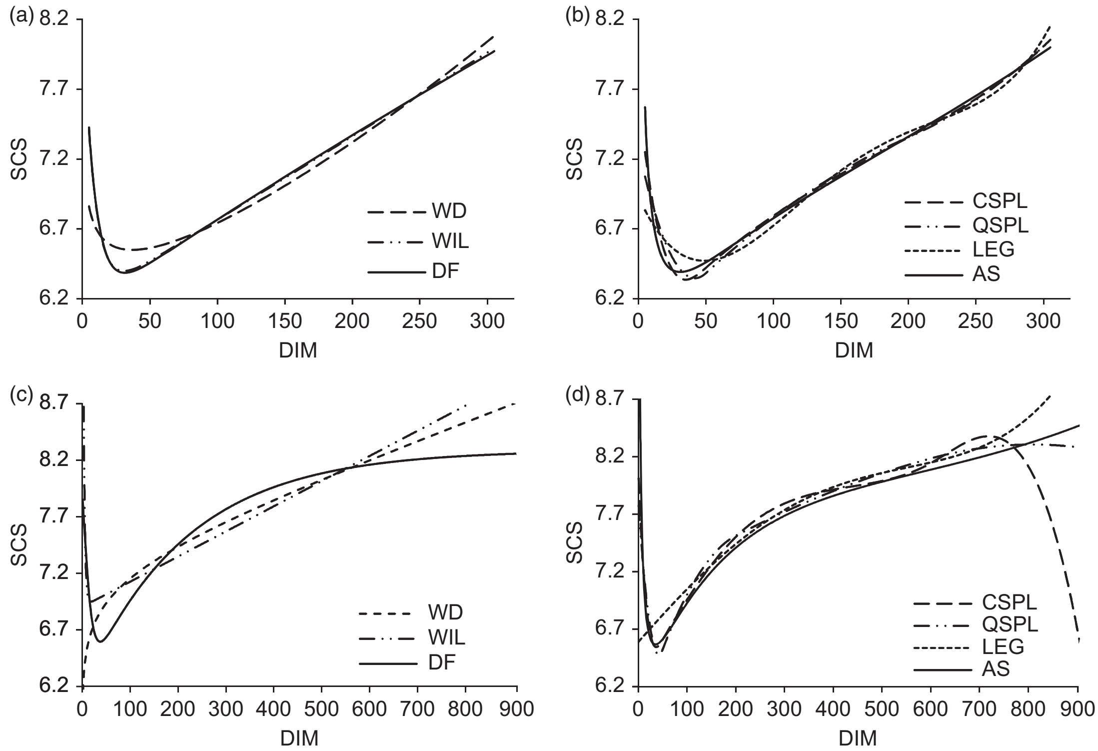

Figure 2 Estimated somatic cell score (SCS) lactation curves for the second-parity group, by Wood (WD), Wilmink (WIL), modified Dijkstra function (DF), Legendre polynomials (LEG), Ali and Schaeffer (AS), quadratic splines (QSPL) and cubic splines (CSPL) models, for standard (a, b) and extended (c, d) classes. DIM = days in milk.

first- and second-parity groups, respectively) compared with standard curves (Table 6).

Beyond 300 DIM, it can be observed that the slope changes and MY and contents tend to an asymptotic value (Figure 3). Such a behaviour of the final stage of lactation can be described by variables calculated from parameters of both AS and DF, even though with a different meaning. In particular, the b_0 parameter of DF expresses the asymptotic value of the curve. Its values were much lower for MY and higher for composition traits in standard compared with extended curves, respectively (Table 7). These figures are obviously mathematical artefacts because of the fact that in the standard lactations there are no data after 305 DIM (i.e. when the asymptotical phase is expected to occur). The size of the estimated values of b_0 seems to be related to the degree of slope of the curve. Test day for MY or composition at the last test day of lactation (LTD) can be predicted by the AS model AS $= a_0 + a_1 + a_2$. LTD for milk components were very similar between standard and extended lactations, whereas some differences can be observed for MY (about 7 and 4 kg/day, for the first- and second-parity groups, respectively). These results seem to indicate a constant milk composition for the extended period (i.e. from 300 to 1000 DIM). In general LTD values are slightly lower, except for MY, than those for b_0 in the extended lactations (Table 7).

Differences between parity groups in extended lactations can be observed in Figure 4, which reports patterns for first and second-parity group predicted by DF for average extended lactations. First-parity cows showed lower peak MY and a greater persistency compared with higher parities. This trend confirms what is usually observed in lactations of standard length (Wood, 1968), where it is ascribed to mammary maturation processes that are still in progress during the first lactation and which counteracts the normal decline in MY at the end of lactation (Stanton *et al.*, 1992).

After ~300 DIM, older cows showed a lower production level compared with first parity and tend to an asymptotic value, in agreement with previous results on extended lactations (Haile-Mariam and Goddard, 2008; Cole and Null, 2009). These results partially differ from those reported by VanRaden *et al.* (2006) who found an asymptotic level of production for all parities of around 20 kg. Asymptotic yield estimated in this study with the DF (Table 7) was positive only for older cows. The time at which the lactation patterns changes the curvature (Tf; calculated for AS solving the following equations: $a_3 + 2a_2X^2 = -2a_4(1 + \log(1/X))$) confirms the differences between parity. Tf has an earlier occurrence in older cows compared with first parity.

Such behaviour is because of the different rates of decline (or increase, for milk components) of yield after the production peak (minimum for milk components). Curves for MY and SCS of primiparous cows show the classical pattern, whereas older cows exhibit an earlier change of curvature. Patterns of PP and FP are similar across parities (Figure 4 and Table 7).

With regard to lactation length, the second derivative had no solution for MY and FP in the 305-day class (Figure 3 and Table 7) presumably because there is no point of inflection after peak yield. In the case of PP and SCS, Tf was estimated

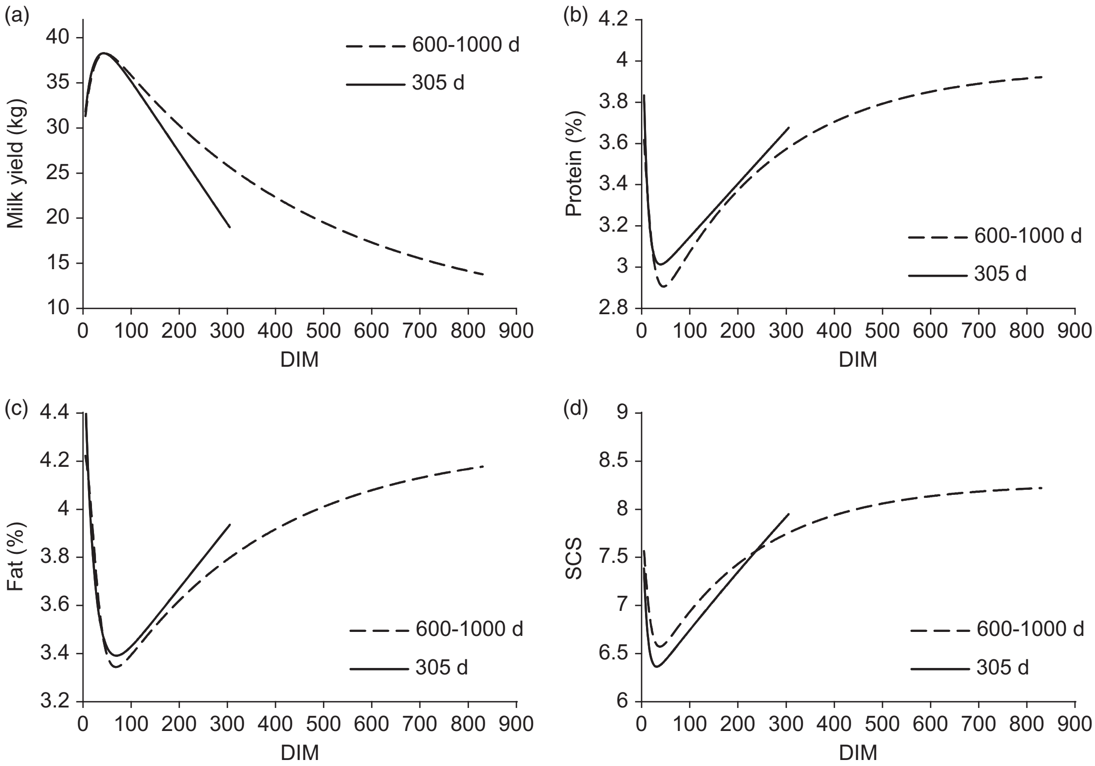

Figure 3 Estimated lactation curves by modified Dijkstra function (DF), for the second-parity group for (a) milk yield (MY), (b) protein percentage (PP), (c) fat percentage (FP) and (d) somatic cell score (SCS) according to length class. DIM = days in milk.

Table 7 *Asymptotic level (b_0) predicted by modified Dijkstra function and LTD and Tf, predicted by Ali and Schaeffer model, according to parity group and length classes for MY, PP, FP and SCS*

		b_0		LTD		Tf	
Trait	Parity	Standard	Extended	Standard	Extended	Standard	Extended
MY	1	−876.5	−10.40	21.07	14.37	–	300
	2	−1434.4	8.76	18.29	14.53	–	132
PP	1	30.2	4.00	3.66	3.69	167	108
	2	66.6	3.97	3.74	3.64	123	111
FP	1	69.9	4.26	4.05	3.62	–	218
	2	49.6	4.27	4.01	3.73	–	189
SCS	1	16.1	8.46	7.14	8.51	252	92
	2	18.1	8.29	8.00	8.88	182	69

LTD = last test day; Tf = time at inflection point; AS = Ali and Schaeffer model; MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score.

around 150 and 200 days, respectively. It is interesting to notice that for these two traits the minimum level of production occurs earlier than for MY and FP (Tables 4 and 5). Results reported on Table 7 highlight an anticipated increase of the curve after the minimum for PP and SCS compared with FP.

Differences in Tf found between different classes of length (very large in the case of SCS) may be related to the fact that standard lactation patterns are more regular than those observed in extended ones. Thus, the change of curvature tends to occur later on in standard lactation. A shorter Tf may indicate a tendency of the animal to reach the asymptotic trend of production early. An anticipated occurrence of inflection in the lactation curve may be of help in an earlier evaluation of the opportunity to keep milking the cow. However, this may need further investigation as the possible influence of mathematical artefacts due to the interaction between properties of the mathematical model and data structure cannot be excluded.

As expected, milk components showed an opposite trend compared with MY (Figure 4). In particular, FP and PP did not show a significant variation among parities, whereas differences were detected for SCS, with a higher level for older cows. Similar SCS patterns were observed by Haile-Mariam and Goddard (2008) and Cole and Null (2009). Moreover, FP and PP had the tendency to reach a plateau around 500 to

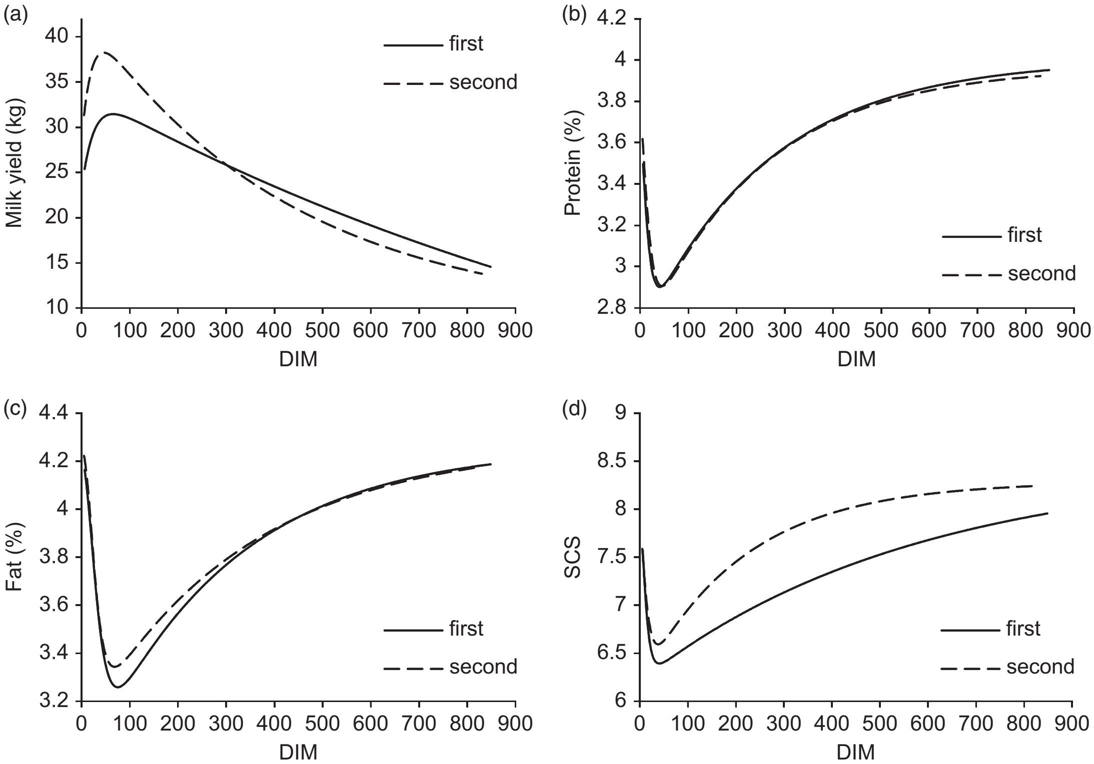

Figure 4 Estimated lactation curves by modified Dijkstra function (DF) model, for (a) milk yield (MY), (b) protein percentage (PP), (c) fat percentage (FP) and (d) somatic cell score (SCS) according to parity class. DIM = days in milk.

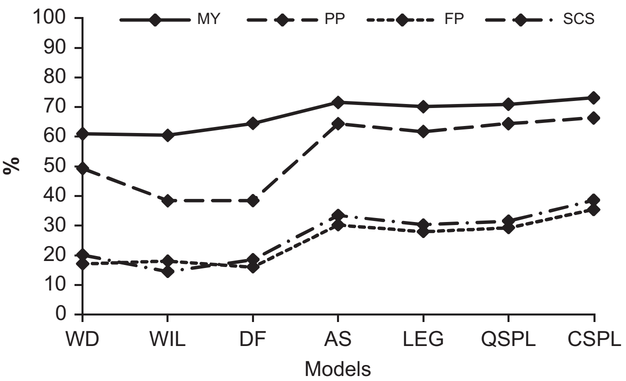

Figure 5 Percentage of individual standard curves having an adjusted coefficient of determination (R^2_{adj}) higher than 0.70 for Wood (WD), Wilmink (WIL), modified Dijkstra function (DF), Legendre polynomials (LEG), Ali and Schaeffer (AS), quadratic splines (QSPL) and cubic splines (CSPL) models. MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score.

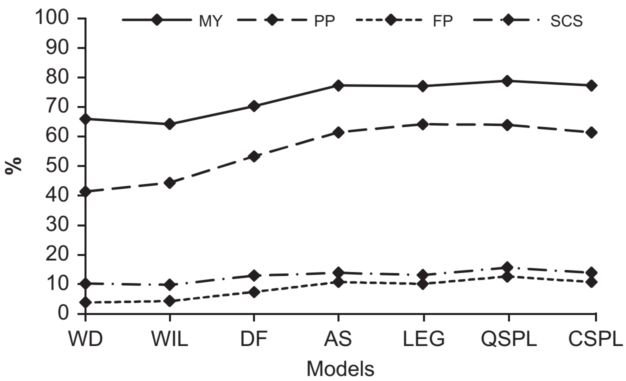

Figure 6 Percentage of individual extended curves having an adjusted coefficient of determination (R^2_{adj}) higher than 0.70 for Wood (WD), Wilmink (WIL), modified Dijkstra function (DF), Legendre polynomials (LEG), Ali and Schaeffer (AS), quadratic splines (QSPL) and cubic splines (CSPL) models. MY = milk yield; PP = protein percentage; FP = fat percentage; SCS = somatic cell score.

600 DIM. SCS showed a continuously increasing trend. The estimated Tf value (Table 7) indicated that the change of the slope occurred earlier for PP and SCS than for FP and MY.

Individual curves

Percentage of individual lactation curves having an R^2_{adj} higher than 0.7 are reported in Figure 5 for standard and Figure 6 for extended lactations, respectively. Goodness of fit is markedly lower compared with average curves, thus confirming the higher variability of individual patterns. Similarly to average curves, poor fitting performances were obtained for FP and SCS, especially in extended lactations.

As expected, functions with a higher number of parameters show better fitting performances compared with three and four parameter functions. This difference was greatest for FP and SCS in standard lactations. QSPL and CSPL showed best fits for all traits and length classes. However AS and LEG models had very similar performances. DF showed intermediate results because of frequent convergence problems, which occurred in many cases (among

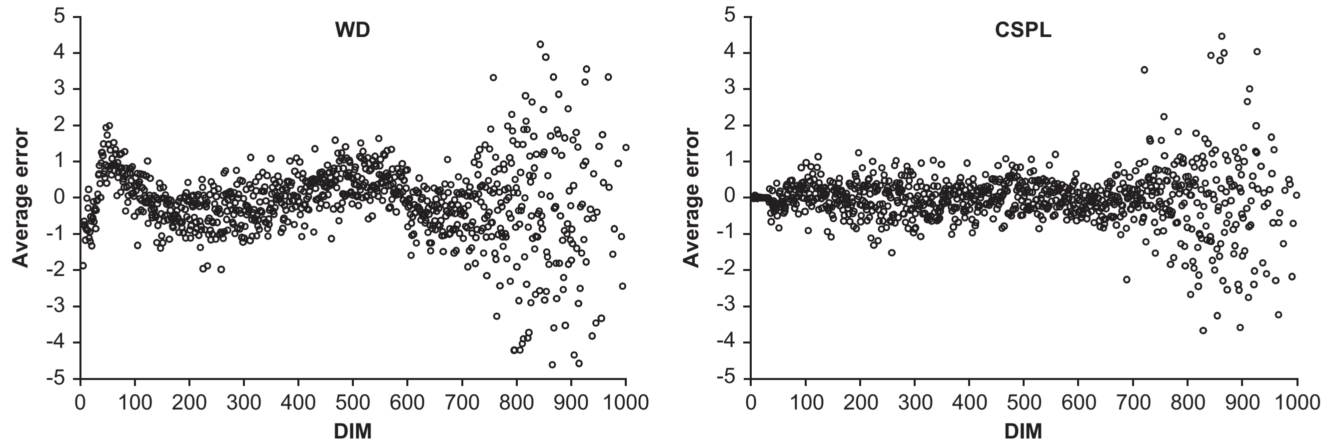

Figure 7 Distribution of average residuals (kg/day) for individual extended curves fitted by Wood (WD), and cubic splines (CSPL) models. DIM = days in milk.

7% to 10% of cases, depending on the trait), especially for standard patterns.

Figure 7 reports mean residuals (difference between observed and predicted) along the lactation for individual patterns in extended curves. Only an example of the best (CSPL) and the worst (WD) fit for MY are shown for brevity. A high variability of residuals was found, especially in the last part of lactation (i.e. beyond 700 DIM). A trend can be observed for WD throughout lactation (similar results were obtained by WIL and DF), whereas for LEG the trend was found in the first 150 DIM. CSPL showed a random distribution of residuals (similar results were obtained by QSPL and AS).

Results obtained in this study for individual lactations of standard length agree with previous reports for dairy cattle (Olori *et al.*, 1999; Macciotta *et al.*, 2005; Silvestre *et al.*, 2006), with a better fit for polynomials (AS, LEG and SPL) compared with models with three parameters. No studies are currently available in the literature on fitting individual extended lactation curves.

For genetic purposes, it may be of interest to have a first look at individual values of extended lactation curve traits. Medians of lactation curve traits for MY in the 600- to 1000-day class are reported in Table 8. Curve traits were similar in most of cases with corresponding values calculated for average curves (Tables 5 and 7). In general they are characterized by a high variability, as expected due to fitting performances.

According to the value of asymptotic production estimated by the DF model, that is, b_o, individual lactation curves can be grouped into three classes (Figure 8). The first, rather frequent for first-parity cows, is characterized by a decreasing phase after the peak yield with constant slope. The second and the third have a pattern characterized by a variation of the slope occurring at ~300 to 500 days after calving, and by a tendency to reach an asymptotic level of production. The main difference between these two groups is the level of asymptotic yield. However, it must be remembered that this research, as most of studies on this topic, has been carried out on an archive where extended lactations were in the majority of cases the result of reproductive failures and not the consequence of management choices. Consequently, results obtained in this study need to be validated with an appropriate experimental plan, that also takes account of appropriate management of extended lactation (i.e. specific feeding strategy).

Table 8 *Median of Yp, Tp, asymptotic yield (b_0) and production at the LTD of individual lactation curves, estimated by AS and modified DF for 600- to 1000-day class length*

Model	Parity	Yp	Tp	b_0	LTD
AS	1	31.68	74.00		−0.17
	2	38.83	56.00		5.44
DF	1	32.54	73.91	−14.00	
	2	39.80	47.69	9.22	

Yp = production at the peak; Tp = time to the peak; LTD = last test day; AS = Ali and Schaeffer; DF = Dijkstra function.

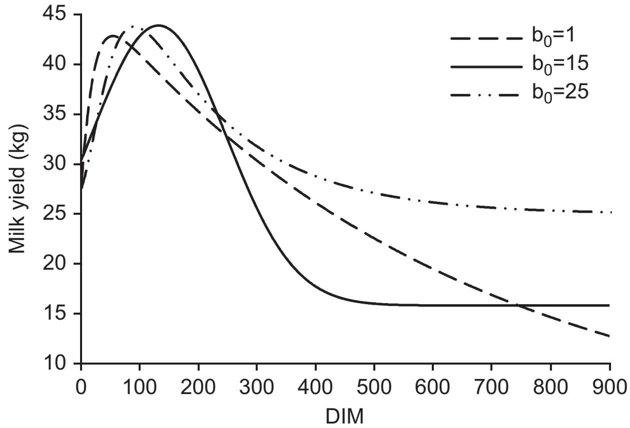

Figure 8 Individual patterns of milk yield (MY) lactation curves with different values of b_0 parameter estimated by modified Dikstra function. DIM = days in milk.

Conclusions

The comparison among seven different functions suggests that models used for describing lactations of standard length can be used also for extended patterns, provided they possess sufficient flexibility. In this study, models with a limited number of parameters were found to be inadequate to fit extended lactations in agreement with results of different authors. On the other hand, a poor prediction ability at the end of extended patterns was highlighted for regression splines, because of their intrinsic ability to fit random perturbation. AS and DF were found to be able to adequately describe both average and individual patterns in standard and extended lactations of Italian Holsteins for MY and PP. A reduced fitting ability was observed for FP and SCS, due to the wide variability of these traits both in standard and extended lactations. The DF was effective in modelling average curves, whereas it experienced computational problems when dealing with individual patterns. This yields a consistent number of lactation curves with very poor fit. These results confirm the reported difficulty for non-linear models to describe individual patterns. Nevertheless, the DF is able to calculate the asymptotic level of production that is of help for classifying extended lactation shape and could provide useful technical information. The study of genetic basis of these traits may be of great interest for being included as possible goals in breeding programs aimed at modifying the lactation curve shape in an economically desirable direction.

Patterns found for extended lactations in the first phase of production do not differ from those observed in standard lactations, with level and occurrences of peak yield, which did not show substantial variation among lactation length. The second phase of production in lactations exceeding 400 to 500 days tends to an asymptotic level, especially for older cows. These results, confirming previous reports on dairy cattle obtained in other countries, seem to indicate the extension of lactation length as a valid management strategy.

Acknowledgements

Authors wish to acknowledge Prof. Cappio Borlino and the two anonymous referees for their helpful comments and suggestions.

References

Ali TE and Schaeffer LR 1987. Accounting for covariances among test day milk yields in dairy cows. Canadian Journal of Animal Science 67, 637–644.

Auldist MJ, O'Brien G, Cole D, Macmillan KL and Grainger C 2007. Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. Journal of Dairy Science 90, 3234–3241.

Bertilsson J, Svennersten-Sjaunja K, Wiktorsson H, Berglund B and Ratnayake G 1997. Optimising lactation cycles for the existing high-yielding dairy cow. A European perspective. Livestock Production Science 50, 5–13.

Brotherstone S, Thompsonb R and Whitea IMS 2004. Effects of pregnancy on daily milk yield of Holstein–Friesian dairy cattle. Livestock Production Science 87, 265–269.

Butler ST, Shalloo L and Murphy JJ 2010. Extended lactations in a seasonal-calving pastoral system of production to modulate the effects of reproductive failure. Journal of Dairy Science 93, 1283–1295.

Cole JB and Null DJ 2009. Genetic evaluation of lactation persistency for five breeds of dairy cattle. Journal of Dairy Science 92, 2248–2258.

Dematawewa CMB, Pearson RE and VanRaden PM 2006. Modeling extended lactations in Holsteins. Journal of Dairy Science 89 (Suppl. 1), 96–97.

Dematawewa CMB, Pearson RE and VanRaden PM 2007. Modeling extended lactations of Holsteins. Journal of Dairy Science 90, 3924–3936.

Dijkstra J, France J, Dhanoa MS, Maas JA, Hanigan MD, Rook AJ and Beever DE 1997. A model to describe growth patterns of the mammary gland during pregnancy and lactation. Journal of Dairy Science 80, 2340–2354.

Dobson H, Smith RF, Royal MD, Knight CH and Sheldon IM 2007. The high-producing dairy cow and its reproductive performance. Reproduction in Domestic Animals 42, 17–23.

Erb HN, Smith RD, Oltenacu PA, Guard CL, Hillman RB, Powers PA, Smith MC and White ME 1984. Path model of reproductive disorders and performance, milk fever, mastitis, milk yield, and culling in Holstein cows. Journal of Dairy Science 68, 3337–3349.

Gonzalez-Recio O, Perez-Cabal MA and Alenda R 2004. Economic value of female fertility and its relationship with profit in Spanish dairy cattle. Journal of Dairy Science 87, 3053–3061.

Grossman M and Koops WJ 2003. Modeling extended lactation curves of dairy cattle: a biological basis for the multiphasic approach. Journal of Dairy Science 86, 988–998.

Haile-Mariam M and Goddard M 2008. Genetic and phenotypic parameters of lactations longer than 305 days (extended lactations). Animal 2, 325–335.

Knight H 2005. Extended lactation: turning theory into reality. Advances in Dairy Technology 17, 113–124.

Lucy MC 2001. Reproductive loss in high-producing dairy cattle: Where will it end? Journal of Dairy Science 84, 1277–1293.

Macciotta NPP, Vicario D and Cappio-Borlino A 2005. Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. Journal of Dairy Science 88, 1178–1191.

Macciotta NPP, Miglior F, Dimauro C and Schaeffer LR 2010. Comparison of parametric, orthogonal, and spline functions to model individual lactation curves for milk yield in Canadian Holsteins. Italian Journal of Animal Science 9, 460–464.

Muir BL, Kistemaker G and Van Doormaal BJ 2004. Estimation of genetic parameters for the Canadian Test Day Model with Legendre polynomials for Holsteins based on more recent data. Report to the Dairy Cattle Breeding and Genetics Committee and the Genetic Evaluation Board, March 2004. Retrieved September 8, 2011, from http://cgil.uoguelph.ca/

Muir BL, Kistemaker G, Jamrozik J and Canavesi F 2007. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. Journal of Dairy Science 90, 1564–1574.

Olori VE, Brotherstone S, Hill WG and McGuirk BJ 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. Livestock Production Science 58, 55–63.

Reinhardt F, Liu Z, Bünger A, Dopp L and Reents R 2002. Impact of application of a random regression test-day model to production trait genetic evaluations in dairy cattle. In Proceedings of the Interbull Meeting 27, Interlaken, Switzerland, pp. 103–108.

Rook AJ, France J and Dhanoa MS 1993. On the mathematical description of lactation curves. Journal of Agricultural Science 121, 97–102.

Schaeffer LR 2004. Application of random regression models in animal breeding. Livestock Production Science 86, 35–45.

Schaeffer LR, Jamrozik J, Kistemaker GJ and Van Doormaal BJ 2000. Experience with a test day model. Journal of Dairy Science 83, 1135–1144.

Silvestre AM, Petim-Batista F and Colaco J 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sampling schemes. Journal of Dairy Science 89, 1813–1821.

Stanton TL, Jones LR, Everett RW and Kachman SD 1992. Estimating milk, fat, and protein lactation curves with a test day model. Journal of Dairy Science 75, 1691–1700.

VanRaden PM, Dematawewa CMB, Pearson RE and Tooker ME 2006. Productive life including all lactations and longer lactations with diminishing credits. Journal of Dairy Science 89, 3213–3220.

Vargas B, Koops WJ, Herrero M and van Arendonk JAM 2000. Modeling extended lactations of dairy cows. Journal of Dairy Science 83, 1371–1380.

Veerkamp RF, Emmans GC, Cromie AR and Simm G 1995. Variance components for residual feed intake in dairy cows. Livestock Production Science 41, 111–120.

Wilmink JBM 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Livestock Production Science 16, 335–348.

Wood PDP 1967. Algebraic models of the lactation curves in cattle. Nature 216, 164–165.

Wood PDP 1968. Factors affecting persistency of lactation in cattle. Nature 218, 894.