

REVIEW ARTICLE

The mathematical description of lactation curves in dairy cattle

Nicolò P.P. Macciotta, Corrado Dimauro, Salvatore P.G. Rassu, Roberto Steri, Giuseppe Pulina

Dipartimento di Scienze Zootecniche, Università di Sassari, Italy

Abstract

This review gives an overview of the mathematical modelling of lactation curves in dairy cattle. Over the last ninety years, the development of this field of study has followed the main requirements of the dairy cattle industry. Non-linear parametric functions have represented the preferred tools for modelling average curves of homogeneous groups of animals, with the main aim of predicting yields for management purposes. The increased availability of records per individual lactations and the genetic evaluation based on test day records has shifted the interest of modellers towards more flexible and general linear functions, as polynomials or *splines*. Thus the main interest of modelling is no longer the reconstruction of the general pattern of the phenomenon but the fitting of individual deviations from an average curve. Other specific approaches based on the modelling of the correlation structure of test day records within lactation, such as mixed linear models or principal component analysis, have been used to test the statistical significance of fixed effects in dairy experiments or to create new variables expressing main lactation curve traits. The adequacy of a model is not an absolute requisite, because it has to be assessed according to the specific purpose it is used for. Occurrence of extended lactations and of new productive and functional traits to be described and the increase of records coming from automatic milking systems likely will represent some of the future challenges for the mathematical modelling of the lactation curve in dairy cattle.

Introduction

Why modelling the lactation curve?

The mathematical representation of milk production during the lactation period repre-

sents one of the most successful applications of mathematical modelling in agriculture (France and Thornley, 1984). Tools that can mimic fundamental processes of milk production in different scenarios are of great help for physiologists, nutritionists and geneticists for studying and testing hypothesis on the behavior of the mammary gland machinery. Models able to forecast future milk yields supply useful information for management decisions dealing with time: lactation curve functions are currently implemented in dairy farm management softwares (De Vries, 2006). At cow level, lactation curve modeling is of help for monitoring individual yields for diet planning, early detection of diseases before the appearance of clinical signs and for selecting animals to be culled (Gipson and Grossman, 1989). Beside management, breeding is another important field of application. Lactations in progress of daughters of young bulls running progeny test can be projected up to 305-d in order to anticipate first crop evaluation (Schaeffer and Jamrozik, 1996). Finally, lactation curve functions are used to fit additive genetic effects in random regression test day models (Schaeffer, 2004).

Since the early work of Brody *et al.* (1923) who proposed a decreasing exponential function to fit the declining phase of the lactation pattern in dairy cattle, studies on lactation curve modelling have experienced a rather waving pattern. Periods of great scientific contribution and discussion have been followed by relative decreases of interest. The development of the different approaches has been driven by specific needs of the dairy industry, by the scientific and technological achievements of lactation physiologists that provided the biological and theoretical foundations, and by computational sciences that supplied the mathematical tools.

Main features of the variable to be modelled

Milk production pattern along the lactation is the result of physiological processes of synthesis and captation from the blood stream by specialized epithelial cells of the mammary gland (Mephram, 1987). The core process of milk production is represented by the variation of mammary epithelial cell number and secretion activity (Capuco *et al.*, 2003). Cell number variation is the result of two events: cell proliferation, that starts in early pregnancy reaching a maximum immediately after parturition; cell remodelling and apoptosis that begins in early lactation and lasts till the dry-off (Hurley, 1989; Knight and Wilde, 1993). For a healthy cow

Corresponding author: Prof. Giuseppe Pulina, Dipartimento di Scienze Zootecniche, Università degli Studi di Sassari, via E. De Nicola 9, 07100 Sassari, Italy.

Tel. +39.079.229207 - Fax: +39.079.229302.

E-mail: gpulina@uniss.it

Key words: Lactation curve, Mathematical models, Lactation curve traits, Dairy cows.

Acknowledgments: the authors would like to thank Ana Helena Dias Francesconi for revising and editing the manuscript.

Received for publication: 15 July 2011.

Revision received: 14 October 2011.

Accepted for publication: 14 October 2011.

This work is licensed under a Creative Commons Attribution NonCommercial 3.0 License (CC BY-NC 3.0).

©Copyright N.P.P. Macciotta *et al.*, 2011

Licensee PAGEPress, Italy

Italian Journal of Animal Science 2011; 10:e51

doi:10.4081/ijas.2011.e51

within a given milking regime, the variation of the cell secretory activity appears to be very little, although studies on this aspect of lactation are very limited (Knight *et al.*, 1998).

The phenotypic expression of these biological processes is represented by the so called standard shape of the lactation curve (Figure 1). It is characterized by a first ascending phase from parturition till a maximum, the lactation peak, followed by a second declining slope that ends with the dry off of the animal. The curve depicted in Figure 1 represents the pattern that can be reconstructed by using the standard availability of data (usually one daily record per month, for an average of 8-10 per lactation).

Basic traits of the curve are represented by the rate of increase of milk production in the first phase, time at peak occurrence, peak production and the rate of decline of milk yield in the second phase. The term *lactation curve* usually refers to the above mentioned deterministic, regular and continuous pattern depicted by the solid line in Figure 1. However, a second stochastic and less predictable component, represented by individual deviations from the standard shape, exists. It can be ascribed to several factors as genetic variation between animals, feeding, farming system, environmental conditions and health status of the animal (Macciotta *et al.*, 2008a).

Lactation curve shapes different from the standard may occur, especially when individual patterns are fitted: i) a reversed shape, with an

initial decreasing phase to a minimum followed by an increase, that is common for fat and protein contents; ii) a continuously increasing curve; a continuously decreasing pattern without the lactation peak i.e., the so-called atypical curve for milk yield (Congleton and Everett, 1980). The most frequent (about 20-30% of cases) is the atypical shape (Macciotta *et al.*, 2005; Olori *et al.*, 1999; Rekik and Ben Gara, 2004). Actually, it represents mainly a computational issue due to the interaction between the mathematical structure of the model used and the combinations of test day values and distribution along the lactation trajectory (Congleton and Everett, 1980; Macciotta *et al.*, 2005). In the most common dairy recording systems the milk is recorded once a month and peak yield can be easily missed. When daily data are available, as in the case of automatic milking systems, the occurrence of atypical curves is expected to be markedly smaller. However, a role of the biological differences among cows should not be excluded. In fact, an effect of calving season and parity on atypical curves occurrence has been reported (Rekik and Ben Gara, 2004). Another exception to the standard shape is represented by the existence of a second lactation peak in cows calving in autumn in pasture-based farming systems (García and Holmes, 2001). In this case, the variation in shape has to be ascribed mainly to seasonal effects, in particular to improved pasture availability in spring.

Describing and/or explaining the milk production pattern: empirical vs mechanistic approach. Does this differentiation make sense?

Mathematical models are a subclass of theoretical models, i.e. those aimed at representing a specific sector of reality based on its basic properties and relationships. When system features are represented by quantities, the model becomes a mathematical model. In applied biology, models are often classified into empirical or mechanistic. An empirical model has a theory that refers just to the level of reality at which the considered phenomenon is expressed, whereas a mechanistic model is characterized by deeper theoretical assumptions. Thus, in the former the descriptive component is prevalent whereas in the latter the explicative aspects are the most important. Such a basic taxonomy is sometimes misused

in agriculture, as in the specific case of lactation curve modelling. It is often assumed that empirical models are just a mere quantitative representation of the phenomenon, without any theoretical assumption, whereas all the theory is reserved to mechanistic models. Actually, some authors pointed out that a rigid distinction cannot be made (Nestorov *et al.*, 1999; Vetharanim *et al.*, 2003) since each model, also the mathematically simplest and description-oriented, is based on theoretical assumptions.

A part from the theoretical foundation, that should be clearly stated, other features of a mathematical model of the lactation curve are: i) a specific mathematical structure; ii) a well defined field of application; iii) a set of potentialities but also defined borders, out of which the model can be not useful or is even misleading.

Modelling the lactation curve

Most of mathematical functions proposed to fit lactation patterns in dairy cattle are mainly aimed at describing the phenomenon. Their basic assumption is that lactation is characterized by a continuous and deterministic component with an increasing phase till a maximum followed by a decreasing slope. The mathematical tool used in this approach is represented by an analytical function of time:

$$y_t = f(t)$$

where y_t is the daily milk production recorded at time t . Table 1 reports some of the most popular mathematical functions used to fit lactation curve data. An increase in the number of parameters, a shift from non-linear to linear

equations and from models specifically conceived to fit the standard shape of the lactation curve to more general functions can be observed across years (Table 1). Such an evolution has been probably driven by several reasons, as advancements in computing or larger availability of data, although the most relevant is represented by the specific requirements of the dairy cattle industry.

Early models paid more attention to the deterministic component of the lactation pattern, being essentially aimed at describing average lactation curves of homogeneous groups of animals for management purposes. An efficient model was therefore asked to disentangle the general frame of the process from environmental perturbations and to predict milk yield with good accuracy. Due to the fairly large number of records available per average curves and to the consequent regular pattern, non-linear estimation procedures could be quite easily implemented.

The Wood incomplete gamma function (Wood, 1967) is probably the most popular empirical model of the lactation curve. It generates the standard shape of the lactation

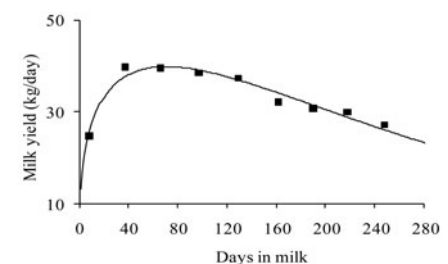


Figure 1. Standard shape of the lactation curve for dairy cattle.

Table 1. Some of the most popular empirical models used to fit the lactation curve in dairy cattle.

Function	Parameters	Author
$y = ae^{-bt}$	2	Brody <i>et al.</i> (1923)
$y_t = at^b e^{ct}$	3	Wood (1967)
$y_t = ae^{-bt} - ae^{ct}$	3	Cobby and Le Du (1978)
$y_t = a + be^{kt} + ct$	4	Wilmink (1987a)
$y_t = a_0 + a_1t + a_2t^2 + a_3 \log(1/t) + a_4(\log(1/t))^2$	5	Ali and Schaeffer (1987)
$y_t = \sum_{i=1}^n [a_i b_i [1 - \tanh^2(b_i(t - c_i))]]$	3 per phase	Grossman and Koops (1988)
$y_t = \sum_{j=1}^m \alpha_j P_j$	5	Brotherstone <i>et al.</i> (2000)
$y_t = \sum_{i=1}^n b_i \alpha_i \cdot t^{i-1} + \sum_{j=1}^m b_j \cdot (t - t_j)^n$	14	White <i>et al.</i> (1999)

y_t , daily milk production at time t from parturition; a, b, c , function parameters.

curve as the product of a constant, a power function and an exponential decay function (Figure 2). The partitioning of Wood equation into its components underlines the direct relationship of its parameters with main features of lactation curve shape (Table 2): *a* has a scale meaning and tends to increase with parity, ie. according to the production level of animals; *b* controls the rate of increase to the lactation peak; finally, *c* expresses the rate of decline after the peak and its absolute value tends to increase with parity. Moreover, Wood parameters can be used to estimate lactation curve traits, time at peak (*t_m*), peak yield (*y_m*) and lactation persistency (*p*):

$$t_m = -\frac{b}{c} \quad y_m = \frac{a}{\left(\frac{c}{b}\right) \exp^b} \quad p = -(b+1) \ln c$$

Values for lactation curve traits reported in Table 2 confirm some assessed results about relationships between lactation curve characteristics and parity number in cattle: first calving cows show a later peak occurrence, a lower peak production and a higher persistency of lactation, compared to later parities.

Limitations of the Wood model are also well known. Overestimation of daily milk yield in the first part of the curve, underestimation around and after the peak have been reported by several authors (Congleton and Everett, 1980; Sherchand *et al.*, 1995; Dematawewa *et al.*, 2007; Dijkstra *et al.*, 2010). Essentially, due to its multiplicative structure, the model is characterized by a high degree of correlation among parameters (ranges 0.70-0.90, Macciotta *et al.*, 2005) that results in a marked rigidity and in a great sensitivity to data distribution (Silvestre *et al.*, 2006). On the other hand, the model possesses some other interesting features as the ability to fit four different shapes including the atypical curves. Several modifications of the incomplete gamma function have been proposed to improve the ability of describing lactation peaks close to parturition (Cappio-Borlino *et al.*, 1995) or to adequately account for seasonal variations (Grossman *et al.*, 1986). In spite of assessed limitations, the Wood model is still the most widely used function for modelling lactation curves (Dijkstra *et al.*, 2010). Moreover it has been used to describe traits other than milk yield as for example milk fatty acid pattern (Craninx *et al.*, 2008), and has been used also for emerging issues, as the fitting of extended lactations (Dematawewa *et al.*, 2007). Around the 1980's, some important changes occurred in the dairy cattle industry.

Individual feeding, health status monitoring, development of farm management softwares and, later on, the use of daily records instead of cumulated lactation yields for genetic evaluations, put emphasis on individual curve modelling. Mean lactation patterns of large groups of animals usually have a shape that is sufficiently regular to be fitted with a parametric function. On the contrary, large variation that occurs in individual curve shapes requires larger flexibility, easy computation, and, specifically for genetic models, an easy implementation in mixed model framework. (Olori *et al.*, 1999; Verbyla *et al.*, 1999; White *et al.*, 1999). Furthermore, the huge individual variation and the reduced number of records per subject (8-10 tests) resulted in a difficult use of non linear estimation procedures and thus linear or linearizable algorithms have represented an almost obliged choice, at least for routine implementations. The combined exponential and linear model of Wilmink (1987a) and the polynomial regression of Ali and Schaeffer (1987) (Table 1) can be regarded as a transition between early and newer models. The Wilmink function consists of three terms that are combined additively, thus enhancing flexibility. Moreover, it can be easily linearized by setting the *k* parameter to a suitable fixed value (Brotherstone *et al.*, 2000) and its parameters still maintain a relationship with lactation curve shape. The Ali and Schaeffer model has a larger number of coefficients that allow for fitting a wider range of shapes although its parameters have no longer a technical meaning. Both models have been successfully used to fit individual curves (Macciotta *et al.*, 2005; Silvestre *et al.*, 2006;

Olori *et al.*, 1999) and implemented in the earlier versions of random regression models (Druet *et al.*, 2003; Schaeffer *et al.*, 2000; Schaeffer, 2004). These two functions have been also used for modelling traits other than milk yield as, for example, dry matter and water intake in Holsteins (Kramer *et al.*, 2009) and to estimate gene effect on dairy traits (Strucken *et al.*, 2011). Although they usually outperform the Wood function, especially in different scenarios of data distribution (Silvestre *et al.*, 2006), these two models tend to yield mathematical artifacts such as negative or too high predicted values of milk yield at the beginning or at the end of lactation (Druet *et al.*, 2003; Macciotta *et al.*, 2005; Silvestre *et al.*, 2006). Above mentioned failures in describing specific stages of the lactation trajectory can be analyzed from different standpoints. First of all, it should be pointed out that less records are usually available at the edges of the lactation compared to middle stages. Under these conditions, shortcomings of

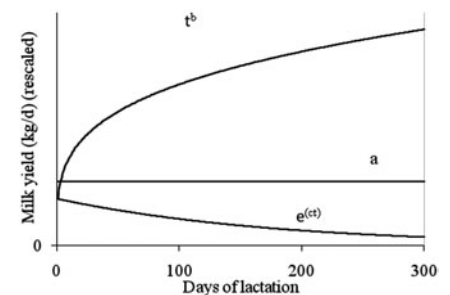


Figure 2. Decomposition of the Wood incomplete gamma function.

Table 2. Parameters and lactation curve traits for dairy cattle of different parities estimated with the Wood model (lactation traits have been calculated on the basis of parameter values reported in the articles).

Parity	a	b	c	y _m , kg/d	y _t ,d	Persistency	Author
First	23.7	0.12	0.061	22.8	58	3.13	Kellogg <i>et al.</i> (1977)
Second	35.7	0.25	0.160	31.1	47	2.29	
Third	39.1	0.32	0.186	33.8	52	2.22	
First	12.6	0.17	0.002	22.6	85	7.27	Val-Arreola <i>et al.</i> (2004)
Second	19.1	0.15	0.004	28.3	37.5	6.35	
Third or >	15.1	0.21	0.005	26.8	42	6.41	
First	15.7	0.21	0.002	33.5	104	7.51	Dematawewa <i>et al.</i> (2007)
Third or >	23.1	0.21	0.004	43.9	57	6.82	
First	13.89	0.25	0.004	31.2	69	7.02	Rekik and Ben Gara (2004)
Second	17.46	0.24	0.005	35.6	53	6.69	
Third	19.56	0.23	0.005	37.9	48	6.57	
First	13.01	0.27	0.003	33.5	89	7.34	Cole and Null (2009)
Second or >	22.01	0.22	0.004	42.6	54	6.74	

y_m, peak production; y_t, peak occurrence.

extrapolating relationships between two variables beyond their known range of validity may occur when models are fitted (Mead *et al.*, 2002). On the other hand, Grossman and Koops (1988) stressed a further theoretical issue, common to several lactation models, represented by the consideration of the whole lactation as a single process. Starting from the assumption that lactation is the sum of different overlapping phases, these authors developed a model by adding different logistic functions (Table 1). The diphasic model, that fits two lactation phases described by two linear and two quadratic logistic functions, has been indicated as the most suitable for describing mean lactation curves for milk yield. Estimated theoretical durations of the two overlapping phases were of approximately 200 and 410 days. In general, multiphasic models are characterized by a large number of parameters (three for each phase) and therefore they require a greater number of tests to be conveniently fitted, as in case of average curves and extended lactations (Dematawewa *et al.*, 2007; Vargas *et al.*, 2000).

Looking for extra-flexibility: polynomials and non-parametric functions

The need for fitting individual deviations from a mean pattern led animal scientists towards more flexible mathematical tools. A common criticism to parametric models, i.e. those whose coefficients show a direct relationship with curve traits, is their high sensitivity to data distribution which may markedly affect the whole shape of the curve. As possible alternatives to parametric models, general functions not specifically conceived to fit lactation curve shape have been proposed from the early 90'. Legendre orthogonal polynomials (LP) (Kirkpatrick *et al.*, 1990) or regression splines (White *et al.*, 1999) (Table 1) are characterized by a large number of parameters, a linear additive mathematical structure and a remarkable capacity to fit a great range of shapes. Theoretically, a LP of order n is able to describe $2n$ different shapes: for example, a fourth-order LP has been able to recognize 18 out of the 32 theoretical shapes, represented by specific deformations of the two basic shapes, standard and atypical (Macciotta *et al.*, 2005). Legendre polynomials are currently used to fit random effects (genetic additive and permanent environment) in random regression model. In any case they still show a remarkable sensitivity to

data distribution, especially at the edges of the lactation trajectory (border effects) (Druet *et al.*, 2003; López-Romero and Carabaño, 2003). Splines are a type of segmented regression in which the curve is divided into different intervals of the dependent variable, joined at points named knots, each fitted with different polynomials (Guo and White, 2005). Thus individual observations influence only a part of the function (Jamrozik *et al.*, 2010). Cubic splines regressions resulted a good compromise between fitting performances, data sensitivity, smoothness and parametrisation in fitting average curves (Druet *et al.*, 2003; Silvestre *et al.*, 2006). Moreover, they are able to fit local sudden variations as in the case of individual somatic cell pattern around an event of clinical mastitis (de Haas *et al.*, 2002). A technical issue in spline fitting is represented by the optimization of the number and location of knots. Some authors recommend that knots should be as many as possible, placed at points of maximum concentration of records (Silvestre *et al.*, 2006; Druet *et al.*, 2003; Misztal, 2006), even if such a criteria necessarily increases the number of function parameters. In several papers, knot number and position are fixed a priori, usually evenly spaced (Druet *et al.*, 2003; Hickson *et al.*, 2006; Silvestre *et al.*, 2006). As an alternative, knot positions can be estimated by considering them as additional independent variables in a non linear estimation procedure (Fadel, 2004; Macciotta *et al.*, 2010). Table 3 reports results of a non linear estimation of knot position for linear, quadratic and cubic splines fitted to average lactation curves of first and third calving Canadian Holsteins (Macciotta *et al.*, 2008c). Knots are located mainly in the first part of lactation, around the peak, i.e. where main curvatures occurs. Such non linear estimation, however, becomes problematic for higher order splines and large number of knots. A recent paper of Jamrozik *et al.* (2010) pointed out that

the optimal knot position varies markedly depending on the effect to be modelled (genetic, permanent, environmental), the trait and the population considered. If the main aim of empirical modelling is to disentangle the deterministic component of lactation pattern, the rush for extra flexibility can have a further theoretical implication. Figure 3 reports the lactation curve of a third calving cow fitted with the

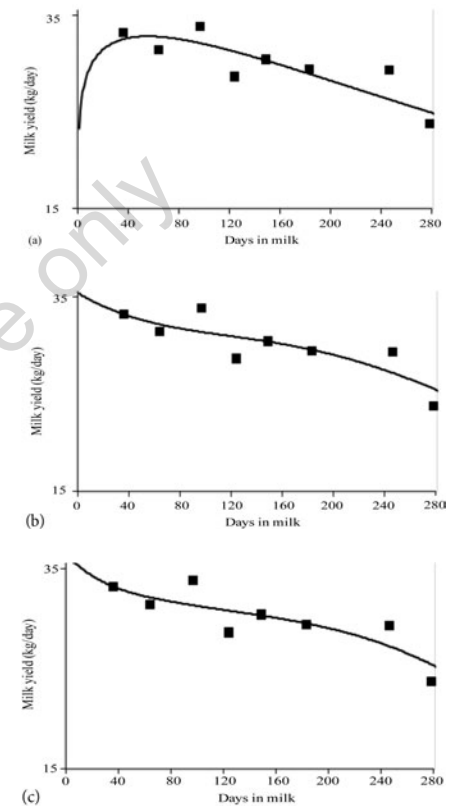


Figure 3. Lactation curve of a dairy cow fitted with the Wood model (a), a fourth order Legendre polynomials (b) and a cubic regression spline (c) with three knots placed at 16, 69 and 252 DIM, respectively.

Table 3. Knot positions (days in milking) and Mean Squared Error of average lactation for regression splines of different orders fitted to average lactation curves of cows of three parities by using a non linear regression procedure.

Spline order	No. of knots					
	1		2		3	
	Position	MSE	Position, d	MSE	Position, d	MSE
1 st parity						
Linear	37	0.09	35, 219	0.06	Not conv.	
Quadratic	46	0.02	17, 53	0.009	20, 59, 202	0.006
Cubic	62	0.01	21, 75	0.005	26, 57, 103	0.004
3 rd parity						
Linear	34	0.09	18, 41	0.03	Not conv.	
Quadratic	46	0.03	16, 51	0.02	11, 38, 59	0.01
Cubic	60	0.02	26, 67	0.01	16, 69, 252	0.007

Wood model (a), a fourth order Legendre polynomials (b) and a cubic regression spline (c) with three knots placed at 16, 69 and 252 days in milk (DIM), respectively. It can be noticed that the Wood function, although not able to fit all the local variation of test records, is able to reconstruct the shape of the curve. On the other hand, Legendre polynomials and splines are slightly efficient in fitting waves in the middle of lactation but their ability to disentangle the main traits of the phenomenon is rather questionable.

Evaluation of factors affecting curve shape: the use of mixed linear models

Variation in lactation curve shape could be caused by several external factors such as calving season, feeding, health status, climate (Dijkstra *et al.*, 2010). Under specific farming conditions, as in the case of pasture-based systems of temperate regions, seasonal and physiological effects on lactation curve shape are confounded (García and Holmes, 2001). In the functional approach, effects of environmental factors are evaluated either by estimating average curves of cows grouped according to sources of variation (i.e. different calving seasons or areas of production) or by comparing parameter values and (co)variances of individual curves (Wood, 1976; Shanks *et al.*, 1981; Ferris *et al.*, 1985; Strucken *et al.*, 2011). However, results of these analyses have to be carefully handled, considering the great variation that characterizes parameter values estimated on individual data (Olori *et al.*, 1999) and the different meaning of the parameters within different curve shapes (Macciotta *et al.*, 2005).

Moreover, an implicit assumption of continuous functions of time is that environmental factors average out over lactation. This is correct if effects are the same for all test day records but it is well known that some factors of variation, such as pregnancy status, weather conditions, change of management groups, are not constant throughout the lactation (Jamrozik and Schaeffer, 1997). Mixed linear models represent a mathematical tool more suitable for evaluating environmental effects, being able to account for factors that could affect each TD differently. The basic structure of mixed linear models used to fit daily milk yields along the lactation trajectory, usually named as test day models, is the following:

$$y = \text{HTD} + F + \text{DIM} + L + e \quad [1]$$

where y is the daily milk yield, HTD is the interaction between the herd and the date of the test which accounts for the effects peculiar to a specific date; F represents some kind of fixed factors (i.e. calving season, region of production, parity); DIM is the fixed effect of days in milk classes, whose least squares solutions allow to generate lactation curves corrected for other effects included in the model (Stanton *et al.*, 1992); L is the random effect of the individual cow and it is associated to a variance component σ^2_L ; e is the random residual, associated at the variance component σ^2_e .

The use of levels of a fixed factor (DIM) to model lactation curve has some advantages. No previous assumptions on the shape of the curve are made and the influence of data is local because a specific TD record in a certain class does not affect the estimate for any other class (Druet *et al.*, 2003). Specific curves for different factors can be estimated by nesting the DIM class into the specific source of variation (Stanton *et al.*, 1992). For example, Figure 4A represents average lactation curves of dairy cattle of three different parities estimated with test day model in which the DIM factor is nested within parity. The difference with the continuous functional approach can be clearly observed in Figure 4B, that reports the same curves reconstructed by fitting the Wood model to daily yields averaged by parities.

The mathematical modelling of lactation curve by mixed linear models supplied the theoretical frame for genetic evaluations based on daily yields developed during the 90's. Structure [1] with the animal (L) effect split into genetic additive and permanent environment, represents the basic form of the test day repeatability models used for dairy cattle (Ptak and Schaeffer, 1993). The further inclusion of continuous functions (as the Wilmink model or LP polynomials) to fit random additive and environmental effects resulted in the development of random regression models (Schaeffer, 2004). Pregnancy has a negative effect on milk production in dairy cattle. When comparing cows managed for 12 or 18 calving intervals (i.e. with pregnancies occurring at about 85 and 250 DIM, respectively), Bertilsson *et al.* (1997) found similar milk production patterns in the first 24-32 weeks. Beyond this period, the milk production of cows at the 12-month cycle tended to drop quickly, whereas that of cows at the 18-month cycle maintained a constant slope. Different lactation curve shapes have also been reported by Brotherstone *et al.* (2004) for cows within different days open period. The lactation curve patterns for preg-

nant cows at different DIM and non-pregnant ones were similar until around 200 DIM. Afterwards, between 200 and 250 DIM, the decrease of production was more rapid in cows with short calving interval than in those with longer calving intervals. This confirms the reduction in milk yield and the change in milk composition that occur around month 6 of pregnancy. Indeed, it has been estimated a reduction of about 5 kg of ECM/d at day 305 of lactation for cows which had become pregnant at 85 DIM compared to those pregnant at 225 DIM (Atzori *et al.*, 2010).

A part from genetic evaluations, the use of mixed linear models to fit the lactation curve has had other applications. One is represented by the prediction of daily milk yields from a limited number of recorded TDs (Pool and Meuwissen, 2000; Vasconcelos *et al.*, 2004). Mixed models are also used in longitudinal studies carried out on dairy animals (Tempelman, 2009), where they are commonly used to test research hypothesis and to evaluate effects of feeding treatments, genotypes at candidate loci, health status on milk production traits (Wilson *et al.*, 2004; Macciotta *et al.*, 2008b; Smith *et al.*, 2009).

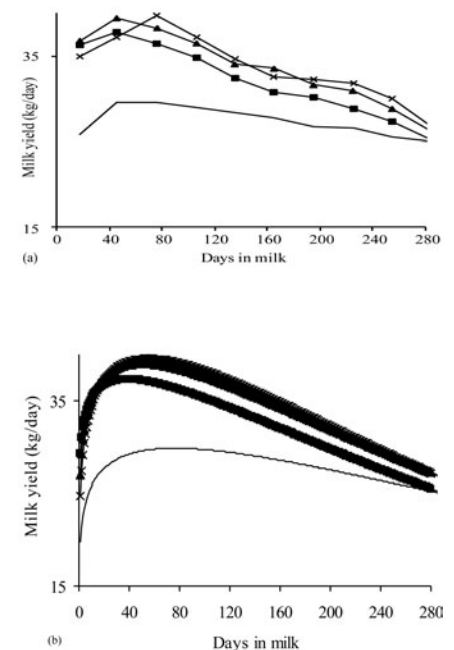


Figure 4. Average lactation curves of dairy cattle of three different parities estimated with test day model in which the DIM factor is nested within parity (a) or reconstructed by fitting the Wood model to daily yields averaged by parities (b).

Exploring the correlation structure among test day records along the lactation

From a statistical standpoint, daily milk records taken at different time intervals from parturition are a case of repeated measures design. In particular, they can be defined as a split-plot in time design (Diggle *et al.*, 1994): the animal is the whole plot and time intervals at which milk measurements are taken represent the sub-plots. Such a data frame is characterized by a complex (co)variance structure within and between subjects, as depicted in Table 4. Milk yields close in time are more correlated than those far apart in time. Such a covariance pattern can be explained, at least in part, by common environmental factors that affect adjacent tests as weather conditions and feed (Ali and Schaeffer, 1987). Moreover, measures recorded on the same cow are more correlated than those recorded on different animals because they share a common contribution from the same individual.

Variances express a rather waving behaviour, with highest values at the lactation edges. This pattern has been frequently highlighted in genetic (co)variance functions estimated with random regression models on milk yield (Jamrozik *et al.*, 1997; Jaffrezic *et al.*, 2000). Correlations among TD pairs at different time distance show a decreasing pattern, although values lagged the same number of observation apart show different values. Actually, heterogeneity of variance and serial correlation are common features of longitudinal data (Littell *et al.*, 1998; Núñez-Antón and Zimmerman, 2000).

The complex covariance structure among test day yields must be adequately modelled because it may affect the rate of type I and II errors of tests for fixed effects (Littell *et al.*, 1998; Wang and Goonewardene, 2004). Variances and correlations reported in Table 4 are estimated without assumptions of the underlying model. Such an unstructured (co)variance requires the estimation of $t(t+1)/2$ parameters, being t the number of time intervals considered. Several models have been used to fit adequately the data reducing at the same time model parametrisation. The most common structure, known as compound symmetry (CS), assumes constant variances and covariances of test day records along the lactation. Such a structure implies the estimation of 2 (co)variance parameters: σ^2_L (between subjects) and σ^2_e (within subjects). The individual lactation variance can be used

for calculating of the average correlation, or repeatability, among TD records within lactation as $(\sigma^2_L / (\sigma^2_L + \sigma^2_e))$. Repeatability values for daily milk yield in cows have been estimated to be between 0.3-0.6 (Stanton *et al.*, 1992; Macciotta *et al.*, 2008b). These figures are of the same magnitude of repeatabilities estimated with BLUP models that use the pedigree relationship matrix to structure the genetic covariance between animals. Actually in this approach the σ^2_L component could be approximately considered as the sum of the genetic effect and the permanent environmental effect.

Some authors have proposed the use of first order autoregressive structure (AR1) to model permanent environmental variance along the lactation using test day models (Carvalho *et al.*, 1998; Wilson *et al.*, 2008). In this structure the variance is constant whereas the correlation between two observation lagged k is ρ^k . Núñez-Antón and Zimmerman (2000) indicated the ante-dependence structure, that fits both heterogeneous variances and covariances, as the best model for milk production data. Their results are in agreement with those obtained by Wang and Goonewardene (2004) on growth curves of steers. In Table 5 are reported (co)variance parameters estimated on the same data set of Table 4 but with the above mentioned (co)variance structures together with two popular statistics of good-

ness of fit, the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

As far as the goodness of fit is concerned, the best value (i.e. the lowest) of the AIC is for the UN structure because, obviously, it gives the closest representation of data. The ante-dependence structure gets the best score if the BIC criterion, more penalizing for the number of parameters, is considered. The value of the autoregressive parameter reported in Table 5 is of the same order of magnitude of those reported by Vasconcelos *et al.* (2004) for Portuguese Holstein cattle.

The correlation structure between TD pairs along the lactation has been largely exploited for prediction purposes. Methods currently used to predict total lactation yields for genetic evaluation purposes are based on correlation between tests in different stages of lactation (Norman *et al.*, 1999). Table 6 reports correlation between actual and predicted lactation yields in first parity cattle by using different number of tests available and different approaches. Methods considered are a multiple trait approach based on Wood's model (WMTP) (Schaeffer and Jamrozik, 1996), an autoregressive moving average model (ARMA) (Macciotta *et al.*, 2002) and an autoregressive test day model (ATD) (Vasconcelos *et al.*, 2004). It can be seen that all approaches are able to predict cumulated yields highly corre-

Table 4. Sample variances (kg²), on the diagonal, and correlations, off diagonal, of test day records for milk yield along the lactation estimated with a mixed linear model without any assumption of (co)variance structure.

38.14										
0.58	24.61									
0.57	0.66	19.82								
0.44	0.55	0.65	20.05							
0.43	0.50	0.64	0.69	19.66						
0.33	0.50	0.57	0.74	0.77	17.64					
0.20	0.34	0.44	0.66	0.66	0.80	19.61				
0.13	0.29	0.44	0.59	0.58	0.68	0.76	19.65			
0.01	0.24	0.35	0.46	0.62	0.65	0.67	0.82	18.67		
0.02	0.13	0.25	0.39	0.53	0.48	0.50	0.59	0.64	24.46	

Table 5. Comparison between different (co)variance structure used in mixed models to fit milk test day yields using Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Covariance structure	Covariance Parameter	N	Estimate	AIC	BIC
Unstructured		55	See Table 4	11051.9	11281.2
Compound symmetry	σ^2_L	2	13.9505	11318.2	11326.5
	σ^2_e		7.996		
First order autoregressive	σ^2	2	21.4145	11200.0	11208.4
	ρ		0.7679		
Ante-dependence		19		1124.4	11203.7

lated with actual yields (>0.90) also when two or three tests are available at the beginning of lactation.

The correlation structure between milk yield test records has been investigated also using multivariate statistical techniques. The marked difference observed between values of partial and Pearson correlations can be interpreted as an evidence of an underlying latent structure of TD records (Macciotta *et al.*, 2004). The principal component analysis performed on the correlation matrix of TD records is able to extract two new variables, able to explain about 90% of the original variance, whose meaning can be inferred by structure of the correspondent eigenvectors (Figure 5). The first component, explaining 75% of the original variance, is equally related to all milk tests whereas the second (about 15% of the original variance) has negative relationship with the first and positive with the second part of the lactation respectively. These new variables can be used as lactation curve traits, expressing average lactation yield and lactation curve shape, respectively (Macciotta *et al.*, 2006; Wilmink, 1987b) The same result has been also obtained at genetic level, by performing eigenvalue decomposition on genetic covariance matrix of TD (Van der Werf *et al.*, 1998; Druet *et al.*, 2003) or the genetic (co)variance matrix of random regression test day models (Jamrozik *et al.*, 2002). A part from supplying new variables that may be of help for the genetic improvement of lactation curve traits, results of the multivariate analysis that underline the existence of a latent structure made of two latent variable seems to give an indirect confirmation of the hypothesis of Grossman and Koops (1988) of a diphasic structure of the lactation pattern.

Extended lactations, a challenge for lactation curve modelling

During the last decades, an increase in the average milk yield per lactation in specialized dairy breeds has occurred due to genetic selection and improved management practices (Lucy, 2001; Sorensen *et al.*, 2008). One side effect of the assessed strategy to increase milk yield by maximizing lactation peak and minimizing calving intervals (Grossman and Koops, 2003) has been the modification of the lactation curve shape. Actually high producing breeds tend to have higher lactation peaks, a later peak occurrence and show high persis-

tency (Table 7).

Production improvement, however, has resulted in an increase of fertility and health problems. The period around lactation peak is of a high risk for culling due to health problems (Bohmanova *et al.*, 2009) and the negative energy balance affects reproductive efficiency (Friggens *et al.*, 2007). A straightforward consequence is the increase of average lactation length observed in many countries, with more than 50% of cows exceeding the 305 d period (Vargas *et al.*, 2000; González-Recio *et al.*, 2006; Cole *et al.*, 2009).

Till the 90's, modelers paid attention almost exclusively to lactations developed within the standard length of 305 days. Actually, the extension of the lactation period does not seem to alter basic curve traits. No substantial differences in peak yields and peak occurrence were observed between extended and standard lactations (VanRaden *et al.*, 2006; Dematawewa *et al.*, 2007; Steri *et al.*, 2009). Moreover, a high genetic correlation (>0.70) between milk yield in the first 5 month of lactation and production after 300 days has been reported (Haile-Mariam and Goddard, 2008). Individual patterns of extended lactations usually show a relevant variability (Steri *et al.*, 2009), as reported for 305d curves. Figure 6 reports individual extended lactation curve patterns of Italian Holstein (Figure 6 A,B) and Italian

Simmental (Figure 6 C,D) cows. It can be observed the occurrence of both flat (Figure 6 A,C), standard (Figure 6B) and a sort of descending plus asymptotic (Figure 6D) patterns. It must be said that all these results have been obtained on extended lactations that have been caused essentially by fertility problems.

The adequacy of mathematical models to describe extended lactations has been evaluated essentially on fitting performances. Vargas *et al.* (2000) recommended the use of the diphasic function to fit extended lactations, even though they underline a relevant variability of parameter estimates. Grossman and

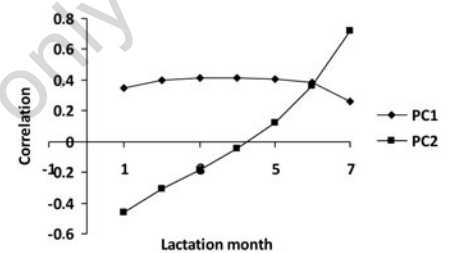


Figure 5. Plot of the eigenvalues of the first two principal components extracted from the correlation matrix of the TD records along the lactation (from Macciotta *et al.*, 2006).

Table 6. Correlations among actual and predicted lactation yields in first parity cows with different approaches and different plans of missing data.

No. of available tests	Method		
	ARMA	WMTP	ATD
2	0.85	0.80	0.86
3	0.92	0.87	0.91
4	0.96	0.92	0.94
5	0.98	0.95	0.96
6	0.99	0.97	0.98
7		0.98	0.99
8		0.99	0.99
9		0.99	0.99

ARMA, Autoregressive Moving Average Model; WMTP, Multiple Trait Model based on Wood function; ATD, Autoregressive Test Day Model.

Table 7. Lactation curve traits estimated in different cattle breeds.

Breed	ym ^o , kg	yt [†] , d	Persistency	Author
Barrosa	3.5 (5.3)	23 (56)	5.59	Silvestre <i>et al.</i> (2010)
Holstein x Zebu crosses	14.9	25	6.14	Cunha <i>et al.</i> (2010)
Guzerà	12.22	25	6.13	Cruz <i>et al.</i> (2009)
Sindi	11.80	22	6.17	Cruz <i>et al.</i> (2009)
Milking Shortorn	33.8	49	6.31	Cole and Null (2009)
Italian Simmental	27.3	29.3	6.41	Unpublished
Norwegian Red	30.2	35.3	-	Andersen <i>et al.</i> (2011)
US Holstein	43.9	59	6.82	Dematawewa <i>et al.</i> (2007)

^oYield at lactation peak; [†]days at lactation peak.

Koops (2003) developed a specific multiphasic approach, linking a first ascending phase of milk yield with a series of three descending phases of yield. A more recent comparison performed by Dematawewa *et al.* (2007) confirmed the good fitting performances of the multiphasic approach but also the computational problems related to the estimation of a high number of parameters. Authors concluded in favour of simpler models, as the Wood or the Rook (Table 1) function, at least for routine use.

Even though most of extended lactations are the result of reproduction failures, a voluntary delay of insemination may increase the profitability of cows both in conventional and pasture-based farming systems (Rotz *et al.*, 2005; Butler *et al.*, 2010). Thus, in this specific field of application, a suitable mathematical model should be able to support decision, for example, on the economical convenience in keep milking a cow. An example is represented by the use of a modified version of the Dijkstra model implemented by VanRaden *et al.* (2006) in the calculation of the lifetime profit for US Holstein cows:

$$y_t = \beta_0 + \beta_1 \cdot e^{\left[\frac{\beta_2(1 - e^{-k\beta_3 t})}{\beta_3} - \beta_4 t \right]}$$

In this formula, the β_0 parameter estimates a baseline that represents a sort of threshold for milk yield profitability. Thus, the suitability of the model for extended lactations depends not only on its general goodness of fit but also on the ability of describing the asymptotic phase that often occurs approximately after 350 DIM (Figure 6). Values of this parameter have been estimated in about 20 kg and 19 kg 12 for US and Italian Holsteins respectively (Steri *et al.*, 2009; VanRaden *et al.*, 2006).

Lactation curve modelling for milk composition

Lactation curves for dairy traits have received little attention compared to milk yield. However they represent an important aspect of milk production trait not only for milk production but also because milk constituents are related to the energy expenditure and has a relevant importance for dairy cattle management especially at the debut of lactation. A trait of interest is the fat:protein ratio proposed by Butcherit *et al.* (2010) as a trait able to differentiate between cows that can or cannot

adapt to the challenge of early lactation.

Also for composition traits the correlation between different stages of lactation show a progressively decreasing pattern (Caccamo *et al.*, 2008).

Conclusions

During about ninety years of documented career, the mathematical description of the lactation curve in cattle has played a fundamental role in supporting animal scientists and technicians to face several scenarios of data structure for different purposes. In general, model evolution has been characterized by a relative increase of the complexity of the mathematical structure of the algorithms used. A driven force has been probably represented by the increase of data available per individual pattern that has shifted the focus of modelling from average patterns to individual deviations. However, it is clear that a function suitable for all the fields of applications hardly exists and that model validity should be assessed on the basis of its sustainability for a particular purpose (Tedeschi, 2006). In any case, continuous functions with a limited number of parameters and specifically conceived for describing the standard shape of the lactation curve are still preferred in many cases.

The prediction of future challenges for lactation curve modelling is an hard task to perform. Most of literature deal with milk yield

whereas other conventional dairy traits as fat, protein and somatic cell have received little attention. However there have been examples of modelling of new traits related to milk nutritional quality or health status, as fatty acid composition and water intake (Craninx *et al.*, 2008; Kramer *et al.*, 2009), that are becoming of great importance for breeding and management strategies. Being difficult and expensive to be measured routinely, tools able to describe underlying patterns from few data and with a relevant predictive ability will be presumably required for their modelling.

The importance of the genetic modification of the lactation curve shape in an economically desirable direction has been pointed out by several researchers (Dekkers *et al.*, 1998; Togashi and Lin, 2003; Weller *et al.*, 2006). Cows with curves characterized by lower peaks and higher persistency are less exposed to health and fertility problems and can exploit efficiently cheaper feeds (Jakobsen *et al.*, 2002). Several models for measuring these traits have been proposed in literature but none of them has been widely accepted (Macciotta *et al.*, 2006; Rekaya *et al.*, 2001). Lactation curve traits are characterised by low to moderate values of heritability, even though a relevant variation could be observed in estimates reported in literature. Ali and Schaeffer, for example, reported h^2 values of 0.12, 0.17 and 0.23 for the three parameters of the Wood model, respectively, using a maximum likelihood multiple trait approach. Varona *et al.* (1998) for the same parameters found values of 0.43, 0.17 and 0.40 using a Bayesian proce-

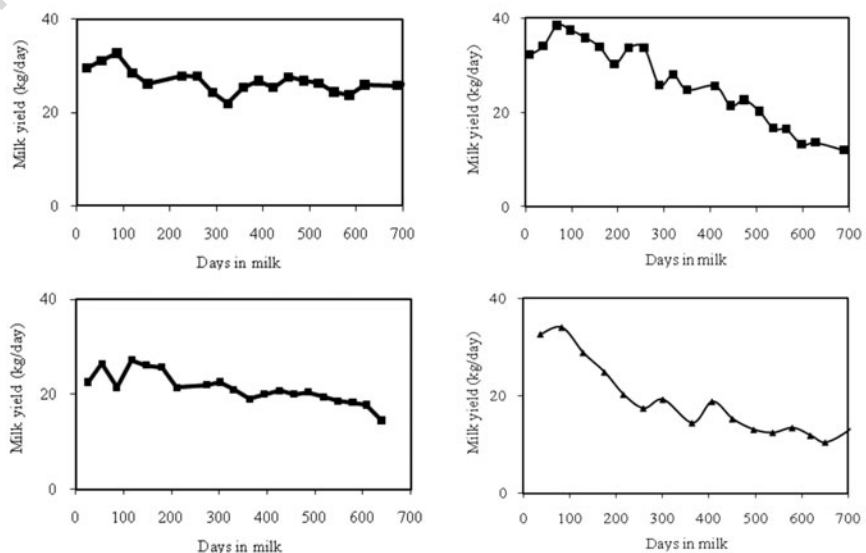


Figure 6. Patterns of extended individual lactations of Italian Holstein (up) and Simmental (down) cows.

ture. The issue is not only mathematical but it involves breeding strategies. The increased occurrence of extended lactations may offer a different perspective to the problem. Knight (2005) demonstrated that the milk yielded in three years with two (extended) or three (conventional) lactations is the same if persistency is improved of the 1% in extended lactations. Thus models able to give an early estimate of persistency in lactation in progress may represent useful tools for both breeding and management strategies.

A huge increase of daily records per animal is likely to occur due to the diffusion of automatic milking systems in cattle farming (Nixon *et al.*, 2009). The modelling of large amount of individual data will undoubtedly increase the amount of information available with positive consequences for management and breeding applications. However, modelling will be asked to address new issues: control of large variation within individuals, data correction for the great variation in milk intervals and frequencies between cows, evaluation of individual deviations from expected patterns for monitoring nutritional and health status. Some early papers used simple model as the Wilmink function (Nielsen *et al.*, 2010). However other non conventional methods, like time series analysis, with their ability to disentangle deterministic and periodic components from the white noise of random variation may represent a suitable alternative for addressing these issues (Deluyker *et al.*, 1990; Macciotta *et al.*, 2002).

References

- Ali, T.E., Schaeffer, L.R., 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637-644.
- Andersen, F., Osteras, Ø., Reksen, O., Gröhn, Y.T., 2011. Mastitis and the shape of the lactation curve in Norwegian dairy cows. *J. Dairy Res.* 78:23-31.
- Atzori, A.S., Steri, R., Rotondo, F., Dimauro, C., Macciotta, N.P.P., Cannas, A., Pulina, G., 2010. Bio-economic model for lactation length simulations in dairy cows. Proc. 26th World Buiatrics Congr., Santiago, Chile. Available from: <http://www.kenes.com/buiatrics/cd/pdf/536.pdf>
- Bertilsson, J., Berglund, B., Ratnayake, G., Svennersten-Sjaunja, K., Wiktorsson, H., 1997. Optimising lactation cycles for the high-yielding dairy cow. A European perspective. *Livest. Prod. Sci.* 50:5-13.
- Bohmanova, J., Miglior, F., Jamrozik, J., 2009. Use of test-day records beyond three hundred five days for estimation of three hundred five-day breeding values for production traits and somatic cell score of Canadian Holsteins. *J. Dairy Sci.* 92:5314-5325.
- Brody, S., Ragsdale, A.C., Turner, C.W., 1923. The rate of decline of milk secretion with the advance of the period of lactation. *J. Gen. Physiol.* 5:441-444.
- Brotherstone, S., Thompson, R., White, I.M.S., 2004. Effects of pregnancy on daily milk yield of Holstein-Friesian dairy cattle. *Livest. Prod. Sci.* 87:265-269.
- Brotherstone, S., White, I.M.S., Meyer, K., 2000. Genetic modelling of daily milk yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70:407-415.
- Butler, S.T., Shalloo, L., Murphy, J.J., 2010. Extended lactations in a seasonal-calving pastoral system of production to modulate the effects of reproductive failure. *J. Dairy Sci.* 93:1283-1295.
- Buttchereit, N., Stamer, E., Junge, W., Thaller, G., 2010. Evaluation of five lactation curve models fitter for fat-protein ratio of milk and daily energy balance. *J. Dairy Sci.* 93:1702-1712.
- Caccamo, M., Veerkamp, R.F., de Jong, G., Pool, M.H., Petriglieri, R., Licitra, G., 2008. Variance components for test-day milk, fat and protein yield, and somatic cell score for analyzing management information. *J. Dairy Sci.* 91:3268-3276.
- Cappio-Borlino, A., Pulina, G., Rossi, G., 1995. A non-linear modification of Wood's equation fitted to lactation curves of Sardinian dairy ewes. *Small Ruminant Res.* 18:75-79.
- Capuco, A.V., Ellis, S.E., Hale, S.A., Long, E., Erdman, R.A., Zhao, X., Paape, M.J., 2003. Lactation persistency: insights from mammary cell proliferation studies. *J. Anim. Sci.* 81(Suppl. 3):18-31.
- Carvalho, J.G.V., Blake, R.W., Pollak, E.J., Quaas, R.L., Duran-Castro, C.V., 1998. Application of an autoregressive process to estimate genetic parameters and breeding values for daily milk yield in a tropical herd of Lucerna cattle and in United States Holstein herds. *J. Dairy Sci.* 81:2738-2751.
- Cobby, J.M., Le Du, Y.L.P., 1978. On fitting curves to lactation data. *Anim. Prod.* 26:127-133.
- Cole, J.B., Null, D.J., 2009. Genetic evaluation of lactation persistency for five breeds of dairy cattle. *J. Dairy Sci.* 92:2248-2258.
- Cole, J.B., Null, D.J., VanRaden, P.M., 2009. Best prediction of yields for long lactations. *J. Dairy Sci.* 92:1792-1810.
- Congleton, W.R.Jr., Everett, R.W., 1980. Error and bias in using the incomplete gamma function to describe lactation curves. *J. Dairy Sci.* 63:101-108.
- Craninx, M., Steen, A., Van Laar, H., Van Nespen, T., Martín-Tereso, J., De Baets, B., Fievez, V., 2008. Effect of lactation stage on the odd- and branched-chain milk fatty acids of dairy cattle under grazing and indoor conditions. *J. Dairy Sci.* 91:2662-2677.
- Cruz, G.R.B., Ribeiro, M.N., Pimenta Filho, E.C., 2009. Estimativas de parâmetros de curva de lactação de bovinos. *Arch. Zootec.* 58:695-704.
- Cunha, D.N.F.V., Pereira, J.C., Silva, F.F., Campos, O.F., Braga, J.L., Martuscello, J.A., 2010. Selection of models of lactation curves to use in milk production simulation systems. *Rev. Bras. Zootec.* 39:891-902.
- de Haas, Y., Barkema, H.W., Veerkamp, R.F., 2002. The effect of pathogen-specific clinical mastitis on the lactation curve for somatic cell count. *J. Dairy Sci.* 85:1314-1323.
- Dekkers, J.C.M., Ten Hag, J.H., Weersink, A., 1998. Economic aspects of persistency of lactation in dairy cattle. *Livest. Prod. Sci.* 53:237-252.
- Deluyker, H.A., Shumway, R.H., Wecker, W.E., Azari, A.S., Weaver, L.D., 1990. Modeling daily milk yield in Holstein cows using time series analysis. *J. Dairy Sci.* 73:539-548.
- Dematawewa, C.M.B., Pearson, R.E., VanRaden, P. M., 2007. Modeling extended lactations of Holsteins. *J. Dairy Sci.* 90:3924-3936.
- De Vries, A., 2006. Economic value of pregnancy in dairy cattle. *J. Dairy Sci.* 89:3876-3885.
- Diggle, P.J., Liang, K.Y., Zeger, S.L., 1994. Analysis of longitudinal data. Clarendon Press, Oxford, UK.
- Dijkstra, J., Lopez, S., Bannink, A., Dhanoa, M.S., Kebreab, E., Odongo, N.E., Fathi Nasri, M.H., Behera, U.K., Hernandez-Ferrer, D., France, J., 2010. Evaluation of a mechanistic lactation model using cow, goat and sheep data. *J. Agr. Sci.* 148:249-262.
- Druet, T., Jaffrézik, F., Boichard, D., Ducroq, V., 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86:2480-2490.
- Fadel, J.G., 2004. Technical note: estimating parameters of nonlinear segmented models. *J. Dairy Sci.* 87:169-173.
- Ferris, T.A., Mao, I.L., Anderson, C.R., 1985. Selecting for lactation curve and milk yield in dairy cattle. *J. Dairy Sci.* 68:1438-1448.
- France, J., Thornley, J.H.M., 1984. Mathematical models in agriculture. Butterworths, London, UK.

- Friggens, N.C., Berg, P., Theilgaard, P., Korsgaard, I.R., Ingvarsen, K.L., Løvendahl, P., Jensen, J., 2007. Breed and parity effects on energy balance profiles through lactation: evidence of genetically driven body energy change. *J. Dairy Sci.* 90:5291-5305.
- García, S.C., Holmes, C.W., 2001. Lactation curves of autumn- and spring-calved cows in pasture-based dairy systems. *Livest. Prod. Sci.* 68:189-203.
- Gipson, T.A., Grossman, M., 1989. Diphasic analysis of lactation curves in dairy goats. *J. Dairy Sci.* 72:1035-1044.
- González-Recio, O., Alenda, R., Chang, Y.M., Weigel, K.A., Gianola, D., 2006. Selection for female fertility using censored fertility traits and investigation of the relationship with milk production. *J. Dairy Sci.* 89:4438-4444.
- Grossman, M., Koops, W.J., 1988. Multiphasic analysis of lactation curves in dairy cattle. *J. Dairy Sci.* 71:1598-1608.
- Grossman, M., Koops, W.J., 2003. Modeling extended lactation curves of dairy cattle: a biological basis for the multiphasic approach. *J. Dairy Sci.* 86:988-998.
- Grossman, M., Kuck, A.L., Norton, H.W., 1986. Lactation curves of purebred and crossbred dairy cattle. *J. Dairy Sci.* 69:195-203.
- Guo, Q., White, R.E., 2005. Cubic spline regression for the open-circuit potential curves of a lithium-ion battery. *J. Electrochem. Soc.* 152:A343-A350.
- Haile-Mariam, M., Goddard, M.E., 2008. Genetic and phenotypic parameters of lactations longer than 305 days (extended lactations). *Animal* 2:325-335.
- Hickson, R.E., Lopez-Villalobos, N., Dalley, D.E., Clark, D.A., Holmes, C.W., 2006. Yields and persistency of lactation in Friesian and Jersey cows milked once daily. *J. Dairy Sci.* 89:2017-2024.
- Hurley, W.L., 1989. Mammary gland function during involution. *J. Dairy Sci.* 72:1637-1646.
- Jaffrezic, F., White, I.M.S., Thompson, R., Hill, W.G., 2000. A link function approach to model heterogeneity of residual variances over time in lactation curve analyses. *J. Dairy Sci.* 83:1089-1093.
- Jakobsen, J.H., Madsen, P., Jensen, J., Peersen, J., Cristiensen, L.G., Sorensen, D.A., 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *J. Dairy Sci.* 85:1607-1616.
- Jamrozik, J., Bohmanova, J., Schaeffer, L.R., 2010. Selection of locations of knots for linear splines in random regression test-day models. *J. Anim. Breed. Genet.* 127:87-92.
- Jamrozik, J., Kistemaker, G.J., Dekkers, J.C.M., Schaeffer, L.R., 1997. Comparison of possible covariates for use in a random regression model for analyses of test day yields. *J. Dairy Sci.* 80:2550-2556.
- Jamrozik, J., Schaeffer, L.R., 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *J. Dairy Sci.* 80:762-770.
- Jamrozik, J., Schaeffer, L.R., Weigel, K.A., 2002. Estimates of genetic parameters for single and multiple-country test-day models. *J. Dairy Sci.* 85:3131-3141.
- Kellogg, D.W., Scott Urquhart, N., Ortega, A.J., 1977. Estimating Holstein lactation curves with a Gamma Curve. *J. Dairy Sci.* 60:1308-1315.
- Kirkpatrick, M., Lofsvold, D., Bulmer M., 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124:979-993.
- Knight, C.H., 2005. Extended lactation: turning theory into reality. *Adv. Dairy Technol.* 17:113-123.
- Knight, C.H., Peaker, M., Wilde, C.J., 1998. Local control of mammary development and function. *Rev. Reprod.* 3:104-112.
- Knight, C.H., Wilde, C.J., 1993. Mammary cell changes during pregnancy and lactation. *Livest. Prod. Sci.* 35:3-19.
- Kramer, E., Stamer, E., Spilke, J., Thaller, G., Krieter, J., 2009. Analysis of water intake and dry matter intake using different lactation curve models. *J. Dairy Sci.* 92:4072-4081.
- Littell, R.C., Henry, P.R., Ammerman, C.B., 1998. Statistical analysis of repeated measures data using SAS procedures. *J. Anim. Sci.* 76:1216-1231.
- López-Romero, P., Carabaño, M.J., 2003. Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. *Livest. Prod. Sci.* 82:81-96.
- Lucy, M.C., 2001. Reproductive loss in high-producing dairy cattle: where will it end? *J. Dairy Sci.* 84:1277-1293.
- Macciotta, N.P.P., Dimauro, C., Steri, R., Cappio-Borlino, A., 2008a. Mathematical modelling of goat lactation curves. In: A. Cannas and G. Pulina (eds.) *Dairy goats feeding and nutrition*. CAB International, Wallingford, UK, pp 31-46.
- Macciotta, N.P.P., Mele, M., Conte, G., Serra, A., Cassandro, M., Dal Zotto, R., Cappio-Borlino, A., Pagnacco, G., Secchiari, P., 2008b. Association between a polymorphism at the stearoyl CoA desaturase locus and milk production traits in Italian Holsteins. *J. Dairy Sci.* 91:3184-3189.
- Macciotta, N.P.P., Miglior, F., Cappio-Borlino, A., Schaeffer, L.R., 2008c. Issues in modelling lactation curves with regression splines. *J. Dairy Sci.* 91(E-Suppl.1):544 (abstr.).
- Macciotta, N.P.P., Miglior, F., Dimauro, C., Schaeffer, L.R., 2010. Comparison of parametric, orthogonal, and spline functions to model individual lactation curves for milk yield in Canadian Holsteins. *Ital. J. Anim. Sci.* 9:460-464.
- Macciotta, N.P.P., Vicario, D., Cappio-Borlino, A., 2005. Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. *J. Dairy Sci.* 88:1178-1191.
- Macciotta, N.P.P., Vicario, D., Cappio-Borlino, A., 2006. Use of multivariate analysis to extract latent variables related to level of production and lactation persistency in dairy cattle. *J. Dairy Sci.* 89:3188-3194.
- Macciotta, N.P.P., Vicario, D., Dimauro, C., Cappio-Borlino, A., 2004. A multivariate approach to modelling shapes of individual lactation curves in cattle. *J. Dairy Sci.* 87:1092-1098.
- Macciotta, N.P.P., Vicario, D., Pulina, G., Cappio-Borlino, A., 2002. Test day and lactation yield predictions in Italian simmental cows by ARMA methods. *J. Dairy Sci.* 85:3107-3114.
- Mead, R., Curnow, R.N., Hasted, A.M., 2002. *Statistical methods in agriculture and experimental biology*. Chapman and Hall/CRC, London, UK.
- Mepham, T.B., 1987. *Physiology of lactation*. Open University Press, Philadelphia, PA, USA.
- Misztal, I., 2006. Properties of random regression models using linear splines. *J. Anim. Breed. Genet.* 123:74-80.
- Nestorov, I., Rowland, M., Hadjitodorov, S.T., Petrov, I., 1999. Empirical versus mechanistic modelling: comparison of an artificial neural network to a mechanistically based model for quantitative structure pharmacokinetic relationships of a homologous series of barbiturates. *AAPS Pharmsci.* 1:5-13.
- Nielsen, P.P., Pettersson, G., Svennersten-Sjaunja, K.M., Norell, L., 2010. Technical note: variation in daily milk yield calculations for dairy cows milked in an automatic milking system. *J. Dairy Sci.* 93:1069-1073.
- Nixon, M., Bohmanova, J., Jamrozik, J., Schaeffer, L.R., Hand, K., Miglior, F., 2009. Genetic parameters of milking frequency and milk production traits in Canadian Holsteins milked by an automated milking system. *J. Dairy Sci.* 92:3422-3430.

- Norman, D.H., VanRaden, P.M., Wright, J.R., Smith, L.A., 1999. Mathematical representations of correlations among yield traits. *J. Dairy Sci.* 82:2205-2211.
- Núñez-Antón, V., Zimmerman, D.L., 2000. Modeling nonstationary longitudinal data. *Biometrics* 56:699-705.
- Olori, V.E., Brotherstone, S., Hill, W.G., McGuirk, B.J., 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.* 58:55-63.
- Pool, M.H., Meuwissen, T.H.E., 2000. Reduction of the number of parameters needed for a polynomial random regression test day model. *Livest. Prod. Sci.* 64:133-145.
- Ptak, E., Schaeffer, L.R., 1993. Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.* 34:23-34.
- Rekaya, R., Weigel, K.A., Gianola, D., 2001. Hierarchical nonlinear model for the persistency of milk yield in the first three lactations of Holsteins. *Livest. Prod. Sci.* 68:181-187.
- Rekik, B., Ben Gara, A., 2004. Factors affecting the occurrence of atypical lactations for Holstein-Friesian cows. *Livest. Prod. Sci.* 87:245-250.
- Rotz, C.A., Zartman, D.L., Crandall, K.L., 2005. Economic and environmental feasibility of a perennial cow dairy farm. *J. Dairy Sci.* 88: 3009-3019.
- Schaeffer, L.R., 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86:35-45.
- Schaeffer, L.R., Jamrozik, J., 1996. Multiple-trait prediction of lactation yields for dairy cows. *J. Dairy Sci.* 79:2044-2055.
- Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J., Van Doormaal, B.J., 2000. Experience with a test-day model. *J. Dairy Sci.* 83:1135-1144.
- Shanks, R.D., Berger, P.J., Freeman, A.E., Dickinson, F.N., 1981. Genetic aspects of lactation curves. *J. Dairy Sci.* 64:1852-1860.
- Sherchand, L., McNew, R.W., Kellogg, D.W., Johnson, Z.B., 1995. Selection of a mathematical model to generate lactation curves using daily milk yields of Holstein cows. *J. Dairy Sci.* 78:2507-2513.
- Silvestre, A.M., Petim-Batista, F., Colaço, J., 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. *J. Dairy Sci.* 89:1813-1821.
- Silvestre, A.M.D., Almeida, J.C.M., Santos, V.A.C., Fontes, P.J.P., Alves, V.C., 2010. Modeling lactation curves of "Barrosã" beef cattle with Wood's model. *Ital. J. Anim. Sci.* 9:243-247.
- Smith, R.L., Grohn, Y.T., Pradhan, A.K., Whitlock, R.H., Van Kessel, J.S., Smith, J.M., Wolfgang, D.R., Schukken, Y.H., 2009. A longitudinal study on the impact of Johne's disease status on milk production in individual cows. *J. Dairy Sci.* 92:2653-2661.
- Sorensen, A., Muir, D.D., Knight, C.H., 2008. Extended lactation in dairy cows: effects of milking frequency, calving season and nutrition on lactation persistency and milk quality. *J. Dairy Res.* 75:90-97.
- Stanton, T.L., Jones, L.R., Everett, R.W., Kachman, S.D., 1992. Estimating milk, fat, and protein lactation curves with a test day model. *J. Dairy Sci.* 75:1691-1700.
- Steri, R., Nicolazzi, E.L., Gaspa, G., Canavesi, F., Dimauro, C., Macciotta, N.P.P., 2009. Modeling technical parameters of individual extended lactation curves in Italian Holsteins. *J. Dairy Sci.* 92(A-Suppl.1):124 (abstr.).
- Strucken, E.M., de Koning, D.J., Rahmatalla, S.A., Brockmann, G.A., 2011. Lactation curve models for estimating gene effects over a timeline. *J. Dairy Sci.* 49:442-449.
- Tedeschi, L.O., 2006. Assessment of the adequacy of mathematical models. *Agr. Syst.* 89: 225-247.
- Tempelman, R.J., 2009. Invited review: Assessing experimental designs for research conducted on commercial dairies. *J. Dairy Sci.* 92:1-15.
- Togashi, K., Lin, C.Y., 2003. Modifying the lactation curve to improve lactation milk and persistency. *J. Dairy Sci.* 86:1487-1493.
- Val-Arreola, D., Kebreab, E., Dijkstra, J., France, J., 2004. Study of the lactation curve in dairy cattle on farms in central Mexico. *J. Dairy Sci.* 87:3789-3799.
- Van der Werf, J.H.J., Goddard, M.E., Meyer, K., 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J. Dairy Sci.* 81:3300-3308.
- VanRaden, P.M., Dematawewa, C.M.B., Pearson, R.E., Tooker, M.E., 2006. Productive life including all lactations and longer lactations with diminishing dredits. *J. Dairy Sci.* 89:3213-3220.
- Vargas, B., Koops, W. J., Herrero, M., Van Arendonk, J.A., 2000. Modeling extended lactations of dairy cows. *J. Dairy Sci.* 83: 1371-1380.
- Varona, L., Moreno, C., Garcia Cortes, L.A., Altarriba, J., 1996. Bayesian analysis of Wood's lactation curve for spanish dairy cows. *J. Dairy Sci.* 81:1469-1478.
- Vasconcelos, J., Martins, A., Petim-Batista, M.F., Colaço, J., Blake, R.W., Carvalheira, J., 2004. Prediction of daily and lactation yields of milk, fat, and protein using an autoregressive repeatability test day model. *J. Dairy Sci.* 87:2591-2598.
- Verbyla, A.P., Cullis, B.R., Kenward, M.G., Welham, S.J., 1999. The analysis of designed experiments and longitudinal data by using smoothing splines. *Appl. Stat.-J. Roy. St. C* 48: 269-311.
- Vetharaniam, I., Davis, S.R., Upsdell, M., Kolver, E.S., Pleasants, A.B., 2003. Modeling the effect of energy status on mammary gland growth and lactation. *J. Dairy Sci.* 86:3148-3156.
- Wang, Z., Goonewardene, L.A., 2004. The use of MIXED models in the analysis of animal experiments with repeated measures data. *Can. J. Anim. Sci.* 84:1-11.
- Weller, J.I., Ezra, E., Leitner, G., 2006. Genetic analysis of persistency in the Israeli Holstein population by the multitrait animal model. *J. Dairy Sci.* 89:2738-2746.
- White, I.M.S., Thompson, R., Brotherstone, S., 1999. Genetic and environmental smoothing of lactation curves with cubic splines. *J. Dairy Sci.* 82:632-638.
- Wilmink, J.B.M., 1987a. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* 16: 335-348.
- Wilmink, J.B.M., 1987b. Comparison of different methods of predicting 305-day milk yield using means calculated from within-herd lactation curves. *Livest. Prod. Sci.* 17:1-17.
- Wilson, D.J., González, R.N., Hertl, J., Schulte, H.F., Bennet, G.J., Schukken, Y.H., Gröhn, Y.T., 2004. Effect of clinical mastitis on the lactation curve: a mixed model estimation using daily milk weights. *J. Dairy Sci.* 87: 2073-2084.
- Wilson, D.J., Grohn, Y.T., Bennett, G.J., González, R.N., Schukken, Y.H., Spatz, J., 2008. Milk production change following clinical mastitis and reproductive performance compared among J5 vaccinated and control dairy cattle. *J. Dairy Sci.* 91:3869-3879.
- Wood, P.D.P., 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164-165.
- Wood, P.D.P., 1976. Algebraic models of the lactation curves for milk, fat and protein production, with estimates of seasonal variation. *Anim. Prod.* 22:35-40.