Modelling Lactation Trends in Dairy Cattle

R.E. Kamidi

International Livestock Research Institute, PO Box 30709, Nairobi, Kenya

Tel.: +254-20-4223000; Fax: +254-20-4223001.
E-mail address: r.kamidi@cgiar.org

Abstract

Lactation curve models are usually not uniformly applicable to the variety of lactation trends for individual cows. In this study, common measures of lactation trend are derived from a simple, uniformly applicable fit of cumulative milk yield data to quadratic curves. The problem of auto-correlated residuals is in most cases adequately addressed by fitting a first order autoregressive model. Model fit pertaining to variables of major interest; peak yield, persistency and total yield is compared to that of the commonly used Wood's incomplete gamma function. Data used were obtained from a total of 194 lactations of 169 cows in 5 mixed-breed herds under conditions varying from a smallholding herd in the tropics to research herds under temperate conditions. The proposed model is plausible and appropriate for a complete description of lactation trends of individual cows based on weekly test-day milk yield data.

Keywords: Lactation curve; Persistency; Peak milk yield; Dairy cattle; Linear models;

1. Introduction

Variations in the shape of the lactation curve for dairy cows are believed to stem from both genetic and environmental factors (Wood, 1980; 1970; 1969; 1968). Olori *et al*. (1999) fitted data obtained from a single uniformly managed herd to five different models, none of which adequately described individual lactations. They inferred that the suitability of empirical models of the lactation curve does not depend on the mathematical form of the function alone but rather, more on the biological nature of the lactation. Seasonal patterns in pasture availability for instance, is a major environmental factor contributing to atypical lactations that cannot be adequately described by standard models. Of particular significance are animals in smallholder rain-fed production systems that are common in tropical highlands where farmers are largely constrained in terms of management, lack of supplementary feeds, diseases and seasonal forage shortages.

In dairy cattle, milk production typically rises to a peak 2 to 8 weeks post partum and steadily declines thereafter. This paper focuses on peak yield, persistency and lactation yield that are clearly the parameters of major interest in individual lactation curves. In the weighted factor analysis model of Zwald et al, 2003, peak yield emerged the most important variable determining factor 1 among 13 management, climatic and genetic variables. It was thus, rated the best predictor of genotype by environment interaction between herds for lactation performance. Forecasting of production traits and genetic evaluation models are however outside the main focus of this study that seeks to demonstrate a plausible model for a universal description of the lactation curve.

2. Methods

2.1. The Model

The cumulative milk yield model first described by Kamidi (2005) is a simple quadratic curve,

$$y = \beta t + \gamma t^2 + \varepsilon$$

where y is cumulative milk yield after t days in milk and β and γ are constants while ε is a random error term. This simple linear model (in the parameters) is amenable to standard linear statistical analyses and it is relatively easy to assign biological interpretation to model parameters. The deceleration constant γ is strictly negative over a complete lactation and directly proportional to percent persistency $P = 100(1 + 2\gamma)$ (Kamidi, 2005).

It is easily shown that model parameter β is a measure of peak yield. The rate of incline in cumulative yield at lactation day t is given by the first derivative;

$$\frac{dy}{dt} = \beta + 2\gamma t$$

Deceleration in yield is the value of the second derivative;

$$\frac{d^2y}{dt^2} = 2\gamma$$

By elementary calculus, deceleration is zero at peak yield which implies $\gamma = 0$. Substitution obtains estimated cumulative yield at peak $\hat{y} = \hat{\beta}t$, the subsequent first derivative with respect to days in milk is the change in cumulative yield per day which is equivalent to estimated yield at peak, $\hat{\beta}$.

A third parameter of interest is total lactation yield that breeders usually benchmark at 305 days in milk. Estimated lactation yield derives from this lactation model as $\hat{Y} = 305\hat{\beta} + 93025\hat{\gamma}$

2.2. Data analysis

Three published datasets (Kamidi, 2005) were used for model validation. The first data set comprised empirical data extracted from milk records obtained from a smallholder 12-hectare farm located near Kitale, Kenya. Milk records taken between May 1995 and February 2003 for 11 mixed–breed cows covering a total of 31 lactations were used. The second data set comprised data extracted from milk records kept at a large-scale farm located a few kilometres from the smallholding. The data were derived from lactations of 96 cows in a mixed-breed herd that calved between May 1984 and September 1985. A total of 101 lactations lasting 38 weeks or longer were considered with data from longer lactations being limited to the first 44 weeks of lactation. A third data set comprised data obtained from cows of Italian brown and Italian Holstein breeds in three herds reared under Italian conditions. A total of 62 lactations recorded between October 2001 and September 2003 lasting 32 weeks or more were considered but limited to the first 44 weeks of lactation. A few gaps in the data were linearly interpolated without loss of general trend. All data were fitted to cumulative quadratic models using the AUTOREG procedure (SAS, 1999). First order autoregressive models of the general form, $y_t = \beta t + \gamma t^2 + \phi y_{t-1} + \varepsilon_t$ were fitted.

Data were also fitted to the incomplete gamma function (Wood, 1967) that is probably the most widely used model. Using the non-linear curve $y_t = a^b e^{-ct}$ to represent milk yield y_t on day t, Wood (1967) defined persistency as $s = -(b+1)\log_e c$.

Peak yield is estimated by $y_p = a\left(\frac{b}{c}\right)^b e^{-b}$. These expressions are nevertheless difficult to interpret biologically. Estimated lactation yields were obtained by summation over the number of days in milk. For 305 days, estimated lactation yield $\hat{Y} = \sum_{t=1}^{305} \hat{a}^{\hat{b}} e^{-\hat{c}t}$

Parameters b and c are positive for typical curves. Curves with negative values of b fail to display any peaks, depicting non-typical lactations. Similarly, it is not possible to estimate persistency for curves with negative or zero values of parameter c. Data were fitted using the NLIN procedure (SAS, 1999).

3. Results and Discussion

Results obtained from analysis of the first data set are presented in Table 1. Perfect fits of cumulative yield data to the curvilinear model were obtained for all cows and lactations.

Cow	Lactation number	Lactation days	Model R^2	Fitted peak, β (kg)	Peak deviation	Fitted lactation yield, Y (kg)	Lactation yield deviation	Persistency, P (%)	SE P
Adema	1	305	0.9959	10.3	0.7	2611.01	21.0	98.9	0.2
Amina	3	305	0.9975	12.5	0.0	2849.0	20.0	97.9	0.2
	4	266	0.9972	14.5	0.0	2308.9	-6.5	95.6	0.2
	5	271	0.9996	15.3	-0.3	2518.9	-19.4	95.6	0.1
	6	224	0.9994	13.5	-1.0	1908.9	-24.6	95.5	0.2
Anita	1	305	0.9983	11.6	1.4	2558.9	20.6	97.9	0.1
	2	305	0.9867	9.3	3.7	2335.9	16.3	98.9	0.3
	3	293	0.9998	15.2	0.8	2553.6	-8.7	95.6	0.1
	4	305	0.9049	8.2	5.3	1493.7	11.2	97.9	0.5
	5	245	0.9981	13.6	3.4	3134.6	20.1	97.8	0.2
Baraka	1	305	0.9996	12.0	0.5	2334.5	19.6	97.1	0.1
Fatuma	2	305	0.9991	11.5	1.0	2471.8	4.3	97.8	0.1
	3	305	0.9905	10.4	2.1	2194.4	23.1	97.9	0.3
	4	305	0.9957	11.4	1.6	2382.8	10.9	97.6	0.2
	5	301	0.9729	9.1	4.4	1175.8	11.7	96.6	0.3
	6	305	0.9976	10.2	1.8	1958.4	17.0	97.5	0.1
Jahenda	1	305	0.9994	8.2	1.3	2120.1	32.2	99.2	0.1
Kaliyesa	5	305	0.9993	15.0	1.0	3257.7	27.7	97.1	0.1
	6	305	0.9742	12.0	3.5	2698.3	4.9	98.0	0.4
	7	305	0.9988	11.7	0.3	2430.7	16.8	97.6	0.1
Kamonya	1	305	0.9826	8.5	3.5	2181.1	18.7	99.1	0.3
Marion	1	305	0.9980	9.8	0.2	2097.7	3.8	98.1	0.1
	2	305	0.9929	9.2	2.3	1786.4	10.1	97.8	0.2
	3	305	0.9978	11.9	1.1	2446.6	22.0	97.4	0.2
	4	305	0.9628	7.6	4.9	1380.9	22.8	98.0	0.3
	5	305	0.9989	10.8	1.2	2408.8	15.2	98.1	0.1
Minayo	1	305	0.9921	7.0	2.0	1290.2	15.4	98.2	0.2
	2	218	0.9996	14.2	-0.7	2361.5	-34.7	96.9	0.2
Sadaka	1	305	0.9991	8.7	2.3	1838.2	20.5	98.2	0.1
	2	305	0.9220	6.5	5.5	1267.2	20.9	98.5	0.4
	3	305	0.9962	11.5	1.5	2526.7	7.7	97.9	0.2

Table 1. Values of peak yield and lactation yield, their deviations and lactation persistency obtained from cumulative milk yield data fitted to quadratic curves for various lactations of 11 cows of data set 1.

Fitted curves resemble projectiles whose curvature is proportionately smaller for the more persistent lactations (see Figure 1).

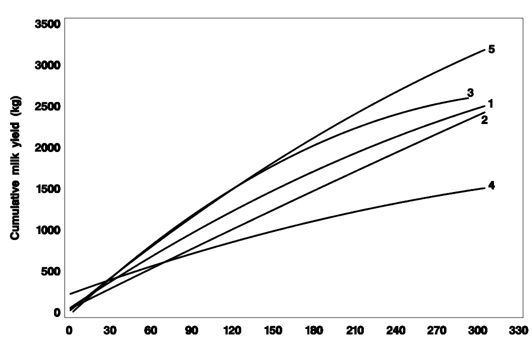

Figure 1. Fitted quadratic curves corresponding to cumulative milk yield obtained from Anita’s five lactations.

As expected for this type of data structure, residuals were characterised by a high positive auto-correlation. Although the parameter estimates obtained by fitting a regular cumulative quadratic model are unbiased, error variances are grossly underestimated as a result of which tests of significance are liberal and widths of confidence intervals deflated. Fitting first order autoregressive models was in most cases found to adequately address the problem of auto-correlation as revealed by examination of residual plots. Figure 2 depicts a typical residual plot obtained by fitting a curvilinear ordinary least squares regression curve to Kaliyesa's 7th lactation data while Figure 3 shows the residual plot from a first order autoregressive fit of the same data. Better estimates of error variance were achieved whereas changes in estimated parameter values were insignificant.

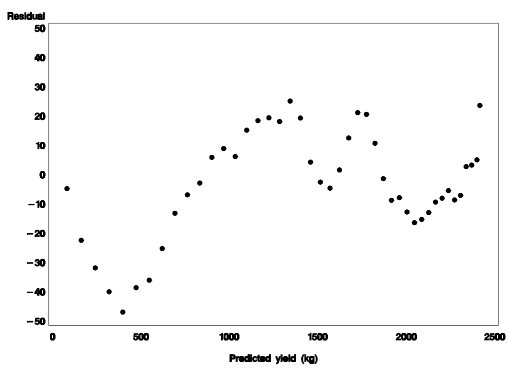

Figure 2. Residual plot obtained from an OLS fit to Kaliyesa's 7th lactation data.

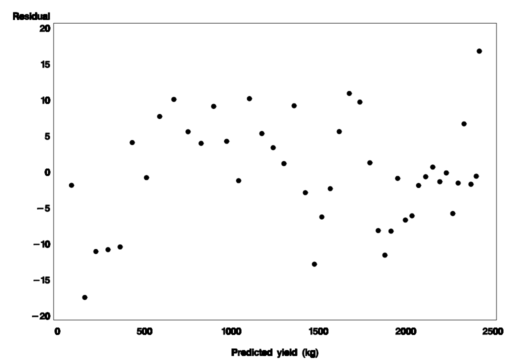

Figure 3. Residual plot obtained from a first order autoregressive fit to Kaliyesa's 7th lactation data

Larger deviations from observed peaks were observed for lactations during which cows suffered extreme stress such as major illness or depleted forage. Lactation persistency was also observed to be relatively higher under stressful conditions. Summary statistics for results obtained from fits of first order autoregressive quadratic curves to cumulative milk yield and corresponding absolute deviations are presented in Table 2 below.

	N	Lactation yield (kg) Cumulative quadratic Fitted values	Deviations	Incomplete gamma Fitted values	Deviations	Peak yield (kg) Cumulative quadratic Fitted values	Deviations	Incomplete gamma Fitted values	Deviations
Data set 1 (32 lactations)									
Mean	43	2218.9	17.0	2249.1	20.9	11.0	1.9	12.9	0.8
Minimum	31	1188.8	3.8	1213.5	10.0	6.5	0.0	9.2	0.1
Maximum	44	3262.3	34.7	3288.1	40.5	15.3	5.5	14.9	1.5
Fit correlation		0.99955		0.99934		0.73229		0.96164	
Non-estimable		0		0		0		19	
Data set 2 (101 lactations)									
Mean	43	4342.3	19.2	4340.9	33.0	20.1	2.6	19.9	3.2
Minimum	38	2705.2	0.2	2773.5	1.2	11.3	0.1	11.5	0.6
Maximum	44	6453.4	151.6	5863.7	79.8	32.8	5.9	26.3	7.3
Fit correlation		0.99938		0.99882		0.91593		0.91804	
Non-estimable		0		0		0		24	
Data set 3 (62 lactations)									
Mean	30	7346.0	22.4	7332.3	47.1	33.8	1.4	32.6	2.1
Minimum	24	4643.9	0.2	4640.3	0.8	21.5	0.0	19.9	0.1
Maximum	38	11212.3	138.1	11231.7	250.1	43.2	6.5	40.3	6.0
Fit correlation		0.99979		0.99895		0.96095		0.96745	
Non-estimable		0		0		0		1	

Table 2. Summary statistics for peak and lactation yields and corresponding absolute deviations obtained by Fitting cumulative milk yield to first order auto regressive quadratic curves and daily milk yield to incomplete gamma curves for various data sets.

Better estimates for lactation yield were obtained from fitted quadratic cumulative curves evident in the smaller mean absolute deviations and higher fit correlations than corresponding incomplete gamma curve values for all data sets. The incomplete gamma function was also found to be inappropriate for modelling lactations of tropical herds of datasets 1 and 2. Due to resultant non-positive estimates for model parameter b or c, the model failed to provide estimates of peak yield for 19 out of the 32 lactations in data set 1 and 24 out of 101 lactations of data set 2. One curve fitted to lactations pertaining to the temperate climate herds also failed to provide an estimate for peak yield. On the other hand, curvilinear fits provided peak yield estimates for all lactations modelled. A slight tendency to under-estimate both peak yield and total yield among lower yield tropical cows (data set 1) was evident from a proportionately larger number of non-negative residuals. On the other hand, the incomplete gamma function tended to over-estimate total yield for this category of cattle. Overall, quadratic curves fitted to cumulative yield provided equally good or better fit to all data sets compared to Wood's incomplete gamma function curves that were inappropriate for most 'atypical' lactations.

References

Kamidi, R.E., 2005. A parametric measure of lactation persistency in dairy cattle. Livest. Prod. Sci. 96, 141-148.

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.

SAS, 1999. The SAS® System, Version 8. SAS Institute, Cary, NC.

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

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

Wood, P.D.P., 1969. Factors affecting the shape of the lactation curve in cattle. Anim. Prod. 11, 307-316.

Wood, P.D.P., 1970. A note on the repeatability of parameters of the lactation curve in cattle. Anim. Prod. 12, 535-538.

Wood, P.D.P., 1980. Breed variations in the shape of the lactation curve of cattle and their implications for efficiency. Anim. Prod. 31, 133-141.

Zwald, N.R., Weigel, K.A., Fikse, W.F., Rekaya, R., 2003. Application of a multiple herd cluster model for genetic evaluation of dairy sires from seventeen countries. J. Dairy Sci. 86, 376-382.