Iranian Journal of Applied Animal Science

Comparison of Different Lactation Curve Models to Describe Lactation Curve in Moroccan Holstein-Friesian Dairy Cows

Research Article

I. Boujenane[1*]

[1] Department of Animal Production and Biotechnology, Institut Agronomique et Veterinaire Hassan II, P.O. Box 6202 Rabat-Instituts, 10101 Rabat, Morocco

Received on: 8 Feb 2013
Revised on: 12 Mar 2013
Accepted on: 16 Apr 2013
Online Published on: Dec 2013

*Correspondence E-mail: i.boujenane@iav.ac.ma

Online version is available on: www.ijas.ir

ABSTRACT

In this study, the incomplete gamma function, an exponential function, a mixed-log function and a polynomial function were evaluated to describe the lactation curve in Moroccan Holstein-Friesian dairy cows. Data from 1990 to 1999, comprising 77130 monthly milk yields of 6029 dairy cows in 280 dairy herds, were used. Edits were carried out by considering the lactation length (<305 d), days in milk (DIM) at which the first test-day was recorded (>5 d and <50 d) and the interval between two consecutive recordings (<60 d). The models were compared based on the mean square error, the adjusted coefficient of determination and the estimated difference between actual and predicted milk yields. The estimated mean square errors of the models were slightly equal differing only in the first decimal for the incomplete gamma function (33.8) and for the other three models (33.7). The adjusted coefficient of determination was estimated to be 0.18 in the exponential function, mixed-log function and polynomial function but 0.91 for the incomplete gamma function. The estimated difference between actual and predicted milk yields was zero for all the models. It was concluded that all models provided an acceptable level of accuracy in describing the shape of the lactation curve for Moroccan Holstein-Friesian dairy cows, but Wood model is observed to be the most suitable.

KEY WORDS coefficient of determination, dairy cattle, lactation curve, Morocco, Wood model.

INTRODUCTION

It is documented that milk yield increases from calving to the peak production, which is attained between 20 and 70 days post-partum, there after decreases smoothly until the end of lactation (Scott *et al.* 1996; Val-Arreola *et al.* 2004). Knowledge of lactation curves in dairy cattle is important for decision making on herd management and selection strategies, and it is also a key element in determining optimum strategies for insemination and replacement of dairy cows (Olori *et al.* 1999; Koçak and Ekiz, 2008). Describing lactation curve by using mathematical models goes back to 1920s (Cobby and Le Due, 1978). The incomplete gamma function is one of the most popular models (Wood, 1967) used to describe the lactation curve. The day at peak production, peak yield and persistency of milk yield can be estimated using the incomplete gamma function (Atashi *et al.* 2006; Boujenane and Hilal, 2012). However, in addition to incomplete gamma function, an exponential function proposed by Wilmink (1987), a mixed-log function proposed by Guo and Swalve (1995) and a polynomial function proposed by Ali and Schaeffer (1987) were evaluated for describing the lactation curve of dairy cows. In Morocco, approximate population size of Holstein-Friesian dairy cows is 400000 heads, raised in different management systems and environments. Herd size usually varies from 3 to more than 100 dairy cows and the mean 305-d milk yield is of 5353 kg (Boujenane, 2002; Boujenane and Hilal, 2012). No comprehensive study has been carried out concerning lactation curve in Moroccan Holstein-Friesian dairy cows.

The objective of this study was to compare the goodness of fit of four mathematical functions to describe the lactation curve of Holstein-Friesian cows in Morocco using field data from dairy farms.

MATERIALS AND METHODS

General

Data comprising of 77130 monthly milk yields of 6029 cows in 280 dairy herds involved in the official milk recording were used. The data were collected from 1990 to 1999 in 280 farms located at different regions of the country and enrolled in the official milk recording classified as type A (cows were milk recorded by a technician once a month after calving until drying off). The data structure is shown in Table 1.

Table 1 Structure of data analysed

Item	
Number of cows	6029
Number of herds	280
Number of test-day records	77130
Average age at calving (months)	39.2±11.5
Average daily milk yield (kg)	17.9±6.41
Average lactation length (days)	361.2±49.5
Average 305-d milk yield (kg)	5486.1±1496.0
Average 1^{st} lactation 305-d milk yield (kg)	5225.2±1332.7
Average 2^{nd} lactation 305-d milk yield (kg)	5654.7±1556.1
Average 3^{rd} lactation 305-d milk yield (kg)	5885.6±1659.7

Data analyses

Data analysed were obtained from cows at their 1^{st}, 2^{nd} and 3^{rd} parity, with age ranging from 24 to 48, 36 to 60 and 48 to 72 months, respectively. Records from cows at 1^{st}, 2^{nd} and 3^{rd} parity represented 49.4%, 31.8% and 18.8%, respectively.

Data were restricted to records for which the first milk recording had occurred between 5 and 50 days (Wilmink, 1987), spacing of consecutive sample days was not more than 60 days and lactation length was not greater than 305 days.

However, when the lactation of a cow was longer than 305 days, records were truncated at 305 days in milk (DIM). Furthermore, milk yield less than 3 kg or greater than 60 kg per day was discarded.

Data were analysed by the non-linear regression using the NLIN procedure and the Marquardt iterative method (SAS, 2002). Starting grids were specified such that all solutions fell within the outer limits of the search grid.

The mathematical models fitted were:

1. Wood incomplete gamma function (WOD) (Wood, 1967).

$$y_t = a \times t^b \times \exp^{-ct}$$

2. Wilmink exponential function (WIL) (Wilmink, 1987).

$$y_t = a + b \times \exp^{-0.5t} + c \times t$$

3. Guo and Swalve mixed-log-model (G and S) (Guo and Swalve, 1995).

$$y_t = a + b \times t^{0.5} + c \times \ln(t)$$

4. Ali and Schaeffer polynomial function (A and S) (Ali and Schaeffer, 1987).

$$y_t = a + b \times (t/305) + c \times (t/305)^2 + d \times [\ln(305/t)] + g \times [\ln(305/t)]^2$$

Where:
t: is day in milking.
y_t: is daily milk yield on day t.
a, b, c, d and g: are the regression coefficients.

These models are the most frequently used in terms of lactation curve prediction fit.

Comparison criteria for functions

The criteria used to compare models were:
1. Mean square error (MSE)

$$MSE = SSE / (n-p-1)$$

Where:
SSE: is error sum of squares.
n: is number of observations.
p: is number of parameters in each function.

Some authors have used mean square prediction error (MSPE=SSE/n) to measure error variation in absolute terms without recognizing its variation through the lactation because MSE is influenced by the number of parameters (Val-Arreola *et al.* 2004). However, in this study, both MSE and MSPE were very similar due to the large number of observations present.

2. Adjusted square of correlation coefficient between actual milk yield and predicted milk yield according to the model or adjusted coefficient of determination R^2_{adj}:

$$R^2_{adj} = 1 - \{[SSE / (n-p-1)] / [SST / (n-1)]\}$$

Where:
SSE: is error sum of squares.
SST: is total sum of squares.
n: is number of observations.
p: is number of parameters in each function.

R^2_{adj} was used instead of $R^2 = 1 - (SSE/SST)$ for two reasons: a. R^2_{adj} is adjusted for the number of parameters in the model (p) to make accurate comparison of models. b. Ratkowsky (1990) reported that the R^2 is no longer useful in non-linear regression. However, in this study both R^2 and R^2_{adj} were very similar due to the large number of observations present.

3. Mean of absolute residuals is the average of absolute difference between actual and predicted lactation milk yields (RES) (Neter *et al.* 1985; Guo and Swalve, 1995) was determined by using the following formula:

$$RES = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Where:
n: is number of observations.
y_i: is actual milk yield.
$\hat{y}_i$: is predicted milk yield.

The mathematical function with the best fit according to previous criteria was selected to fit the lactation curve of dairy cattle. Thus, models resulting in smaller MSE, smaller RES and higher R^2_{adj} were considered to be superior.

RESULTS AND DISCUSSION

Prediction of average lactation curve

Four mathematical functions were fitted with 77130 monthly milk records from 6029 cows. Figure 1 represents the average actual and predicted lactation curves in the range from DIM 5 to 305 for WOD, WIL, G and S; and A and S models. The lactation curve of the Moroccan Holstein-Friesian cows was characterized by milk production increasing during early lactation, attaining a peak at approximately 20 to 40 days post-partum followed by a gradual decline until the cow is drying off. Moreover, the results showed that the actual and predicted curves were almost similar for WIL, G and S; and A and S functions, whereas the WOD function slightly overestimated milk yield up to day 20, followed by an underestimation period up to day 130, then a good fit between day 131 and day 260 and finally underestimated it up to the end of lactation.

These results are in agreement with those reported by Scott et *al.* (1996) who found that WOD function overestimated milk yield after calving up to 10 weeks and underestimated milk yield from 10 to 20 weeks of lactation. Likewise, Grossman and Koops (1988) reported that a systematic deviation from actual milk yield is observed especially at the beginning and at the end of longer (more than 305 days) lactations. Moreover, Cobby and Le Due (1978) and Olori *et al.* (1999) reported that the bias in predicting daily milk yield with non-linear models depends on the stage of lactation.

Comparative study of the models

The estimated coefficients and goodness of fit criteria from all functions used are represented in Table 2. The MSE of the four models were almost similar, and ranged from 33.7 for A and S model to 33.8 for WOD model. Olori *et al.* (1999) reported smaller RMSE in the A and S model than in the WOD, WIL and; G and S models. Based on the mean square prediction error, Quinn *et al.* (2005) reported that the A and S model gave the best fit to their data compared to the WIL and; G and S models.

Table 2 Estimated parameters of average lactation curve, mean square error (MSE), adjusted coefficient of determination R^2_{adj} and lactation milk deviation (RES) for different functions

Function	Predicted functional form	MSE	R^2_{adj}	RES
WOD	$\hat{y}_t = 20.002t^{0.040}\exp^{-0.002t}$	33.8	0.91	-0.01
WIL	$\hat{y}_t = 23.485 - 3.930\exp^{-0.05t} - 0.034t$	33.7	0.18	0.00
G and S	$\hat{y}_t = 15.536 - 1.560\sqrt{t} + 4.352\ln(t)$	33.7	0.18	0.00
A and S	$\hat{y}_t = 20.212 - 6.146(\frac{t}{305}) - 1.076(\frac{t}{305})^2 + 2.397[\ln(\frac{305}{t})] - 0.637[\ln(\frac{305}{t})]^2$	33.7	0.18	0.00

Depending on the model fitted, the R^2_{adj} ranged from 0.18 for WIL, A and S and; G and S models to 0.91 for WOD model. The very high estimate of R^2_{adj} from WOD model is slightly surprising because it differs tremendously from those of the other models.

Nevertheless, the results of the current study are in perfect agreement with those of Gantner *et al.* (2010) who found that the R^2_{adj} for Wood's model ranged from 0.952 till 0.972 according to the subgroup, while those of linear regression models were considerably lower, particularly in Wilmink's model. Contrary to the current study, Olori *et al.* (1999) reported that the R^2_{adj} estimates from the WOD, WIL and; G and S models were similar and equal to 0.66, 0.69 and 0.67, respectively. Silvestre *et al.* (2006) reported that accuracy of WOD, WIL and; A and S models was affected by increasing the interval between tests and the interval between calving and the first test day. Moreover, Rekik and Ben Gara (2004) showed that the probability of occurrence of atypical curves increases by 4% for each day that the first test-day date is delayed. These conclusions showed that data collection during the initial phase of lactation is crucial for correct estimation of lactation curve shape, especially before the peak yield. Silvestre *et al.*

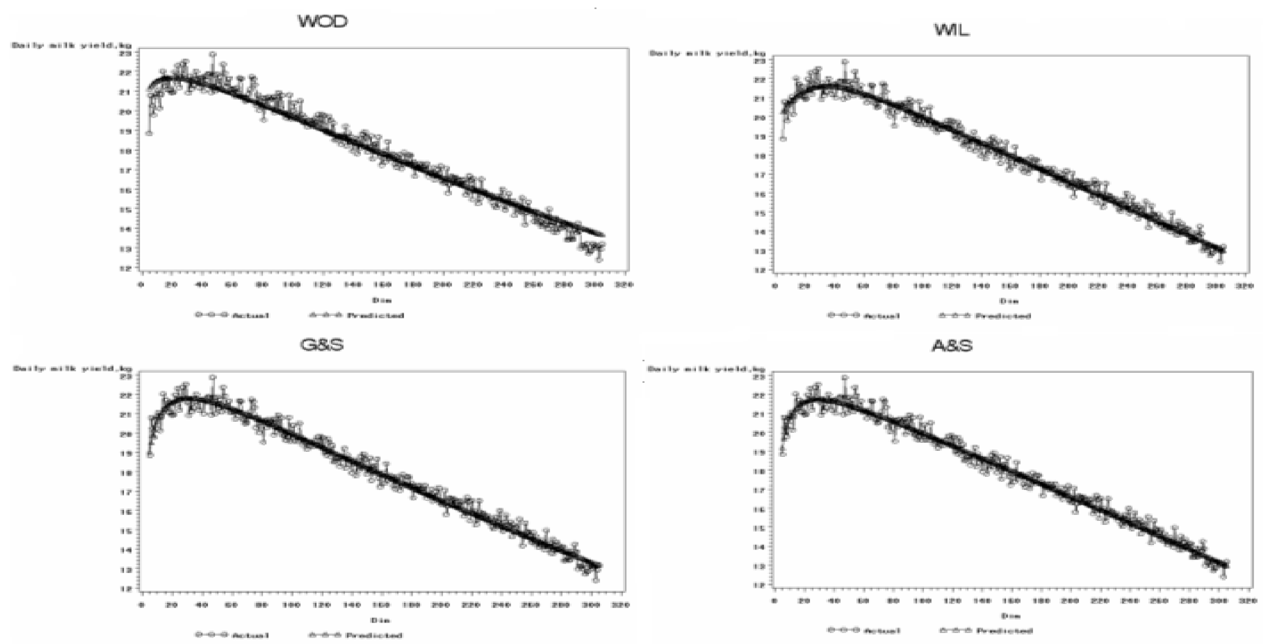

Figure 1 The actual and predicted lactation curves fitted from Wood, Wilmink, Guo and Swalve and Ali and Schaeffer functions

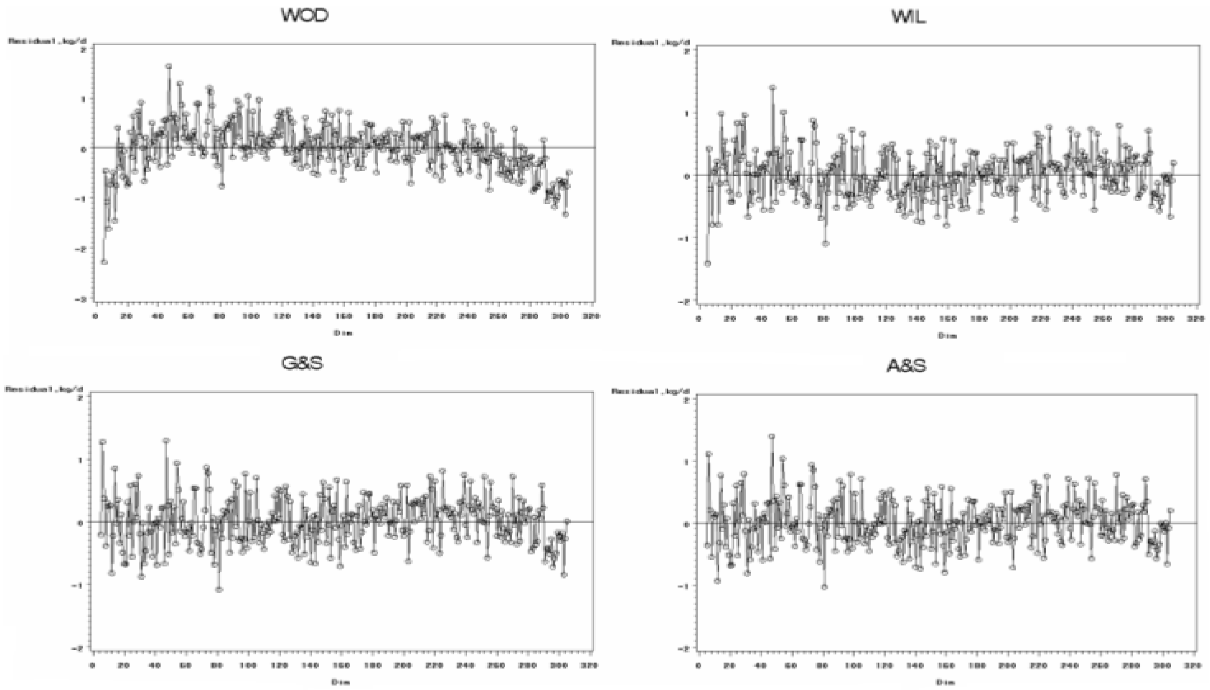

Figure 2 Distribution of average absolute residuals for DIM 5 to 305 for Wood, Wilmink, Guo and Swalve and Ali and Schaeffer functions

(2006) reported that when the first test occurred after 2 months of lactation, the best models for fitting the lactation curve are the non-parametric models. However, Koncagul and Yazgan (2008) reported that accuracy of WOD, WIL and; A and S models was not affected by increasing the interval between tests. Nevertheless, in the present study the maximum interval calving-first test was 50 days, as proposed by Wilmink (1987). Figure 2 shows the average absolute residual distribution (differences between actual and predicted milk yield) in the range from DIM 5 to 305 for WOD, WIL, G and S; and A and S models. A common pattern was observed in the 4 models, in which average absolute residuals oscillated between negative and positive during lactation. For the fitted curve to be an adequate representation of the data, the residuals should show no evidence of trends but should be randomly scattered about the horizontal axis. From the figures, it could be seen that the WOD function under predicted milk yield during early and late lactation. As the lactation progressed, all functions behaved in a similar manner. Therefore, it was possible to see that the pattern of residual distribution from WIL, G and S and; A and S functions varied from -1.41 to 1.39 kg/d, -0.88

to 1.29 kg/d and -1.02 to 1.39 kg/d, respectively, whereas that of WOD model ranged from -2.28 to 1.63 kg/d, especially at the beginning and the end of lactation, with an average equal to zero for all models in the range from DIM 5 to 305. From Figure 2, it seems that RES from WOD model has positive autocorrelations which result in obtaining biased model parameters. However, the value of Durbin-Watson statistic, used to test for the presence of both positive and negative correlations among the errors, was found to be near 2 indicating that there is no correlation. Thus, all models performed consistently through the lactation, indicating more satisfactory description of the lactation curve by these functions. The results obtained were similar to those of Olori *et al.* (1999) and Kitpipit *et al.* (2008) who found that A and S and WIL provided residuals error smaller than WOD function. Druet *et al.* (2003) reported that researcher has to decide to which criteria more emphasis should be given. In the present study, more weight was given to MSE. Afterwards, the emphasis was given to the mean residual followed by the R^2_{adj}. Ranking functions ability to describe lactation curves from the best to the worst on the basis of the previous criteria were WOD then the other three functions WIL, G and S and; A and S at the same level. Hence, WOD performed better biologically and statistically, as it was more closely similar to the underlying process in this population. This result agrees with that of Gantner *et al.* (2010) who concluded that considering obtained results, in terms of total variance explanation the non-linear Wood's model showed superiority above the linear ones (Wilmink's, Ali-Schaeffer's and Guo-Swalve's). However, it does not agree with that of Kitpipit *et al.* (2008) and Koçak and Ekiz (2008) who reported that the A and S model gave better fit than the WOD, WIL, and; G and S models. Moreover, Kitpipit *et al.* (2008) suggested that A and S could be used to describe the lactation curve in dairy cattle which were raised under the same conditions and composed with many breed structure. Likewise, Olori *et al.* (1999) suggested that the polynomial model gave the best fit in a farm-based study, which is not the case in the present study. Thus, except the high estimate of R^2_{adj} obtained from WOD, the other estimates in the current study for different models were generally in the range of those reported in the literature (Kitpipit *et al.* 2008). Therefore, based on MSE and RES criteria, WOD, WIL, G and S and; A and S models could be used to describe the lactation curve of Holstein-Friesian cows in Morocco. However, when R^2_{adj} is considered, WOD function is found slightly more accurate than other functions.

CONCLUSION

Four empirical functions have been compared and all of them allowed a suitable description of the shape of the lactation curve of Holstein-Friesian cows for dairy herds at first, second and third lactation. Thus, the better understanding of the lactation curve of Moroccan dairy cows will be used as a tool for better management and selection.

ACKNOWLEDGEMENT

I am grateful to the staff of the Livestock Service of the Ministry of Agriculture who made these data available.

REFERENCES

Ali T.E. and Schaeffer L.R. (1987). Accounting for covariance among test day milk yields in dairy cows. *Can. J. Anim. Sci.* **67,** 637-644.

Atashi H., Moradi Shahrbabak M. and Abdolmohammadi A. (2006). Study of some suggested measures of milk yield persistency and their relationships. *Int. J. Agric. Biol.* **8(3),** 387-390.

Boujenane I. (2002). Estimates of genetic and phenotypic parameters for milk production in Moroccan Holstein-Friesian cows. *Revue Elev. Méd Vét Pays Trop.* **55(1),** 63-67.

Boujenane I. and Hilal B. (2012). Genetic and non genetic effects for lactation curve traits in Holstein-Friesian cows. *Arch. Tierz.* **55(5),** 450-457.

Cobby J.M. and Le Due Y.L.P. (1978). On fitting curves to lactation data. *Anim. Prod.* **26,** 127-133.

Druet T., Jaffrezie F., Boichard D. and Ducrocq V. (2003). Modelling lactation curves and estimation of genetic parameters for first lactation test day records of French Holstein cows. *J. Dairy Sci.* **86,** 2480-2490.

Gantner V., Jovanovac S., Raguz N., Solic D. and Kuterovac K. (2010). Non-linear *vs.* linear regression models in lactation curve prediction. *Bulgarian J. Agric. Sci.* **16(6),** 794-800.

Grossman M. and Koops W.J. (1988). Multiphasic analysis of lactation curves in dairy cattle. *J. Dairy Sci.* **71,** 1598-1608.

Guo Z. and Swalve H.H. (1995). Modelling of the lactation curve as a sub-model in the evaluation of test day records. *Int. bull Bulletin 11, Int. Bull Eval. Serv., Uppsala, Sweden.*

Kitpipit W., Sopannarath P., Buaban S. and Tumwasorn S. (2008). Comparison of five mathematical functions for prediction of monthly yield in Thai multibreed dairy cattle population. *Kasetsart J. Nat. Sci.* **42,** 246-255.

Koçak O. and Ekiz B. (2008). Comparison of different lactation curve models in Holstein cows raised on a farm in the south-eastern Anatolia region. *Arch. Tierz. Dummerstorf.* **4,** 329-337.

Koncagul S. and Yazgan K. (2008). Modelling first lactation of Holstein cows on herds in the southeast regions of Turkey. *J. Anim. Vet. Adv.* **7(7),** 830-840.

Neter J., Wasserman W. and Kutner M.H. (1985). Methods Applied Statistical. Richard D. Irwin, Inc., Homewood, Illinois. USA.

Olori V.E., Brotherstone S., Hill W.G. and Mc Guirk 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.

Quinn N., Killen L. and Buckley F. (2005). Empirical algebraic modelling of lactation curves using Irish data. *Irish J. Agric. Food Res.* **44,** 1-13.

Ratkowsky D.A. (1990). Handbook of Nonlinear Regression Models. Marcel Dekker, New York, USA.

Rekik B. and Ben Gara A. (2004). Factors affecting the occurrence of atypical lactations for Holstein-Friesian cows. *Livest. Prod. Sci.* **87,** 245-250.

SAS Institute. (2002). SAS User's Guide, Statistics, Version 9 Ed, SAS Institute Inc., Cary, NC, USA.

Scott T.A., Yandell B., Zepeda L., Shaver R.D. and Smith T.R. (1996). Use of lactation curves for analysis of milk production data. *J. Dairy Sci.* **79,** 1885-1894.

Silvestre A.M., Petim-Batista F. and Calaço J. (2006). The accuracy of seven mathematical functions in modelling dairy cattle lactation curves based on test-day records from varying sample schemes. *J. Dairy Sci.* **89,** 1813-1821.

Val-Arreola D., Kebreab E., Dijkstra J. and France J. (2004). Study of the lactation curve in dairy cattle on farms in central Mexico. *J. Dairy Sci.* **87,** 3789-3799.

Wilmink J.B.M. (1987). Adjustment of test day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* **16,** 335-348.

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