



Fitting mathematical models to lactation curves from holstein cows in the southwestern region of the state of Parana, Brazil

ABÍLIO G.T. FERREIRA¹, DOUGLAS S. HENRIQUE^{1*}, RICARDO A.M. VIEIRA²,
EMILYN M. MAEDA¹ and ALTAIR A. VALOTTO³

¹Universidade Tecnológica Federal do Paraná/UTFPR, Estrada para Boa Esperança, Km 04, Comunidade São Cristóvão, 85660-000 Dois Vizinhos, PR, Brasil

²Universidade Estadual do Norte Fluminense Darcy Ribeiro/UENF, CCTA, LZNA, Avenida Alberto Lamego, 2000, Parque Califórnia, 28013-602 Campos dos Goytacazes, RJ, Brasil

³Associação Paranaense de Criadores de Bovinos da Raça Holandesa, Rua Professor Francisco Dranka, 608, Bairro Orleans, 81200-404 Curitiba, PR, Brasil

Manuscript received on December 20, 2013; accepted for publication on June 5, 2014

ABSTRACT

The objective of this study was to evaluate four mathematical models with regards to their fit to lactation curves of Holstein cows from herds raised in the southwestern region of the state of Parana, Brazil. Initially, 42,281 milk production records from 2005 to 2011 were obtained from “Associação Paranaense de Criadores de Bovinos da Raça Holandesa (APCBRH)”. Data lacking dates of drying and total milk production at 305 days of lactation were excluded, resulting in a remaining 15,142 records corresponding to 2,441 Holstein cows. Data were sorted according to the parity order (ranging from one to six), and within each parity order the animals were divided into quartiles (Q25%, Q50%, Q75% and Q100%) corresponding to 305-day lactation yield. Within each parity order, for each quartile, four mathematical models were adjusted, two of which were predominantly empirical (Brody and Wood) whereas the other two presented more mechanistic characteristics (models Dijkstra and Pollott). The quality of fit was evaluated by the corrected Akaike information criterion. The Wood model showed the best fit in almost all evaluated situations and, therefore, may be considered as the most suitable model to describe, at least empirically, the lactation curves of Holstein cows raised in Southwestern Parana.

Key words: dairy cattle, milk yield, mathematical models selection, non-linear models.

INTRODUCTION

The lactation curve is a graphic representation of the milk production from an animal throughout a defined period (Yadav et al. 1977). Accurate fitting of lactation curves contributes towards making decisions about management and breeding

of animals that can be adjusted according to predictions of milk production at some stage of lactation (Cobuci et al. 2000, 2012).

The first mathematical representation of the lactation curve of cattle was made by (Brody et al. 1923). This first model (equation 1) describes a declining curve whose peak lactation is expected to occur on the day of parity. Therefore, this model is not suitable to describe the entire lactations of cows

Correspondence to: Douglas Sampaio Henrique
E-mail: douglas@utfpr.edu.br
*Tutor do grupo PET Zootecnia

that are highly specialized in milk production, but it is useful to outline only the declining phase of the curve (Brody et al. 1924, Cobby and Le Du 1978):

$$M = Ae^{-k_1t} \quad (1)$$

wherein:

M = the milk flow in time "t" after the start of the process;

A = theoretical value of maximum milk flow;

k_1 = constant of declining or specific reaction rate (t^{-1}).

Subsequently, Brody et al. (1924) proposed another model (equation 2) with two phases: an initial rising phase proceeded by a declining phase, which Brody and co-workers assumed to occurs after the second month of lactation:

$$M = Ae^{-k_1t} - Be^{-k_2t} \quad (2)$$

wherein:

k_2 = constant of decline of the rising phase of the curve of milk secretion;

the subtraction $A - B$ indicates the milk flow at the time of parturition.

Wood (1967) proposed the use of the incomplete gamma function to describe the lactation curve and obtained a better fit of the ascending and descending phases of the curve, than the models used until that time. According to Guimarães et al. (2006), the incomplete gamma function, as suggested by Wood (1967), is the most common model used to estimate the lactation curves. The wide use of this model is, probably due to its ability to fit relatively well to a broad kind of data such as animals of different breeds with diverse production potential, raised in varied production systems and in several locations (Cobby and Le Du 1978, Papajcsik and Bodero 1988, Dijkstra et al. 1997, Landete-Castillejos and Gallego 2000, Cobuci et al. 2000, Pollott 2000, Cunha et al. 2010, Gloria et al. 2012). Another important feature of Wood's model is its simplicity and a reduced number of parameters (Wood 1967).

Despite the good fit of the model presented by Wood (1967), its parameters have no direct biological explanation. These models are known as empirical models, because they only describe the relationships between the variables without establishing cause-and-effect relationships. Over the last decades, improvements on computer processing power has allowed the development of more complex models whose parameters may represent biochemical and physiological aspects of the processes to be modeled. These models are called mechanistic models and, unlike empirical models, are formulated in order to establish the cause-and-effect relationships between the variables.

The first model of the lactation curve with parameters that have direct physiological explanation and whose numerical values can be estimated with data obtained in field situations were proposed by Dijkstra et al. (1997). This mechanistic model was developed to describe the growth pattern of the mammary glands during pregnancy and lactation. The model consists of a single pool (cell of mammary tissue) that is represented by DNA accumulation during the periods of pregnancy and lactation. The cell proliferation and death represents influxes and effluxes from the single pool, respectively. The authors adjusted the model to the data of mammary gland development (represented by the accumulation of DNA) in several species (mouse, rat, guinea pig, hamster and goat) and then made some redefinitions of parameters of the original model to represent the lactation curve in cattle.

Another mechanistic model has been proposed by Pollott (2000) based on the fit of two logistic curves to represent the major biological processes involved in the lactation process such as parenchyma cell proliferation, differentiation into cells with ability to secrete milk, and reduction in number of cells due to apoptosis (Pollott and Gootwine 2000). According to Pollott (2000), in addition to the advantage of biological meaning of

the parameters, the model can be modified by using additional terms that allow predicting the effect of factors that change the course of milk production, such as improvements in the cow nutrition or the beginning of pregnancy. However, Pollott's model is more complex than Dijkstra's model and has a larger number of parameters.

Several authors have assessed the quality of fit of the mentioned models at different scenarios (Papajcsik and Bodero 1988, Dijkstra et al. 1997, Cobuci et al. 2000, Val-Arreola et al. 2004, Guimarães et al. 2006, Cunha et al. 2010, Souza et al. 2014). The results from these tests varied, mainly due to differences in the level of milk production and the order of parity, thus the ranking of models according to quality of fit differed significantly from one study to another (Papajcsik and Bodero 1988, Dijkstra et al. 1997, Cobuci et al. 2000, Val-Arreola et al. 2004, Guimarães et al. 2006, Cunha et al. 2010).

Some authors evaluating the model of Brody (Papajcsik and Bodero 1988, Cobuci et al. 2000, Guimarães et al. 2006, Cunha et al. 2010) represented the lactation curve using the following equation:

$$M = Ae^{-k_1 t} - Ae^{-k_2 t} \quad (3)$$

Equation (3) predicts milk yield equal to zero at day of parity, as stated in Papajcsik and Bodero (1988). However, the original model of Brody (Equation 2) does not produce this result. Simulations of lactation curves using the values from the parameters of equation (2) estimated by Brody et al. (1924) with data from Holstein cows milked two, three or four times daily, predicted initial yield of 5.9, 7.3 and 18.1 kg day⁻¹, respectively. Therefore, the evaluated equation (3) actually differs from that originally published by Brody et al. (1924) and this mistake has been repeated for two decades at least!

Therefore, the question of which model would apply to all situations cannot be answered with the results obtained until now, and possibly, this

model does not exist. However, the models must be evaluated in each condition in which it is intended to be used. A particular situation is the evaluation of the model of Brody et al. (1924), which has, so far, been described erroneously and certainly invalidates the results described in literature. In this sense, a data set of milk production obtained in the southwestern region of Parana was used to evaluate four (Brody et al. 1924, Wood 1967, Dijkstra et al. 1997, Pollott, 2000) of the five mathematical models described previously, to establish which one would be more suitable to be used in selection programs of dairy cows in the region, which is currently the largest milk producer in the state of Parana (SEAB 2012), but still lacks technical information to improve productivity per cow (IPARDES 2009).

MATERIALS AND METHODS

Data provided by "Associação Paranaense de Criadores de Bovinos da Raça Holandesa" (APCBRH) from dairy herds located in the southwestern region of the state of Parana, with 42,281 records between the years 2005-2011, were obtained. The records contained milk production at the date of daily control, recorded monthly, the animal identification, dates of birth of the cow in addition to parity, order of parity, partial production and total production at 305 days of lactation. Information from animals that did not have the date of dry off and the total production at 305 days were excluded from the data set to be analyzed, resulting in 15,142 records concerning 2,441 Holstein cows (Table I).

The data were ranked according to the order of parity of the cows (which ranged from one to six) and within each parity order, the animals were divided into quartiles (Q25%, Q50%, Q75% and Q100%) according to total milk yield at 305 days. This classification was performed using the univariate procedure of SAS (version 9, SAS Institute Inc., Cary, NC, USA). Four mathematical models were fitted to each quartile belonging to

TABLE I
Maximum and minimum milk productions at 305 days,
number of cows and number of records for each parity
order (OP) and production level quartile (Q).

OP	Q	N. of cows	Milk production at 305 days		N. of records
			Minimum	Maximum	
1	25%	343	19	4508	1261
1	50%	342	4512	6442	1705
1	75%	343	6443	8327	2214
1	100%	342	8330	16103	2422
2	25%	162	165	5656	815
2	50%	161	5658	7667	1179
2	75%	161	7679	9329	1211
2	100%	161	9340	14397	1237
3	25%	66	302	5783	325
3	50%	66	5906	8224	489
3	75%	65	8284	10487	555
3	100%	65	10492	14658	533
4	25%	29	420	6029	145
4	50%	28	6062	8200	203
4	75%	28	8285	10432	225
4	100%	28	10530	13677	221
5	25%	9	1580	6176	38
5	50%	9	6524	9616	82
5	75%	9	9856	11756	85
5	100%	9	12090	13537	72
6	25%	4	5775	7566	24
6	50%	4	7915	9430	33
6	75%	4	9494	11009	48
6	100%	3	11206	13409	20
TOTAL		2441			15142

each parity order: two empirical models (Brody and Wood) and two mechanistical models (Dijkstra and Pollott). The Marquardt's algorithm implemented in the NLIN procedure was used to fit the models to the data (version 9, SAS Institute Inc., Cary, NC, USA).

The description of the models used in this work, and their parameters, were as similar as possible to the ones found in the original papers.

Model of Brody:

Previously described in Equation (2).

Model of Wood:

$$M = At^b e^{-ct} \quad (4)$$

M = milk production;

A, b, c = constants (as described by Wood, 1967);

b and c must be smaller than 1;

t = time in days (in the original paper, time was given in weeks).

Model of Dijkstra:

$$M = M_0 \exp\{\mu_T [1 - \exp(-k_2 t)] / k_2 - \lambda\} \quad (5)$$

M = milk production in time "t" (d^{-1});

M_0 = theoretical initial milk production ($kg \text{ day}^{-1}$);

μ_T = specific rate of cell proliferation at parturition;

k_2 = decay parameter;

λ = specific rate of cell death;

t = time in days.

Model of Pollott:

M = daily milk production;

$$M = \left[\frac{MS}{1 + \left(\frac{1 - P_0}{P_0} \right) \exp(-Gt)} \right] \times \left[\frac{1}{1 + \left(\frac{1 - Q_0}{Q_0} \right) \exp(-Dt)} \right] \quad (6)$$

P_0 = the total proportion of parenchymal cells which become active during early lactation since the first day of lactation;

G = relative growth rate in the number of active cells;

t = time in days;

Q_0 = proportion of the total parenchyma cells that are dead, dying on the first day of lactation;

D = mortality rate for cells;

MS = total milk secretion potential during lactation.

The quality of fit were verified by the Akaike information criterion (Akaike 1974, Burnham and Anderson 2004) using the recommendations suggested by Vieira et al. (2012).

After the estimation of parameters, the total milk production over the lactation was estimated, by means of the definite integral of milk production over the time interval [0, 305] for each mathematical model in each order of parity (OP1 to OP6) and quartile (Q25% to Q100%) as described in the general integral function below:

$$\int_0^{305} f(t) dt \quad (7)$$

The yield at the peak and the time of peak production were also calculated for the situations described above, by determining the point of maximum of the lactation curves and the value of t in the point of maximum.

RESULTS

The Pollott model used in this study was modified to reduce the number of parameters from seven to five, according to the suggestions of Pollott (2000) and Pollott and Gootwine (2000). Despite the simplification procedure, the model did not fit the data in any of the parity orders (OP) or quartiles (Q).

The model of Brody (1924) obtained values of $\Delta r \subset [0, 2]$ in 11 cases (Table II). Brody's model

was similar to that of Wood's and Dijkstra's in six situations (OP1 Q50%; OP2 Q25%; OP3 Q25%; OP3 Q50%; OP3 Q75% and OP4 Q50%) to reproduce data behavior and minimize loss of information (Table II). Brody's model was similar exclusively to the model of Dijkstra in two circumstances (OP1 Q25% and OP1 Q75%) and in two other cases it was similar only to the model of Wood (OP5 Q50% and OP6 Q75%). The model of Brody failed to reproduce the behavior of the observed data ($\Delta r > 10$ - Table II) in only one situation (OP1 Q100%). However, observing the values of w_r (Table II - all less than 0.8) the model of Brody was not considered a likelihood representation of reality in any situation. However, that model has an intermediate probability ($w_r \subset (0.5, 0.8]$) to represent reality in two cases (OP1 Q75% and OP5 Q75%). Brody model was considered the best choice among the models tested ($ER_r = 1$ - Table II) in OP1 Q75% and OP5 Q75% and the worst choice ($ER_r > 20$ - Table II) in OP1 Q100%, OP2 Q75% and OP3 Q100%, presenting an intermediate performance in other cases, with the exception of OP4 Q75%, OP4 Q100% and OP6 Q50% in which the model failed to converge.

The model of Dijkstra et al. (1997) presented similar performance to the model of Brody et al. (1924). The model of Dijkstra et al. (1997), obtained $\Delta r \subset [0, 2]$ in ten cases. This model was similar to both models (Brody and Wood) in mimicking observed data (Table II) in six situations (OP1 Q50%; OP2 Q25%; OP3 Q25%; OP3 Q50%; OP3 Q75%; OP4 Q50%), being similar exclusively to the model of Brody in two other cases (OP1 Q25% and OP1 Q75%), and exclusively to the model of Wood in another two (OP4 Q75% and OP4 Q100%). Just like the model of Brody, the

model of Dijkstra failed to emulate the observed data in Q100% OP1 ($\Delta > 10$ - Table II) and was not considered a likelihood representation of reality in any studied situation ($w_r < 0.8$ - Table II). Only in OP4% Q75 the model of Dijkstra presented intermediate probability $w_r \in (0.5, 0.8]$ for representing the reality. This model was the best choice to represent the data in OP1 Q25% and OP4 Q75% ($ER_r = 1$). The model did not converge in OP5 Q50%, OP5 Q75% and OP6 Q75%.

The model of Wood (1967) obtained $\Delta_r \in [0, 2]$ in 21 of the 24 cases studied (Table II) and in the three other situations presented $\Delta_r \in (2, 10]$ (Table II).

The model of Wood was the only model capable of representing the data with likelihood, i.e., values of $w_r > 0.8$. This occurred at OP1 Q100%, OP2 Q75%, OP2 and OP3 Q100% (Table II). Moreover, the Wood's model was considered the best choice ($ER_r = 1$) in 20 occasions (Table II) and the worst choice in only one case (OP1 Q75% - Table II).

The functions to estimate total milk yield at 305 days, peak production and time to peak production (days) for the model of Wood is represented by equations (8), (9) and (10) respectively:

TABLE II
Difference among AICc_r values (Δ_r), likelihood probabilities (w_r), evidence ratio or relative likelihood (ER_r) and root of mean square prediction error (RMSPE) for each parity order (OP) and production level quartile (Q).

OP	Q	Brody				Wood				Dijkstra			
		Δ_r	w_r	ER_r	RMSPE*	Δ_r	w_r	ER_r	RMSPE*	Δ_r	w_r	ER_r	RMSPE*
1	25%	0.024	0.464	1.012	4.99	3.941	0.065	7.175	5.00	0.000	0.470	1.000	4.99
1	50%	0.200	0.337	1.105	4.57	0.000	0.372	1.000	4.58	0.496	0.291	1.282	4.59
1	75%	0.000	0.537	1.000	4.34	6.348	0.022	23.90	4.35	0.397	0.440	1.220	4.34
1	100%	17.07	0.000	5090	5.26	0.000	1.000	1.000	5.25	20.40	3.7×10^{-5}	26876	5.27
2	25%	1.813	0.223	2.476	6.92	0.000	0.553	1.000	6.92	1.816	0.223	2.479	6.92
2	50%	3.023	0.154	4.534	4.93	0.000	0.698	1.000	4.92	3.102	0.148	4.715	4.93
2	75%	8.681	0.013	76.73	4.81	0.000	0.979	1.000	4.80	9.596	0.008	121.2	4.81
2	100%	5.789	0.051	18.07	6.14	0.000	0.913	1.000	6.13	6.459	0.036	25.26	6.14
3	25%	1.902	0.218	2.588	7.50	0.000	0.564	1.000	7.50	1.910	0.217	2.599	7.50
3	50%	0.339	0.302	1.185	5.46	0.000	0.358	1.000	5.47	0.104	0.340	1.053	5.46
3	75%	0.968	0.291	1.623	5.05	0.000	0.472	1.000	5.06	1.375	0.237	1.989	5.05
3	100%	6.822	0.031	30.30	5.36	0.000	0.945	1.000	5.34	7.390	0.023	40.25	5.36
4	25%	2.457	0.186	3.417	8.39	0.000	0.635	1.000	8.38	2.528	0.179	3.540	8.39
4	50%	1.601	0.235	2.226	4.96	0.000	0.523	1.000	4.96	1.532	0.243	2.152	4.96
4	75%	-	-	-	-	1.208	0.353	1.829	5.19	0.000	0.647	1.000	5.16
4	100%	-	-	-	-	0.000	0.565	1.000	5.71	1.258	0.301	1.876	5.70
5	25%	2.423	0.186	3.359	9.08	0.000	0.626	1.000	9.08	2.409	0.188	3.336	9.08
5	50%	1.598	0.262	2.223	5.77	0.000	0.582	1.000	5.77	-	-	-	-
5	75%	0.000	0.541	1.000	5.75	3.657	0.087	6.225	5.90	-	-	-	-
5	100%	2.889	0.162	4.240	4.87	0.000	0.686	1.000	4.85	3.004	0.153	4.491	4.87
6	25%	2.732	0.170	3.920	6.80	0.000	0.665	1.000	6.81	2.781	0.166	4.017	6.80
6	50%	-	-	-	-	0.000	0.817	1.000	5.23	3.066	0.176	4.632	5.29
6	75%	1.085	0.297	1.720	6.62	0.000	0.511	1.000	6.63	-	-	-	-
6	100%	2.408	0.188	3.334	5.13	0.000	0.625	1.000	5.34	2.412	0.187	3.340	5.12

*kg day⁻¹.

$$M_{tot} = A \int_0^{305} t^b \exp(-ct) dt \quad (8)$$

$$M_{max} = A(b/c)^b e^{-b} \quad (9)$$

$$t_p = b/c \quad (10)$$

in which M_{max} is the maximum milk production and t_p is the day of peak of lactation. The other parameters were previously defined.

The same estimates can be obtained for the model of Dijkstra by equations (11), (12) and (13):

$$M_{tot} = M_0 \int_0^{305} \exp \left\{ \frac{\mu [1 - \exp(-k_2 t)]}{k_2} - \lambda t \right\} dt \quad (11)$$

$$M_{max} = M_0 \left(\frac{\lambda}{\mu T} \right)^{\lambda/k_2} \exp [(\mu T - \lambda)/k_2] \quad (12)$$

$$t_p = [\ln(\mu T/\lambda)]/k_2 \quad (13)$$

And for the model of Brody by equations (14), (15) and (16):

$$M_{tot} = \int_0^{305} (Ae^{-k_1 t} - Be^{-k_2 t}) dt \quad (14)$$

$$M_{max} = A \cdot \exp \left[\frac{-k_1}{k_2 - k_1} \times \ln \left(\frac{Bk_2}{Ak_1} \right) \right] - B \cdot \exp \left[\frac{-k_1}{k_2 - k_1} \times \ln \left(\frac{Bk_2}{Ak_1} \right) \right] \quad (15)$$

$$t_p = \ln \left(\frac{B \cdot k_2}{A \cdot k_1} \right) \times \frac{1}{k_2 - k_1} \quad (16)$$

Estimates of peak production and time of peak (days) obtained with the three models (Brody, Wood and Dijkstra) were fairly similar (Table III), as well as the estimate of total milk yield at 305 days (Table III). The values estimated by Brody and Dijkstra models were always closer to one another than the values estimated by the model of Wood (Table III).

The data of 1606 cows were used to estimate the persistency factor according to the exponential law that described the lactation curve between the peak production and the mid-gestation

(Lucas, 1960). This data was used to estimate the correlation between persistency, M_{max} , M_{tot} , and M_1 (milk production in the first day of lactation).

High correlation values (Pearson correlation coefficients) were estimated between M_{max} and M_{tot} , M_{max} and M_1 and M_{tot} and M_1 (Table IV), but low values for M_{max} and persistency, M_{tot} and persistency, and M_1 and persistency (which was the only relationship that had a negative and very low correlation value). Nevertheless, all the correlations tested were significant ($P < 0.05$).

Figures 1, 2 and 3 present the three lactation curves generated by the three models (Brody, Wood and Dijkstra) for data of two different order of parity (OP1 and OP3), the curves obtained by the models of Dijkstra and Brody (Fig. 1 and 3) are very similar, whereas the model of Wood produced slightly different curve shapes (Fig. 2).

DISCUSSION

According to Burnham and Anderson (2004), the Akaike information criterion (AIC) has solid statistical and philosophical basis for assessing the quality of fit of mathematical models and for assisting in the selection of the most suitable model. The AIC_c (corrected AIC) is a version of the AIC to be used in practice because it can be used even when the number of observations (n) is not sufficiently large (unlike the AIC) and AIC_c converges to AIC as "n" increases. The AIC_c is determined from the sum of squares of residuals, the number of observations and the number of model parameters (including the parameter of the error variance). The individual value of AIC_c does not have an interpretation. To select the best models, we must calculate the Δ_r (the difference among AIC_c), the likelihood probabilities (w_r) and the relative likelihood (ER_r), which is the ratio between the maximum w_r and the w_r of the model in concern (Burnham and Anderson 2004, Vieira et al. 2012). Δ_r values between 0 and 2 indicate that the models are similar in reproducing the observed data behavior

TABLE III
Time to peak production, days (t_p), peak production, kg/d (M_{max}) and total milk yield at 305 days, kg (M_{tot}) estimated by Brody, Wood and Dijkstra models for each parity order (OP) and production level quartile (Q).

OP	Q	Milk production estimates								
		Brody			Wood			Dijkstra		
		t_p	M_{max}	M_{tot}	t_p	M_{max}	M_{tot}	t_p	M_{max}	M_{tot}
1	25%	21	17.54	4087	19	17.02	4093	21	17.56	4089
1	50%	43	23.83	6195	44	23.38	6173	45	24.19	6297
1	75%	58	28.74	7609	68	28.18	7585	58	28.82	7620
1	100%	77	35.85	9794	78	35.60	9759	77	35.97	9819
2	25%	21	23.52	5402	16	23.38	5378	21	23.50	5400
2	50%	45	28.53	7256	39	28.59	7214	46	28.53	7256
2	75%	49	34.34	8738	46	34.11	8700	49	34.42	8750
2	100%	56	41.81	10851	52	41.58	10796	56	41.86	10858
3	25%	29	26.06	5773	28	25.64	5735	29	25.78	5710
3	50%	41	31.40	7572	35	31.44	7533	42	31.43	7584
3	75%	59	37.59	9543	50	37.79	9501	59	37.65	9553
3	100%	52	45.49	11817	49	45.19	11762	52	45.53	11818
4	25%	42	30.18	6531	40	30.03	6473	43	30.16	6541
4	50%	51	31.29	7545	48	31.02	7497	52	31.32	7541
4	75%	-	-	-	41	39.02	9656	41	38.15	9721
4	100%	-	-	-	24	45.36	11675	24	44.95	11737
5	25%	34	34.39	7430	25	34.80	7455	34	34.47	7445
5	50%	30	33.79	8774	23	33.37	8745	-	-	-
5	75%	13	43.84	9271	43	40.85	10460	-	-	-
5	100%	47	51.09	12903	55	49.91	12864	46	51.22	12897
6	25%	41	37.64	7910	54	36.02	7759	41	37.67	7957
6	50%	-	-	-	46	38.76	8981	53	38.91	9032
6	75%	8	46.65	675	42	42.98	10522	-	-	-
6	100%	39	47.37	12341	58	46.01	12271	39	47.39	12342

TABLE IV
Pearson product-moment coefficients.

Variable	Variable ¹			
	M_{max}	M_{tot}	Persistency	M_I
M_{max}		0.95***	0.13***	0.82***
M_{tot}			0.20***	0.70***
Persistency				-0.05*
M_I				

¹ M_{max} = milk production at peak, kg day⁻¹; M_{tot} = total milk yield at 305 days, kg; Persist = persistency factor; M_I = Milk production at the first day of lactation kg day⁻¹.

*Significant (P < 0.05), ***Significant (P < 0.001).

and reduce the loss of information (in this case the model with the fewest number of parameters must be preferred). Δ_r values greater than 2 and less than or equal to 10 mean that the model performance was not as good and values greater than 10 indicate that the model either failed to reproduce the data and minimize the loss of information (Burnham and Anderson 2004, Vieira et al. 2012). Regarding w_r , when its value exceeds 0.8 the model can be considered a plausible representation of the reality; for a w_r greater than 0.5 and less than or equal to 0.8 the model is not so much representative; and

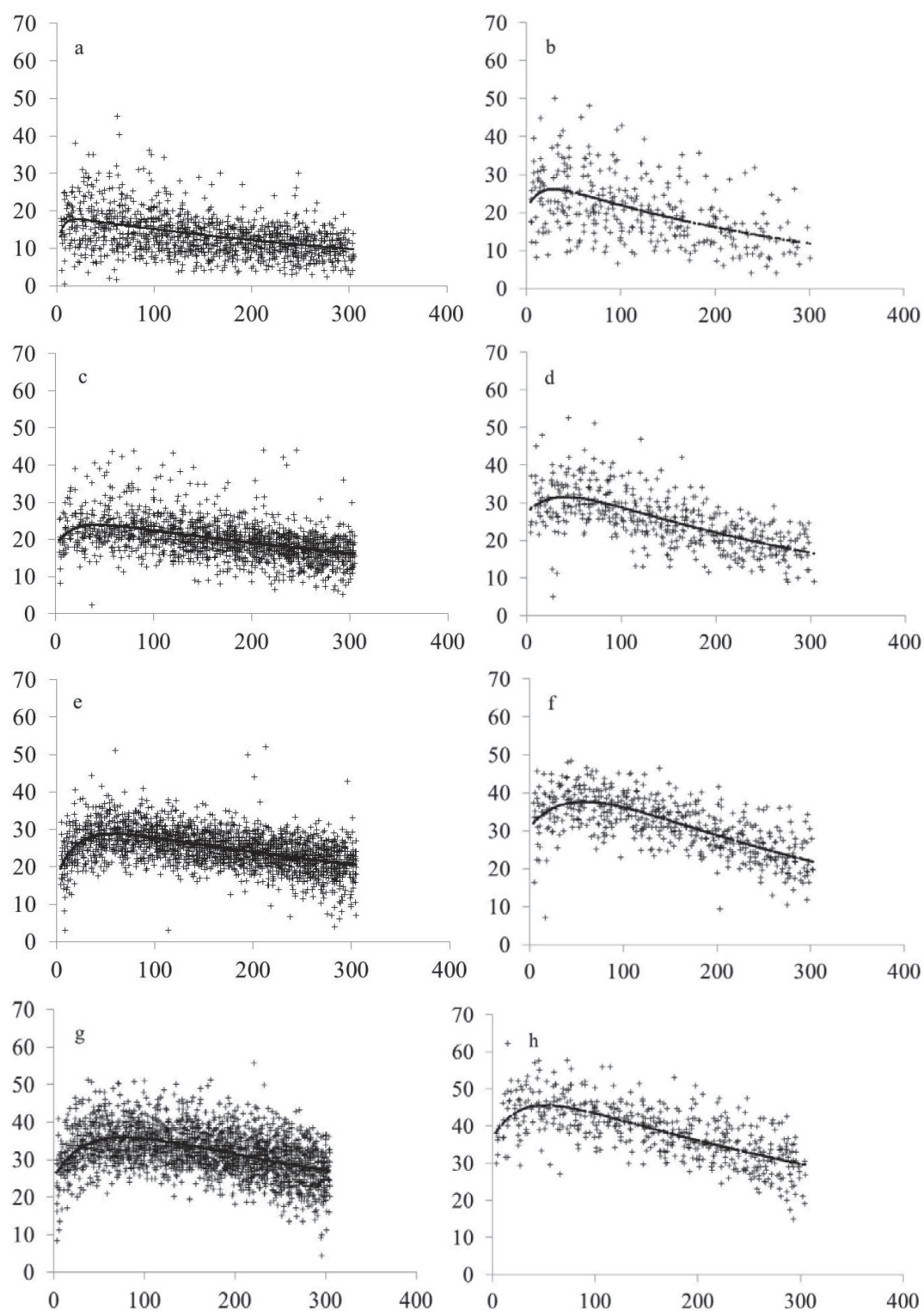


Figure 1 - X axis – days of lactation and Y axis – milk production (kg/d). Observed data and lactation curve estimated by the model of Brody to 1st (a, c, e, g) and 3rd parity order (b, d, f, h) at quartiles: 25% (a, b), 50% (c, d), 75% (e, f) and 100% (g, h).

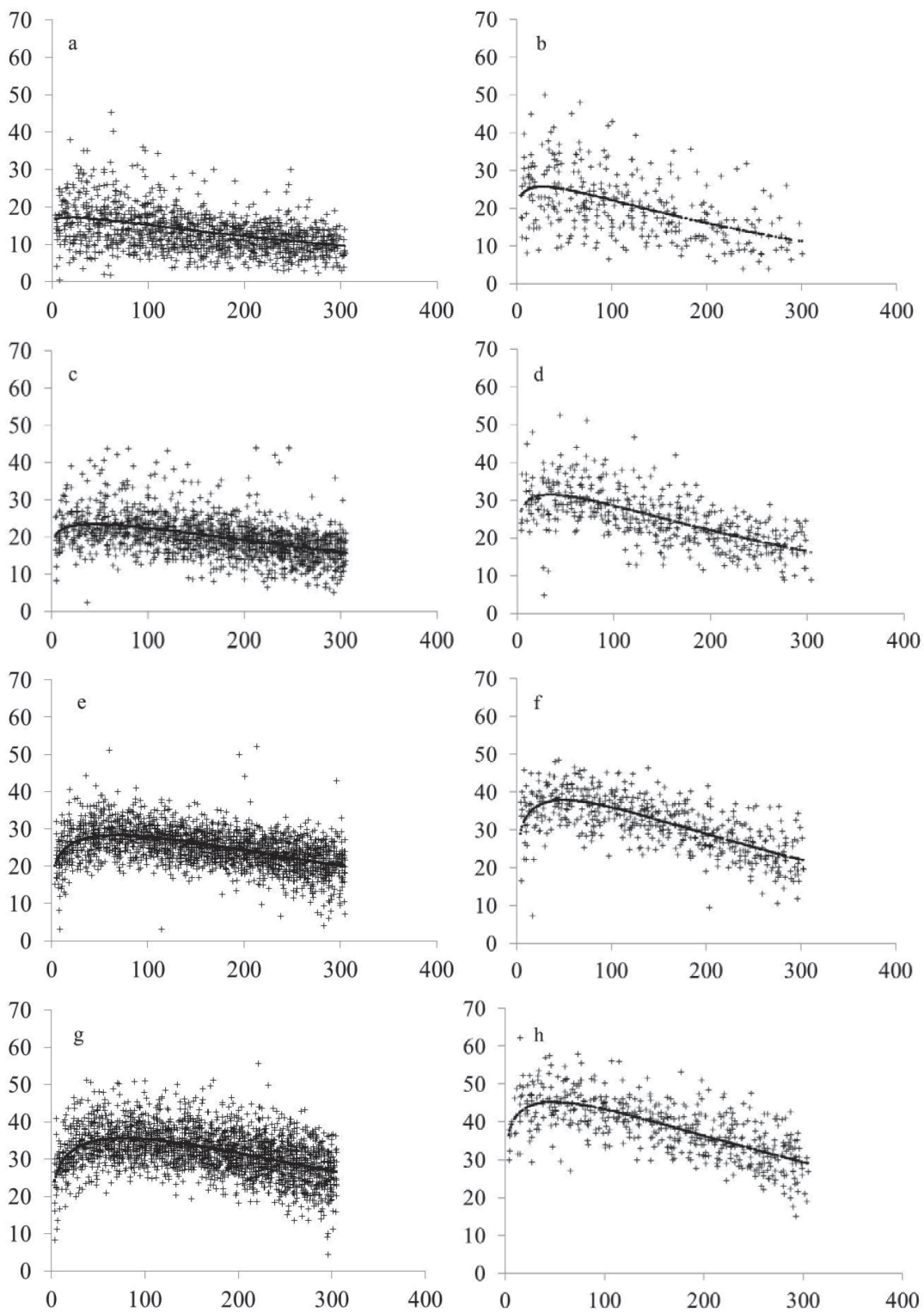


Figure 2 - X axis – days of lactation and Y axis – milk production (kg/d). Observed data and lactation curve estimated by the model of Wood to 1st (a, c, e, g) and 3rd parity order (b, d, f, h) at quartiles: 25% (a, b), 50% (c, d), 75% (e, f) and 100% (g, h).

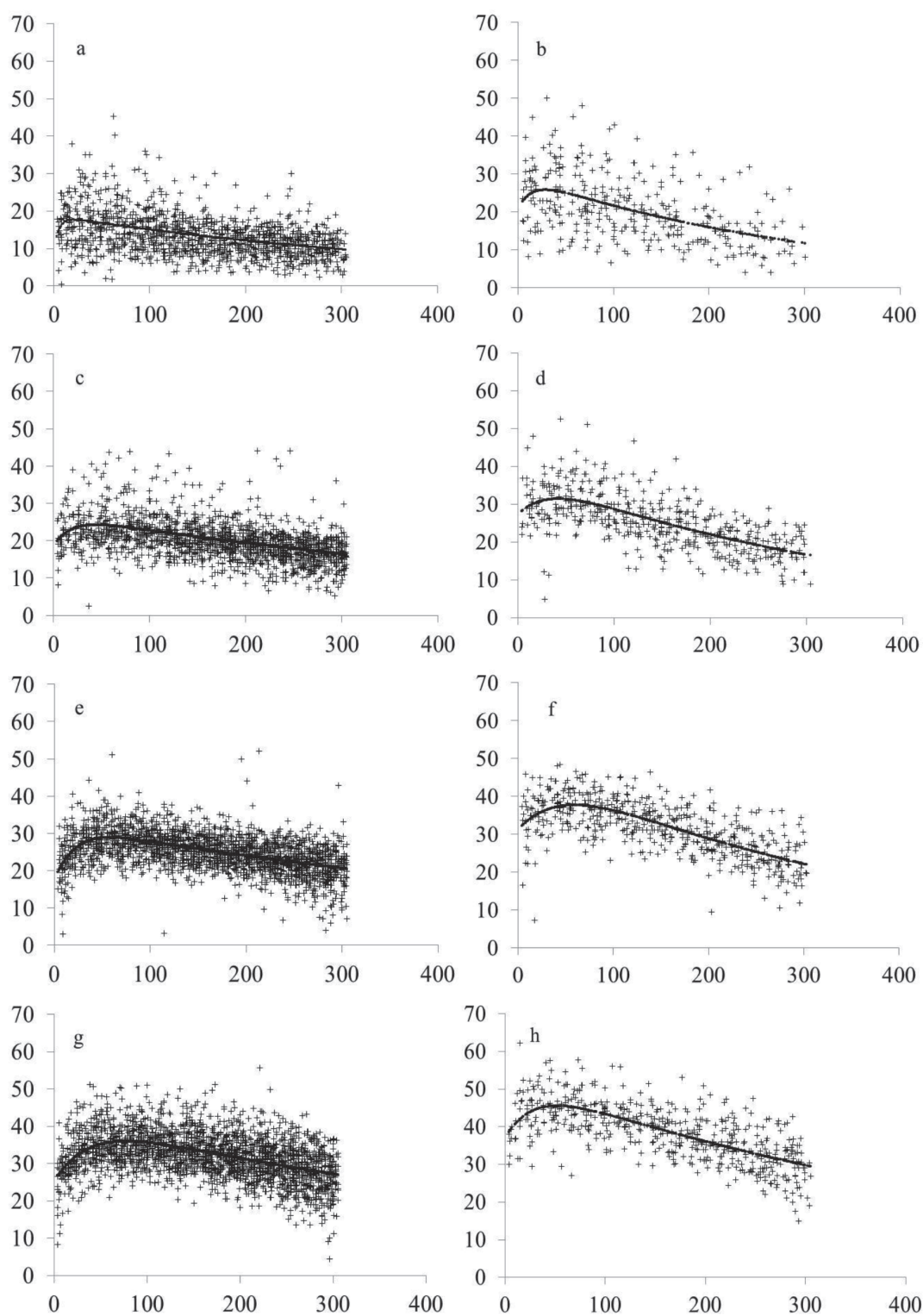


Figure 3 - X axis – days of lactation and Y axis – milk production (kg/d). Observed data and lactation curve estimated by the model of Dijkstra to 1st (a, c, e, g) and 3rd parity order (b, d, f, h) at quartiles: 25% (a, b), 50% (c, d), 75% (e, f) and 100% (g, h).

when w_r is less than or equal to 0.5 the model is not feasible. When choosing the best model, the one that results in $ER_r = 1$ will be the best choice, models in which ER_r values are greater than 1 and less than or equal to 20 are considered less likely and those whose ER_r exceeds the value 20 are the worst choices (Vieira et al. 2012).

Quantification of some biological processes, such as the rate of milk secretion (S) in kg per cell per day, are hard to be gathered in the field, thus Pollott suggested some modifications to his model in order to enable an adjustment to production data obtained from dairy farms (Pollott 2000, Pollott and Gootwine 2000). Furthermore, other adaptations were proposed that allowed for the prediction of changes in the lactation curve according to external factors, as for example improvements in nutrition or pregnancy (Pollott 2000). After these modifications, the model was successfully adjusted to milk production data from dairy cattle (Pollott 2000), and sheep (Pollott and Gootwine 2000). Pollott (2000) also compared this simplified form of his model with those proposed by Wood, Dijkstra and Morant (the first two were also tested in this work), and found lower values for the mean squared error in comparison to the model proposed by him. However, in the present work, the model proposed by Pollott failed to converge to any set of data, and so, it was not possible to compare it with the other models studied.

Dijkstra et al. (1997) used data of entire lactations from 23 animals to compare their model with that proposed by Wood, in its ability of adjustment based on the mean square of the residuals, R2 and Durbin-Watson statistics. Their mathematical model fitted better to lactation curves characterized by a sharp peak production, whereas the model of Wood fitted better to whole lactations in which the approach to peak production and the subsequent decline were smoother. Results from their work showed that, in general, the residuals from the model of Wood were positively auto correlated (Durbin-Watson statistics between 0.31 and 2.13), whereas the model of

Dijkstra presented less pronounced auto correlation, indicating that Dijkstra equation successfully described the trends of lactation curves (reported statistical Durbin-Watson between 0.59 and 2.5). Nonetheless, correlation among residuals is better accounted for with the use of techniques that model repeated measures over time (Vonesh, 2012).

The model of Dijkstra presented similar performance to the model of Brody, in the present work, despite the 73 years of difference between the dates of publication of the papers. While Brody et al. (1923, 1924) formulated the first descriptions of the lactation curve, Dijkstra et al. (1997), developed the first mechanistic, biologically based, description of the lactation curve, simple enough to be used in practical situations. However, the models of Dijkstra and Brody presented the lowest performances when compared to the model of Wood (1967) to mimic the data set we used (Table II).

The model of Wood is the best to describe an entire lactation, and because of that it is widely used for practical purposes e.g., selection of superior animals, in comparison with other lactation curve models (Molento et al. 2004, Cobuci and Costa 2012, Gloria et al. 2012). Papajcsik and Bodero (1988) compared 20 models of the lactation curve and concluded that Wood's model and a modified form of it provided the best representations of the lactation curve of Holstein cows in southwest Queensland (Australia). Whereas Val-Arreola et al. (2004), in central Mexico, obtained better fits to lactation curves of cows from two contrasting management systems (small scale and intensive) divided into three orders of parity, with the model of Dijkstra.

Cunha et al. (2010) fitted eight different models of lactation curve to milk production data from crossbred Holstein-Zebu sorted by lactation order (1st, 2nd, 3rd and above 3rd) and from cow herd farms classified as low, medium and high production. These authors obtained good fits with the model of Wood for cows in groups of low and medium milk production and for the group of high production, the

model of Dijkstra fitted better. The model of Pollott was not the best choice in any of the groups studied and presented fitting problems in groups of 1st and 2nd lactation order from low and high production levels (Cunha et al. 2010). According to Cunha et al. (2010) it is possible that mechanistic models, such as the Dijkstra and Pollott, do not fit well to lactation curves of animals characterized by low milk production potential. However, in the present study the parameters of the model of Pollott did not converge even to curves of high production level cows (Q100%) and the model of Dijkstra fitted well to some groups of low production level (Q25% - Table II), being the best option for the group OP1 Q25% (Table II). Therefore, the assumption of Cunha et al. (2010) is not supported by our results.

The estimates for peak day of lactation calculated by the models of Brody et al. (1924), Wood (1967) and Dijkstra et al. (1997) fitted by Cunha et al. (2010) are quite divergent, for example, the peak day calculated by the model of Brody was closer to the beginning of lactation, followed by Wood, and the later peak is estimated by the model of Dijkstra (Cunha et al. 2010). But, as previously mentioned, these authors, as well as several others, described the model of Brody differently from the original paper, resulting in mistaken inferences about it. Estimates made in the present work regarding the day of milk production peak, calculated with the three models (Brody et al. 1924, Wood 1967, Dijkstra 1997) were quite similar (Table III). The models of Brody and Dijkstra had the most similar values (Table III and Figs. 1 and 3), and in almost all situations the day of peak milk yield calculated with the model of Wood was closer to the parity date than the other two models (Table III and Fig. 2).

None of the models analysed in the present study showed a best fit to some specific situations, such as, a better fit to data of high production animals (as found by Cunha et al. 2010 for the model of Dijkstra) or to a smoother peak production curve

shape (as the result found by Dijkstra et al. 1997 for the model of Wood). The most relevant and general results, in summary, were: the model of Pollott was not a good fit to any data, regardless of the parity order (OP) or the production level (Q); the model of Wood was the best choice in most situations (Table II) and the models of Brody and Dijkstra presented intermediate and similar quality of fit (Table II), with estimates of milk production peak, day of peak and total milk yield at 305 days, very close (Table III). The number of parameters of each model (five to Pollott, four to Brody and Dijkstra, and three to Wood) was the only aspect that influenced the results obtained by the Akaike criterion; apparently, the quality of fit was inversely proportional to the number of parameters. And, although the model of Wood presented the best performance in almost all studied situations, and had the additional advantage of fitting very well to the data with a reduced number of parameters (only two), this model did not have parameters with biological meaning.

The model of Wood was the best choice to fit the data in most circumstances evaluated in the present work, probably due to its simplicity and smaller number of parameters. Therefore, when the biological description of parameters is not required for inferences, it is the most recommended to fit the lactation curve in practical situations. Nevertheless, functions of the parameters may be of interest and have biological interpretation, such as those described by equations (8), (9), and (10). A very important point is that the quality of fit of available models must always be evaluated according to the different situations.

ACKNOWLEDGMENTS

The authors acknowledge Dr. José Augusto Horst, Manager of “Programa de Análise de Rebanhos Leiteiros” from “Associação Paranaense dos Criadores de Bovinos da Raça Holandesa” APCBRH for supplying data that enabled the accomplishment of this work.

RESUMO

O objetivo do presente estudo foi avaliar quatro modelos matemáticos quanto ao seu ajuste a curvas de lactação de vacas Holandesas pertencentes a rebanhos da região sudoeste do estado do Paraná, Brasil. Inicialmente 42.281 controles de produções leiteiras referentes aos anos de 2005 a 2011, foram da “Associação Paranaense de Criadores de Bovinos da Raça Holandesa (APCBRH)”. Os dados que não possuíam a data de secagem e a produção total aos 305 dias foram excluídos, restando 15.142 controles referentes a 2.441 vacas Holandesas. Os dados foram classificados de acordo com a ordem de parição (variando de um a seis), e dentro de cada ordem de parição os animais foram separados em quartis (Q25%, Q50%, Q75% e Q100%) conforme a produção total aos 305 dias. Para cada ordem de parição, em cada quartil foram ajustados quatro modelos matemáticos, dois predominantemente empíricos (modelo de Brody e de Wood) e dois com características mais mecanicistas (modelo de Dijkstra e de Pollott). A qualidade de ajuste foi avaliada pelo critério de Akaike corrigido. O modelo de Wood apresentou o melhor ajuste em quase todas as situações avaliadas e, portanto, deve ser considerado como o melhor modelo para descrever, pelo menos empiricamente, as curvas de lactação de vacas Holandesas criadas no Sudoeste do Paraná.

Palavras-chave: gado leiteiro, produção de leite, seleção de modelos matemáticos, modelos não lineares.

REFERENCES

- AKAIKE H. 1974. A new look at the statistical model identification. *IEEE T Automat Contr* 19: 716-723.
- BEEVER DE. 1997. A model to describe growth patterns of the mammary gland during pregnancy and lactation. *J Dairy Sci* 80(10): 2340-2354.
- BRODY S, RAGSDALE AC AND TURNER CW. 1924. The relation between the initial rise and the subsequent decline of milk secretion following parturition. *J Gen Physiol* 6: 541-545.
- BRODY S, TURNER CW AND RAGSDALE AC. 1923. The rate of decline of milk secretion with the advance of the period lactation. *J Gen Physiol* 5: 442-444.
- BURNHAM KP AND ANDERSON DR. 2004. Multimodel inference understanding AIC and BIC in model selection. *Sociol Method Res* 33: 261-304.
- COBBY JM AND LE DU YLP. 1978. On fitting curves to lactation data. *Anim Prod* 26(2): 127-133.
- COBUCI JA AND COSTA CN. 2012. Persistency of lactation using random regression models and different fixed regression modeling approaches. *Rev Bras Zootecn* 41(9): 1996-2004.
- COBUCI JA, EUCLYDES RF, VERNEQUE RS, TEODORO RL, LOPES PS AND ALMEIDA E SILVA M. 2000. Curva de lactação na raça Guzerá. *Rev Bras Zootecn* 29(5): 1332-1339.
- CUNHA DNFV, PEREIRA JC, FONSECA E SILVA F, CAMPOS OF, BRAGA JL AND MARTUSCELLO JA. 2010. Selection of models lactation curves to use in milk production simulation systems. *Rev Bras Zootecn* 39(4): 891-902.
- DIJKSTRA J, FRANCE J, DHANOA MS, MAAS JA, HANIGAN MD, ROOK AJ AND BEEVER DE. 1997. A model to describe growth patterns of the mammary gland during pregnancy and lactation. *J Dairy Sci* 80(10): 2340-2354.
- GLORIA JR, BERGMANN JAG, QUIRINO CR, RUAS JRM, PEREIRA JCC, REIS RB, COELHO SG AND SILVA MDE. 2012. Environmental and genetic effects on the lactation curves of four genetic groups of crossbred Holstein-Zebu cows. *Rev Bras Zootecn* 41(11): 2309-2315.
- GUIMARÃES VP, RODRIGUES MT, SARMENTO JLR AND ROCHA DT. 2006. Utilização de funções matemáticas no estudo da curva de lactação em Caprinos. *Rev Bras Zootecn* 35(2): 535-543.
- IPARDES. 2009. Caracterização sócio econômica da atividade leiteira no Paraná: Sumário Executivo. Curitiba: Instituto Paranaense de Desenvolvimento Econômico e Social e Instituto Paranaense de Assistência Técnica e Extensão Rural, 29 p.
- LANDETE-CASTILLEJOS T AND GALLEGO L. 2000. Technical note: The ability of mathematical models to describe the shape of lactation curves. *J Dairy Sci* 78: 3010-3013.
- LUCAS HL. 1960. Critical features of good dairy feeding experiments. *J Dairy Sci* 43: 193-212.
- MOLENTO CFM, MONARDES H, RIBAS NP AND BLOCK E. 2004. Curvas de lactação de vacas holandesas do Estado do Paraná, Brasil. *Cienc Rural* 34(5): 1585-1591.
- PAPAJSIK IA AND BODERO J. 1988. Modeling lactation curves of Friesian cow in a subtropical climate. *Anim Prod* 47: 201-207.
- POLLITT GE. 2000. A biological approach to lactation curve analysis for milk yield. *J Dairy Sci* 83(11): 2448-2458.
- POLLITT GE AND GOOTWINE E. 2000. Appropriate mathematical models for describing the complete lactation of dairy sheep. *Anim Sci* 71: 197-207.
- SEAB - SECRETARIA DA AGRICULTURA E ABASTECIMENTO DO ESTADO DO PARANÁ. 2012. Disponível em: www.agricultura.pr.gov.br/arquivos/files/deral...leite_2012.pdf acesso em 01/10/2012. Curitiba: Tecnologia e Ensino Superior – Fundo Paraná.
- SOUZA R, ALCALDE CR, OLIVEIRA CAL, MOLINA BSD, MACEDO FDF, GOMES LC, HYGINO B AND POSSAMAI APS. 2014. Lactation curves and economic results of Saanen goats fed increasing dietary energy levels obtained by the addition of calcium salts of fatty acids. *Rev Bras Zootecn* 43(2): 73-79.

- 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
- VIEIRA RAM, CAMPOS PRSS, SILVA JFC, TEDESCHI LO AND TAMY WP. 2012. Heterogeneity of the digestible insoluble fiber of selected forages in situ. *Anim Feed Sci Tech* 171(2): 154-166.
- VONESH EF. 2012. Generalized linear and nonlinear models for correlated data: theory and applications using SAS[®], Cary: SAS Institute Inc., 538 p.
- WOOD PDP. 1967. Algebraic of the lactation curve in cattle. *Nature* 216: 164-165.
- YADAV MC, KATPATAL BC AND KAVSHIK SN, 1977. Components of gamma type function of a lactation curve, and factors affecting them in Haryana and its Friesian crosses. *Indian J Anim Sci* 49(9): 502-505.

