

Uso de polinômios segmentados para o estudo da curva de lactação de búfalos da raça Murrah e seus mestiços em sistema de criação semi-intensivo, no estado de São Paulo

M Muñoz Berrocal **, R Aspilcueta Borquis *, **, N Hurtado-Lugo*, ** e H Tonhati*

Facultad de Zootecnia, Universidad Nacional Agrária de la Selva. Tingo Maria, Peru

milthon_munoz@yahoo.es

** São Paulo State University: Faculdade de Ciências Agrárias e Veterinária, Jaboticabal, 14884-900, SP, Brasil*

*** Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP*

Resumo

Considerando a necessidade de melhor compreender a curva de lactação e as poucas informações desta natureza com búfalos, o objetivo deste estudo foi avaliar o uso de Polinômios Segmentados, dados pelos modelos Quadrático-Quadrático (QQ) e com três segmentos quadráticos (QQQ), no ajuste da curva de lactação média de animais da raça Murrah e seus mestiços em um sistema de criação semi-intensivo. Para este sistema, onde são realizadas duas ordenhas ao dia, foram usados registros de produção de três rebanhos, totalizando 3.267 lactações, perfazendo um total de 27.600 controles de produção de leite. O controle leiteiro foi feito em intervalos de 30 dias, no período de 1990 a 2001. As lactações utilizadas continham no mínimo quatro, e no máximo nove controles. Para o QQ, a curva média alcançou o “Nó” no segundo mês da lactação, já para o QQQ os “Nós” foram estimados para o segundo e quarto mês da lactação. Os parâmetros e cada função foram estimados por processos iterativos, usando o PROC REG do pacote estatístico SAS. Os critérios utilizados para verificar a qualidade do ajuste para cada função foram: Coeficiente de Determinação Ajustado (R^2_A), Teste de *Lack of Fit* (LOF) e gráficos de distribuição de resíduos.

Neste sistema os QQ e QQQ, acompanharam as variações na produção de leite ao longo da lactação, sendo não significativos os testes de *Lack of Fit*. O polinômio segmentado QQQ, estimou os menores resíduos e apresentou um $R^2_A = 99.98\%$.

Palavras-chave: búfalos, curvas de lactação, funções matemáticas, polinômios segmentados

Use of polynomials targeted for the study of lactation curves of Murrah buffaloes and their crosses in a semi-intensive production system in Sao Paulo state

Abstract

Considering the need to better understand the curve of lactation and the little information of this nature with buffalo, the purpose of this study was to evaluate the use of Polynomials Segments, data from models Quadratic-Quadratic (QQ) and three segments squared (QQQ), the adjustment of lactation curve average of the animals and their crossbred Murrah in a

system of farm semi-intensive. For these systems, which are carried twice millings day, records were used to produce three herds, totaling 3.267 lactations for a total of 27.600 test day. The test-day was done at intervals of 30 days, from 1990 to 2001. The lactations used contained at least four and no more than nine controls. For the QQ model, the average curve reached a "node" in the second month of lactation, as the QQQ model for the "node" have been estimated for the second and fourth month of lactation. The parameters and each function were estimated by iterative process, using the procedure PROC REG of SAS statistical package. The criteria used for evaluating the quality of adjustment for each function were: Coefficient of Determination Adjusted (R^2A), Test Lack of Fit (LOF) and graphics distribution residues.

In this production system the QQ and QQQ models, accompanied the changes in yield production during the lactation, with no significant tests LOF and the R^2A estimated for both function was over 99%. The QQQ models, estimated minors residues and presented a $R^2A=99.98\%$.

Keywords: buffaloes, lactation curves, mathematical functions, segmented polynomials

Introdução

A curva da lactação é a representação gráfica da produção diária, semanal ou mensal e pode ser usada para estimar a produção de leite em qualquer período ou na extensão da lactação. Na literatura encontra-se varias funções matemáticas utilizadas para descrever a produção de leite usando funções matemáticas e, entre elas, estão os polinômios segmentados.

Os polinômios segmentados foram definidos por Rice (1969) como “segmentos de polinômios de grau p , conectados em pontos chamados “nós”, os quais, tomam as primeiras derivadas contínuas de grau $p-1$ ”. Os segmentos que compõem estes polinômios são definidos por processos visuais, mediante um diagrama de pontos, com o fim de detectar os pontos de mudança do comportamento da curva (“nó”) (Schenkel 1989).

Seu uso surgiu com o intuito de solucionar problemas apresentados pelos polinômios ordinários e pelas funções não lineares, com a finalidade de estimar o melhor modelo para as curvas de lactação em gado leiteiro. Conhece-se que, os polinômios ordinários podem levar a problemas de multicolinearidade e de desuniformidade ao longo da curva; a dependência do comportamento da função, na área onde se concentra o maior número de pontos (Schenkel 1989) além de dificuldades na interpretação biológica dos parâmetros, quando estes são de grau elevado (Rice 1969). Os polinômios segmentados promovem quebra de multicolinearidade pela inclusão de zeros entre as variáveis explanatórias, o que, favorece a diminuição da variância dos parâmetros. Porém, apesar de solucionar problemas desta ordem, a estimação dos pontos de junção (“nós”) representam problemas quando se trabalha com polinômios segmentados, principalmente, quando estes são de grau elevado (Schenkel 1989).

Faro (1996) comparou três polinômios segmentados para o ajuste de curvas de lactação de vacas bovinas da raça Caracu: Quadrático-Linear, Quadrático-Quadrático e Quadrático com quatro segmentos e concluíram que este último promoveu o melhor ajuste, gerando desvios médios entre a produção total observada e estimada de -0.01% .

O objetivo do presente estudo foi avaliar os Polinômios Segmentados: Quadrático-Quadrático e Quadrático com três segmentos, para verificar o melhor ajuste da curva de lactação média uma população de bubalinos da raça Murrah e seus mestiços em um sistema semi-intensivo de criação.

Material e métodos

Os registros de produção de leite utilizados neste sistema de criação referem-se a 27.600 controles de produção de leite, correspondente a 3.267 lactações, de 840 búfalas da raça Murrah e seus mestiços, controladas no período de 1989 a 2001, em três fazendas localizadas no estado de São Paulo, Brasil.

Os animais alimentavam-se de Capim *Brachiaria decumbens* e *Brachiaria humidicola*, na época chuvosa (outubro a março). Na época seca (abril a setembro), a dieta era complementada com silagem de Capim do gênero *Braquiaria* e capim elefante (*Pennisetum purpureum*), milho (*Zea mays*) e cana picada (*Saccharum officinarum*). Durante todo o ano as fêmeas em lactação recebiam concentrada a base de algodão e/ou de soja e cevada oriunda de cervejaria. O sal mineral era fornecido à vontade em cochos.

Cada propriedade tinha em média 120 búfalas em ordenha, distribuídas em várias categorias de idade. A reprodução era realizada por monta natural e/ou inseminação artificial. A ordenha, mecânica, era realizada duas vezes ao dia, com bezerro ao pé. O controle leiteiro foi implantado a partir de 1989, com duas ordenhas ao dia. Os valores da produção de leite foram anotados em fichas, que depois foram cadastrados no sistema de informação Interherd, onde também foram feitas anotações referentes à vida reprodutiva de cada animal.

O controle sanitário, realizado em intervalos de quatro meses, era constituído de vacinação contra a febre aftosa, carbúnculo sintomático, raiva e brucelose, somente para as fêmeas. O controle de ecto e endoparasitos foram feitos, com uma a duas aplicações semestrais.

Foram eliminadas as lactações que iniciaram o controle de produção acima de 30 dias após o parto, lactações com intervalo de controles maior do que 45 dias, lactações com menos que quatro controles, de modo a ter lactações com, no mínimo, 115 dias de duração. O número máximo de controles observados numa lactação foi nove, sendo o período máximo observado de 270 dias, onde as lactações foram truncadas. Os dados foram classificados em meses de 30 dias, de acordo com a periodicidade com que os controles eram realizados.

Polinômios segmentados usados para o ajuste das lactações

Foram testados dois Polinômios Segmentados para ajustar as lactações para a curva média. O ajuste foi feito usando todas as produções em cada mês, considerando os modelos descritos a seguir:

- Polinômio Segmentado Quadrático-Quadrático (QQ)

$$y = b_0 + b_1X + b_2X^2 + (c_1 - b_2)Z$$

Onde:

$$Z=0, \text{ se } X \leq K;$$

$$Z=(X-K)^2, \text{ se } X > k.$$

- Polinômio Segmentado com Três Segmentos Quadrático (QQQ)

$$y = b_0 + b_1X + b_2X^2 + c_1Z_1 + c_2Z_2$$

Onde:

$$Z_1=(X-K_1)^2, \text{ se } X > K_1,$$

$$Z_2 = (X - K_2)^2, \text{ se } X > K_2.$$

Nos modelos descritos, Y é a produção de leite diária, X é o mês da lactação b_0, b_1, b_2, c_1, c_2 , são parâmetros de cada função K, K_1, K_2 , são os “nós” ou pontos de junção entre os segmentos dos polinômios segmentados.

Estimação do “nó” ou ponto de junção

A estimação do “Nó” foi feita, a princípio, por inspeção, utilizando um diagrama de pontos. Esse procedimento auxilia a visualização de quantos segmentos compõem a reta, além dos prováveis pontos onde ocorrem mudanças na curvatura. Em uma segunda etapa, valores iniciais obtidos através do diagrama foram continuamente experimentados no PROC REG (SAS 1995), de maneira a ser encontrado o ponto onde a soma de quadrados do resíduo (SQE) foi minimizada. O “Nó” estimado indica o ponto onde a mudança é mais significativa, porque representa a maior proporção de pontos; não sendo este, necessariamente, a melhor localização, sob o ponto de vista estratégico, o qual, no caso de curvas de lactação é o pico de produção. Depois de determinar o comportamento e a forma da curva da produção de leite em função dos meses da lactação, foram modeladas várias curvas, ajustadas por polinômios segmentados. A opção tomada foi pelas funções QQ, e QQQ, por apresentarem as melhores representações da curva média observada.

Estatística usada para a escolha da melhor função

Para escolha da melhor função foi empregado o coeficiente de determinação que mede a proporção da variação da variável dependente (produção de leite) que é explicada pelo modelo de regressão. Quando modelos estatísticos com número de parâmetros diferentes são comparados, é adequado utilizar o coeficiente de determinação ajustado (R^2_A), ao número de parâmetros de cada modelo para que estes sejam comparáveis. Por outro lado, o aumento do R^2_A , só ocorre quando a perda no número de graus de liberdade do resíduo é compensada pelo aumento da soma de quadrados do modelo de regressão, em consequência da inclusão de maior número de variáveis no modelo.

Segundo Draper e Smith (1981) a fórmula é a seguinte:

$$R^2_A = \frac{(n - 1)R^2 - p}{n - p - 1}$$

Onde:

p = Número de parâmetros do modelo de regressão

n = Número de observações

R^2 = Coeficiente de Determinação.

Gráfico de distribuição dos resíduos

Foram utilizados dois tipos de gráficos de distribuição de resíduos para mostrar a qualidade do ajuste proporcionado para cada função e a presença de auto-correlação residual para a curva média. O primeiro mostra a dispersão dos resíduos em função do tempo. Se os resíduos apresentam-se distribuídos aleatoriamente em relação à variável classificatória X , é indicativo de bom ajuste. Porém, se estes se apresentam em blocos de valores positivos ou negativos, é indicativo de auto-correlação residual positiva

ou negativa, o que, indica mal ajuste. O segundo tipo são os gráficos de *lag* de resíduos (Draper e Smith 1981). O gráfico mostra a existência ou não de auto-correlação serial e relaciona o (i)ésimo com o (i-1)ésimo resíduo, o que, permite a constatação da presença ou não de auto-correlação entre os resíduos sucessivos. Quando ocorre auto-correlação positiva o gráfico apresenta uma reta ascendente e se negativa, apresenta uma reta descendente. Se estes se apresentam dispersos sem mostrar alguma tendência, é indicativo de bom ajuste.

Teste do “Lack of Fit”

Segundo Searle (1971) e Draper e Smith (1981), este teste se baseia na suposição de que para um conjunto de dados onde há medidas repetidas da variável independente X , para diferentes medidas da variável dependente Y , é possível particionar a soma de quadrados do resíduo em dois termos. Um é o chamado <<Erro Puro>> e o outro <<Lack of Fit>> do modelo (Tabela 1).

Tabela 1. O quadro de análise de variância do Teste do Lack of Fit

Fonte de Variação	G.L.*	Soma de Quadrados	Quadrado Médio	F _c
Modelo de Regressão	$r-1$	SQR	$QMR=SQR/r-1$	
Erro Puro	$N-p$	SQEP	$QMEP=SQEP/N-p$	
Lack of Fit	$p-r$	SQLOF	$QMLOF=SQLOF/p-r$	$QMLOF/QMEP$

* N =número total de observações; r =número de parâmetros do modelo de regressão; p =número de classes da variável classificatória X ; $QMEP$ = quadrado médio do erro puro, estimado por $SQEP/(N-p)$, o qual é um estimador da variância residual (s^2); $QMLOF$ = quadrado médio devido ao “Lack of Fit”, representado por $SQLOF/(p-r)$, que estima s^2 se o modelo é correto; caso contrário, estima s^2 +vício, se o modelo é inadequado (Draper e Smith 1981)

Para um determinado conjunto de dados, onde X é o dia da lactação e Y a produção de leite correspondente, tem-se que X_1, X_2, \dots, X_p , são os p valores distintos de X , onde cada X_i ocorre n_i vezes, com n_i valores correspondentes de Y ; denominados, então, Y_{ij} , para $j=1, 2, \dots, n_i$ e para $i=1, 2, \dots, p$. Para todo i , desde que $i \geq 1$, pode-se considerar que:

$$n = \sum_{i=1}^p n_i = N$$

n = número total de observações

Então, a Soma de Quadrados do Erro será:

$$SQE = \sum_{i=1}^p \sum_{j=1}^{n_i} Y_{ij}^2 - \hat{Y}_{ij}^2 = \sum_{i=1}^p \sum_{j=1}^{n_i} Y_{ij}^2 - \hat{b}' X' Y$$

com $(N-p)$ graus de liberdade, a qual pode ser particionada em soma de quadrados do erro puro ($SQEP$) e soma de quadrados do “Lack of Fit” ($SQLOF$), dadas por:

$$SQEP = \sum_{i=1}^p \left[\sum_{j=1}^{n_i} Y_{ij}^2 - n_i \left(\bar{Y}_i \right)^2 \right]$$

$$SQLOF = SQE - SQEP$$

Com $(N-p)$ e $(p-r)$ graus de liberdade, respectivamente.

Estimativas das variáveis Tempo do Pico (TP) e Produção no Pico da lactação (PP)

Considerando-se que o pico de lactação ocorre nos primeiros meses após do parto, o tempo do pico foi estimado para os polinômios segmentados, em função do primeiro segmento quadrático de cada modelo. Assim o ponto de máximo ou mínimo é determinado de acordo com as restrições de uma equação quadrática. As produções no pico são estimadas em função do tempo do pico, também para cada primeiro segmento. Por exemplo, para um polinômio segmentado Quadrático-Quadrático, estarão dados:

$$y = b_0 + b_1X + b_2X^2 + (c_1 - b_2)Z$$
$$TP = -b_1/2b_2$$
$$PP = b_0 + b_1(TP) + b_2(TP)^2$$

Resultados e discussão

Curva de lactação média observada

A curva de lactação média observada para o sistema de criação em estudo pode ser observada na Figura 1.

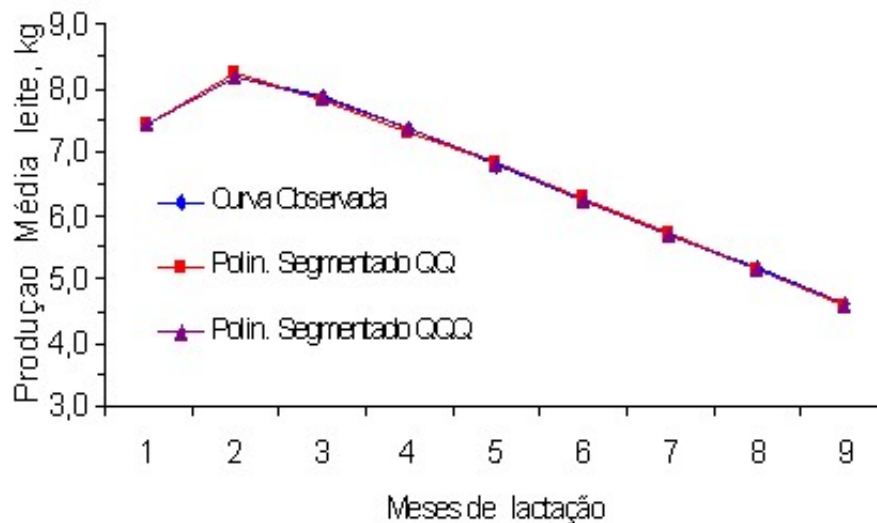


Figura 1. Curvas de lactação observadas e estimadas pelos Polinômios Segmentados QQ e QQQ, para búfalas Murrah e seus mestiço.

A produção média iniciou-se com 7,44 kg; o pico foi alcançado no segundo mês de lactação com produção de 8,16kg. A produção total de leite da lactação, estimada para a curva média foi de 2281,50 kg., em 270 dias de lactação. Para estes rebanhos o pico de lactação ocorre no segundo mês de lactação. Uma característica interessante dessa curva é que ela não apresenta um declínio brusco da produção de leite. Observa-se maior persistência durante a lactação, o que, poderia dever-se ao sistema de manejo dos animais recebendo melhor alimentação em comparação ao sistema extensivo (Faro 1996; Oliveira 1998; Neiva 1998). A curva média apresentou mudança de curvatura, aproximadamente, no segundo mês de lactação e uma pequena mudança na curvatura no sétimo mês de lactação. Isto pode ser explicado pelo o fato do final da lactação coincidir com a época das chuvas, onde os animais têm maior disponibilidade de pastagens.

Curvas de lactação estimadas

Para o polinômio segmentado quadrático com dois segmentos (QQ), composto por quatro parâmetros, os valores médios estimados foram de $4,19286 \pm 0,33496$; $4,47488 \pm 0,39523$; $-1,22773 \pm 0,0,10517$ e $1,25500 \pm 0,10839$, respectivamente, para b_0 , b_1 , b_2 e c_1 . O polinômio segmentado quadrático com três segmentos (QQQ), foi composto por cinco parâmetros, com valores médios de $4,82247 \pm 0,13858$ para b_0 , $3,5652 \pm 18036$ para b_1 , $-0,94767 \pm 0,05216$ para b_2 , $0,86231 \pm 0,06126$ para c_1 , e $0,08778 \pm 0,01272$ para c_2 .

Para o QQ o ponto de junção (Nó), foi estimado no segundo mês de lactação, depois a produção estimada apresentou uma tendência quadrática. A curva estimada (Figura 1) pelo QQ, superestimou a produção nos dois primeiros meses da lactação, e subestimou do 3^o ao 4^o mês, acompanhando muito bem o restante da lactação. Foram estimados resíduos que variaram entre $-0,07$ kg a $0,08$ kg (Figura 2).

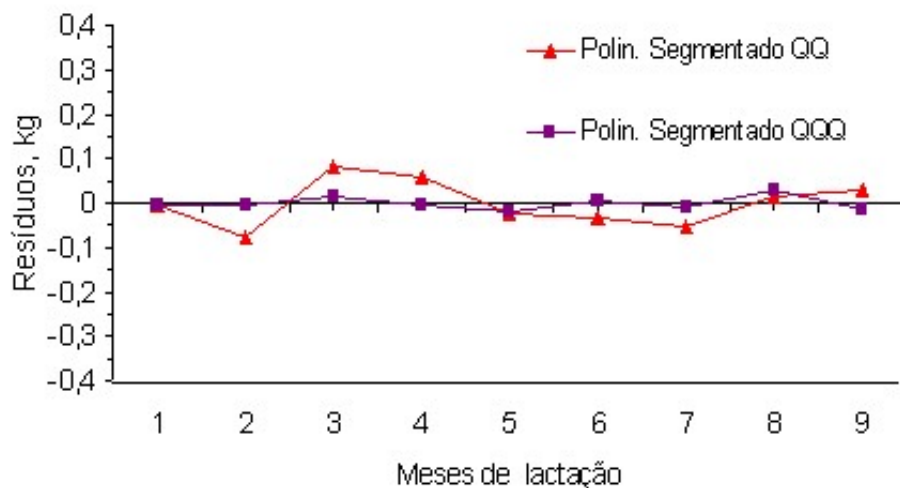


Figura 2. Distribuição dos resíduos estimados pelos Polinômios Segmentados QQ e QQQ, para búfalas da raça Murrah e seus mestiços

Os desvios foram maiores nos os primeiros meses da lactação, com exceção do 1^o mês. A disposição dos mesmos em relação aos meses de lactação sugere ausência de autocorrelação residual. Para o QQQ os pontos de junção entre os segmentos (Nó) foram estimados no segundo e quarto mês de lactação. Nas curvas estimadas (Figura 1) pelo QQQ, observa-se que a inclusão do terceiro segmento proporciona ao modelo maior flexibilidade e faz que a curva média estimada acompanhe as oscilações da produção ao longo da lactação, os resíduos estimados foram quase nulos, com valores de $-0,02$ kg a $0,03$ kg, o que, indica valores estimados muito próximos dos observados (Figura 2).

Os Coeficientes de Determinação Ajustado (R^2_A) estimados pela função foi superior a 99% e o test de Lack of fit (Tabela 2) resulta não significativo. Estas condições indicam que os modelos QQ e QQQ proporcionam bom ajuste da curva média de lactação.

Tabela 2. Média estimadas de Tempo do Pico (TP) e produção no Pico (PP); e teste estatístico utilizados para medir a qualidade de ajuste de cada função estudada.

Função	TP, mês	PP, kg	Lack of fit	R^2_A , %
--------	---------	--------	-------------	-------------

QQ	1,8	8,17	1,90 NS	99,74
QQQ	1,9	8,20	0,14NS	99,98

NS=Não significativo

Comparação entre as funções

As funções QQ e QQQ acompanharam a curva média observada (Figuras 1 e 2), e observa-se que o QQQ, por ter um segmento a mais, proporcionou ajuste quase perfeito. Observa-se na Tabela 2, que o QQQ acompanhou as oscilações da produção e de essa maneira foi a função que melhor descreveu a forma da curva de lactação média desses rebanhos. Segundo Faro (1996), a forma da curva de lactação de um rebanho deve ser um dos critérios considerados para a escolha de uma função. Uma maneira prática de fazer isso, além da visualização através de gráficos, é estimar os componentes da curva para cada função, principalmente, produção inicial, tempo no pico e produção no pico.

O QQQ apresentou resíduos mínimos e isto pode ser atribuído ao maior número de parâmetros, mostrando-se mais flexível e gerando desvios menores (Figuras 2 e 3). Ambos os modelos em estudo apresentaram ausência de auto-correlação residual.

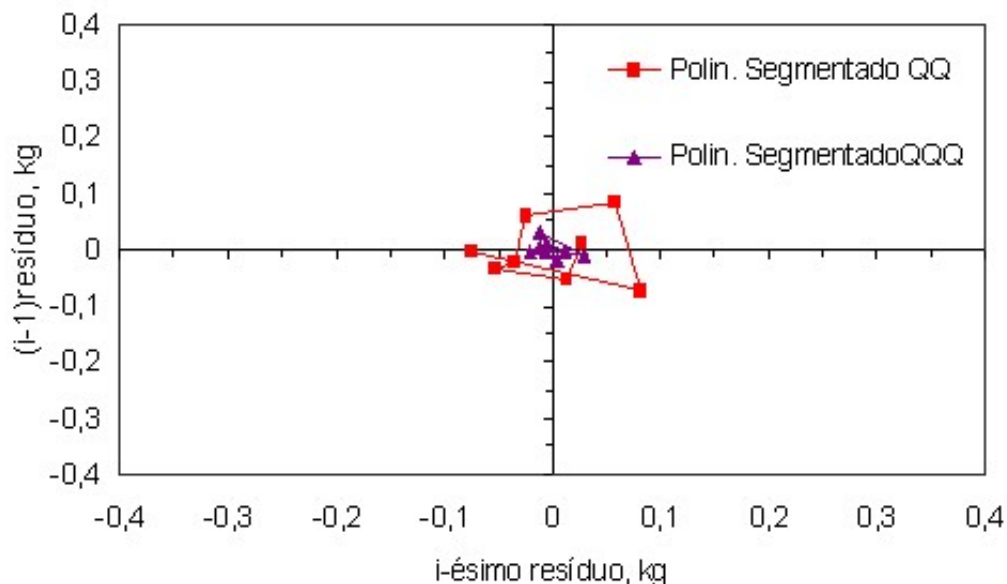


Figura 3. Gráfico de lag de resíduos, pelos Polinômios Segmentados QQ e QQQ, para búfalas da raça Murrah e seus mestiços

Segundo Glasbey (1998), isto ocorre quando o modelo é adequado. Ao comparar os modelos em função dos testes estatísticos utilizados para medir a qualidade do ajuste proporcionado pelas funções estudadas (Tabela 2), observa-se que ambas as funções (QQ e QQQ), apresentaram altos valores de R^2_A . Os testes de Lack of Fit não foram significativos para os dois modelos de polinômios segmentados estudados, e juntamente com a análise de resíduos, indicam que ambos modelos são adequados (Draper e Smith 1981). Tomando como base o coeficiente de determinação ajustado, o teste de Lack of Fit, distribuição dos resíduos, lag de resíduos, ambos os modelos podem ser utilizados para descrever a curva média, já que, apresentaram aproximações similares da produção ao longo de toda a lactação.

Ambos os modelos estudados, apresentaram coeficientes de determinação ajustado (R^2_A), semelhantes, e altos, sendo de 99,74% e 99,98%, para os QQ e QQQ respectivamente. Isto indica que qualquer modelo

dos polinômios segmentados estudados, poderia ser utilizado para o ajuste da curva de lactação nas condições consideradas. Estes modelos apresentaram aproximações muito boas entre Y e \hat{y} , já que seus resíduos foram pequenos. Isto pode dever-se ao fato de que a inclusão de segmentos foi compensada pela melhoria na qualidade do ajuste, para a curva média (Faro 1996; Ribeiro 1997).

Conclusões

- Os Polinômios Segmentados Quadrático-Quadrático e o Polinômio Segmentado com Quadrático-Quadrático-Quadrático apresentaram bom ajuste para a curva média de lactação de búfalas leiteiras.
- O coeficiente de determinação ajustado, teste de *Lack of Fit*, o gráfico de resíduos entre as produções observadas e estimadas e o *lag* de resíduos, podem contribuir na tomada de decisão quanto à qualidade de ajuste proporcionada por uma função.

Literatura citada

Draper N M and Smith H 1981 Applied Regression Analysis, 2nd edition, Wiley, New York

Faro L E 1996 Estudo da curva de lactação de um rebanho da raça caracu. Dissertação (Mestrado em Genética e Melhoramento Animal), UNESP - FCAVJ, SP. . 179p.

Glasbey C A 1998 Examples of regression with serially correlated errors. The Statistician. Edinburgh, v. 37, p. 277-291,

Neiva R S 1998 Produção de bovinos leiteiros. 1^a edição. Lavras: UFLA, 534p.

Oliveira M D S 1998 Pecuária leiteira. Jaboticabal: Funep, 70p.

Ribeiro M N 1997 Estudo da curva de lactação de um rebanho caprino no Estado da Paraíba. Tese (Doutorando em zootecnia), UNESP – FCAVJ, SP. . 91p.

Rice J R 1969 The approximation of functions. V. II. Addison-Wesley, Massachusetts.

Schenkel F S 1989 Utilização de Polinômios Segmentados na Pesquisa Zootécnica: Considerações Teóricas e Práticas, Universidade do Rio Grande do Sul, Porto Alegre, Brasil, (Notas)

Searle S R 1971 Linear Models, John Wiley and Sons, Inc., New York

SAS system for regression 1995 SAS Institute Inc., Cary, North Caroline .

Received 7 April 2008; Accepted 27 May 2008; Published 5 August 2008

[Go to top](#)