

Uso de polinômios segmentados para o estudo da curva de lactação em Búfalos Murrah e seus mestiços em sistema de criação extensivo, no Estado de São Paulo

M. Muñoz-Berrocal¹, H. Tonhati², M.F. Cerón-Muñoz³, L.O. Seno², A.R. Otaviano²

Facultad de Zootecnia, Universidad Nacional Agraria de la Selva. Apto. 156, Tingo María, Perú.

Use of segmented polynomials to study the lactation curve of Murrah and their crossed breeds of buffaloes under an extensive production system in the State of São Paulo

ABSTRACT. The objective of this study was to evaluate the use of segmented polynomials, described by the two models quadratic-quadratic (PSQQ) and that with three quadratic segments (PSQQQ), for adjustment of the mean lactation curve of Murrah females and their crosses under a system of extensive production, with once daily milking. Studies of this nature in buffaloes are practically non-existent. Milk production data from twelve herds, including 1280 lactations and 8839 test-day milk weights were analyzed. Test days were spaced at 30-day intervals, during the period from 1995 to 2001. Lactations included in the analysis were based on a minimum of four and a maximum of nine test days. For the PSQQ model, the mean curve reached its «Knot» (junction point between the segments) in the second month of lactation, while for the PSQQ model the «Knots» were estimated to be in the second and fourth months. The parameters of each function were estimated by interactive procedures, using PROC REG of SAS. Criteria used to verify the goodness of fit for each function were: adjusted coefficient of determination (R^2_A); lack of fit test (LOF); and graphic distribution of residuals and of residual lag. The two models closely described variations in milk production throughout the lactation with the lack of fit tests not significant. The coefficients of determination were practically identical for both models under study ($R^2_A = 91.3\%$).

Key words: Buffaloes, Lactation curves, Mathematical functions, Segmented polynomials.

© 2005 ALPA. Todos los derechos reservados

Arch. Latinoam. Prod. Anim. 2005. 13(1): 24-29

RESUMO. O objetivo deste estudo foi avaliar o uso de Polinômios Segmentados, dados pelos modelos Quadrático-Quadrático (PSQQ) e com três segmentos quadráticos (PSQQQ), no ajuste da curva de lactação média de animais Murrah e seus mestiços sobre um sistema de criação extensiva com uma ordenha por dia. Estudos desta natureza com búfalos praticamente não existem. Trabalhou-se com doze rebanhos, utilizando-se 1280 lactações, referentes a 8839 controles de produção de leite. O controle leiteiro foi feito em intervalos de 30 dias, no período de 1995 a 2001. As lactações utilizadas continham no mínimo quatro, e no máximo nove controles. Para o PSQQ, a curva média alcançou o «Nó» ponto de junção entre os segmentos no 2º mês da lactação, já para o PSQQQ os «Nós» foram estimados para o 2º e 4º mês da lactação. Os parâmetros de cada função foram estimados por processos iterativos, usando-se o PROC REG do 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), gráficos de distribuição de resíduos e o gráfico de Lag de resíduos. Os PSQQ e PSQQQ, acompanharam as variações na produção de leite ao longo da lactação, sendo não significativos os testes de Lack of Fit. Os coeficientes de determinação praticamente foram idênticos para os modelos em estudo ($R^2_A = 91,3\%$).

Palavras-chave: Búfalos, Curva de lactação, Funções matemáticas, Polinômios segmentados.

Recibido Enero 23, 2003. Aceptado: Falta información

¹Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo- FAPESP. E-mail: milthon_munoz@hotmail.com

²Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal (UNESP) - SP, Brasil. E-mail: tonhati@fcav.unesp.br

³Grupo Grica-Universidad de Antioquia-Colombia

Introdução

Os polinômios segmentados foram definidos por Rice (1969) como «segmentos de polinômios de grau p , conectados em pontos chamados «nós», que tomam as primeiras derivadas contínuas de grau $p-1$.» Os segmentos que o compõem são definidos por processo visual, com um diagrama de pontos, de forma a detectar-se os pontos de mudança de comportamento da curva («nó») (Schenkel, 1989).

Seu uso surgiu como o intuito de solucionar problemas apresentados pelos polinômios ordinários e pelas funções não lineares. Sabe-se que, os polinômios ordinários podem levar a problemas de multicolinearidade, desuniformidade ao longo da curva e à 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») é um problema ao se trabalhar com polinômios segmentados, principalmente, quando estes são de grau elevado (Schenkel, 1989).

El Faro *et al.*, (1996) comparam três polinômios segmentados para o ajuste de curvas de lactação de vacas 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 em alguns rebanhos Bubalinos da raça Murrah e seus Mestiços em os sistemas de criação extensivo.

Material e Métodos

Os registros de produção de leite utilizados neste sistema de criação extensivo, referem-se a 8839 controles de produção de leite, correspondentes a 1280 lactações de 623 búfalos Murrah e seus mestiços, de composição genética variada, controladas no período de 1995 a 2001, em doze fazendas localizadas na região do Vale do Ribeira, São Paulo.

Os animais alimentavam-se de capins do gênero *Brachiaria* como *B. decumbens* e *B. humidicola* principalmente, e pastagens nativas e naturalizadas como

Paspalum sp., capim-nobre ou angola, grama folhavelha e taquarinha.

Cada propriedade tem em média 20 búfalas em lactação, sendo animais de várias categorias de idade, os animais não recebem nenhuma suplementação ao longo do ano. O sal mineral é fornecido à vontade em cochos. Utiliza-se o sistema de monta natural e inseminação artificial. A ordenha é manual, realizada uma vez ao dia, com bezerro ao pé, sendo que o leite de um quarto do úbere é reservado à alimentação do bezerro, até o término da lactação, que coincide com o desmame.

O controle sanitário era realizado em intervalos de 4 meses, para a febre aftosa, além das vacinações contra carbúnculo sintomático e a raiva, e somente para as fêmeas, contra brucelose. O controle de ectoparasitos e endoparasitos era feito, com 1 a 2 aplicações semestrais.

O controle leiteiro foi implantado nas fazendas no ano de 1995, com uma ordenha ao dia. Os valores da produção de leite foram anotados em fichas, que depois são cadastrados no sistema de informação Interherd, onde também são feitas anotações referentes à vida reprodutiva de cada animal.

Para a análise estatística descritiva foram usados os dados colhidos das doze fazendas, que tinham o mesmo sistema de criação. Foram eliminadas as lactações que iniciaram o controle de produção 30 dias após o parto, lactações com controles maiores do que 45 dias entre eles e lactações com menos que 4 controles, de modo a ter lactações com, no mínimo, 115 dias de duração. O número máximo de controles observados numa lactação foi de 9, 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 vários Polinômios Segmentados, sendo os PSQQ e PSQQQ os que melhor ajustaram as lactações para a curva média. O ajuste foi feito usando-se todas as produções em cada mês, considerando os modelos descritos a seguir:

- Polinômio Segmentado Quadrático-Quadrático (PSQQ)

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

onde: $Z=0$, se $X \leq K$;

$Z=(X-K)^2$, se $X > K$.

- Polinômio Segmentado com Três Segmentos Quadrático (PSQQQ)

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

onde: $Z_1=(X-K_1)^2$, se $X > K_1$

$Z_2=(X-K_2)^2$, 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 que se encontre o ponto onde a soma de quadrados do resíduo (SQE) é minimizada. O «Nó» estimado indicará o ponto onde a mudança é mais significativa, porque representa a maior proporção de pontos. Isso não indica, necessariamente, que estrategicamente é a melhor localização. 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, modelou-se várias curvas ajustadas por polinômios segmentados. Obteve-se pelas funções PSQQ, e PSQQQ, por apresentarem as melhores representações da curva média observada.

Estatística Usada para a escolha da Melhor Função

Quando compara-se modelos com diferentes números de parâmetros, é adequado utilizar o coeficiente de determinação ajustado ao número de parâmetros de cada modelo, para que estes sejam comparados em iguais condições. 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

O coeficiente de determinação mede a proporção da variação da variável dependente (produção de leite), que é explicada pelo modelo de regressão. Quando diversos modelos com números de parâmetros diferentes são comparados, é usual ajustar o R^2 ao número de parâmetros de cada modelo. Assim, 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.

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 autocorrelação 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. O segundo tipo é o gráfico de *lag* de resíduo (DRAPER e SMITH, 1981). O gráfico mostra a existência ou não de autocorrelação serial e relaciona o (i)ésimo com o (i-1)ésimo resíduo, o que permite a constatação da presença ou não de autocorrelação entre os resíduos sucessivos. Quando ocorre autocorrelação 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»

SEARLE (1971) e DRAPER e SMITH (1981), indicam que o teste baseia-se 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 (Apêndice: Quadro 1).

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 $i > 1$, pode-se considerar que:

$$n = \sum_{i=1}^p n_i = N = \text{número total de observações}$$

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:

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

$$SQLOF = SQE - SQEP$$

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

Resultados e Discussão

Para o PSQQ o ponto de junção entre os segmentos (Nó) foi estimado no segundo mês, depois a produção estimada apresentou tendência quadrática. As curvas médias observada e estimadas por este polinômio são mostradas na Figura 1. Esta função acompanhou as oscilações da curva observada ao longo da lactação, e produziu resíduos que variaram

Quadro 1. Apêndice: O quadro de análise de variância do Teste do Lack of Fit, representado por:

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	$QMEP=SQEP/N-p$	
<i>Lack of Fit</i>	$p-r$	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).

entre -0,03 kg a 0,12 kg (Figura 2). Os desvios, foram maiores nos primeiros meses de lactação, apresentando-se dispostos maioritariamente em blocos de valores positivos (Figura 3), não observando-se tendência de autocorrelação residual positiva.

O teste de *Lack of Fit* não foi significativo, e o coeficiente de determinação Ajustado (R^2_A) estimado pela função foi de 91,34%, indicando que o modelo mostra um bom ajuste para o PSQQ.

As curvas médias observadas e estimadas para o PSQQQ encontram-se na Figura 1. Esta função, similarmente ao PSQQ, também estimou bem as produções de leite ao longo de toda a lactação, e sendo os desvios dos resíduos muitos pequenos, variando -0,03 kg a 0,14 kg. Os resíduos apresentaram-se dispostos em blocos positivos (Figura 3). Entretanto, estes valores foram superiores aos obtidos para a Função PSQQ. Na Figura 4, não se observou tendência de autocorrelação residual positiva, e o teste de *Lack of Fit* não foi significativo. De modo geral, as aproximações proporcionadas pelos PSQQ e PSQQQ foram idênticas, sendo os coeficientes de determinação ajustados exatamente os mesmos. Se-

gundo RIBEIRO (1997), é esperado, entretanto, que a inclusão de um outro segmento tornaria o modelo mais flexível, proporcionando maior capacidade de acompanhar as oscilações da produção ao longo de toda a lactação.

Comparação entre as funções

Os Polinômios segmentados testados, PSQQ e PSQQQ, acompanharam de perto os segmentos de curva observada (Figura 1), representou bem as oscilações que ocorrem na produção até o final da lactação. Ambos os polinômios apresentaram resíduos muitos próximos e reduzidos. Segundo Glasbey (1988), isto ocorre quando o modelo é adequado. Os menores resíduos foram apresentados pelo PSQQ, os quais variaram de -0,02 a 0,12 kg (Figura 2). Valores muitos pequenos, indicam ausência de autocorrelação residual positiva ou negativa (Figuras 3), e mostram bom ajuste da curva média estimada (Ribeiro, 1997).

Ambos os modelos estudados, apresentaram coeficientes de determinação ajustado (R^2_A), semelhantes, e altos, sendo de 91,34% e 91,33%, para os PSQQ e PSQQQ respectivamente. Isto indica que

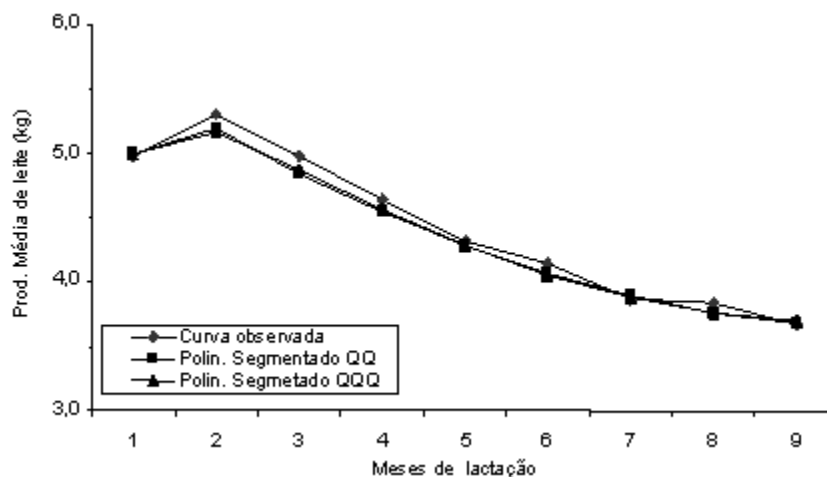


Figura 1. Curvas de lactação observadas e estimadas pelos, Polinômios Segmentados PSQQ e PSQQQ, para búfalas Murrah e seus mestiços.

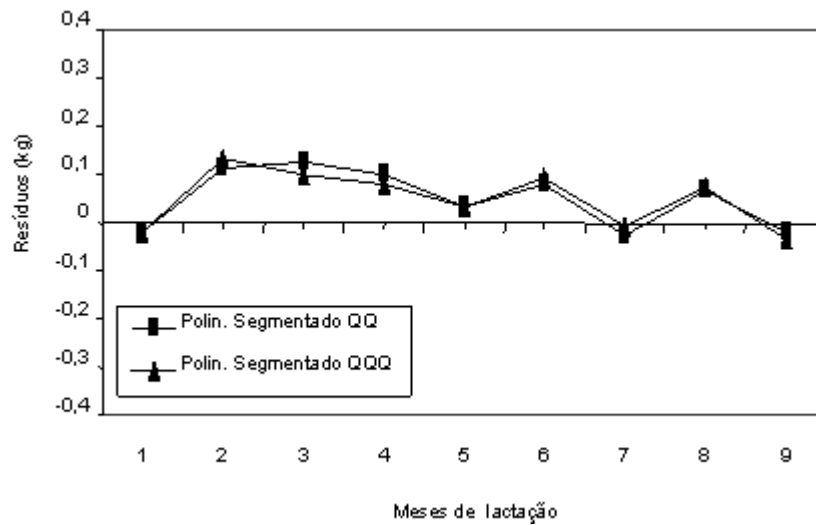


Figura 2. Distribuição dos resíduos estimados, pelos Polinômios Segmentados PSQQ e PSQQQ, para búfalas Murrah e seus mestiços.

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 v , 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 (El Faro, 1996; Ribeiro, 1997).

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).

Na Tabela 1, encontra-se um resumo dos testes estatísticos utilizados para medir a qualidade do ajuste proporcionado pelas funções estudadas.

Tomando como base o coeficiente de determinação ajustado, o teste de *Lack of Fit*, distribuição dos resíduos, *lag* de resíduos, ambos modelos podem ser utilizados para descrever a curva média, já que apresentaram aproximações similares da produção ao longo de toda a lactação.

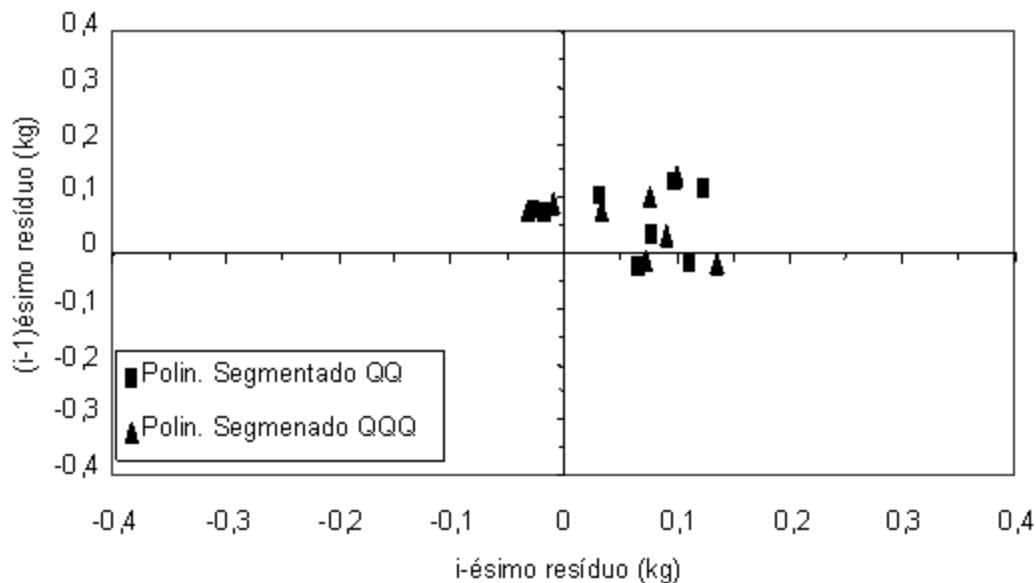


Figura 3. Gráfico de *lag* de resíduos, pelos Polinômios Segmentados PSQQ e PSQQQ, para búfalas Murrah e seus mestiços.

Tabela 1. Testes estatísticos utilizados para medir a qualidade de ajuste de cada função estudada

Função	Lack of fit	R ² _A %
PSQQ	0,50 NS	91,34
PSQQQ	0,53 NS	91,33

NS=Não significativo

Conclusões

De acordo com os resultados deste estudo, pode-se concluir que:

Os Polinômios Segmentados Quadrático-Quadrático e o Polinômio Segmentado com Três Segmentos Quadráticos, apresentaram bom ajustamento para a curva média de lactação em búfalas.

O coeficiente de determinação ajustado, tempo de pico, produção no pico, 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., H. Smith, H. 1981. Applied Regression Analysis, 2.ed., New York, Wiley.
- El Faro, L. 1996. Estudo da curva de lactação de um rebanho da raça Caracu. 1996. 179 f. Dissertação (Mestrado em Zootecnia). Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal.
- Glasbey, C. A. Examples of Regression with sereally correlated errors. The Statistician. Edinburgh, v.37, p.277-291, 1988.
- Ribeiro, M. N. 1997. Estudo da curva de lactação de um rebanho caprino no Estado da Paraíba. Tese (Doutorado em Zootecnia), UNESP - FCAVJ, SP. 91p.
- Rice, J. R. 1969. The approximation of functions. V. II. Addison-Wesley, Massachusetts.
- SAS. 1995. System for Regression. SAS Institute Inc., Cary, North Caroline.
- Searle, S. R. 1971. Linear Models, New York, John Wiley & Sons, Inc.
- Schenkel, F. S. 1989 Utilização de Polinômios Segmentados na Pesquisa Zootécnica: Considerações Teóricas e Práticas, Universidade do Rio Grande do Sul, Porto Alegre. (Notas)