

**ANA PAULA CKROH
ELIANE DE FÁTIMA COIMBRA**

MÉTODOS ESTATÍSTICOS NO AJUSTE DE CURVAS DE LACTAÇÃO

**CURITIBA
2005**

ANA PAULA CKROH
ELIANE DE FÁTIMA COIMBRA

MÉTODOS ESTATÍSTICOS NO AJUSTE DE CURVAS DE LACTAÇÃO

Trabalho de Conclusão de Curso apresentado à disciplina de Laboratório de Estatística do Curso de Graduação em Estatística da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientadora: Prof^a Dr^a Suely Ruiz Giolo

CURITIBA
2005

SUMÁRIO

RESUMO	iii
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	3
2.1 A raça Caracu.....	3
2.2 A curva de lactação.....	6
2.3 Métodos estatísticos.....	9
2.3.1 Histórico.....	9
2.3.2 Modelos lineares generalizados para dados longitudinais	13
2.3.2.1 Modelos lineares generalizadas.....	13
2.3.2.2 Função de ligação	15
2.3.2.3 Dados longitudinais	16
2.3.2.4 Equações de estimação generalizadas.....	16
2.3.2.5 Seleção do modelo	17
2.3.2.6 Teste de hipótese	18
2.3.3 Modelos não-lineares.....	19
3 MATERIAL E MÉTODOS	22
3.1 Material.....	22
3.1.1 Características da fazenda	22
3.1.2 Manejo	23
3.1.3 Alimentação.....	24
3.1.4 Reprodução	24
3.2 Métodos	25
3.2.1 Modelo linear generalizado para dados longitudinais.....	25
3.2.2 Modelo não-linear para descrever a curva de lactação	27
4 RESULTADOS E CONCLUSÕES	30
4.1 Análise descritiva	30
4.2 Modelos lineares generalizados para dados longitudinais	40
4.2.1 Ajuste do modelo	40
4.2.2 Seleção do modelo	40
4.2.3 Modelo ajustado	41
4.2.4 Estimativas dos parâmetros do modelo escolhido	41
4.2.5 Curva de lactação ajustada apenas para combinação de estimativas	44
4.2.6 Diagnóstico do ajuste do modelo	44
4.3 Ajuste de modelos não-lineares.....	46
5 CONCLUSÕES E CONSIDERAÇÕES FINAIS	56
REFERÊNCIAS BIBLIOGRÁFICAS	58
APÊNDICES	62

MÉTODOS ESTATÍSTICOS NO AJUSTE DE CURVAS DE LACTAÇÃO

RESUMO

Este trabalho teve como objetivo avaliar a utilização de modelos estatísticos, lineares e não-lineares, no ajuste de curvas de lactação de vacas Caracu. Foram utilizadas 32.978 observações colhidas semanalmente durante o período de 1991 a 1998 de 825 vacas de um único rebanho da Fazenda Chiqueirão, localizada em Poços de Caldas, MG. Anterior ao ajuste de modelos, foi realizada uma análise descritiva com o intuito de explorar alguns aspectos considerados relevantes para o estudo. Sob o contexto de modelos lineares generalizados para dados longitudinais foram, então, ajustados dois modelos lineares. Para o primeiro modelo foi assumida, para a variável resposta, a distribuição Gama e, para o segundo, a distribuição Normal. Para ambos os modelos foram testadas três estruturas de correlação: a auto-regressiva, a independente e a permutável. A estrutura selecionada, de acordo com o critério de informação de Akaike, foi a auto-regressiva. Diversas covariáveis foram consideradas nos modelos, dentre elas, o mês de retiro e a idade da vaca. Modelos não-lineares foram, a seguir, considerados. Inicialmente, um modelo que considera a função gama incompleta (Wood, 1967) foi ajustado à curva de lactação média. Este mesmo modelo foi também ajustado às curvas de lactação individuais. Devido às limitações computacionais do pacote estatístico R, utilizado nas análises, não foi possível incorporar covariáveis nestes modelos. A partir dos parâmetros estimados para os modelos não-lineares, foi possível a obtenção de estimativas de quantidades de interesse, tais como, o tempo de pico da lactação e a persistência. De modo geral, os modelos ajustados apresentaram-se razoáveis. O modelo linear gama com estrutura de correlação auto-regressiva, escolhido dentre os analisados, mostrou a necessidade de estudos adicionais quanto ao diagnóstico do ajuste, bem como uma maneira de se considerar a informação do parentesco entre os animais. Por outro lado, os modelos não-lineares precisam ser estendidos para que se possa levar em conta as covariáveis e, também, o parentesco. Um esforço computacional se faz necessário nesse sentido.

Palavras-chave: modelos lineares generalizados, modelos não-lineares, raça Caracu.

1 INTRODUÇÃO

No Brasil, a quantidade de leite produzida não atende à demanda da população em relação à necessidade mínima exigida para uma base alimentar satisfatória, já que a produtividade do rebanho bovino leiteiro brasileiro é baixa. Isso se deve ao fato do pouco conhecimento a respeito da pecuária leiteira, bem como a falta de tecnologia e investimentos na atividade. Portanto, para que isto não aconteça, a cada dia que passa aumenta o número de interessados em controlar a produção de seus rebanhos. Esse controle é um fator determinante para o sucesso da exploração leiteira. Ele pode ser feito, na prática, por meio do estudo de curvas de lactação, que nada mais é do que a representação gráfica da produção de leite de uma fêmea no decorrer de sua lactação, em geral de 305 dias ou 10 meses.

O conhecimento das estimativas dos parâmetros da curva de lactação é importante no monitoramento de um rebanho, pois expressam grande quantidade de informação inerentes ao animal durante toda a sua vida. Segundo Keown e Van Vleck (1973), a curva de lactação pode ser dividida em três segmentos:

- 1) Fase desde o dia de parição até o pico da lactação;
- 2) Fase que vai do pico até aproximadamente 270 dias de lactação; e
- 3) Fase próxima ao final da lactação.

A curva de lactação é distinta entre as várias espécies de mamíferos. Animais de porte maior como a vaca, tem uma curva de lactação longa, com o pico de lactação no início do período. A produção de leite e a forma da curva de lactação podem ser influenciadas por diversos fatores, dentre os quais: ambiente, ano de parição, estação de parição, idade ao parto, ordem de lactação, duração dos períodos secos, duração de serviço, duração da gestação, e o grupo genético.

Touchberry (1974) observou que o pico de produção e a própria forma da curva de lactação são dependentes da condição física da vaca ao parto, do seu potencial genético, do seu estado de saúde e do regime alimentar nos vários estágios de lactação.

O estudo sobre curvas de lactação pode contribuir para entender melhor o sistema de produção, uma vez que o conhecimento do comportamento da curva e de suas implicações sobre a produção de leite auxiliam o produtor na previsão da produção de suas vacas, bem como possibilitam a identificação de possíveis falhas no sistema de manejo do rebanho, como alimentação deficiente, instalações inadequadas, patologias não aparentes, entre outras. Além disso, possibilitam a estimativa da produção total de leite, do pico de produção e da persistência da produção.

Curvas de lactação podem ser estudadas por meio de modelos que foram e vem sendo propostos ao longo dos anos para descrever a curva de lactação. Nesse contexto, o presente trabalho foi motivado pela necessidade do pesquisador em extrair informações a respeito do desempenho dos animais e tem como objetivo geral avaliar a utilização de modelos estatísticos, lineares e não-lineares, no ajuste de curvas de lactação de vacas Caracu. Especificamente, os objetivos são:

- 1) Ajustar curvas de lactação para vacas da raça Caracu;
- 2) Avaliar os modelos estatísticos (lineares e não-lineares) ajustados às curvas de lactação de vacas da raça Caracu.

2 REVISÃO DE LITERATURA

2.1 A raça Caracu

De acordo com texto publicado na Revista Agropecuária Tropical, de Agosto/Setembro de 2001, o gado Caracu surgiu por volta de 27 a.C., quando o imperador Júlio César Augusto alcançou o rio Loire, na região da Gália, ali deixando um gado que, mais tarde, chegaria a Portugal e a Hispânia (200 a.C.). Desde o Império Romano, passando pela Idade Média, foram se misturando o Bos Aquitanicus (França), o Bos Ibericus (Portugal e Espanha) e o Bos Batavicus (ancestral do gado Holandês), resultando na formação das ibéricas (Portugal e Espanha): Alentejana, Minhota, Mirandesa, Arauquesa, Galesa etc. Boa parte dos estudiosos acredita que o Caracu seja originário das raças Minhota e Alentejana, ou da mistura de diversos gados ibéricos, ocorrida nos séculos XVI e XVII, no Brasil. A origem científica, no entanto, só será possível quando for realizado o mapeamento do genoma do Caracu.

No Brasil, o Caracu era uma das raças de gado preferida no final do século XIX, dando origem a diversos grupamentos étnicos, como o Jaguanes, o Lageano, o Curraleiro, o Francano (Franqueiro ou Pedreiro), o Junqueiro, o Bruxo, o Baio, o Mocho etc.

A primeira exposição brasileira de bovinos aconteceu em 1908, na capital Rio de Janeiro, comemorando cem anos da abertura dos Portos. Ali, o Caracu brilhou ao lado de diversas raças européias. Foi tanto o sucesso que, em 1909, foi fundado o Posto de Seleção de Nova Odessa (SP), para as raças Caracu e Mocho Nacional. Depois, em 1911, aconteceu a primeira exposição de Uberaba, num recinto planejado e decorado como se fosse a Índia, com a finalidade de promover o gado Zebu. Ali, no auge da festança, o campeão entre todas as raças presentes, foi o touro "Brasil", um Caracu. Durante muitos anos, os mestiços de Caracu com o Zebu deixaram claro qual o melhor caminho da pecuária brasileira. Os mestiços de Caracu com Zebu ocuparam novas fronteiras, mas o Caracu tão importante nos cafezais, era muito caro para esse

mister e, então, os mestiços de Zebu ganharam destaque. Com o passar dos tempos, o Zebu mostrou ser muito rústico e, então, milhares de vacas Caracu foram absorvidas na formação dos azebuados da época que levaram à formação de gados como o Indubrasil, Induaraxá, Indoporã, Induberaba, Indubahia, dentre outros, sempre garantido aptidão leiteira e o grande porte. Neste período, não havia distinção entre as raças zebuínas, tudo era Zebu, desde que tivesse uma giba no dorso.

Em 1929, devido ao "Big Crack" (Grande Quebra) da Bolsa de Nova Iorque, houve uma bancarrota no mundo inteiro. No Brasil, aconteceu a derrocada do setor rural, resultando na incineração de mais de 80 milhões de sacas de café. A queda do "ouro verde" fez com que milhares de fazendas desaparecessem do mapa. Os fazendeiros mudaram-se para regiões longínquas onde passaram a explorar um gado sem especialização. Assim, desapareceu a maioria do Gado Caracu, que era fornecedor do trabalho agrícola, leite e carne nas próprias fazendas de café. De 1930 a 1970, o Caracu sobreviveu nas mãos dos abnegados criadores, passando por um lento e firme aperfeiçoamento racial e zootécnico, aumentando a rusticidade, a habilidade maternal, a produção leiteira e a precocidade.

O Caracu, entre todas as raças do Brasil Colonial, era o "melhorador". O Caracu, também conhecido como "Colonião" ou "Gado Laranja", era então, o "ponto alto" e o "ponto final" desse grande período. Era assim considerado por ser um gado de grande porte, produtor de leite, muito manso e no abate garantia um bom rendimento. Eles sobreviviam a condições ambientais adversas mais que outras raças, uma aptidão selecionada por muitos e muitos anos. Não é à toa que seu nome "caracu" significa, na língua tupi-guarani, "bicho que vive escondido no mato" ou "kara-ku" que significa "bicho forte, de tutano forte, graúdo" (na língua guarani).

Pode-se dizer que o Caracu é, no mundo, a única raça européia selecionada para conviver com o clima tropical e, nesse aspecto, seu valor é incalculável. Existem outras raças Ibéricas, mas a única em evolução positiva é o Caracu.

Segundo a Associação Brasileira de Criadores de Caracu, a seleção natural desta raça provocou modificações anatômicas e fisiológicas que lhe proporcionam algumas características:

- Pêlo curto;
- Resistência ao calor;
- Resistência a endo e ectoparasitas;
- Facilidade de locomoção;
- Cascos resistentes, tanto para solos duros quanto encharcados;
- Umbigo curto e sem prolapso;
- Capacidade de digerir fibras grosseiras;
- Facilidade de parto.

O padrão da raça é chifre alaranjado, com saída para os lados, orelhas pequenas, pelagem nos vários tons de amarelo, sem pêlos ou manchas brancas, estrutura longilínea, linha de dorso plana, com pequena inclinação na garupa, prepúcio curto, vassoura do rabo amarela, mucosa alaranjada, e cascos claros, avermelhados ou rajados.

Em regime exclusivo de pasto, o peso médio das vacas está em torno de 550 a 650kg. Tendo cascos de até 750kg. Os touros pesam ao redor de 1.000kg podendo chegar a 1.200kg. Aos dois anos as novilhas atingem cerca de 400kg, existindo alguns animais que chegam a pesar 500kg. Os bezerros de um ano atingem uma média de 300kg devido à boa habilidade materna das matrizes.

A produção em rebanhos de seleção leiteira está em torno de 2.100 quilos por lactação (inclui novilhas de 1ª cria) em regime de pasto com pequena suplementação. Produz um leite com alto teor de gordura, em torno de 5%, e um extrato seco também elevado. No Brasil, apresenta duração de lactação de, aproximadamente 10 meses, bem superior às observadas nas raças nativas colombianas que apresentam duração de lactação de aproximadamente 5 meses.

As fêmeas são colocadas em reprodução a partir dos 14-15 meses, podendo ser mantidas em reprodução até os 16-17 anos (com casos de parição aos 21 anos), tendo aquelas mais férteis, de 11 a 13 partos. Um touro Caracu, em uma estação de monta normal, serve cerca de 50 fêmeas ou mais, com alto índice de prenhez positiva.

O rebanho Caracu não fica reunido no pasto. Ele se espalha, diminuindo a intensidade de pisoteio, possibilitando um melhor aproveitamento e longevidade das pastagens. Nas regiões frias, o gado penetra nas matas, consumindo folhas e protegendo-se do frio. No cerrado brasileiro, no período de escassez de alimentos (seca), ele complementa sua dieta com vários tipos de arbustivos nativos.

A Figura 2.1 apresenta fêmeas da raça Caracu nas fases bezerra, jovem e adulta.

Figura 2.1 – Imagens de fêmeas da raça Caracu nas fases bezerra, jovem e adulta

(a) bezerra caracu



(b) vaca jovem caracu



(c) vaca caracu adulta



2.2 A curva de lactação

O Brasil possui um grande rebanho bovino leiteiro, entretanto, os níveis de produtividade estão bastante aquém do desejado. Nesse contexto, um dos objetivos dos pesquisadores e produtores de leite tem sido usar uma metodologia de seleção que leve a um aumento dessa produção (Gonçalves et al., 2002).

A baixa produtividade dos rebanhos leiteiros das raças nativas em regiões tropicais é, em parte, decorrente da baixa produção de leite e da curta duração de lactação. O estudo da duração da lactação assume importância considerável por ser uma variável que reflete diretamente na produção total de leite do rebanho e, portanto,

na eficiência econômica da atividade. O período de lactação consiste no tempo decorrido entre o parto e a secagem da vaca. Nas raças especializadas, a lactação apresenta duração, em geral, de 10 meses ou mais. As raças nativas em regiões tropicais apresentam, em geral, lactação mais curta (Quirino et al., 1998)

Para tanto, torna-se fundamental estimar, a partir de registros parciais, a produção total de leite de um animal em lactação. Uma forma prática e consistente de obter este controle é pelo estudo da curva de lactação. O conhecimento das curvas de lactação de um rebanho auxilia na adequação de técnicas de alimentação e manejo, e no descarte e seleção de animais de acordo com um padrão desejável, pré-estabelecido de acordo com a capacidade de produção. Permite, também, que seja acompanhada a evolução de produção leiteira dos animais, com o conhecimento de suas variações ao longo de uma lactação. Dessa maneira, a comparação da forma da curva entre grupos distintos de animais, com diferentes idades ao parto, ano e mês de parto, retiros e outras variáveis de interesse, seria de grande importância, pois, mediante essas comparações, poderiam ser obtidas informações sobre a eficiência desses grupos, estimando sua produção de leite parcial ou total, propiciando um melhor controle de produção.

Em geral, a curva de lactação apresenta uma fase ascendente até o pico de produção, aproximadamente entre 30 a 90 dias de lactação, seguida de decréscimo relativamente consistente até o décimo mês de lactação. Cobuci (2000) define a curva de lactação como sendo a representação gráfica da produção de leite no decorrer da lactação de uma vaca.

As curvas de lactação e os parâmetros calculados a partir delas (pico e persistência da lactação), vêm sendo utilizados há muito tempo para auxiliar o manejo de fazendas leiteiras. A persistência da lactação foi definida na terceira década do século passado (Gaines, 1927); logo após, tanto a persistência quanto o pico de lactação foram relatados como as principais características para se descrever a curva de lactação. Segundo Camargo (2000), a persistência de lactação é definida como a constância na produção de leite durante a lactação, e medida pela queda na produção a

partir do pico na lactação, que ocorre por volta dos 60 dias pós-parto. Uma das características que define uma matriz leiteira especializada é a persistência de lactação. Quanto a este aspecto, uma vaca para ser considerada especializada, deverá apresentar média de produção de leite no décimo mês de lactação equivalente a 60% da média de produção de leite obtida no pico de lactação. Para Santos (2000), vaca de alta persistência é aquela que apresenta queda de produção após o pico de lactação de no máximo 10% a cada 30 dias, isto é, capaz de manter pelo menos 90% da produção a cada 30 dias após o pico de lactação. Vacas com alta persistência são capazes de produzir leite por mais de 10 meses, enquanto vacas com baixa persistência normalmente produzem leite por apenas cinco a nove meses.

Dentre os fatores que podem influenciar a produção de leite e, consequentemente, alterar o formato da curva de lactação, com mudanças no pico de produção e declividade da curva, destacam-se o rebanho, o ano de parto, a estação de parto e a idade da vaca ao parto. Estes fatores foram significativos na maioria dos trabalhos consultados. Na raça Guzerá, Cobuci et al. (2000) encontraram efeito significativo do rebanho, ano e idade da vaca ao parto sobre a produção de leite e as estimativas dos parâmetros da curva de lactação. Junqueira et al. (1997), ao estudarem a lactação de vacas da raça Holandesa, verificaram que todos os parâmetros da função gama incompleta foram influenciados pela idade da vaca ao parto.

Em outros estudos, verificou-se, ainda, que vacas na primeira lactação apresentam maior persistência do que as que possuem mais idade. Vacas que parem no verão tendem a produzir menos leite do que vacas paridas em outras estações. E, finalmente, vacas mais velhas tendem a ser mais severamente afetadas pela estação do ano do que as que estão em sua primeira lactação.

2.3 Métodos estatísticos

As funções utilizadas para o ajuste de curvas de lactação podem ser classificadas em relação aos seus parâmetros como: lineares, intrinsecamente lineares e não-lineares. As lineares são as mais utilizadas, pois seus parâmetros apresentam apenas relações aditivas entre si, podem ser estimadas por regressões lineares, além de terem interpretação biológica ou estatística mais fácil (Faro et al., 2002).

2.3.1 Histórico

Um dos primeiros trabalhos a propor um modelo para descrever a curva de lactação foi o de Brody et al. (1923), o qual, segundo Wood (1967), não proporciona um ajuste adequado, pois não considera o pico de lactação.

Na literatura, a função mais conhecida para o ajuste de curvas de lactação é a gama incompleta (Wood, 1967), seja na sua forma não-linear ou linearizada. Muitas das funções propostas na literatura são modificações da função de Wood. Outra função conhecida é a polinomial inversa, que é também uma função intrinsecamente linear. Bianchini Sobrinho (1984) também propôs duas funções lineares denominadas linear hiperbólica e quadrática logarítmica. Entretanto, alguns pesquisadores afirmam que o modelo mais adequado será diferente para cada região geográfica e para cada situação climática.

Recentemente, alguns trabalhos foram realizados com o intuito de comparar a qualidade do ajuste em relação às curvas de lactação, proporcionadas por funções propostas na literatura. Cobuci et al. (2000), em um estudo sobre “Curva de Lactação na Raça Guzerá”, estimaram os parâmetros de 22 funções por meio de regressão não-linear, em função do número de dias do parto até a avaliação do controle leiteiro. Na Tabela 2.1 podem ser visualizadas as funções que apresentaram melhor ajuste. A análise para a escolha dos modelos que melhor se ajustaram às produções de leite se baseou no percentual de lactação em que os parâmetros das funções foram

significativos e no número de curvas atípicas encontradas para cada modelo. Foram avaliadas também as influências dos fatores não genéticos sobre a produção de leite total. Coeficientes de herdabilidade e repetibilidade foram estimados para as características: produção de leite total, produção inicial de leite e taxa de declínio da produção de leite.

Tabela 2.1 – Funções que apresentaram melhor ajuste (Cobuci et al., 2000)

Função	Autor
$y = ane^{-cn}$	Papajcsik e Boderó (1988)
$y = a - cn + \ln(n)$	Cobuci et al. (2000)
$y = a - cn$	Madalena et al. (1979)
$y = ae^{-cn}$	Brody et al. (1923)

Faro et al. (2002) utilizaram quatro funções no ajuste de curvas de lactação de vacas da raça Caracu, as quais são apresentadas na Tabela 2.2.

Tabela 2.2 – Funções lineares e não-lineares utilizadas no ajuste de curvas de lactação

Função	Função linear	Autor
$Y_x = a_0 + a_1X + a_2X^{-1}$	Linear Hiperbólica	Bianchini Sobrinho (1984)
$Y_x = a_0 + a_1X + a_2X^2 + a_3 \ln(X)$	Quadrática Logarítmica	Bianchini Sobrinho (1984)
Função	Função não-linear	Autor
$Y_x = a_0X^{a_1}e^{-a_2X}$	Gama Incompleta	Wood (1967)
$Y_x = X(a_0 + a_1X + a_2X_2)^{-1}$	Polinomial Inversa	

As funções dos parâmetros como: produção inicial (PI), tempo de pico (TP), produção no pico (PP) e persistência (S) foram estimadas sob cada modelo conforme apresentado na literatura. Para comparação das funções, os autores utilizaram: coeficiente de determinação ajustado ao número de parâmetros (R_A^2), estatística de Durbin-Watson (DW), médias e coeficientes de variação para os parâmetros e funções

dos parâmetros. Baseado nestes critérios, concluíram que as quatro funções indicaram bons R_A^2 para a curva de lactação média. Já para os ajustes das curvas de lactação individuais, as funções apresentaram falhas na descrição das curvas. A função quadrática logarítmica apresentou melhor resultado quanto ao R_A^2 e estatística de DW. Contudo, a grande variação ocorrida nos parâmetros e funções dos parâmetros estimados indicaram pouca coerência na interpretação biológica dos resultados.

Gonçalves et al. (2002) também estudaram, para a raça Holandesa, o modelo quadrático logarítmico, a função gama incompleta, Regressão Múltipla e Multifásicos (Monofásico e Difásico). Para a escolha do modelo/função que melhor se ajustou às curvas de lactação, foram considerados os critérios: menor desvio-padrão residual, coeficiente de determinação ajustado ao número de parâmetros (R_A^2), auto-correlação de resíduos de 1ª ordem e o teste de Durbin-Watson. Foram avaliadas, para cada modelo, as funções das estimativas dos parâmetros, produção inicial, pico de produção, produção em 305 dias, tempo até o pico e duração da fase. Segundo os critérios para a escolha da curva, evidenciou-se que os modelos estudados se ajustaram bem aos controles mensais de produção de leite. Entretanto, verificou-se que a produção de leite no início da lactação foi subestimada pelas funções quadrático-logarítmica, gama incompleta e a de regressão múltipla. A função monofásica foi considerada inadequada, visto que o tempo até o pico foi quase três vezes superior ao dos demais modelos. Contudo, os autores constataram que a função difásica foi a que apresentou resíduos menores, mais simétricos e menos correlacionados que as demais. A curva desta função foi a que melhor estimou a produção de leite, ajustando-se adequadamente aos dados.

Outro estudo realizado sobre curvas de lactação foi o de Silva et al. (2005). Neste estudo, o objetivo dos autores foi o de utilizar o método Bayesiano no ajuste do modelo de Wood em dados de produção de leite de cabras da raça Saanen de primeira e segunda lactação. Segundo os autores, os modelos não-lineares normalmente utilizados na descrição de curvas de lactação, utilizam meios visando linearizar o modelo mediante transformação logarítmica. Porém, em muitos casos, devido às

irregularidades nos dados longitudinais e às correlações existentes entre os parâmetros do modelo, esses métodos produzem estimativas irreais, levando a confecção de curvas de lactação atípicas. Observaram, também, que em outros estudos o método Bayesiano foi utilizado com sucesso, uma vez que considera todos os parâmetros como variáveis aleatórias, o que reduz, substancialmente, o número de curvas atípicas, além de necessitar de um número menor de dados de produção por animal. Realizado o estudo, concluíram que realmente o método Bayesiano é eficiente no estudo das curvas de lactação quando os parâmetros são correlacionados entre si.

No mesmo ano, Muñoz-Berrocal et al. (2005) realizaram um estudo sobre o “uso de modelos lineares e não-lineares para o estudo da curva de lactação em Búfalas Murrah e seus mestiços em sistema de criação semi-extensivo, no Estado de São Paulo”. Na Tabela 2.3 pode-se observar as cinco funções ajustadas pelos autores à curva de lactação média.

Tabela 2.3 – Funções lineares e não-lineares ajustadas à curva de lactação média

Função	Função linear	Autor
$Y_x = b_0 + b_1X + b_2X^{-1}$	Linear Hiperbólica	Bianchini Sobrinho (1984)
$Y_x = b_0 + b_1X + b_2X^2 + b_3 \ln(X)$	Quadrática Logarítmica	Bianchini Sobrinho (1984)
Função	Função não-linear	Autor
$Y_x = b_0e^b 1^{x+b} 2^{x^2}$	Parabólica Exponencial	
$Y_x = b_0X^{b_1}e^{b_2X}$	Gama Incompleta	Wood (1967)
$Y_x = X(b_0 + b_1X + b_2X_2)^{-1}$	Polinomial Inversa	

Para a escolha da melhor função foi utilizado o coeficiente de determinação (R_A^2). Para mostrar a qualidade do ajuste proporcionado para cada função e a presença de autocorrelação residual para a curva média, foram utilizados dois tipos de gráficos de distribuição de resíduos. Ao comparar os modelos concluíram que as funções lineares; quadrática logarítmica e linear hiperbólica foram aquelas que apresentaram melhor ajuste à curva média de lactação.

2.3.2 Modelos lineares generalizados para dados longitudinais

2.3.2.1 Modelos lineares generalizados

Na maioria dos trabalhos consultados, foram utilizados, além dos modelos não-lineares, os modelos normais lineares para descrever a curva de lactação. No entanto, nos casos em que eram utilizados os modelos normais lineares, em que a suposição de normalidade para a resposta não era razoável, tentava-se algum tipo de transformação no sentido de alcançar a normalidade procurada.

Com o avanço das técnicas computacionais, os modelos que exigiam a utilização de processos iterativos para a estimação dos parâmetros começaram a ser mais utilizados. A proposta mais interessante foi apresentada por Nelder e Wedderburn (1972), que propuseram os modelos lineares generalizados, que podem ser vistos como uma generalização do modelo tradicional de regressão linear, em que ao invés da suposição da variável resposta com distribuição normal, é assumido que a mesma pertence à família exponencial de distribuições. A ligação entre a média e o preditor linear, não é necessariamente mais a identidade, podendo assumir qualquer forma monótona não-linear.

Segundo Paula (2002), os modelos lineares generalizados podem ser definidos da seguinte forma: suponha Y_1, \dots, Y_n variáveis aleatórias independentes, cada uma com função densidade de probabilidade dada por

$$f_y(y_i; \theta_i, \phi) = \exp[\phi\{y\theta_i - b(\theta_i)\} + c(y_i, \phi)], \quad (1)$$

em que: $E(Y_i) = \mu_i = b'(\theta_i)$, $\text{Var}(Y_i) = \phi^{-1}V_i$, com $V = d\mu/d\theta$ a função de variância e $\phi^{-1} > 0$ o parâmetro de dispersão.

A função de variância desempenha um papel importante na família exponencial, uma vez que ela caracteriza a distribuição. Isto é, dada a função de variância, tem-se uma classe de distribuição correspondente, e vice-versa. Essa

propriedade permite a comparação de distribuições por meio de testes simples para a função de variância. Uma propriedade interessante envolvendo a distribuição de Y e a função de variância é a seguinte:

$$\sqrt{\phi}(Y - \mu) \xrightarrow{d} N(0, V(\mu)) \text{ quando } \phi \rightarrow \infty.$$

Ou seja, para ϕ grande, Y converge em distribuição para a normal de média μ e variância $V(\mu)/\phi$. Essa abordagem assintótica foi introduzida por Jorgensen (1987).

Os modelos lineares generalizados são definidos pela expressão em (1) e pelo componente sistemático

$$g(\mu_i) = \eta_i,$$

em que: $\eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ é o preditor linear, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $p < n$, é o vetor de parâmetros desconhecidos, $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^T$ denota os valores de p variáveis explicativas e $g(\cdot)$ é uma função monótona e diferenciável, denominada função de ligação.

Dois exemplos de distribuições pertencentes à família exponencial são: a distribuição Normal e a distribuição Gama.

a) Distribuição Normal

Para a distribuição normal tem-se a função densidade de probabilidade

$$f(y/\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}, \text{ com } y \in \mathfrak{R}, \mu \in \mathfrak{R} \text{ e } \sigma > 0.$$

Então, fazendo-se $\theta = \mu$ e $\phi = \sigma^2$ segue que

$$f(y/\theta, \phi) = \exp\left\{\frac{y\theta - \theta^2/2}{\phi} - \frac{1}{2}\left[\frac{y^2}{\phi} + \ln(2\pi\phi)\right]\right\},$$

tal que: $a(\phi) = \phi$, $b(\theta) = \frac{\theta^2}{2}$, $c(y, \phi) = -\frac{1}{2}\left[\frac{y^2}{\phi} + \ln(2\pi\phi)\right]$, $E(Y) = \mu = \theta$ e $Var(Y) = \phi$.

b) Distribuição Gama

Para a distribuição Gama tem-se que a função de densidade de probabilidade é

$$f(y/\alpha, n) = \frac{\alpha^n}{\Gamma(n)} e^{-\alpha y} y^{n-1}, \text{ com } \alpha > 0, n > 0 \text{ e } y > 0.$$

Então, considerando: $\theta = -\frac{\alpha}{n}$ e $\phi = n$, segue que

$$f(y/\theta, \phi) = \exp\left\{\frac{y\theta + \ln(-\theta)}{\phi^{-1}} + (\phi - 1)\ln(y\phi) + \ln(\phi) - \ln(\Gamma(\phi))\right\}$$

Tal que: $a(\phi) = \phi^{-1}$, $b(\theta) = -\ln(-\theta)$, $c(y, \phi) = (\phi - 1)\ln(y\phi) + \ln(\phi) - \ln(\Gamma(\phi))$, $E(Y) = -\frac{1}{\theta}$

e $Var(Y) = \frac{1}{\phi\theta^2}$.

2.3.2.2 Função de ligação

A função de ligação relaciona a combinação linear das variáveis explicativas, isto é, o preditor linear, com o valor esperado da variável resposta.

Como caso particular, para cada distribuição, tem-se a função de ligação canônica quando $\eta_i = \theta_i$, para $i = 1, 2, \dots, n$. Em relação às distribuições Normal e Gama citadas anteriormente tem-se:

- Para a distribuição Normal: $\eta = \theta \Leftrightarrow g(\mu) = \mu$, ou seja, a função de ligação canônica é a identidade.
- Para a distribuição Gama: $\eta = \theta \Leftrightarrow g(\mu) = -\mu^{-1}$.

Nota-se que a função de ligação canônica assegura que os valores obtidos para μ sejam sempre admissíveis.

2.3.2.3 Dados longitudinais

Segundo Costa (2003), os estudos longitudinais constituem um caso especial dos estudos de medidas repetidas, em que os tratamentos são alocados às unidades experimentais e os dados são coletados mais de uma vez para cada unidade experimental. Desta forma, têm-se dois fatores a serem estudados, tratamento e tempo, em que tratamento é o fator entre-indivíduos e tempo o fator intra-indivíduos.

A característica distinta de estudos longitudinais é a dimensão ordenada com que os dados são coletados e o fato de que as observações repetidas para um indivíduo tendem a ser correlacionadas. Tal correlação pode ser modelada por meio de uma estrutura de correlação dos dados observados. Modelar uma estrutura de correlação apropriada é essencial para que as inferências sobre as médias sejam válidas.

Como visto nos estudos anteriores, o controle da produção de leite pode ser feito em dias, semanas e meses, durante todo o período de lactação. Este controle feito no decorrer da lactação é um caso típico de estudo longitudinal.

2.3.2.4 Equações de estimação generalizadas

Para obter as estimativas associadas ao vetor de parâmetros de regressão β , faz-se necessário obter as equações de estimação generalizadas (GEE) definidas por

$$\sum_{l=1}^n \frac{\partial \mu_l}{\partial \beta} \text{cov}(Y_l)^{-1} \{y_l - \mu_l\} = 0.$$

Para levar em consideração a correlação entre as respostas de cada unidade, Liang e Zeger (1986) definiram a matriz de correlação de trabalho $R(\alpha)$, de modo que:

$$\text{cov}(Y_l) = V_l^{\frac{1}{2}} R(\alpha) V_l^{\frac{1}{2}}.$$

Entre algumas das principais estruturas de correlação propostas estão: a auto-regressiva, a independente e a permutável, definidas a seguir.

- *auto-regressiva*, $\text{corr}(Y_{1j}, Y_{1j+k}) = \alpha^k$, para $j = 1, \dots, n_1$ e $k = 1, \dots, n_1$ e igual a zero se $j+k > n_1$. Se a distribuição da variável resposta é normal, esta estrutura de correlação corresponde à do modelo autoregressivo de ordem k . Por exemplo, no modelo autoregressivo de ordem 1 tem-se:

$$R(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{pmatrix}.$$

- *independente*, considera as respostas independentes numa mesma unidade. Isto implica que R é a matriz identidade, ou seja, assume a forma:

$$R(\alpha) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}_{n_1 \times n_1}.$$

- *permutável*, $\text{corr}(Y_{1j}, Y_{1j+k}) = \alpha$, significa que a correlação entre quaisquer variável é a mesma e, sendo assim:

$$R(\alpha) = \begin{pmatrix} 1 & \alpha & \dots & \alpha \\ \alpha & 1 & \dots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \dots & 1 \end{pmatrix}_{n_1 \times n_1}.$$

2.3.2.5 Seleção do modelo

Para escolher a estrutura de correlação que melhor se ajusta aos dados é necessário utilizar algum critério de seleção. Dentre os critérios, destaca-se o Critério de Informação de Akaike (AIC), que considera o valor do logaritmo da função de verossimilhança penalizado para o número de parâmetros estimados, isto é,

$$AIC = -2\ell + 2p$$

sendo l o máximo do logaritmo da função de verossimilhança e p o número de parâmetros do modelo. A estrutura de correlação que apresentar o menor valor para o AIC é considerada a mais adequada aos dados.

2.3.2.6 Teste de hipótese

Segundo Demétrio (2002), os métodos de inferência para os modelos lineares generalizados baseiam-se, fundamentalmente, na teoria de máxima verossimilhança. De acordo com esta teoria, existem três estatísticas para testar hipóteses relativas aos parâmetros β 's que são deduzidas de distribuições assintóticas de funções adequadas das estimativas dos β 's. São elas: i) razão de verossimilhanças; ii) Wald; e iii) escore, as quais são assintoticamente equivalentes e, sob H_0 e para ϕ conhecido, convergem para uma variável com distribuição χ_p^2 , sendo, porém, a razão de verossimilhanças, o critério que define um teste uniformemente mais poderoso.

Quando se tem um vetor de parâmetros, muitas vezes há interesse no teste de hipóteses de apenas um subconjunto deles. Seja, então, uma partição do vetor de parâmetros β dada por

$$\beta = [\beta_1^T \beta_2^T]^T$$

em que β_1 , de dimensão q , é o vetor de interesse e β_2 , de dimensão $(p - q)$, o vetor *nuisance*. De forma semelhante, tem-se a partição da matriz do modelo $X = [X_1 X_2]$, do vetor escore $U = \phi^{-1} X^T W \Delta (y - \mu) = [U_1^T U_2^T]^T$ com $U_1 = \phi^{-1} X_1^T W \Delta (y - \mu)$ e da matriz de informação de Fisher para $\hat{\beta}$

$$\mathfrak{I} = \frac{1}{\phi} X^T W X = \begin{bmatrix} \mathfrak{I}_{11} & \mathfrak{I}_{12} \\ \mathfrak{I}_{21} & \mathfrak{I}_{22} \end{bmatrix},$$

sendo que $\mathfrak{I}_{12} = \mathfrak{I}_{21}^T$.

Usando resultados conhecidos de álgebra de matrizes, envolvendo partição de matrizes, tem-se, para amostras grandes, a variância assintótica de $\hat{\beta}_1$

$$\text{Var}\left(\hat{\beta}_1\right) = \left(\mathfrak{I}_{11} - \mathfrak{I}_{12}\mathfrak{I}_{22}^{-1}\mathfrak{I}_{21}\right)^{-1} = \phi\left[X_1^T W^{1/2} (I - H_2) W^{1/2} X_1\right]^{-1},$$

sendo $H_2 = W^{1/2} X_2 (X_2^T W X_2)^{-1} X_2^T W^{1/2}$.

Sejam as hipóteses $H_0 : \beta_1 = \beta_{1,0}$ versus $H_0 : \beta_1 \neq \beta_{1,0}$, sendo $\beta_{1,0}$ um valor específico para β_1 . Seja $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1^T & \hat{\beta}_2^T \end{bmatrix}^T$ o estimador de máxima verossimilhança para β sem restrição e $\hat{\beta}_0 = \begin{bmatrix} \beta_{1,0}^T & \hat{\beta}_{2,0}^T \end{bmatrix}^T$ com $\hat{\beta}_{2,0}$ o estimador de máxima verossimilhança para β_2 sob H_0 .

O teste de Wald é baseado na distribuição normal assintótica de $\hat{\beta}$, sendo uma generalização da estatística t de Student. É, geralmente, o mais usado no caso de hipóteses relativas a um único coeficiente β_j . Assintoticamente $\hat{\beta} \sim N_p(\beta, \mathfrak{I}^{-1})$, assim a estatística para esse teste é dada por

$$W = \left(\hat{\beta}_1 - \beta_{1,0}\right)^T \left[\text{Var}\left(\hat{\beta}_1\right) \right]^{-1} \left(\hat{\beta}_1 - \beta_{1,0}\right),$$

sendo $\text{Var}\left(\hat{\beta}_1\right)$ a $\text{Var}\left(\hat{\beta}_1\right)$ avaliada em $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1^T & \hat{\beta}_2^T \end{bmatrix}^T$. Para amostras grandes, rejeita-

se H_0 , a um nível de significância de $\alpha\%$, se $W > \chi_{q,1-\alpha}^2$.

2.3.3 Modelos não-lineares

Existem, muitas situações nas quais não é desejável, ou mesmo possível, descrever um fenômeno por meio de um modelo de regressão linear.

Como visto na literatura, os modelos não-lineares são muito utilizados para descrever curvas de lactação, sendo que um dos primeiros modelos propostos foi a

Gama Incompleta (Wood, 1967). A partir disso, outras funções foram propostas, mas a grande maioria são modificações da função de Wood.

Nos modelos não-lineares, em vez de se fazer uma descrição puramente empírica do fenômeno em estudo, pode-se, a partir de suposições importantes sobre o problema (frequentemente dadas através de uma ou mais equações diferenciais), trabalhar no sentido de obter uma relação teórica entre as variáveis observáveis de interesse. O problema, diferentemente do caso linear, é que os parâmetros entram na equação de forma não-linear, o que dificulta a estimação dos parâmetros do modelo. Outra vantagem dos modelos não-lineares é a obtenção de parâmetros que são facilmente interpretáveis.

Em muitas situações, necessitam-se menos parâmetros nos modelos não-lineares do que nos lineares, o que simplifica e facilita a interpretação. Alguns aspectos do uso de modelos não-lineares são citados a seguir:

- Os modelos não-lineares têm uma base teórica, os parâmetros dos modelos fornecem um maior conhecimento sobre o fenômeno em estudo do que os modelos lineares.
- Os modelos não-lineares, geralmente fornecem um bom ajuste, com menos parâmetros do que os modelos lineares.
- A transformação de um modelo não-linear em um modelo linear nos parâmetros, se por um lado facilita o processo de ajuste, implica em fazer suposições não realísticas sobre o termo dos erros (distribuição normal com variância constante); além disso, perde-se informação sobre os erros- padrão dos parâmetros originais.
- Além disso, existem modelos que são intrinsecamente não-lineares, isto é, não podem ser linearizados por transformação.
- Embora se usem variáveis contínuas como variáveis independentes, não há razão para que as variáveis independentes, nos modelos não-lineares, sejam contínuas. Ao contrário, pode-se fazer uso de variáveis *dummy* para indicar

a presença ou ausência de um grupo, ou codificar diferenças entre indivíduos (dados de medidas repetidas).

Por outro lado, pode ser citado um ponto negativo desses modelos, os mesmos não levam em consideração a estrutura de correlação. Contudo, nos estudos de curvas de lactação, os controles são feitos diariamente, semanalmente ou mensalmente e as observações repetidas para um indivíduo tendem a ser correlacionadas.

A forma geral do modelo não-linear é dada da seguinte forma:

$$Y_i = f(X_i, \gamma) + \varepsilon_i \quad (2)$$

em que:

$$X_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ \cdot \\ \cdot \\ X_{iq} \end{bmatrix} \quad \gamma = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \cdot \\ \cdot \\ \gamma_{p-1} \end{bmatrix}.$$

A função $f(X_i, \gamma)$ é não-linear; os erros ε_i tem média zero, variância constante e são não correlacionados. Assume-se, em geral, que os erros apresentam distribuição normal, são independentes e com variância constante, com γ o vetor de parâmetros desconhecidos do modelo.

Nos modelos não-lineares não é possível encontrar formas analíticas para os estimadores de mínimos quadrados ou máxima verossimilhança. Sendo assim, métodos numéricos devem ser usados juntamente com os métodos referidos, o que requer cálculos computacionais intensivos.

A função de Wood, que é a mais citada na literatura, é expressa por

$$Y_x = a_0 X^{a_1} e^{-a_2 X} + \varepsilon_i, \quad i = 1, \dots, n,$$

em que $\varepsilon_i \sim N(0, \sigma^2)$; Y_x é a variável dependente que representa a produção de leite na semana X de lactação e a_0 , a_1 e a_2 são parâmetros desconhecidos a serem estimados.

3 MATERIAL E MÉTODOS

3.1 Material

Os registros de lactação utilizados no presente estudo foram obtidos do rebanho da raça Caracu da fazenda Chiqueirão, localizada em Poços de Caldas, MG, e referem-se ao período de 1978 a 1998. Do arquivo inicial, com 86.598 controles de 2.155 vacas em primeira lactação, foram utilizados 32.978 controles de 825 vacas, do período de 1991 a 1998, pois foi verificada a falta de registros no ano de 1990. Por este fato, optou-se por utilizar a segunda parte do banco, visto serem informações mais recentes.

Os controles foram registrados semanalmente, provenientes de ordenha manual, realizada duas vezes ao dia, com a presença do bezerro ao pé da vaca. O intervalo entre uma ordenha e outra foi de aproximadamente 9 horas. Os registros de lactação das vacas apresentaram entre 1 e 43 controles leiteiros, totalizando 32.978 controles.

O banco de dados comporta, ainda, a origem do animal, bem como os códigos do pai e da mãe, que são apresentados para cada vaca. O número de retiros para ordenha que a fazenda possui vão de 1 a 15, sendo que durante a lactação o animal passa por até três retiros. A idade de cada vaca ao parto também é apresentada, variando de 26,8 a 50 meses. Outro fator que compõe o banco de dados é o mês de controle, variando de janeiro a dezembro, sendo que, em geral, a lactação de cada vaca dura em média 10 meses, podendo iniciar em qualquer estação do ano e terminando aproximadamente 10 meses após seu início. Os animais são considerados de médio e grande porte, sendo que as vacas adultas pesam aproximadamente 600kg.

3.1.1 Características da fazenda

A fazenda Chiqueirão situa-se a 21° 47' de latitude sul e 48° 34' de longitude oeste, na divisa dos estados de Minas Gerais e São Paulo, nos municípios de Poços de Caldas - MG e São Sebastião da Gramma - SP. A altitude varia de 1400 a 1700 metros,

a temperatura anual média fica em torno de 20°C e a precipitação anual média é em torno de 1600 mm. Possui atualmente cerca de 4.800 hectares, sendo 2.700 destinados ao rebanho. Na parte paulista, o solo é rico e a temperatura mais elevada, sendo possível a cultura de café. Nessa área, as pastagens predominantes são de capim-gordura (*Melinis minutiflora*, Pal. de Beauv). Na parte mineira, o solo é bem mais pobre, rico em alumínio, o relevo menos montanhoso, onde predominavam os campos nativos. São feitas culturas de batata inglesa, milho e, nas terras mais acidentadas, eucalipto. As pastagens predominantes são de *Panicum maximum*, consorciados com leguminosas como o trevo branco e o cornichão. Durante o inverno, quando as temperaturas são baixas, planta-se azevém para pastoreio. Atualmente, a área destinada ao rebanho vem diminuindo significativamente com o plantio de eucalipto. No entanto, as novas tecnologias de formação de pastagem, juntamente com o aumento da fertilidade do solo devido à cultura da batata, têm permitido manter o rebanho estável em número de animais.

3.1.2 Manejo

A fazenda possui duas maternidades e quinze retiros somente para ordenha, que comportam entre 50 e 75 vacas. As vacas parem a pasto e são levadas para as maternidades, quando então são separadas dos bezerros, só estando juntas dos mesmos no momento da ordenha. São feitas duas ordenhas diárias, às 5 horas da manhã e às 14 horas da tarde. A ordenha é feita com bezerro ao pé da vaca. As várias categorias de animais são mantidas em pastos, separados por idade e sexo. Por ser uma raça bastante rústica, o controle de endoparasitas é bastante fácil, sendo que em torno de 3 controles por ano em animais abaixo de 2 anos de idade são suficientes. O controle de ectoparasitas deve ser sistemático, principalmente da mosca do berne (*Dermatobia hominis*), muito abundante na região. A incidência de carrapatos é baixa, não sendo necessário um controle específico. As vacinações contra febre aftosa, carbúnculo sintomático, raiva (hidrofobia) e brucelose são feitas periodicamente. Também é feita uma pré-imunização com sangue aos 4 e 9 meses de idade, visando o controle da

tristeza bovina causada principalmente por *Babesia bovis* e *Anaplasma marginali*. Atualmente, a utilização de vacinas contra diarreia neonatal em vacas pré-parto tem ajudado na imunização de suas crias contra essa doença. Para facilitar o manejo de ordenha, todas as fêmeas são descornadas aos 30 dias de idade.

3.1.3 Alimentação

A alimentação do rebanho da fazenda Chiqueirão é feita basicamente a pasto, exceção feita para o período de seca (junho a setembro), quando se utiliza silagem de milho, dada a ocorrência de geadas. Para as vacas em lactação é fornecido cerca de 2 kg/dia de ração concentrada até aproximadamente o quarto mês de lactação. Após esse período, uma menor quantidade de ração é fornecida, dependendo da produção do lote de animais. Aos bezerros é deixada uma teta pela manhã e uma teta à tarde, e fornecido cerca de 1,5 kg/dia de ração concentrada até o décimo mês de idade. Todos os animais têm acesso a misturas naturais.

3.1.4 Reprodução

O manejo reprodutivo dos animais da fazenda é feito utilizando-se apenas monta natural. Utiliza-se uma estação de monta bastante longa, que se estende de abril a fevereiro do ano subsequente, evitando-se que ocorram partos nos meses de dezembro a janeiro, dada a grande intensidade das chuvas. As novilhas entram na estação de monta a partir de 18 meses de idade. Utiliza-se um touro para cada lote de 25 a 30 novilhas. O touro é normalmente contemporâneo das novilhas em idade, mas evita-se o acasalamento entre os animais com parentesco próximo. Nos retiros de ordenha, a entrada dos touros se dá aproximadamente 2 meses após o parto. Apenas um touro é colocado por retiro e, se necessário, o mesmo é substituído, dando-se um intervalo de 30 dias da saída para a substituição. Procura-se não utilizar touros com idade acima de oito anos, evitando-se uma endogamia mais estreita.

3.2 Métodos

As análises estatísticas foram efetuadas com a utilização do *software* estatístico R, de distribuição livre, que pode ser encontrado em <http://www.r-project.org>. Inicialmente, foi feita uma análise descritiva dos dados com o intuito de explorar alguns aspectos considerados relevantes para o estudo. Em seguida, foram ajustados aos dados modelos lineares e não-lineares para descrever a curva de lactação.

3.2.1 Modelo linear generalizado para dados longitudinais

Para dados longitudinais, como é o caso dos dados analisados neste trabalho, tem-se: $\tilde{y}_i = (y_{i1}, y_{i2}, \dots, y_{in_i})'$ uma sequência de n_i medidas observadas para a i -ésima vaca ($i = 1, \dots, m$) e t_j ($j = 1, \dots, n_i$) os tempos em que estas medidas são tomadas para cada vaca. Associado a cada y_{ij} tem-se, também, valores x_{ijk} ($k = 1, \dots, p$) de p covariáveis. Devido à estrutura longitudinal dos dados é esperado que exista correlação entre as observações medidas em uma mesma vaca.

Assumimos que os y_{ij} são realizações de variáveis aleatórias Y_{ij} com distribuição de probabilidade pertencentes à família exponencial, de modo que a esperança para o modelo marginal, denotada por $E(Y_{ij})$, seja escrita como função das covariáveis atendendo as seguintes suposições:

i) $E(Y_{ij}) = \mu_{ij}$ depende das covariáveis \tilde{x}_{ij} por $g(\mu_{ij}) = \tilde{x}_{ij}'\beta$ em que g é uma função de ligação conhecida, tal como a recíproca, $\frac{1}{\mu_{ij}}$, para respostas com distribuição gama e a identidade, μ_{ij} , para respostas com distribuição normal;

ii) a variância depende da média de modo que $V(Y_{ij}) = V(\mu_{ij})\phi^{-1}$, em que V é uma função de variância conhecida e ϕ é o parâmetro de dispersão, que pode ser necessário estimar;

iii) a covariância entre Y_{ij} e Y_{ik} é uma função de parâmetros desconhecidos, α , isto é, $Cov(Y_{ij}, Y_{ik}) = V_i(\alpha)$, com $V_i(\alpha) = A_i^{1/2} R(\alpha) A_i^{1/2}$, em que A_i é a matriz diagonal da função de variância para cada indivíduo e $R(\alpha)$ é a, assim denominada, matriz de correlação de trabalho. Exemplos de matrizes de correlação de trabalho são, dentre outras, a independente, a AR(1) e a permutável, descritas na Seção 2.3.2.4.

No estudo analisado neste trabalho, a variável resposta é contínua (produção de leite) e, desse modo, as distribuições Normal (gaussiana) e Gama, disponíveis no *software* R, são assumidas para a resposta na forma da família exponencial, de modo que as funções de densidade de probabilidade da gama e da normal são dadas, respectivamente, por

$$f(y/\theta, \phi) = \exp\left\{\frac{y\theta + \ln(-\theta)}{\phi^{-1}} + (\phi - 1)\ln(y\phi) + \ln(\phi) - \ln(\Gamma(\phi))\right\}$$

e

$$f(y/\theta, \phi) = \exp\left\{\frac{y\theta - \theta^2/2}{\phi} - \frac{1}{2}\left[\frac{y^2}{\phi} + \ln(2\pi\phi)\right]\right\}.$$

As estruturas de correlação AR(1), independente e permutável são as estruturas analisadas para a escolha da que melhor representa a estrutura dos dados sob análise.

Para a estimação dos parâmetros foi feito uso das equações de estimação generalizadas (GEE), com o auxílio do pacote estatístico R, usando o “geepack”. Para mais detalhes ver Liang e Zeger (1986), dentre outros. O critério de informação de Akaike (*AIC*) foi utilizado para seleção da estrutura de correlação que melhor se adequar aos dados dos modelos ajustados. Ainda, a estatística utilizada para testar hipóteses relativas aos parâmetros β_j ($j = 1, \dots, p$) foi a estatística de Wald, que segue distribuição assintótica normal, sendo uma generalização da estatística t de Student.

No contexto de MLG, a qualidade do ajuste é usualmente verificada por meio de alguns gráficos de diagnóstico, os quais podem auxiliar a detectar possíveis afastamentos das suposições feitas para o modelo, bem como verificar a existência de

observações extremas com alguma interferência desproporcional nos resultados do ajuste. Como neste estudo os dados são desbalanceados, não será possível a obtenção do diagnóstico, pois este recurso ainda não se encontra implementado no pacote estatístico utilizado e, como forma alternativa foram calculados, a partir do modelo ajustado, alguns valores estimados e, estes foram comparados aos seus respectivos valores observados.

Caso seja verificada a existência de diferenças pequenas entre os pontos, haverá indícios de que o ajuste é satisfatório. Com isso, o modelo que melhor se ajusta aos dados é encontrado e, ainda, os fatores significativos são detectados.

3.2.2 Modelo não-linear para descrever a curva de lactação

a) Modelo não-linear para a curva de lactação média das 825 vacas

O ajuste à curva de lactação média das 825 vacas baseou-se na função gama incompleta, proposta por Wood (1967), expressa por

$$Y_x = a_0 X^{a_1} e^{-a_2 X}, \quad (3)$$

sendo Y_x a produção média de leite na semana X de lactação; a_0 um parâmetro desconhecido a ser estimado, que representa a produção de leite no início da lactação, responsável por baixar ou elevar a curva de lactação; a_1 um parâmetro desconhecido a ser estimado, que está associado à base de ascensão da produção de leite no início da lactação; e a_2 um parâmetro desconhecido a ser estimado que está associado à fase de declínio da produção de leite após o pico.

No procedimento de estimação dos parâmetros da função utilizada neste estudo, expressa em (3), foi obtido um sistema de equações não-lineares, sendo que para a solução desse sistema foi necessário um método numérico iterativo, o qual necessita de valores iniciais para a_0 , a_1 e a_2 . Os valores iniciais considerados neste estudo para

a_0 , a_1 e a_2 foram 10, 0,05 e 0,004, respectivamente. Tais valores foram retirados de um estudo anterior realizado por Faro et al. (2002). A partir disso, ajustou-se o modelo à curva de lactação média das 825 vacas obtendo-se o gráfico das curvas ajustada e observada da produção semanal média das vacas Caracu da fazenda Chiqueirão, bem como os valores estimados dos parâmetros a_0 , a_1 e a_2 .

As funções dos parâmetros a_0 , a_1 e a_2 , como tempo de pico (TP), produção no pico (PP) e persistência (S), foram estimadas de acordo com Wood (1967). Tais funções estão apresentadas na Tabela 3.1.

Tabela 3.1 – Funções dos parâmetros: tempo de pico, produção no pico e persistência

<i>Função</i>	<i>Fórmula</i>
Tempo de Pico	$TP = \frac{a_1}{a_2}$
Produção no Pico	$PP = a_0 \left(\frac{a_1}{a_2} \right)^{a_1} e^{-a_1}$
Persistência	$S = -(a_1 + 1) \ln a_2$

Os critérios utilizados para avaliação da qualidade do ajuste do modelo não-linear da curva de lactação média, foram às médias, desvios-padrão e coeficiente de variação para os parâmetros e funções dos parâmetros. O coeficiente de determinação ajustado (R_A^2), usado como critérios de avaliação em outros estudos, não foi utilizado devido às limitações do *software* estatístico utilizado.

b) Modelo não-linear para a curva de lactação de cada vaca

Para os dados de produção de leite de cada vaca foi ajustada a função gama incompleta (Wood, 1967), expressa por

$$Y_x = a_{0j} X^{a_{1j}} e^{-a_{2j} X}, \quad j = 1, \dots, 825 \quad (4)$$

sendo Y_x a produção de leite da vaca na semana X ; a_{0j} um parâmetro desconhecido,

que representa a produção de leite no início da lactação, responsável por baixar ou elevar toda a curva de lactação; a_{1j} um parâmetro desconhecido, que está associado à base de ascensão da produção de leite no início da lactação; e a_{2j} um parâmetro desconhecido que está associado à fase de declínio da produção de leite após o pico.

Os parâmetros da função utilizada para ajustar a curva de lactação individual, expressa em (4), foram estimados por meio de regressão não-linear. O procedimento de estimação dos parâmetros segue o mesmo método utilizado na estimação dos parâmetros da curva de lactação média das 825 vacas. Porém, no ajuste das curvas de lactação individual, o procedimento é efetuado 825 vezes, pois é necessária a obtenção dos parâmetros relativos a lactação de cada vaca.

A partir disso, identificou-se as vacas com produção típica, ou seja, que não apresentaram nenhum dos parâmetros a_0 , a_1 e a_2 negativos, e as vacas com produção atípica, que são aquelas que apresentam pelo menos um dos parâmetros a_0 , a_1 ou a_2 negativos. As curvas ajustadas e observadas foram traçadas para 4 vacas de cada tipo e, com isso, observou-se o comportamento da curva em cada situação. Para as quatro vacas que apresentaram produção típica, cujas curvas foram traçadas, apresentou-se os valores estimados dos parâmetros a_0 , a_1 e a_2 , bem como os valores correspondentes ao tempo de pico (TP), a produção no pico (PP) e a persistência (S), todos estimados de acordo com Wood (1967) (Tabela 3.1).

Medidas como média, desvio-padrão e coeficiente de variação (CV) também foram calculados para as funções dos parâmetros, isto é, para o tempo de pico (TP), a produção no pico (PP) e a persistência (S). Neste caso, foram considerados apenas os parâmetros das vacas com produção típica.

Os critérios utilizados para avaliação da qualidade de ajuste do modelo não-linear das curvas de lactação individual foram: as médias, desvios-padrão e coeficiente de variação para os parâmetros e funções dos parâmetros e, ainda, o percentual de curvas atípicas encontradas no modelo.

4 RESULTADOS E CONCLUSÕES

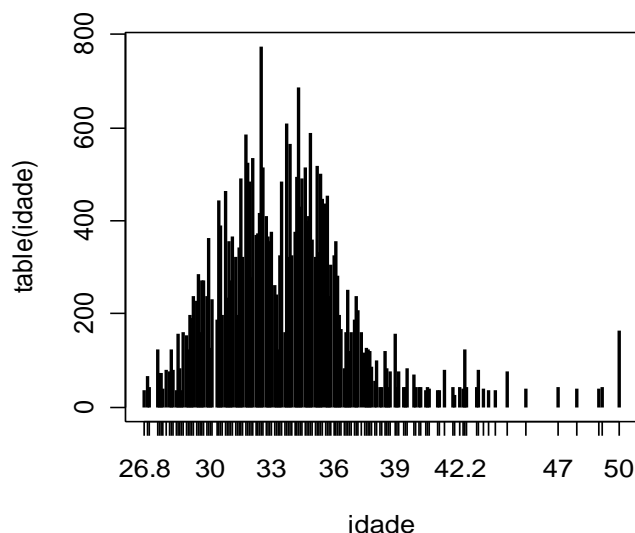
A seguir, são apresentados os resultados obtidos inicialmente por meio de análise descritiva dos dados e, em seguida, o ajuste dos modelos linear e não-linear.

4.1 Análise descritiva

Por meio da análise descritiva foi possível ter uma ideia geral do desempenho produtivo das 825 vacas que compõem o banco de dados analisado, bem como visualizar o comportamento da curva de produção média, ter uma ideia do perfil individual das vacas (neste caso de 50 delas), observar os períodos mais produtivos (semanas, meses e anos), além dos locais (retiros) que mais se destacaram.

Na Figura 4.1 é possível observar, a partir das frequências das idades das 825 vacas, que a maior parte delas apresentava de 30 a 36 meses, indicando que esta parte do rebanho era constituída por vacas mais jovens.

Figura 4.1 - Gráfico de frequências das vacas por idade



Observando os valores médios de lactação, mostrados na Tabela 4.1, bem como a Figura 4.2, que mostra a curva de lactação média para o rebanho pertencente à

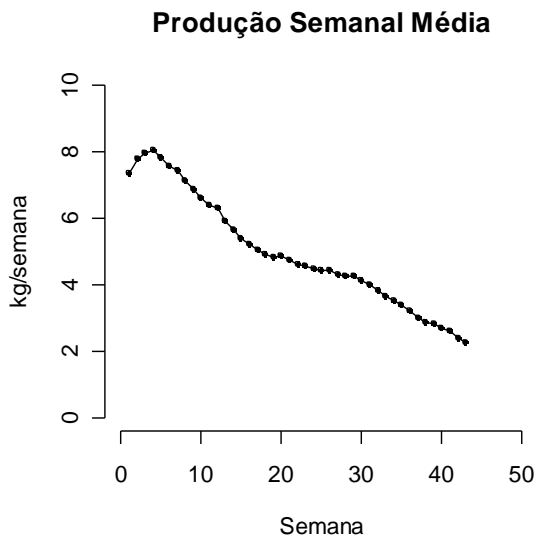
Fazenda Chiqueirão, Poços de Caldas - MG no período de 1991 a 1998, nota-se que nas três primeiras semanas a produção média leiteira está em ascensão, iniciando com 7,35kg. Na 4ª semana pós-parto ocorreu o que comumente se chama “pico de lactação” (quando a produção média atinge seu valor mais alto), neste caso 8,04kg. A partir da 5ª semana, a produção começou a decair até atingir, na 43ª semana, a menor produção que foi de 2,25kg. Uma característica da curva é que ela apresenta um declínio gradual da produção de leite.

Notou-se que a 17ª semana foi a que reuniu o maior número de registros dos animais (total de 821). A produção média nesta semana foi de aproximadamente 62% da produção média no pico.

Tabela 4.1 - Produção semanal média (kg) de vacas Caracu no período de 1991 a 1998

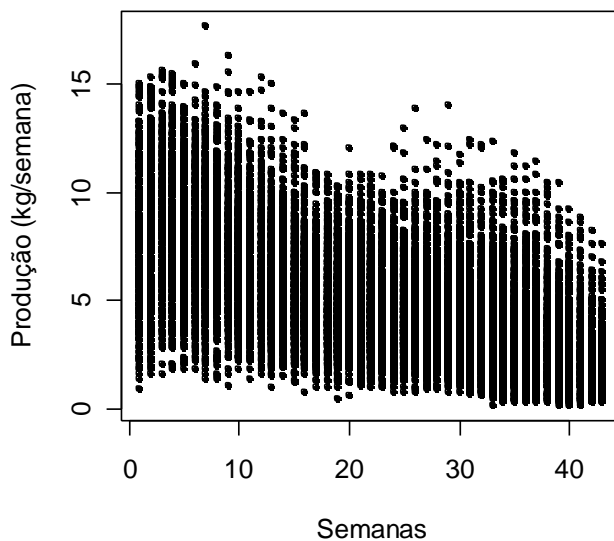
<i>Semana</i>	<i>Nº de registros na semana</i>	<i>Média</i>	<i>Desvio-padrão</i>	<i>Semana</i>	<i>Nº de registros na semana</i>	<i>Média</i>	<i>Desvio-padrão</i>
1	642	7,35	2,68	23	817	4,56	1,68
2	769	7,81	2,61	24	816	4,47	1,67
3	756	7,97	2,49	25	817	4,45	1,67
4	769	8,04	2,49	26	818	4,42	1,73
5	774	7,83	2,39	27	818	4,29	1,67
6	791	7,58	2,33	28	819	4,25	1,72
7	801	7,44	2,33	29	817	4,25	1,80
8	804	7,13	2,15	30	814	4,15	1,81
9	806	6,90	2,14	31	813	3,99	1,80
10	802	6,61	2,11	32	811	3,83	1,77
11	806	6,41	2,04	33	811	3,64	1,78
12	802	6,29	2,04	34	803	3,51	1,81
13	798	5,92	1,97	35	792	3,39	1,81
14	805	5,66	1,88	36	775	3,20	1,78
15	813	5,38	1,88	37	762	3,00	1,73
16	817	5,20	1,74	38	728	2,86	1,63
17	821	5,04	1,67	39	680	2,80	1,57
18	815	4,93	1,71	40	634	2,67	1,46
19	816	4,84	1,74	41	573	2,58	1,40
20	817	4,85	1,73	42	475	2,39	1,29
21	815	4,75	1,72	43	330	2,25	1,22
22	816	4,62	1,69				

Figura 4.2 – Curva da produção semanal média



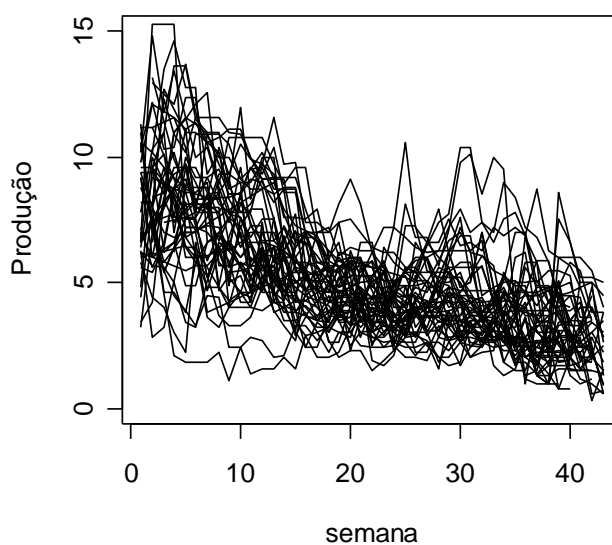
A Figura 4.3, que exibe a dispersão em torno da curva de produção semanal média, aponta a existência de grande variabilidade em torno dessa média. As curvas de lactação de determinadas vacas devem, portanto, estar mais distantes da curva de produção semanal média do que as de outras vacas. Essa dispersão pode ter sido causada, por exemplo, por fatores como a genética do animal e manejo do rebanho

Figura 4.3 - Dispersão da produção observada por semana e produção semanal média



Os perfis individuais de 50 vacas estão representados na Figura 4.4, em que se nota a existência de grande variabilidade ao redor da produção semanal média. Há curvas de produção individual que estão distantes da curva média e outras que estão mais próximas, indicando que pode existir diferenças genéticas entre as 50 vacas, bem como influência de outros fatores que podem ter afetado a produção de cada vaca.

Figura 4.4 – Perfis individuais da produção semanal média de 50 vacas



Quanto à produção média por classe de idade da vaca, mostradas na Tabela 4.2 e Figura 4.5, observa-se que vacas nas faixas etárias de 26,8 a 28,7 e 44,8 a 46,7 meses foram as que menos produziram. A produção média leiteira ficou em 4,67kg para as mais novas e em 4,25kg para as na faixa etária de 44,8 a 46,7.

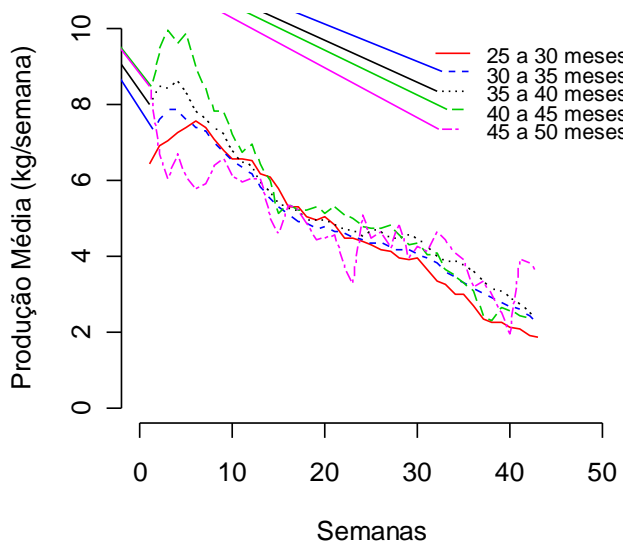
A maior produção média leiteira foi observada para as vacas de 40,8 a 42,7 meses (5,78kg). Para as vacas mais velhas, neste caso de 48,8 a 50,7 meses, registrou-se uma peculiaridade, já que normalmente produzem menos. Para elas foi registrada uma produção média de 5,49kg, que foi considerada boa para esta faixa etária.

As faixas etárias que concentraram animais com maior produção média foram: 34,8 a 36,7; 36,8 a 38,7; 38,8 a 40,7; e 42,8 a 44,7 meses, com a produção variando de 5,18kg a 5,78kg.

Tabela 4.2 – Produção média (por semana) por classe de idade da vaca

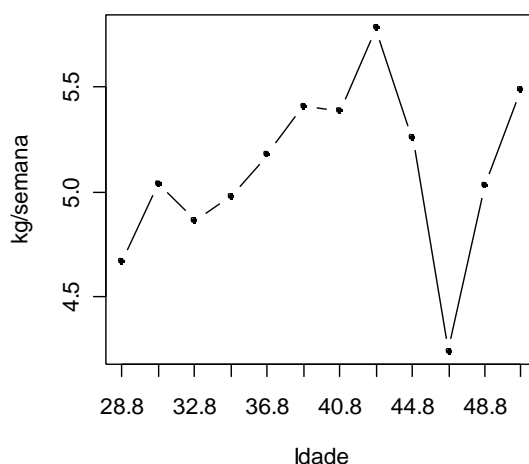
<i>Idade da vaca</i>	<i>Produção</i>
26,8 - 28,7	4,670296
28,8 - 30,7	5,039102
30,8 - 32,7	4,866919
32,8 - 34,7	4,976043
34,8 - 36,7	5,182353
36,8 - 38,7	5,406452
38,8 - 40,7	5,388492
40,8 - 42,7	5,779828
42,8 - 44,7	5,259935
44,8 - 46,7	4,245000
46,8 - 48,7	5,031325
48,8 - 50,7	5,486939

Figura 4.5 - Produção semanal média por classe de idade da vaca



Conforme mostra a Figura 4.6, a produção de leite semanal aumentou à medida que os animais avançaram em idade, isto é, animais mais jovens apresentaram uma produção semanal menor. Porém, nota-se que as vacas que estavam com 40,8 meses tiveram praticamente a mesma produção dos animais com 38,8 meses, ou seja, a produção foi praticamente a mesma, ficando a curva de produção semanal quase no mesmo patamar. Para os animais com 46,8 meses houve uma queda considerável na produção semanal, que pode ser percebida no gráfico, pois este apresentou um decréscimo acentuado. Esta curva apresentou uma certa peculiaridade, após o pico negativo a produção voltou a aumentar novamente para os animais mais velhos, o que não é comum para vacas mais velhas.

Figura 4.6 - Produção semanal média por idade da vaca



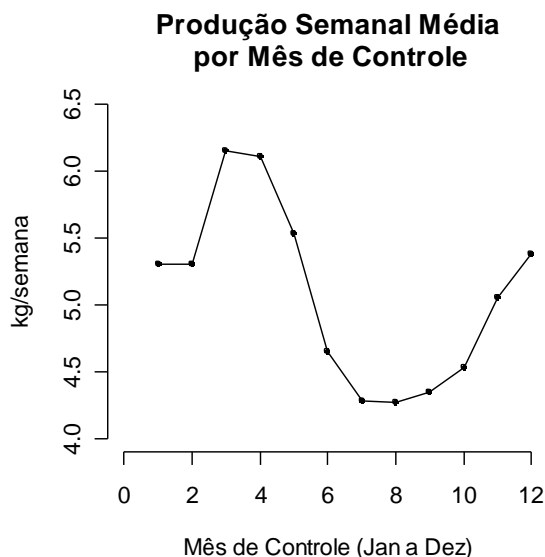
A produção média de leite por mês de controle, mostrada na Tabela 4.3 e Figura 4.7, iniciou-se com aproximadamente 5,3kg no mês de janeiro, permanecendo a mesma em fevereiro. Em março ocorreu a produção máxima, para então começar a decair gradativamente a partir de abril. Em maio houve uma diminuição maior até o mês de agosto, voltando a aumentar em setembro e nos meses seguintes para, finalmente, atingir dezembro com aproximadamente o mesmo patamar de janeiro.

Os meses mais produtivos para esta parte do rebanho (825 vacas) foram: março com 6,15kg, abril com 6,11kg e maio com 5,54kg. De junho a agosto, período de inverno, a produção média diminuiu, voltando a aumentar em novembro e dezembro, 5,05kg e 5,38kg, respectivamente, não atingindo as médias dos meses de março e abril.

Tabela 4.3 - Produção média por mês de controle

<i>Mês de controle</i>	<i>Produção Média</i>
Janeiro	5,30
Fevereiro	5,30
Março	6,16
Abril	6,11
Maió	5,54
Junho	4,65
Julho	4,28
Agosto	4,27
Setembro	4,34
Outubro	4,53
Novembro	5,06
Dezembro	5,38

Figura 4.7 - Produção média por mês de controle 01 a 12 (janeiro a dezembro).



A produção semanal por mês de controle dentro do ano de controle para os anos de 1991, 1992, 1993 e 1994, mostradas na Tabela 4.4 e Figura 4.8 (a), foi, de modo geral, maior no mês de janeiro de 1994 do que em 1992 e 1993 (em 1991 não foi

observada). O ano de 1991 apresentou a produção mais alta nos meses de fevereiro, março e abril, se comparadas aos anos de 1992, 1993 e 1994 no mesmo período, sendo que o maior valor foi registrado em março com uma produção semanal de 9,24kg. O ano de 1992 foi o que apresentou a maior produção no mês de abril em relação aos demais anos analisados. Nos meses de maio, junho, julho e agosto a produção semanal média de leite diminuiu em todos os quatro anos, voltando a aumentar a partir do mês de setembro para os anos de 1992 e 1993, permanecendo baixa e praticamente a mesma para os anos de 1991 e 1994 nesse mês, retornando a crescer em outubro. No período de maio a setembro, que coincide com o inverno, as produções foram mais baixas, mas aumentaram no último trimestre. Para esses quatro anos houve diferença na produção semanal média.

Para os anos de 1995, 1996, 1997, 1998, mostrados na Tabela 4.4 e Figura 4.8 (b), observa-se que a produção semanal média não diferiu muito, somente 1998 apresentou valores menores para a produção semanal média no segundo semestre, o que fez com que a trajetória da curva mudasse em relação às curvas dos outros anos. No ano de 1997 podem-se observar as maiores produções semanais médias.

No geral, nota-se que houve variação na produção semanal média por mês de controle dentro do ano de controle e que se comparados aos quatro primeiros anos dos dados do estudo, os anos de 1995, 1996, 1997 e 1998 foram menos produtivos.

Tabela 4.4 - Produção média por mês de controle para os anos de 1991 a 1999

<i>Mês de controle</i>	<i>1991</i>	<i>1992</i>	<i>1993</i>	<i>1994</i>	<i>1995</i>
Janeiro	---	3,976471	5,703797	5,800000	4,909843
Fevereiro	8,695000	4,440449	5,860697	5,872414	4,476053
Março	9,236709	7,317931	6,043726	6,818028	5,389013
Abril	8,663559	8,961458	5,673408	6,327154	5,512032
Mai	7,711009	7,593878	6,260465	5,882563	5,280102
Junho	5,841364	6,006324	5,227273	4,889299	4,447301
Julho	4,724621	5,129178	4,525225	3,808316	4,400000
Agosto	4,115789	4,604153	4,610502	4,067283	4,561641
Setembro	3,780365	4,561891	4,998897	3,706364	4,425905
Outubro	4,014444	5,052222	5,397865	4,118130	4,300000
Novembro	4,657741	5,093902	5,977316	4,944762	5,226250
Dezembro	4,258621	5,069312	6,521277	5,442518	5,463191

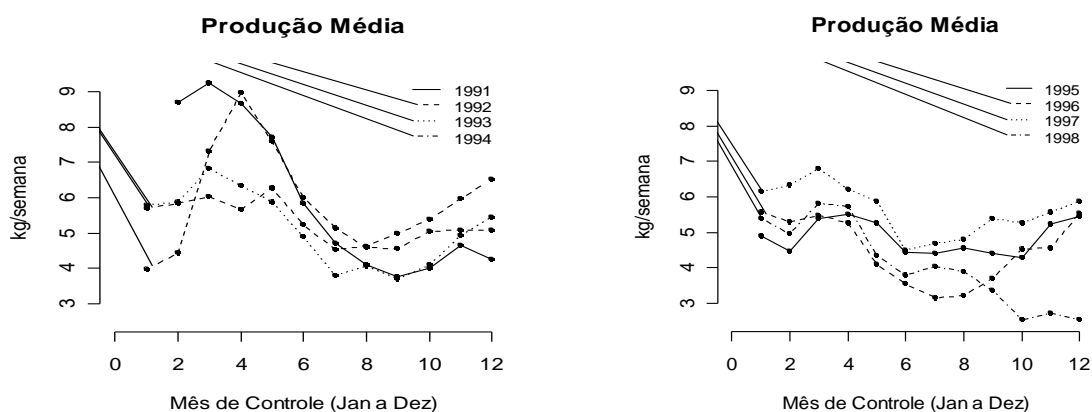
Tabela 4.4 - Produção média por mês para os anos de 1991 a 1999 (continuação)

<i>Mês de controle</i>	<i>1996</i>	<i>1997</i>	<i>1998</i>	<i>1999</i>
Janeiro	5,563476	6,139437	5,382151	2,094118
Fevereiro	5,286683	6,337745	4,968533	--
Março	5,492540	6,789790	5,821776	--
Abril	5,251681	6,218735	5,739726	--
Mai	4,114471	5,872781	4,361807	--
Junho	3,553482	4,517849	3,798267	--
Julho	3,163990	4,671324	4,055524	--
Agosto	3,233548	4,802268	3,876512	--
Setembro	3,710480	5,380623	3,387241	--
Outubro	4,540339	5,257866	2,539357	--
Novembro	4,562543	5,561801	2,733714	--
Dezembro	5,542544	5,871429	2,550000	--

Figura 4.8 - Produção semanal média por mês de controle dentro do ano de controle

(a) de 1991 a 1994

(b) de 1995 a 1998



A partir da Tabela 4.5 e Figura 4.9 (a) e (b), percebe-se que o retiro mais produtivo foi o 5 dentro dos anos de 1992, 1993, 1997, pois a produção semanal média de leite ficou acima dos 8,0kg, e nos anos de 1995, 1996 e 1998, em que a produção semanal média foi acima de 7,0kg de leite. Depois do 5º retiro, o segundo mais produtivo foi o 6º nos anos de 1991 e 1994, este apresentou uma produção média um pouco acima do 7,0kg e nos anos de 1993, 1997 e 1998 apresentou uma produção média semanal em torno dos 6,9kg. Nos demais retiros a produção média semanal girou em torno dos 4,0 e 5,0kg de leite.

Tabela 4.5 - Produção média por retiro em cada ano de controle (retiros de 1 a 15)

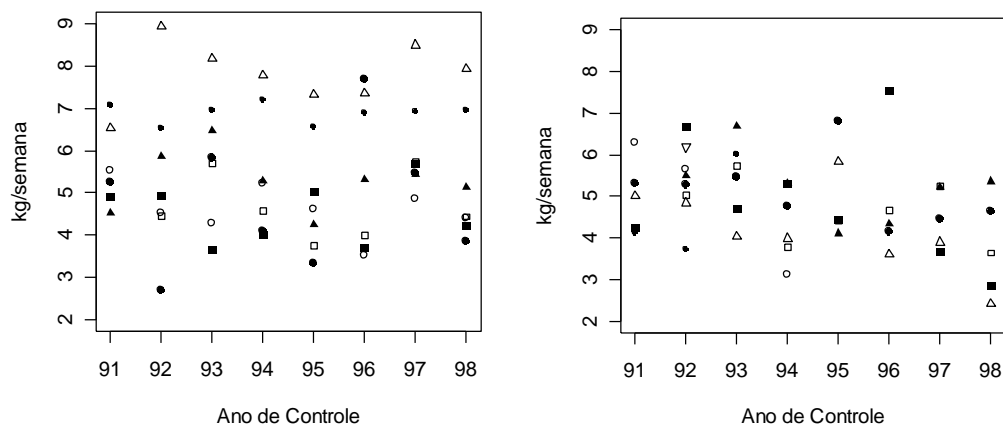
<i>Ano</i> \ <i>Retiro</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>
<i>1991</i>	4,897260	5,259643	4,511864	---	6,534483	7,091176	5,530534	4,240000
<i>1992</i>	4,909146	2,681081	5,875300	4,444382	8,958036	6,538788	4,528700	6,666465
<i>1993</i>	3,623548	5,824242	6,468037	5,694972	8,180952	6,959690	4,278632	4,701250
<i>1994</i>	4,005590	4,085635	5,281459	4,558232	7,790566	7,223438	5,223762	5,282609
<i>1995</i>	5,015741	3,332129	4,237271	3,744207	7,335277	6,576812	4,629014	4,432509
<i>1996</i>	3,704845	7,692941	5,319345	3,999073	7,354450	6,894767	3,523529	7,552174
<i>1997</i>	5,679760	5,479024	5,441871	5,727299	8,510966	6,934122	4,862233	3,680000
<i>1998</i>	4,223820	3,842222	5,130769	4,431164	7,935185	6,960000	4,410504	2,827668

<i>Ano</i> \ <i>Retiro</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>
<i>1991</i>	5,318893	NaN	NaN	5,009959	4,120408	6,288889	--
<i>1992</i>	5,296203	5,488235	5,020816	4,839444	3,713445	5,644318	6,184564
<i>1993</i>	5,467202	6,683451	5,737888	4,033333	6,019753	5,466535	5,613816
<i>1994</i>	4,761224	5,323883	3,765854	3,982533	--	3,117391	--
<i>1995</i>	6,828696	4,110567	4,396849	5,834483	--	--	--
<i>1996</i>	4,142406	4,342254	4,650626	3,602288	--	--	--
<i>1997</i>	4,474215	5,228805	5,242353	3,897436	--	--	--
<i>1998</i>	4,654634	5,364286	3,638889	2,407258	--	--	--

Figura 4.9 - Produção média por retiro em cada ano de controle

(a) retiros de 1 a 8

(b) retiros de 9 a 15



4.2 Modelos lineares generalizados para dados longitudinais

4.2.1 Ajuste do modelo

Foram ajustados dois modelos, conforme descrito na Seção 3.2.1, usando-se a metodologia de modelos lineares generalizados (MLG) proposta por Nelder e Wedderburn (1972), em que foi levada em consideração:

- (a) a natureza contínua da variável resposta; e
- (b) a correlação entre as observações registradas em um mesmo animal.

4.2.2 Seleção do modelo

Para o ajuste do modelo linear foram consideradas as covariáveis “ano”, “mês”, “retiro”, “idade” e “semana”. Neste estudo, o grau de parentesco entre os animais não foi considerado. O modelo selecionado para o ajuste da curva de lactação foi aquele baseado na distribuição Gama cuja função densidade de probabilidade na forma da família exponencial é dada como segue

$$f(y/\theta, \phi) = \exp \left\{ \frac{y\theta + \ln(-\theta)}{\phi^{-1}} + (\phi - 1)\ln(y\phi) + \ln(\phi) - \ln(\Gamma(\phi)) \right\}.$$

Este modelo, considerando a estrutura de correlação auto-regressiva (AR1) apresentou o menor valor do Critério de Informação de Akaike (AIC) conforme pode ser observado na Tabela 4.6.

Tabela 4.6 – AIC para as diferentes estruturas de correlação

Estrutura de Correlação	AIC	
	Gama	Normal
auto-regressiva (AR(1))	30.367,19	32.978,25
independente	36.888,38	32.978,18
permutável	35.494,87	32.978,39

4.2.3 Modelo ajustado

Para o modelo ajustado tem-se que:

$$E(Y) = \hat{\mu} = \frac{1}{\hat{\eta}}$$

com $\hat{\eta} = \frac{1}{\hat{\mu}} = \hat{\beta}_0 + \hat{\beta}_i * ano_i + \hat{\beta}_j * m\hat{e}s_j + \hat{\beta}_k * retiro_k - \hat{\beta}_l * idade + \hat{\beta}_m * semana$, em que: $\hat{\eta} =$

preditor linear, $\hat{\mu}$ = produção média de leite estimada, $\hat{\beta}_0$ = constante, $\hat{\beta}_i$ = estimativas de β_i para os anos de 1992 a 1999, $\hat{\beta}_j$ = estimativas de β_j para os meses de fevereiro a dezembro, $\hat{\beta}_k$ = estimativas de β_k para os retiros de 2 a 15, $\hat{\beta}_l$ = estimativas de β_l para a idade do animal, e $\hat{\beta}_m$ = estimativas de β_m para a semana, tal que:

$$ano_i = \begin{cases} 1 \text{ se ano} = 1991 + i & i = 1, \dots, 8 \\ 0 \text{ c.c.} \end{cases}$$

$$mes_j = \begin{cases} 1 \text{ se m\hat{e}s} = j & j = 2, \dots, 12 \\ 0 \text{ c.c.} \end{cases}$$

$$retiro_k = \begin{cases} 1 \text{ se retiro} = k & k = 2, \dots, 15 \\ 0 \text{ c.c.} \end{cases}$$

idade do animal = 26,8 a 50 meses, e semana = 1 a 43.

Note que as covariáveis “ano”, “mês” e “retiro” foram incorporadas ao modelo por meio de variáveis *dummy*.

4.2.4 Estimativas dos parâmetros do modelo escolhido

A Tabela 4.7 apresenta as estimativas dos parâmetros do modelo escolhido.

Tabela 4.7 – Sumário do modelo Gama ajustado à curva de lactação

β	Parâmetro	Estimativa	Desvio padrão Robusto de β	Z_{obs}	$Pr(Z > Z_{obs})$
β_0	Intercepto	0,1344	0,0166	65,3590	6,66E-16***
β_1	factor(ano)92	0,0048	0,0071	0,4592	4,98E-01
β_2	factor(ano)93	0,0102	0,0069	2,1791	1,39E-01
β_3	factor(ano)94	0,0112	0,0065	3,0102	8,27E-02
β_4	factor(ano)95	0,0156	0,0066	5,5617	1,83E-02*
β_5	factor(ano)96	0,0175	0,0075	5,5191	1,88E-02*
β_6	factor(ano)97	0,0069	0,0063	1,2139	2,70E-01
β_7	factor(ano)98	0,0138	0,0074	3,5106	6,09E-02
β_8	factor(ano)99	0,3048	0,0978	9,7095	1,83E-03*
β_9	factor(mes)2	0,0207	0,0033	39,7480	2,88E-10***
β_{10}	factor(mes)3	0,0181	0,0034	28,2358	1,07E-07***
β_{11}	factor(mes)4	0,0199	0,0036	30,6811	3,04E-08***
β_{12}	factor(mes)5	0,0217	0,0037	34,7644	3,72E-09***
β_{13}	factor(mes)6	0,0316	0,0037	72,9750	0,00E+00***
β_{14}	factor(mes)7	0,0356	0,0037	92,3258	0,00E+00***
β_{15}	factor(mes)8	0,0318	0,0038	69,9153	1,11E-16***
β_{16}	factor(mes)9	0,0255	0,0037	46,6496	8,48E-12***
β_{17}	factor(mes)10	0,0204	0,0034	36,5321	1,50E-09***
β_{18}	factor(mes)11	0,0099	0,0031	10,3856	1,27E-03*
β_{19}	factor(mes)12	0,0030	0,0027	1,1926	2,74E-01
β_{20}	factor(retiro)2	0,0038	0,0081	0,2186	6,40E-01
β_{21}	factor(retiro)3	0,0085	0,0048	3,1251	7,07E-02
β_{22}	factor(retiro)4	0,0141	0,0060	5,4740	1,93E-02*
β_{23}	factor(retiro)5	0,0105	0,0033	10,0869	1,49E-03*
β_{24}	factor(retiro)6	0,0220	0,0033	43,1562	5,05E-11***
β_{25}	factor(retiro)7	0,0034	0,0048	0,4887	4,84E-01*
β_{26}	factor(retiro)8	-0,0091	0,0055	2,6936	1,01E-01*
β_{27}	factor(retiro)9	-0,0092	0,0044	4,3066	3,79E-02*
β_{28}	factor(retiro)10	-0,0020	0,0057	0,1183	7,31E-01
β_{29}	factor(retiro)11	0,0047	0,0071	0,4383	5,07E-01
β_{30}	factor(retiro)12	0,0222	0,0061	13,1932	2,80E-04*
β_{31}	factor(retiro)13	0,0109	0,0096	1,2933	2,55E-01
β_{32}	factor(retiro)14	-0,0105	0,0056	3,5413	5,98E-02
β_{33}	factor(retiro)15	-0,0300	0,0079	14,5217	1,38E-04*
β_{34}	idade	-0,0018	0,0004	16,3308	5,31E-05**
β_{35}	semana	0,0060	0,0001	1874,7306	0,00E+00***
ϕ	escala	0,18609	0,00617	909,42730	-

No MLG ajustado, o ano de 1991, o mês 1 e o retiro 1 foram considerados como a categoria de referência em relação às demais categorias. Ainda, para a análise do valor p foi considerado um nível de significância de 5% e, em seguida, analisou-se a significância de cada parâmetro em relação ao valor fixado.

De acordo com a Tabela 4.7, apresentaram efeito significativo os anos de 1995 e 1996, visto que $p\text{-valor} < 0,05$, portanto rejeita-se a hipótese $\beta_4 = \beta_5 = 0$, concluindo-se, então, que 1995 e 1996 foram anos de parição em que houve uma maior influência na produção média de leite quando comparada com o ano de 1991 (ano de referência). No caso do ano de 1999, que também foi significativo, houve registro de observações da produção apenas para os dois primeiros meses do ano, sendo, em seguida, finalizada a coleta de dados, porém a produção média foi maior do que em 1991.

No caso dos meses, nota-se que apenas o mês de dezembro não foi significativo (não diferiu do mês de referência = janeiro). Para os demais meses a hipótese nula foi rejeitada, evidenciando que o mês em que ocorre a lactação exerce influência no resultado da produção média.

Analisando, ainda, a Tabela 4.7 para os retiros, constatou-se efeito significativo para 5 de um total de 15, sendo eles, 4, 5, 6, 12 e 15, rejeitando-se também a hipótese de que $\beta_{23} = \beta_{24} = \beta_{25} = \beta_{30} = \beta_{33} = 0$. A produção nestes retiros foi maior que nos demais, inclusive em comparação ao retiro base que neste caso foi o de número 1.

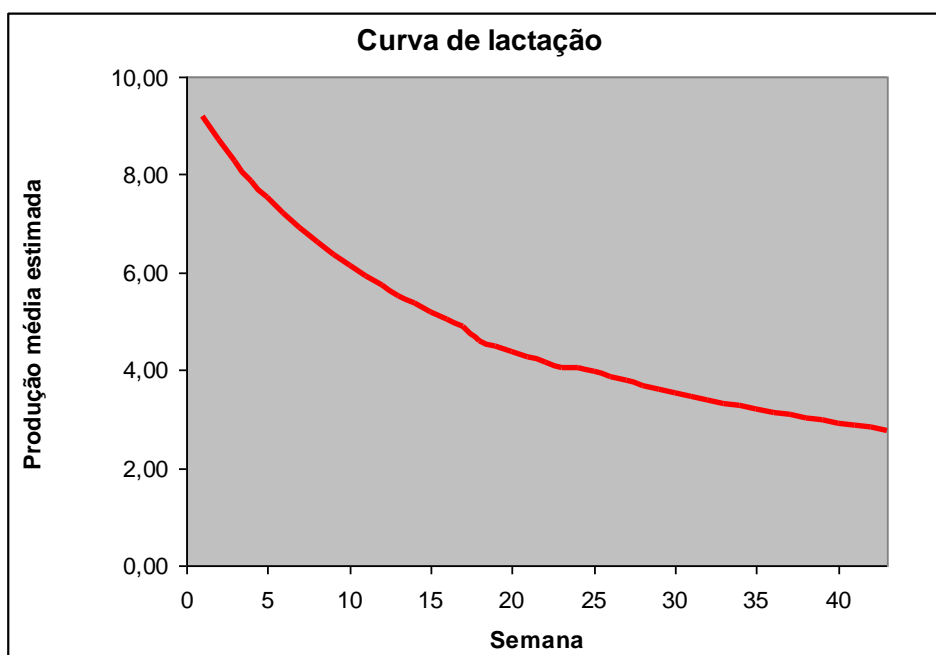
No modelo selecionado, idade e semana são também covariáveis importantes, sendo rejeitadas as hipóteses nulas de efeito não significativo das mesmas.

Observa-se, portanto, que todas as covariáveis utilizadas apresentaram efeito significativo, sendo necessária a permanência de todas no modelo.

4.2.5 Curva de lactação ajustada para combinação de covariáveis

Na Figura 4.10 é apresentada a curva de lactação estimada em função das semanas para a seguinte combinação: ano 1992, mês de março, retiro 5 e idade da vaca igual a 36,2 meses, sendo possível observar que à medida que aumenta o número de semanas diminuiu a produção média estimada. Ressaltando que esta é apenas uma das possíveis combinações entre todas as possíveis.

Figura 4.10 – Gráfico da curva de lactação em função da semana estimada para o ano de 1992, mês 3 (março), retiro 5 e idade da vaca igual a 36,2 meses



4.2.6 Diagnóstico do ajuste do modelo

Devido aos dados serem desbalanceados, isto é, o número de observações registrado para cada animal são diferentes, não foi possível a obtenção de alguns dos gráficos de diagnóstico de influência usados frequentemente para avaliar a qualidade

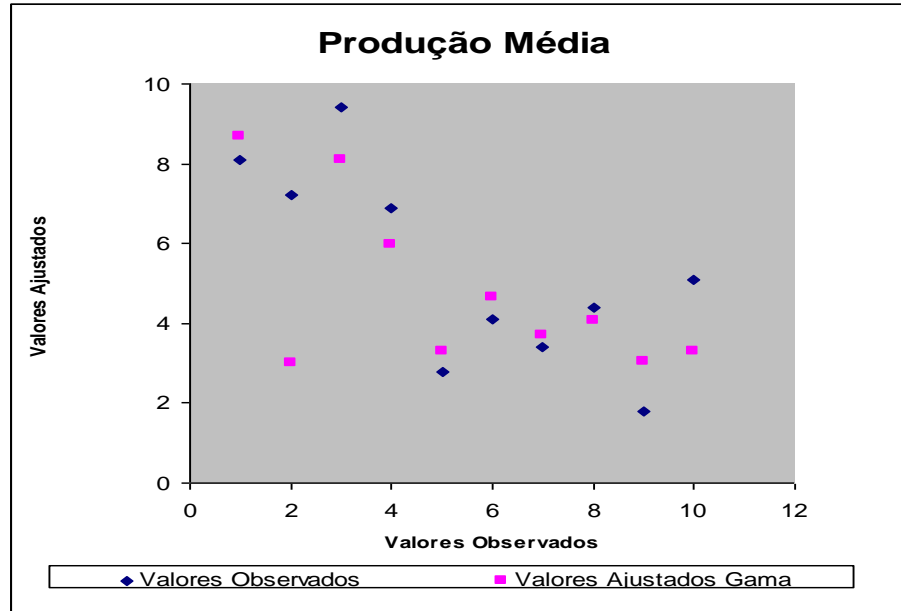
do modelo ajustado, pois a implementação no *software* R ainda não se encontra disponível nesses casos. Uma avaliação um tanto simplista e descritiva, nesse sentido, foi feita por meio de alguns pontos observados e estimados pelo modelo.

Na Tabela 4.8 e Figura 4.11 encontram-se os valores observados e seus respectivos valores ajustados a partir do modelo gama. Uma análise desses valores indica que o ajuste parece ser razoável. É possível observar que os valores ajustados 2, 9 e 10 apresentaram uma diferença maior em relação aos seus respectivos valores observados, indicando, nestes casos, que o ajuste não foi muito bom, o que pode ter ocorrido devido a animais que tenham apresentado lactação atípica. Entre esses pontos, o mais crítico foi o ponto 2. Os demais pontos indicaram que o ajuste tem uma qualidade razoável, uma vez que um diagnóstico de maior precisão não pôde ser feito, devido às limitações apresentadas pelo pacote estatístico utilizado.

Tabela 4.8 – Valores Observados X Valores ajustados Gama

<i>Valores Observados</i>	<i>Valores Ajustados Gama</i>
8,1	8,68595
7,2	3,013474
9,4	8,091347
6,9	5,988451
2,8	3,291164
4,1	4,645172
3,4	3,694003
4,4	4,05871
1,8	3,023534
5,1	3,290945

Figura 4.11 – Gráfico da produção média utilizando a função gama para valores observados x valores ajustados



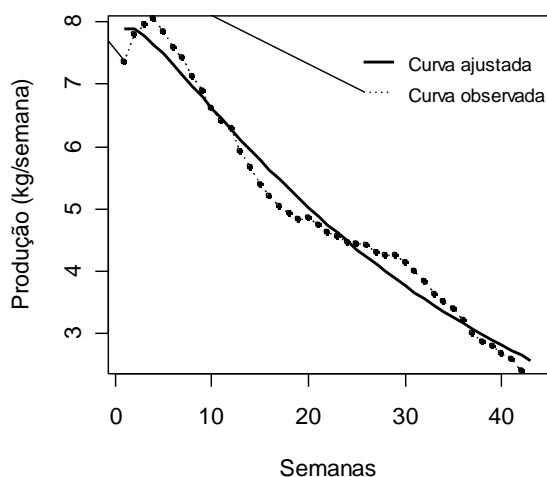
4.3 Ajuste de modelos não-lineares

a) Modelo não-linear para descrever a curva de lactação média das 825 vacas

Na Figura 4.12 estão representadas as curvas de lactação semanal média observada e ajustada pela função gama incompleta, expressa por (3), estimada por meio de regressão não-linear, expressa pelo modelo (2), para as vacas do rebanho em estudo. Verifica-se que o ajuste foi razoável, pois a diferença entre cada ponto observado e o respectivo ponto ajustado é pequena e, ainda, o formato da curva é semelhante ao encontrado em outros estudos sobre curvas de lactação da raça Caracu. A produção média observada inicia-se com 7,35kg e o pico de produção é alcançado no primeiro mês de lactação, com 8,04kg, sendo que a produção média observada no primeiro mês foi de 7,79kg. Após o pico de produção, houve um declínio contínuo; a produção começa a cair gradativamente, atingindo uma produção média observada de

2,40kg no último mês de lactação. Assim, sua curva não apresentou a tendência de uma reta como se observa para vacas zebuínas.

Figura 4.12 – Curva de lactação semanal média observada e ajustada



A Tabela 4.9 mostra um comparativo entre os valores estimados dos parâmetros a_0 , a_1 e a_2 para a lactação média de vacas Caracu, obtidos a partir da função gama incompleta para este estudo e para um estudo anterior realizado por Faro et al. (1996).

Tabela 4.9 – Valores estimados dos parâmetros a_0 , a_1 e a_2 para a lactação média de vacas Caracu, ajustada pela função gama incompleta.

<i>Parâmetro</i>	<i>Valor estimado neste estudo</i>	<i>Valor estimado por Faro et al. (1996)</i>
a_0	8,12	10,59
a_1	0,04	0,008
a_2	0,03	0,027

Observa-se que há divergências entre os valores encontrados nos dois estudos. Verifica-se que valor do parâmetro a_0 encontrado neste estudo é inferior ao obtido por Faro et al. (1996); porém, o valor obtido para o parâmetro a_1 é superior, enquanto os valores encontrados para o parâmetro a_2 em ambos os estudos são praticamente iguais, considerando-se o arredondamento.

Na Tabela 4.10 são apresentados os valores estimados para o tempo de pico (TP), a produção no pico (PP) e a persistência (S). Porém, deve-se salientar que para a obtenção desses valores, vacas com produção atípica foram consideradas, ou seja, aquelas que apresentam um dos parâmetros a_0 , a_1 ou a_2 negativos. Segundo a literatura, valores negativos para um dos parâmetros, pode interferir no resultado, apresentando valores subestimados.

Tabela 4.10 – Valor estimado para o tempo de pico (TP), a produção no pico (PP) e a persistência (S), para a lactação média de vacas Caracu ajustada pela função gama incompleta

<i>Medida de interesse</i>	<i>Valor Estimado</i>
Tempo de Pico = TP (em semanas)	1,45
Produção no Pico = PP (em kg)	7,90
Persistência = S	3,64

A grande maioria dos trabalhos realizados fornecem os valores estimados para TP, PP e S, para a lactação individual das vacas. Neste caso, os valores apresentados na Tabela 4.10 são para a lactação média, com isso, não existem valores de estudos anteriores que possam servir de referência para os valores obtidos neste estudo. Sabe-se apenas, da literatura, que as funções dos parâmetros a_0 , a_1 e a_2 , como produção inicial e produção no pico, auxiliam na determinação da forma da curva de lactação e que a persistência é o período pelo qual a lactação se mantém com altos níveis de

produção de leite. Entretanto, para a função gama incompleta, essa medida não tem unidade, o que dificulta sua interpretação biológica. Porém, segundo Andrighetto et al. (2004), quanto maior os seus valores absolutos, mais persistente é a lactação. Segundo Camargo (2000), a persistência de lactação é uma das características que define uma matriz leiteira especializada. Quanto a este aspecto, uma vaca para ser considerada especializada deverá apresentar média de produção de leite no décimo mês de lactação equivalente a 60% da média de produção de leite obtida no mês do pico de lactação. Observa-se que as vacas do rebanho em estudo apresentaram média de produção de leite no último mês de lactação de aproximadamente 31% da média de produção de leite obtida no pico de lactação. Esta porcentagem é considerada pequena, quando comparada com a pré-estabelecida por Camargo (2000). Porém, deve-se ressaltar que Camargo (2000), considera vacas que apresentam pico de lactação por volta dos 60 dias pós-parto, ou seja, no segundo mês de lactação, e que para as vacas da raça Caracu isso não se observa, elas apresentam normalmente pico de lactação no primeiro mês de lactação, como é o caso das vacas do rebanho em estudo. Todavia, Santos (2000), considera vaca de alta persistência aquela que apresenta queda de produção após o pico de lactação de no máximo 10% a cada 30 dias, isto é, capaz de manter pelo menos 90% da produção a cada 30 dias após o pico de lactação. Menciona, ainda, que vacas com alta persistência são capazes de produzir leite por mais de 10 meses. Com isso, pode-se dizer que as vacas do presente estudo apresentaram as características pré-estabelecidas por Santos (2000), podendo ser consideradas como de alta persistência, pois as mesmas apresentaram, em média, 11 meses de lactação e, ainda, a queda média de produção após o pico de lactação foi de aproximadamente 10% a cada 30 dias.

b) Modelo não-linear para descrever a curva de lactação de cada vaca

Curvas de lactação individuais foram ajustadas pela função gama incompleta, expressa por (4), estimadas por meio de regressão não-linear, expressa pelo modelo (2), para as lactações das 825 vacas do rebanho em estudo. Médias dos parâmetros da

função gama incompleta, assim como os percentuais de curvas típicas e atípicas, foram obtidas e são apresentadas nas Tabelas 4.11 e 4.12.

Tabela 4.11 – Média, desvio-padrão e coeficiente de variação (CV) associados aos parâmetros das curvas de lactação individuais de vacas Caracu ajustadas pela função gama incompleta

<i>Parâmetro</i>	<i>Média</i>	<i>Desvio-Padrão</i>	<i>CV (%)</i>
a_0	8,32	5,95	71,5
a_1	0,11	0,32	278,8
a_2	0,04	0,03	76,0

Os valores médios estimados para os parâmetros a_0 e a_1 por meio da função gama incompleta (Tabela 4.11), apresentaram-se razoavelmente próximos aos encontrados por Faro et al. (2002), que obtiveram os valores 10,82 e 0,05, respectivamente, para as curvas de lactação individuais de vacas Caracu de um rebanho da fazenda Recreio. Porém, o valor obtido para o parâmetro a_2 não foi muito semelhante, sendo 0,04 o encontrado neste estudo e 0,004 o encontrado por Faro et al. (2002). Todavia os valores obtidos neste estudo para os parâmetros a_0 , a_1 e a_2 são inferiores aos encontrados por Cobuci et al. (2000), que em um estudo sobre curvas de lactação na raça Guzerá, encontraram 12,091, 0,369 e 0,188 para os parâmetros a_0 , a_1 e a_2 , respectivamente.

Os desvios-padrão obtidos para os parâmetros a_0 e a_1 (Tabela 4.11) foram inferior aos encontrados por Faro et al. (2002), que obtiveram, respectivamente, 8,69 e 1,15. Já o desvio-padrão do parâmetro a_2 foi praticamente igual, sendo 0,02 o obtido por Faro et al. (2002) e 0,03 o encontrado neste estudo. Os valores do coeficiente de variação (CV) dos três parâmetros (Tabela 4.11), também foram inferiores ao obtidos por Faro et al. (2002), que encontraram altos CV para os parâmetros estimados pela

função gama incompleta para as curvas de lactação individuais de vacas Caracu, sendo para a_0 de 80,2, para a_1 de 2421,4, e para a_2 de 446,2.

Segundo Faro et al. (1996), alguns autores fazem uma observação em relação à interpretação biológica dos parâmetros que compõem cada modelo, a de que “é difícil estabelecer a partir de cada função quais valores obtidos para os parâmetros estão próximos de valores considerados normais”. No caso da função gama incompleta, sabe-se apenas que o parâmetro a_0 está relacionado à produção inicial, e os parâmetros a_1 e a_2 indicam apenas as taxas de aumento pré-pico ou declínio pós-pico.

A forma da curva de lactação é determinada principalmente pelo parâmetro a_0 . Neste estudo, todas as estimativas deste parâmetro foram positivas, apresentando valores entre 0,08 e 146,89. Este intervalo pode ser considerado pequeno, quando comparado ao encontrado por Faro et al. (2002), que foi de -2,54 e 444,76. Esta diferença é uma das explicações da divergência encontrada no CV dos dois estudos.

Tabela 4.12 – Percentual de curvas típicas e atípicas

<i>Curvas</i>	<i>Percentual</i>
Típicas	63,71%
Atípicas	36,29%

Segundo as restrições do modelo, curvas típicas ocorrem se valores positivos são observados para os parâmetros a_0 , a_1 e a_2 . Neste caso, aproximadamente 63,71% das curvas podem ser consideradas típicas (Tabela 4.12). A Figura 4.13 mostra algumas curvas de lactação individuais típicas obtidas neste estudo (selecionadas aleatoriamente), nas quais percebe-se um pico de produção e, em seguida, uma queda gradativa, forma esta bem tradicional, verificada em estudos anteriores.

Segundo a literatura, é esperado que a raça Caracu apresente curvas que sejam atípicas, no sentido de não apresentarem pico de lactação em torno do segundo mês,

como ocorre com a lactação de vacas holandesas. Neste estudo, a raça Caracu apresentou pico de produção, em média, na quarta semana de lactação. Verifica-se, neste caso, que 36,29% das curvas são atípicas (Tabela 4.12), cuja forma tende a ser decrescente como se observa na Figura 4.14, que mostra algumas curvas atípicas individuais selecionadas aleatoriamente. Este percentual é inferior ao obtido por Faro et al. (2002), que ao ajustarem curvas de lactação individuais de vacas Caracu pela função gama incompleta, estimaram 40% de curvas atípicas, que também foi superior ao encontrado por Cobuci et al. (2000), que obtiveram 25,3% de curvas atípicas no ajuste de curvas de lactação de vacas da raça Guzerá.

Figura 4.13 – Curvas de lactação individuais típicas ajustada e observada

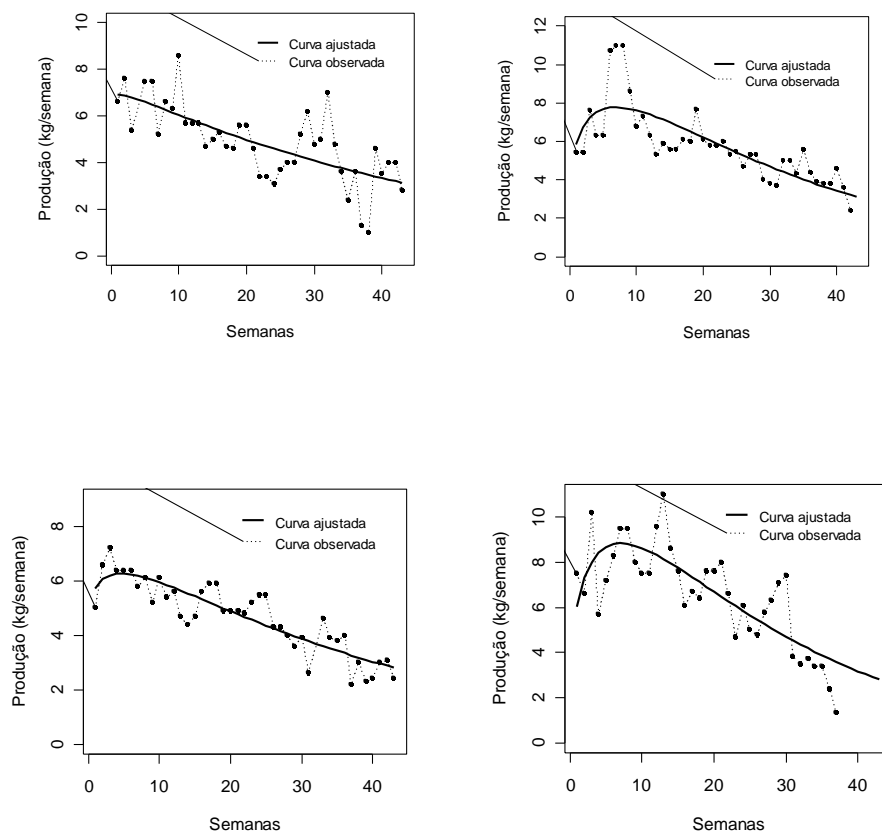
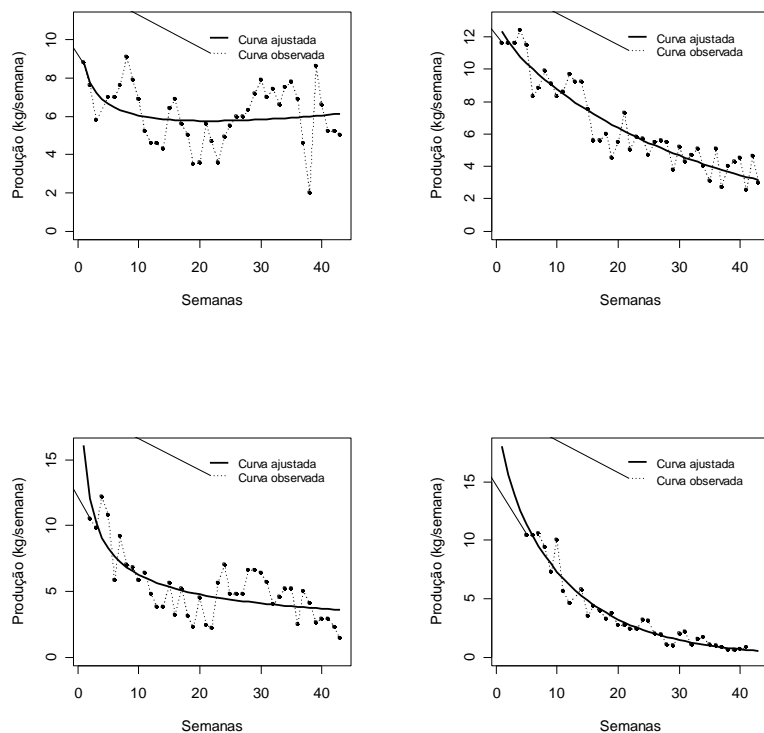


Figura 4.14– Curvas de lactação individuais atípicas ajustada e observada



Na Tabela 4.13 são apresentados os valores estimados para os parâmetros a_0 , a_1 e a_2 para a lactação individual de quatro vacas que apresentaram produção típica, ajustadas pela função gama incompleta, expressa por (4), cujos gráficos com a curva ajustada e observada podem ser visualizados na Figura 4.13.

Tabela 4.13 – Valores estimados para os parâmetros a_0 , a_1 e a_2 , para a lactação individual de quatro vacas Caracu que apresentaram produção típica

Parâmetro	Valor Estimado			
	Vaca 1	Vaca 2	Vaca 3	Vaca 4
a_0	7,08	6,05	5,87	6,34
a_1	0,02	0,27	0,13	0,35
a_2	0,02	0,04	0,03	0,05

Segundo a literatura, é difícil estabelecer a partir de cada função quais valores obtidos para os parâmetros estão próximos de valores considerados normais, com isso, não se pode concluir se os valores estimados na Tabela 4.13 podem ser considerados bons ou ruins e a literatura normalmente não apresenta os valores dos parâmetros estimados para cada vaca, apresentam apenas os valores médios estimados. Sabe-se apenas que o parâmetro a_0 está relacionado à produção inicial, e os parâmetros a_1 e a_2 indicam apenas as taxas de aumento pré-pico ou declínio pós-pico.

Valores estimados para o tempo de pico (TP), a produção no pico (PP) e a persistência (S), para a lactação individual de quatro vacas que apresentaram produção típica, cujos valores estimados para os parâmetros podem ser observados na Tabela 4.13, são apresentados na Tabela 4.14.

Tabela 4.14 – Valores estimados para o tempo de pico (TP), a produção no pico (PP) e a persistência (S), para a lactação individual de quatro vacas Caracu com produção típica, ajustadas pela função gama incompleta.

<i>Variável</i>	<i>Valor Estimado</i>			
	<i>Vaca 1</i>	<i>Vaca 2</i>	<i>Vaca 3</i>	<i>Vaca 4</i>
TP (semana)	0,99	6,87	4,50	7,04
PP (kg)	6,94	7,78	6,27	8,84
S	3,99	4,11	4,02	4,05

Verifica-se na Tabela 4.14, que o tempo de pico da vaca 1 foi bem inferior ao tempo de pico encontrado para as outras três vacas. Já a produção no pico e a persistência das quatro vacas não apresentaram muita diferença, porém, a vaca 1 continua apresentando os menores valores. Como em estudos anteriores os valores estimados para TP, PP e S não são apresentados para cada vaca, pode-se apenas dizer que os valores encontrados neste estudo (Tabela 4.14) são inferiores aos valores médios obtidos em outros estudos.

Como observado nas Figuras 4.13 e 4.14, as curvas típicas são diferentes das curvas atípicas, tanto no seu formato como na sua interpretação, com isso têm-se dois grupos, o das vacas com produção típica e o das vacas com produção atípica.

Na Tabela 4.15 são apresentados: o valor médio estimado para o tempo de pico (TP), a produção no pico (PP) e a persistência (S), porém, para obtenção dessas estimativas apenas as vacas que apresentaram produção típica foram consideradas, isto é, aquelas que apresentaram os parâmetros a_0 , a_1 e a_2 positivos. Segundo a literatura, é conveniente que valores negativos não sejam utilizados neste cálculo, pois estimativas dos parâmetros a_0 , a_1 ou a_2 inferiores a zero com certeza afetariam o resultado, apresentando valores fora do esperado. Sendo assim, as vacas que apresentaram produção atípica, ou seja, que apresentaram pelo menos um dos parâmetros a_0 , a_1 ou a_2 negativos, não foram consideradas.

Tabela 4.15 – Média, desvio-padrão, coeficiente de variação (CV) e valores mínimo e máximo de tempo de pico (TP), produção no pico (PP) e persistência (S) de vacas Caracu com lactação típica

<i>Variável</i>	<i>Média</i>	<i>Desvio-padrão</i>	<i>CV (%)</i>	<i>Mínimo</i>	<i>Máximo</i>
TP (semanas)	5,36	4,43	82,7	0,06	65,75
PP (kg)	8,13	2,03	24,9	3,50	14,83
S	3,86	0,58	15,0	2,73	6,98

Como se observa na Tabela 4.15, a média do tempo de pico (5,36 semanas) e da produção no pico (8,13kg), são semelhantes às encontradas em estudos anteriores, apenas a persistência apresentou valor inferior, sendo 3,86 o valor médio estimado para a persistência neste estudo e 6,04 o valor estimado por Faro et al. (2002). A persistência é o período pelo qual a lactação se mantém com altos níveis de produção

de leite. Entretanto, para a função gama incompleta, essa medida não tem unidade, o que dificulta sua interpretação biológica. Embora não exista na literatura valores ótimos ou pré-estabelecidos, sabe-se que maiores valores estimados para S estão relacionados a maiores persistências de lactação. Contudo, a persistência (S) foi a função do parâmetro que menos variou (CV = 15,0%).

O desvio-padrão de cada uma das funções dos parâmetros foi muito semelhante ao encontrado em outros estudos, bem como o coeficiente de variação (CV) obtido para TP, PP e S.

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Modelos lineares e não-lineares foram considerados neste trabalho para a análise da produção de leite (lactação) observada ao longo do tempo (semanas). Utilizando a teoria de modelos lineares generalizados (MLG) para dados longitudinais, foram ajustados alguns modelos e, dentre estes, o modelo gama com estrutura de correlação auto-regressiva (AR1) foi o que apresentou os melhores resultados. A estrutura de correlação escolhida para este modelo por meio do critério de informação de Akaike mostra que a produção de leite (em kg) observadas ao longo das semanas para uma mesma vaca são de fato correlacionadas e que, esta correlação, é maior para produções mais próximas e menores para produções mais distantes. A partir deste modelo, foi possível verificar que as covariáveis: ano, retiro, mês, idade da vaca e semana de lactação apresentaram efeito significativo e são, portanto, necessárias para explicar a variabilidade que ocorre na produção de leite.

A dificuldade encontrada no ajuste desses modelos foi quanto ao diagnóstico de ajuste. Gráficos de diagnóstico de influência para MLG para dados longitudinais encontram-se disponíveis para dados balanceados. Os dados do estudo aqui analisados são, contudo, desbalanceados e isto requer algumas implementações computacionais para que tal diagnóstico possa ser realizado. Desse modo, o diagnóstico se restringiu

apenas à comparação de alguns valores observados e seus respectivos valores estimados a partir do modelo ajustado. No geral, este tipo de modelagem pareceu ser bastante razoável. Incorporar ao modelo a informação do parentesco dos animais é um aspecto que merece ser estudado.

Com relação aos resultados obtidos por meio do ajuste de modelos não-lineares, para os quais foram utilizados a função gama incompleta, pode-se concluir que esta função parece ser razoável para descrever a curva de lactação média. As curvas de lactação individuais são também razoavelmente descritas por tal função. Como, no entanto, ocorre uma heterogeneidade muito grande entre as vacas, seja entre aquelas com produção típica ou entre aquelas com produção atípica, estes modelos precisam ser melhorados ou estendidos. A heterogeneidade mencionada pode ser devida a diversos fatores observáveis (idade da vaca, por exemplo) e/ou fatores genéticos. A inclusão de covariáveis (fatores) nestes modelos, bem como a inclusão do parentesco entre os animais são, portanto, sugestões a serem consideradas. Devido às limitações computacionais encontradas até o momento, isto não foi possível de ser realizado neste trabalho. Em sendo possível a incorporação dessas e de outras informações, os modelos não-lineares tornam-se bastante atrativos devido a interpretação prática que se tem dos seus parâmetros ou de funções desses parâmetros. O tempo de pico, a produção no pico e a persistência, apresentados neste trabalho, são exemplos de tais funções.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRIGHETTO, C.; PICCININ, A.; GIMENEZ, J.N.; JORGE, A.M.; MÓRI, C. *Curva de lactação de búfalas Murrah ajustadas pela função gama incompleta*. In: V SIMPÓSIO DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 2004, Pirassununga - SP.

BATISTA, E.F. et al. Comparação de métodos no ajuste de curvas de lactação de bovinos por meio de simulação. *Ciência Agrotec.*, Lavras, v. 27, n.4, p. 865-872, 2003.

BERGAMASCO, A.F.; AQUINO, L.H.; MUNIZ, J.A. Ajuste de modelos não-lineares a dados de crescimento de fêmeas da raça Holandesa. *Ciência Agrotec.*, Lavras, v. 25, n.2, p. 235-241, 2001.

BIANCHINI, E.S. *Estudo da curva de lactação de vacas da raça Gir*. Tese de Doutorado, USP, Ribeirão Preto, 1983. 88p.

BRODY, S. *Bioenergetson Growth*. New York: Reinhold, 1945. 1023p.

COBUCCI, J.A.; EUCLYDES, R.F.; VERNEQUE, R.S. et al. Curva de lactação na raça Guzará. *Revista Brasileira de Zootecnia*, Viçosa, v. 29, n.5, p. 1332-1339, 2000.

COBUCCI, J.A.; EUCLYDES, R.F.; COSTA, C.N. et al. Análises da persistência na lactação de vacas da raça Holandesa, usando produção no dia de controle e modelo de regressão aleatória. *Revista Brasileira de Zootecnia*, Viçosa, v. 33, p. 546-554, 2004.

CORDEBELLA, A. *Contagem de células somáticas e produção de leite em vacas Holandesas confinadas*. 2003. Tese (Doutorado) – Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo, Piracicaba, SP.

DURÃES M.C. et al. Curvas de lactação de vacas da raça Holandesa mantidas em confinamento total. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v. 43, n.5, p. 459-460, 1991.

FARO, L. E.; ALBUQUERQUE, L. G. *Comparação dos ajustes da curva de lactação de vacas caracu de acordo com três funções matemáticas*. In: XXXIII REUNIAO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 1996, FORTALEZA. ANAIS DA SBZ. FORTALEZA, 1996. p. 15-17.

FARO, L.E.; ALBUQUERQUE, L.G. Comparação de alguns modelos matemáticos para o ajuste às curvas de lactação individuais de vacas da raça Caracu. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, Belo Horizonte, v. 54, n.3, p. 295-302, 2002.

FARO, L.E.; ALBUQUERQUE, L.G. Utilização de modelos de regressão aleatória para produção de leite no dia de controle, com diferentes estruturas de variâncias residuais. *Revista Brasileira de Zootecnia*, Jaboticabal, v. 32, n.5, p.1104-1113, 2003.

FARO, L.E.; ALBUQUERQUE, L.G. Predição de valores genéticos para produção de leite no dia do controle e para a produção acumulada até 305 dias. *Revista Brasileira de Zootecnia*, Jaboticabal, v. 34, n.2, p. 496-507, 2005.

GONÇALVES, T.M.; OLIVEIRA, A.I.G.; FREITAS, T.F.; PEREIRA, I.G. Curvas de lactação em rebanhos da raça Holandesa no Estado de Minas Gerais. Escolha do modelo de melhor ajuste. *Revista de Zootecnia*, Viçosa, v. 31, n.3, p. 1689-1694, 2002.

JUNQUEIRA, L.V., NEIVA, R.S., VEIGA, R.D. et al. 1997. Estudo das curvas de lactação de vacas Holandesas de alguns rebanhos do Estado de Minas Gerais, por intermédio de uma função gama incompleta. *R. Soc. Bras. Zootec.*, v. 26, n.6, p. 1109-1118.

LEDIC, I.L.; VERNEQUE, R.S.; MARTINEZ, M.L. Estimativas de parâmetros fenotípicos e ambientais das produções de leite no dia do controle e em 305 dias de lactação em vacas da raça Gir. *Revista Ciência Agrotec.*, Lavras, v. 27, n.4, p. 921-926, 2003.

LEDIC, I.L.; SOBRINHO, E.B.; TUPY, O. Comparação de métodos para estimar a produção de leite com base em controles semanais, quinzenais, mensais e bimestrais. *Pesquisa Agropecuária Brasileira*, Brasília, v. 26, n.8, p.1127-1137, 1991.

MACMANUS, C.; FILHO, S.G.; MARIANTE, S.A.; LOUVANDINI, H. Fatores que influenciam os parâmetros das curvas de lactação em Cabras no Distrito Federal. *Revista Brasileira de Zootecnia*, Brasília, v. 32, n.6, p. 1614-1623, 2003.

MOLENTO, C.F.M.; MONARDES, H.; RIBAS, N.P.; BLOCK, E. Curvas de lactação de vacas holandesas do Estado do Paraná, Brasil. *Revista Ciência Rural*, Santa Maria, v. 34, n.5, p. 1585-1591, 2004.

MUÑOZ-BERROCAL, M.; TONHATI, H.; CERÓN-BERROCAL, M.; DUARTE, J.M.C.; CHABARIBERI, R.L. Uso de modelos lineares e não lineares para o estudo da curva de lactação em Búfalas Murrah e seus mestiços em sistema de criação semi extensivo, no Estado de São Paulo. *Archivos Latinoamericanos de Produccion Animal*, Perú, v. 13, n.1, p. 19-23, 2005.

NELDER, J.A.; WEDDERBURN, R.W.M. Generalized linear models. *Journal of the Royal Statistical Society*, Series A, v.135, p. 370-384, 1972

NICOLAU, C.V.J.; SILVA, R.G.; MOTA, L.S.L.S.; VERÍSSIMO, C.J. Características da pele em bovinos da raça Caracu. *Archivos Zootecnicos*, Brasília, v. 53, p. 25-34, 2004.

QUIRINO, C.R.; PEREIRA, J.C.C.; PEREIRA, S.; BERGMANN, J.A.G. Avaliação da duração da lactação e da produção diária média de leite na primeira lactação na raça Caracu utilizando o modelo animal. *Revista de Produção Animal*, Belo Horizonte, v. 6, n.1, p. 71-78, 1998.

SILVA, F.F.; MUNIZ, J.A.; AQUINO, L.H.; SÁFADI, T. Abordagem Bayesiana da curva de lactação de cabras Saanem de primeira e segunda ordem de parto. *Pesquisa Agropecuária Brasileira*, Brasília, v. 40, n.1, p. 27-33, 2005.

WOOD, P.D.P. *Algebraic Model of the Lactation Curve in Cattle*, Nature, London, v. 216, p. 164-165, 1967.

ZEGER, S.L.; LIANG, Y.Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, v. 42, p.121-130.

DOCUMENTOS CONSULTADOS ONLINE

CAMARGO, A.C. *Sistema de produção de leite: conceitos básicos*, disponível em: <http://www.bichoonline.com.br/artigos/bb0030.htm>

COSTA, S.C. *Modelos lineares generalizados mistos para dados longitudinais*. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11134/tde-09052003-164143>

DEMÉTRIO, C.G.B. *Modelos lineares generalizados em experimentação agrônômica*. Disponível em: <http://www.lce.esalq.usp.br/clarice/Apostila.pdf>

DIAS, A.C. *Fazenda Chiqueirão*. Disponível em: <http://www.caracu.agr.br/index.swf>

PAULA, G.A. *Modelos de regressão com apoio computacional*. Disponível em: <http://www.ime.usp.br/~giapaula/livro.html>

PÉREZ, F.L. *Modelos lineares generalizados*. Disponível em: <http://www.est.ufpr.br/~lucambio/Econometria.html> - MLG

REVISTA AGROPECUÁRIA TROPICAL. *O Europeu tropical que veio também nas caravelas*. Disponível em: <http://www.zebus.com.br/at/>

SANTOS, F.A.P. *Intensificação de sistemas de produção de leite*, disponível em: <http://www.bichoonline.com.br/artigos/bb0035.htm>

SLEPAK, A. *Análise de dados longitudinais com aplicação na indústria*. Disponível em: <http://www.est.ufpr.br/~lucambio/Econometria/Adlind.pdf>

APÊNDICES

APÊNDICE 1 - COMANDOS DO SOFTWARE “R”

1.1 Lactação de vacas Caracu da Fazenda Chiqueirão em Poços de Caldas – MG Análise Descritiva

```
caracu<-read.table("c:/caracu.txt",h=T)
dat1<-subset(caracu, ano>90)

write(t(dat1),ncol=11,"c:/Temp/dat1.dat")
attach(caracu)
nrow(caracu)
[1] 32978

a1<-unique(pai)
a2<-unique(mae)
a3<-unique(ano)
```

1.1.1 Gráfico com frequências de vacas por idade

```
plot(table(idade),type="h")
```

1.1.2 Medidas = (n semanal, produção semanal média, desvio padrão)

```
medidas<-matrix(nrow=43,ncol=3)
for(i in 1:43){
  medidas[i,1]<-length(prod[semana==i])
  medidas[i,2]<-mean(prod[semana==i])
  medidas[i,3]<-sd(prod[semana==i])
}
medidas
```

1.1.3 Gráfico da produção semanal média

```
plot(medidas[,2], type = "l", bty="n", ylim=range(c(0,10)),xlim=range(c(0,50)),
     xlab="Semana",ylab="kg/semana")
points(medidas[,2],pch=20)
title("Produção Semanal Média")
```

1.1.4 Dispersão da produção observada por semana e linha da produção semanal média

```
plot(semana,prod,type="p",pch=20,xlab="Semanas", ylab="Produção (kg/semana)")
lines(medidas[,2],pch=21,col=4,lwd=2)
```

1.1.5 Gráfico dos perfis individuais e linha da produção semanal média perfis de n = 50 vacas de um total de nt = 825 vacas

```
a2<-sort(unique(animal))
nt<-length(a2)
n<-50
plot(prod ~ semana, data=caracu, ylab="Produção", ylim=range(0,15),
     subset=animal==a2[1],lty=1,type="l")
for(i in 1:n){
  points(prod ~ semana, data=caracu, subset=animal==a2[i],lty=1,type="l")
}
```

1.1.6 Produção média (por semana) por classe de idade da vaca

```
dat1<- subset(caracu, idade < 30)
prodsm<-matrix(nrow=43,ncol=1)
for(i in 1:43){
  prodsm[i]<-mean(dat1$prod[dat1$semana==i])
}
plot(prodsm, type = "l", bty="n",ylim=range(c(0,10)),xlim=range(c(0,50)),
      xlab="Semanas",ylab="Produção Média (kg/semana)",col=2)
legend(30,10,lty=c(1),col=2,c("25 a 30 meses"),bty="n",cex=0.8)

dat1<- subset(caracu, idade >= 30 & idade < 35)
prodsm<-matrix(nrow=43,ncol=1)
for(i in 1:43){
  prodsm[i]<-mean(dat1$prod[dat1$semana==i])
}
lines(prodsm,lty=2,col=4)
legend(30,9.5,lty=c(2),col=4,c("30 a 35 meses"),bty="n",cex=0.8)

dat1<- subset(caracu, idade >= 35 & idade < 40)
prodsm<-matrix(nrow=43,ncol=1)
for(i in 1:43){
  prodsm[i]<-mean(dat1$prod[dat1$semana==i])
}
lines(prodsm,lty=3,col=1)
legend(30,9,lty=c(3),col=1,c("35 a 40 meses"),bty="n",cex=0.8)

dat1<- subset(caracu, idade >= 40 & idade < 45)
prodsm<-matrix(nrow=43,ncol=1)
for(i in 1:43){
  prodsm[i]<-mean(dat1$prod[dat1$semana==i])
}
lines(prodsm,lty=5,col=3)
legend(30,8.5,lty=c(5),col=3,c("40 a 45 meses"),bty="n",cex=0.8)

dat1<- subset(caracu, idade >= 45 & idade < 50)
prodsm<-matrix(nrow=43,ncol=1)
for(i in 1:43){
  prodsm[i]<-mean(dat1$prod[dat1$semana==i])
}
lines(prodsm,lty=6,col=6)
legend(30,8,lty=c(6),col=6,c("45 a 50 meses"),bty="n",cex=0.8)
```

1.1.7 Produção média por classe de idade da vaca

```
id<-c(26.8,28.8,30.8,32.8,34.8,36.8,38.8,40.8,42.8,44.8,46.8,48.8,50.8)

prodidade<-matrix(nrow=12,ncol=1)
for(i in 1:12){
  prodidade[i]<-mean(prod[idade>id[i]& idade<id[i+1]])
}

prodidade

plot(prodidade,type="b",pch=20, xlab="Idade",ylab="kg/semana",xaxt="n")
axis(1,1:12,c("28.8","30.8","32.8","34.8","36.8","38.8","40.8","42.8",
             "44.8","46.8","48.8","50.8"))
```

1.1.8 Produção média por mês de controle 01 a 12 (janeiro a dezembro)

```
prodmes<-matrix(nrow=12,ncol=1)
```

```

for(i in 1:12){
  dat1<- subset(caracu, mes==i)
  prodmes[i]<-mean(dat1$prod)
}
prodmes

plot(prodmes, type = "l", bty="n",ylim=range(c(4,6.5)),xlim=range(c(0,12)),
      xlab="Mês de Controle (Jan a Dez)", ylab="kg/semana")
points(prodmes,pch=20)
title("Produção Semanal Média\n por Mês de Controle")

```

1.1.9 Produção semanal média por mês de controle dentro do ano de controle

```

prodano1<-matrix(nrow=12,ncol=9)
for(j in 1:9){
  for(i in 1:12){
    dat1<-subset(caracu, ano==90+j)
    prodano1[i,j]<-mean(dat1$prod[dat1$mes==i])
  }
}
prodano1

plot(prodano1[,1], type = "l", bty="n",ylim=range(c(2.5,9.5)),xlim=range(c(0,12)),
      xlab="Mês de Controle (Jan a Dez)", ylab="kg/semana")
points(prodano1[,1],pch=20)
title("Produção Média")

lines(prodano1[,2], type = "l", lty=2)
points(prodano1[,2],pch=20)

for(i in 3:4){
  lines(prodano1[,i], type = "l", lty=i-1)
  points(prodano1[,i],pch=20)
}
legend(9,9.5,lty=c(1,2,3,4),c("1991", "1992", "1993", "1994"),bty="n",cex=0.8)

plot(prodano1[,5], type = "l", bty="n",ylim=range(c(2.5,9.5)),xlim=range(c(0,12)),
      xlab="Mês de Controle (Jan a Dez)", ylab="kg/semana")
points(prodano1[,5],pch=20)
title("Produção Média")

for(i in 6:8){
  lines(prodano1[,i], type = "l", lty=i-4)
  points(prodano1[,i],pch=20)
}
legend(9,9.5,lty=c(1,2,3,4),c("1995", "1996", "1997", "1998"),bty="n",cex=0.8)

```

1.1.10 Produção média por retiro em cada ano

```

ret1<-matrix(nrow=8,ncol=15)
for(j in 1:15){
  for(i in 1:8){
    dat1<-subset(caracu, retiro==j)
    ret1[i,j]<-mean(dat1$prod[dat1$ano==90+i])
  }
}
ret1

plot(ret1[,1],type="p",pch=15, ylim=range(c(2,9)),xlab="Ano de Controle",
      ylab="kg/semana",xaxt="n",cex=0.9)
axis(1,1:8,c("91", "92", "93", "94", "95", "96", "97", "98"))
lines(ret1[,2],type="p",pch=16,lty=2,cex=0.8)

```

```

lines(ret1[,3],type="p",pch=17,lty=3,cex=0.8)
lines(ret1[,4],type="p",pch=22,lty=1,cex=0.8)
lines(ret1[,5],type="p",pch=24,lty=2,cex=0.8)
lines(ret1[,6],type="p",pch=20,lty=3,cex=0.8)
lines(ret1[,7],type="p",pch=21,lty=4,cex=0.8)

plot(ret1[,8],type="p",pch=15,ylim=range(c(2,9)),xlab="Ano de Controle",
      ylab="kg/semana",xaxt="n",cex=0.9)
axis(1,1:8,c("91","92","93","94","95","96","97","98"))
lines(ret1[,9],type="p",pch=16,lty=1,cex=0.8)
lines(ret1[,10],type="p",pch=17,lty=2,cex=0.8)
lines(ret1[,11],type="p",pch=22,lty=3,cex=0.8)
lines(ret1[,12],type="p",pch=24,lty=2,cex=0.8)
lines(ret1[,13],type="p",pch=20,lty=3,cex=0.8)
lines(ret1[,14],type="p",pch=21,lty=1,cex=0.8)
lines(ret1[,15],type="p",pch=25,lty=2,cex=0.8)

```

1.2 Modelo 1: Ajuste do modelo linear generalizado para lactação

$$\text{Função gama: } f(y/\theta, \phi) = \exp \left\{ \frac{y\theta + \ln(-\theta)}{\phi^{-1}} + (\phi - 1)\ln(y\phi) + \ln(\phi) - \ln(\Gamma(\phi)) \right\}$$

```

dados<-read.table("C:/bovinos.txt",header=T)
attach(dados)
names(dados)
library(geepack)
require(geepack)

```

1.2.1 Modelo ajustado com a estrutura de correlação autoregressiva

```

ajuste1<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,
family=Gamma,corstr="ar1",data=dados)
AIC.gee(ajuste1)

```

1.2.2 Modelo ajustado com a estrutura de correlação independente

```

juste2<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,
family=Gamma,corstr="independence",data=dados)
AIC.gee(ajuste2)

```

1.2.3 Modelo ajustado com a estrutura de correlação permutável

```

ajuste3<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,
family=Gamma,corstr="exchangeable",data=dados)
AIC.gee(ajuste3)

```

1.2.4 Sumário do modelo ajustado com a estrutura de correlação que apresentou o menor AIC (autoregressiva)

```
summary(ajuste1)
```

1.2.5 Pontos para o gráfico da curva de lactação para as estimativas do ano de 1992, mês 3, retiro 5, idade 36,2 meses em função da semana

```

x1<-subset(dados, ano==92)
x1
x2<-subset(x1, mes==2)

```

```

x2
x3<-subset(x2, retiro==1)
x3
x4<-subset(x3, idade==39)
x4
x5<-subset(x4, semana==37)
x5

x1<-mean(prod[ano==92 & mes==2 & retiro==1 & idade==39 & semana==37])

```

1.2.5 Valores Observados

```

prod1<-mean(prod[ano==92 & mes==3 & retiro==8 & idade==35.2 & semana==5])
prod1

prod2<-mean(prod[ano==93 & mes==2 & retiro==3 & idade==32.1 & semana==36])
prod2

prod3<-mean(prod[ano==94 & mes==8 & retiro==6 & idade==49 & semana==2])
prod3

prod4<-mean(prod[ano==93 & mes==10 & retiro==4 & idade==34.1 & semana==8])
prod4

prod5<-mean(prod[ano==95 & mes==6 & retiro==2 & idade==30.9 & semana==29])
prod5

prod6<-mean(prod[ano==96 & mes==5 & retiro==9 & idade==28.5 & semana==17])
prod6

prod7<-mean(prod[ano==97 & mes==12 & retiro==3 & idade==44.5 & semana==33])
prod7

prod8<-mean(prod[ano==98 & mes==6 & retiro==10 & idade==32.0 & semana==21])
prod8

prod9<-mean(prod[ano==98 & mes==11 & retiro==9 & idade==32.3 & semana==40])
prod9

prod10<-mean(prod[ano==96 & mes==4 & retiro==4 & idade==27.8 & semana==28])
prod10

```

1.2.6 Valores Ajustados para Gama

```

prod1<-(0.134408784+ 0.004823010+ 0.018093384+(-0.009083686)+(-0.001791540 *35.2)+( 0.005989836 *5))
prod1
1/prod1

prod2<-0.134408784+0.010178464+0.020658063+0.008471910+(-0.001791540*
32.1)+(0.005989836*36)
prod2
1/prod2

prod3<-(-0.134408784+0.011223675+0.031763341+0.021998801+(-0.001791540*
49.0)+(0.005989836*2))
prod3
1/prod3

prod4<-(-0.134408784+0.011223675+0.020401564+0.014126902+(-0.001791540*
34.1)+(0.005989836*8))
prod4
1/prod4

```

```
prod5<-(-0.134408784+0.015649851+0.031646025+0.003792573+(-0.001791540*  
30.9)+(0.005989836*29))
```

```
prod5  
1/prod5
```

```
prod6<-(-0.134408784+0.017536698+0.021726629+(-0.009163156)+(-0.001791540*  
28.5)+(0.005989836*17))
```

```
prod6  
1/prod6
```

```
prod7<-(-0.134408784+0.006897535+0.002989727+0.008471910+(-0.001791540*  
44.5)+(0.005989836*33))
```

```
prod7  
1/prod7
```

```
prod8<-(-0.134408784+0.013831486+0.031646025+(-0.001959887)+(-0.001791540*  
32.0)+(0.005989836*21))
```

```
prod8  
1/prod8
```

```
prod9<-(-0.134408784+0.013831486+0.009934941+(-0.009163156)+(-0.001791540*  
32.3)+(0.005989836*40))
```

```
prod9  
1/prod9
```

```
prod10<-(-0.134408784+0.017536698+0.019881103+0.014126902+(-0.001791540*  
27.8)+(0.005989836*28))
```

```
prod10  
1/prod10
```

1.3 Modelo 1.1: Ajuste do modelo linear generalizado para lactação

$$\text{Função normal: } f(y/\theta, \phi) = \exp \left\{ \frac{y\theta - \theta^2/2}{\phi} - \frac{1}{2} \left[\frac{y^2}{\phi} + \ln(2\pi\phi) \right] \right\}$$

```
dados<-read.table("C:/bovinos.txt",header=T)  
attach(dados)  
names(dados)  
library(geepack)  
require(geepack)
```

1.3.1 Modelo ajustado com a estrutura de correlação autoregressiva

```
ajuste4<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,  
family=gaussian,corstr="ar1",data=dados)  
AIC.gee(ajuste4)
```

1.3.2 Modelo ajustado com a estrutura de correlação independente

```
ajuste5<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,  
family=gaussian,corstr="independence",data=dados)  
AIC.gee(ajuste5)
```

1.3.3 Modelo ajustado com a estrutura de correlação permutável

```
ajuste6<-geese(prod~factor(ano)+factor(mes)+factor(retiro)+idade+semana,id=animal,  
family=gaussian,corstr="exchangeable",data=dados)  
AIC.gee(ajuste6)
```

1.3.4 Sumário do modelo ajustado com a estrutura de correlação que apresentou o menor AIC (autoregressiva)

summary(ajuste4)

1.3.5 Valores Observados

prod1<-mean(prod[ano==92 & mes==3 & retiro==8 & idade==35.2 & semana==5])
prod1

prod2<-mean(prod[ano==93 & mes==2 & retiro==3 & idade==32.1 & semana==36])
prod2

prod3<-mean(prod[ano==94 & mes==8 & retiro==6 & idade==49 & semana==2])
prod3

prod4<-mean(prod[ano==93 & mes==10 & retiro==4 & idade==34.1 & semana==8])
prod4

prod5<-mean(prod[ano==95 & mes==6 & retiro==2 & idade==30.9 & semana==29])
prod5

prod6<-mean(prod[ano==96 & mes==5 & retiro==9 & idade==28.5 & semana==17])
prod6

prod7<-mean(prod[ano==97 & mes==12 & retiro==3 & idade==44.5 & semana==33])
prod7

prod8<-mean(prod[ano==98 & mes==6 & retiro==10 & idade==32.0 & semana==21])
prod8

prod9<-mean(prod[ano==98 & mes==11 & retiro==9 & idade==32.3 & semana==40])
prod9

prod10<-mean(prod[ano==96 & mes==4 & retiro==4 & idade==27.8 & semana==28])
prod10

1.3.6 Valores Ajustados para Normal

prod11<-(-7.14012316+(-0.10316646)+(-0.33565590)+0.37975982+(0.05781883*
35.2)+(-0.14719948*5))
prod11

prod12<-(-7.14012316+(-0.28887338)+(-0.41565947)+(-0.33565590)+(0.05781883*
32.1)+(-0.14719948*36))
prod12

prod13<-(-7.14012316+(-0.38924494)+(-0.90494967)+(-0.82473755)+(0.05781883*
49.0)+(-0.14719948*2))
prod13

prod14<-(-7.14012316+(-0.28887338)+(-0.57347665)+(-0.48304250)+(0.05781883*
34.1)+(-0.14719948*8))
prod14

prod15<-(-7.14012316+(-0.44338649)+(-0.94710983)+(-0.16033037)+(0.05781883*
30.9)+(-0.14719948*29))
prod15

prod16<-(-7.14012316+(-0.48971716)+(-0.62720911)+0.52530975+(0.05781883*
28.5)+(-0.14719948*17))
prod16

```
prod17<-(7.14012316+(-0.07668477)+(-0.10695958)+(-0.22648341)+(0.05781883*
44.5)+(-0.14719948*33))
prod17
```

```
prod18<-(7.14012316+(-0.37330175)+(-0.94710983)+0.10771823+(0.05781883*
32.0)+(-0.14719948*21))
prod18
```

```
prod19<-(7.14012316+(-0.37330175)+(-0.28763303)+0.52530975+(0.05781883*
32.3)+(-0.14719948*40))
prod19
```

```
prod20<-(7.14012316+(-0.48971716)+(-0.46678115)+(-0.48304250)+(0.05781883*
27.8)+(-0.14719948*28))
prod20
```

1.4 Modelo 2: Ajuste de um único modelo de regressão não-linear para lactação

Função gama incompleta: $yt = ao*t^a1*exp(-a2*t)$

```
fm1 <- nls(prod ~ Ao*semana^A1*exp(-A2*semana),data = caracu,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
```

```
fm1
```

```
summary(fm1)
residuos <- resid(fm1)
ajustados <- fitted(fm1)
plot(fitted(fm1),resid(fm1))
qqnorm(resid(fm1))
```

1.4.1 Gráfico das curvas ajustada e observada

```
t<-1:43
a<-coef(fm1)
prode<-a[1]*t^a[2]*exp(-a[3]*t)

plot(t,prode, type="l",xlab="Semanas", ylab="Produção (kg/semana)",lwd=2)
points(t,medidas[,2],pch=20)
lines(t,medidas[,2],lty=3)
legend(24.1,7.7,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(24.1,7.2,lty=3,c("Curva observada"),bty="n",cex=0.8)
```

1.5 Modelo 2.1: Curvas não-lineares ajustadas para a lactação de cada vaca

```
caracu1<-subset(caracu,animal!=9234401)
detach(caracu)
attach(caracu1)

vaca<-unique(animal)
vaca[1:165]

fm1<-matrix(nrow=165,ncol=4)
for(i in 1:165){
  dat1<-subset(caracu1, animal==vaca[i])
  mod<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
  fm1[i,1]<-vaca[i]
  fm1[i,2]<-coef(mod)[1]
  fm1[i,3]<-coef(mod)[2]
  fm1[i,4]<-coef(mod)[3]
}
fm1
```

```

vaca[166:330]

fm2<-matrix(nrow=165,ncol=4)

for(i in 1:165){
dat1<-subset(caracu1, animal==vaca[165+i])
mod<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
fm2[i,1]<-vaca[165+i]
fm2[i,2]<-coef(mod)[1]
fm2[i,3]<-coef(mod)[2]
fm2[i,4]<-coef(mod)[3]
}
fm2

vaca[331:495]

fm3<-matrix(nrow=165,ncol=4)

for(i in 1:165){
dat1<-subset(caracu1, animal==vaca[330+i])
mod<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
fm3[i,1]<-vaca[330+i]
fm3[i,2]<-coef(mod)[1]
fm3[i,3]<-coef(mod)[2]
fm3[i,4]<-coef(mod)[3]
}
fm3

vaca[496:660]

fm4<-matrix(nrow=165,ncol=4)

for(i in 1:165){
dat1<-subset(caracu1, animal==vaca[495+i])
mod<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
fm4[i,1]<-vaca[495+i]
fm4[i,2]<-coef(mod)[1]
fm4[i,3]<-coef(mod)[2]
fm4[i,4]<-coef(mod)[3]
}
fm4

vaca[661:824]

fm5<-matrix(nrow=164,ncol=4)

for(i in 1:164){
dat1<-subset(caracu1, animal==vaca[660+i])
mod<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
start=c(Ao = 10, A1 = 0.05, A2=0.004))
fm5[i,1]<-vaca[660+i]
fm5[i,2]<-coef(mod)[1]
fm5[i,3]<-coef(mod)[2]
fm5[i,4]<-coef(mod)[3]
}
fm5

```

1.5.1 Produção típica

1.5.1.1 Coeficientes da vaca 1

```
dat1<-subset(caracu, animal==vaca[1])
fmv1<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
          start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv1

coef(fmv1)
residuos<-resid(fmv1)
ajustados<-fitted(fmv1)
plot(fitted(fmv1),resid(fmv1))
qqnorm(resid(fmv1))
```

1.5.1.2 Gráfico das curvas ajustada e observada da vaca 1

```
t<-1:43
a<-coef(fmv1)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,10)),ylab="Produção (kg/semana)",lwd=2)
vaca1<-subset(caracu,animal==vaca[1])
points(vaca1$semana,vaca1$prod,pch=20)
lines(vaca1$semana,vaca1$prod,lty=3)
legend(20,9.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,9.0,lty=3,c("Curva observada"),bty="n",cex=0.8)
```

1.5.1.3 Coeficientes da vaca 169

```
dat1<-subset(caracu, animal==vaca[169])
fmv169<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv169

coef(fmv169)
residuos<-resid(fmv169)
ajustados<-fitted(fmv169)
plot(fitted(fmv169),resid(fmv169))
qqnorm(resid(fmv169))
```

1.5.1.4 Gráfico das curvas ajustada e observada da vaca 169

```
t<-1:43
a<-coef(fmv169)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,12)),ylab="Produção (kg/semana)",lwd=2)
vaca169<-subset(caracu,animal==vaca[169])
points(vaca169$semana,vaca169$prod,pch=20)
lines(vaca169$semana,vaca169$prod,lty=3)
legend(20,10.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,10.0,lty=3,c("Curva observada"),bty="n",cex=0.8)
```

1.5.1.5 Coeficientes da vaca 344

```
dat1<-subset(caracu, animal==vaca[344])
fmv344<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv344
```

```

coef(fmv344)
residuos<-resid(fmv344)
ajustados<-fitted(fmv344)
plot(fitted(fmv344),resid(fmv344))
qqnorm(resid(fmv344))

```

1.5.1.6 Gráfico das curvas ajustada e observada da vaca 344

```

t<-1:43
a<-coef(fmv344)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,9)),ylab="Produção (kg/semana)",lwd=2)
vaca344<-subset(caracu,animal==vaca[344])
points(vaca344$semana,vaca344$prod,pch=20)
lines(vaca344$semana,vaca344$prod,lty=3)
legend(20,8.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,8.0,lty=3,c("Curva observada"),bty="n",cex=0.8)

```

1.5.1.7 Coeficientes da vaca 771

```

dat1<-subset(caracu, animal==vaca[771])
fmv771<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv771

```

```

coef(fmv771)
residuos<-resid(fmv771)
ajustados<-fitted(fmv771)
plot(fitted(fmv771),resid(fmv771))
qqnorm(resid(fmv771))

```

1.5.1.8 Gráfico das curvas ajustada e observada da vaca 771

```

t<-1:43
a<-coef(fmv771)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,11)),ylab="Produção (kg/semana)",lwd=2)
vaca771<-subset(caracu,animal==vaca[771])
points(vaca771$semana,vaca771$prod,pch=20)
lines(vaca771$semana,vaca771$prod,lty=3)
legend(20,10.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,10.0,lty=3,c("Curva observada"),bty="n",cex=0.8)

```

1.5.2 Produção atípica

1.5.2.1 Coeficientes da vaca 7

```

dat1<-subset(caracu, animal==vaca[7])
fmv7<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv7

```

```

coef(fmv7)
residuos<-resid(fmv7)
ajustados<-fitted(fmv7)
plot(fitted(fmv7),resid(fmv7))
qqnorm(resid(fmv7))

```

1.5.2.2 Gráfico das curvas ajustada e observada da vaca 7

```
t<-1:43
a<-coef(fmv7)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,11)),ylab="Produção (kg/semana)",lwd=2)
vaca7<-subset(caracu,animal==vaca[7])
points(vaca7$semana,vaca7$prod,pch=20)
lines(vaca7$semana,vaca7$prod,lty=3)
legend(20,10.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,10.0,lty=3,c("Curva observada"),bty="n",cex=0.8)
```

1.5.2.3 Coeficientes da vaca 180

```
dat1<-subset(caracu, animal==vaca[180])
fmv180<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv180

coef(fmv180)
residuos<-resid(fmv180)
ajustados<-fitted(fmv180)
plot(fitted(fmv180),resid(fmv180))
qqnorm(resid(fmv180))
```

1.5.2.4 Gráfico das curvas ajustada e observada da vaca 180

```
t<-1:43
a<-coef(fmv180)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,13)),ylab="Produção (kg/semana)",lwd=2)
vaca180<-subset(caracu,animal==vaca[180])
points(vaca180$semana,vaca180$prod,pch=20)
lines(vaca180$semana,vaca180$prod,lty=3)
legend(20,12.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,12.0,lty=3,c("Curva observada"),bty="n",cex=0.8)
```

1.5.2.5 Coeficientes da vaca 332

```
dat1<-subset(caracu, animal==vaca[332])
fmv332<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
            start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv332

coef(fmv332)
residuos<-resid(fmv332)
ajustados<-fitted(fmv332)
plot(fitted(fmv332),resid(fmv332))
qqnorm(resid(fmv332))
```

1.5.2.6 Gráfico das curvas ajustada e observada da vaca 332

```
t<-1:43
a<-coef(fmv332)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,16)),ylab="Produção (kg/semana)",lwd=2)
vaca332<-subset(caracu,animal==vaca[332])
```

```

points(vaca332$semana,vaca332$prod,pch=20)
lines(vaca332$semana,vaca332$prod,lty=3)
legend(20,15.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,14.8,lty=3,c("Curva observada"),bty="n",cex=0.8)

```

1.5.2.7 Coeficientes da vaca 772

```

dat1<-subset(caracu, animal==vaca[771])
fmv771<-nls(prod ~ Ao*semana^A1*exp(-A2*semana), data = dat1,
           start=c(Ao = 10, A1 = 0.05, A2=0.004))
fmv771

coef(fmv771)
residuos<-resid(fmv771)
ajustados<-fitted(fmv771)
plot(fitted(fmv771),resid(fmv771))
qqnorm(resid(fmv771))

```

1.5.2.8 Gráfico das curvas ajustada e observada da vaca 772

```

t<-1:43
a<-coef(fmv771)
prode<-a[1]*t^a[2]*exp(-a[3]*t)
plot(t,prode,type="l",xlab="Semanas",ylim=range(c(0,18)),ylab="Produção (kg/semana)",lwd=2)
vaca771<-subset(caracu,animal==vaca[771])
points(vaca771$semana,vaca771$prod,pch=20)
lines(vaca771$semana,vaca771$prod,lty=3)
legend(20,17.8,lty=1,lwd=2,c("Curva ajustada"),bty="n",cex=0.8)
legend(20,16.4,lty=3,c("Curva observada"),bty="n",cex=0.8)

```

1.5.3 Matriz dos Parâmetros

```

fm<-as.matrix(rbind(fm1,fm2,fm3,fm4,fm5))
min(fm[,2])
max(fm[,2])

mean(fm[,2])
mean(fm[,3])
mean(fm[,4])

sd(fm[,2])
sd(fm[,3])
sd(fm[,4])

cv1<-(sd(fm[,2])/mean(fm[,2]))
cv1
cv2<-(sd(fm[,3])/mean(fm[,3]))
cv2
cv3<-(sd(fm[,4])/mean(fm[,4]))
cv3

id.neg<-which(apply(fm>=0,1,all))
fmA<-fm[id.neg,]
fmA

```

1.5.4 Funções dos Parâmetros

1.5.4.1 Tempo de Pico

```
tp<-fmA[,3]/fmA[,4]
mean(tp)
sd(tp)

cv<-(sd(tp)/mean(tp))
cv

min(tp)
max(tp)
```

1.5.4.2 Produção no Pico

```
pp<-fmA[,2]*(tp)^fmA[,3]*exp(-fmA[,3])
mean(pp)
sd(pp)

cv<-(sd(pp)/mean(pp))
cv

min(pp)
max(pp)
```

1.5.4.3 Persistência

```
s<-(-(fmA[,3]+1)*log(fmA[,4]))
mean(s)
sd(s)

cv<-(sd(s)/mean(s))
cv

min(s)
max(s)
```

1.5.4.4 Matriz com Tempo de Pico, Produção no Pico e Persistência

```
fc<-cbind(tp,pp,s)
```