

**QUALIDADE DO AJUSTE DA FUNÇÃO
LOGÍSTICA MONOFÁSICA E DIFÁSICA
COM ESTRUTURA DE ERROS
INDEPENDENTES E AUTORREGRESSIVOS**

HUGO ANDRADE MEDEIROS

1999

46490

13243 MFW

HUGO ANDRADE MEDEIROS

**QUALIDADE DO AJUSTE DA FUNÇÃO LOGÍSTICA
MONOFÁSICA E DIFÁSICA COM ESTRUTURA DE
ERROS INDEPENDENTES E AUTORREGRESSIVOS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Mestrado em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof. Luiz Henrique de Aquino

LAVRAS
MINAS GERAIS - BRASIL
1999

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Medeiros, Hugo Andrade

Qualidade do ajuste da função logística monofásica e difásica com estrutura de erros independentes e autorregressivos / Hugo Andrade Medeiros. – Lavras : UFLA, 1999.

73 p. : il.

Orientador: Luiz Henrique de Aquino.

Dissertação (Mestrado) – UFLA.

Bibliografia.

1. Curva de Crescimento. 2. Função logística. 3. Análise de regressão não linear. I. Universidade Federal de Lavras. II. Título.

CDD-311.2

-519.536

HUGO ANDRADE MEDEIROS

**QUALIDADE DO AJUSTE DA FUNÇÃO LOGÍSTICA
MONOFÁSICA E DIFÁSICA COM ESTRUTURA DE ERROS
INDEPENDENTES E AUTORREGRESSIVOS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Mestrado em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

APROVADA em 26 de fevereiro de 1999

Prof. Daniel Furtado Ferreira UFLA

Prof. Eduardo Bearzoti UFLA

Prof.^ª Thelma Sáfyadi UFLA

Luiz Henrique de Aquino
Prof. Luiz Henrique de Aquino
UFLA
(Orientador)

LAVRAS
MINAS GERAIS - BRASIL

À glória de DEUS

Dedico.

AGRADECIMENTOS

Ao Instituto Presbiteriano Gammon e à Universidade Federal de Lavras, em especial ao Departamento de Ciências Exatas, pela oportunidade de realizar este curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

Aos professores Luiz Henrique de Aquino, orientador e amigo, Joel Augusto Muniz, coordenador do curso, Eduardo Bearzoti, Daniel Furtado Ferreira, Thelma Sáfadi e Ruben Delly Veiga, assim como aos demais professores do Departamento de Ciências Exatas, pelo carinho, amizade, apoio e sugestões.

À professora Maria Aparecida Possato, pelo apoio e incentivo.

Aos colegas de trabalho, Paula, Fabyano e Marcelo, pelo valioso auxílio e aos colegas de turma, Janaína, Telde, Marcelo e, em especial à Dileta, pelo grande estímulo e amizade.

Aos meus pais, Jeová e Zélia, pelo suporte emocional e material.

À minha esposa Vivian e à filha Giovana, pela paciência, compreensão, incentivo e inspiração.

SUMÁRIO

	Página
RESUMO.....	i
ABSTRACT.....	iii
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	3
2.1 Curvas de Crescimento.....	3
2.2 Modelos Não Lineares.....	10
3 MATERIAL E MÉTODOS.....	25
3.1 Simulação.....	25
3.2 Exemplo de Aplicação.....	32
4 RESULTADOS E DISCUSSÃO.....	34
5 EXEMPLO DE APLICAÇÃO.....	52
6 CONCLUSÕES.....	60
REFERÊNCIAS BIBLIOGRÁFICAS.....	61
ANEXOS.....	66

RESUMO

MEDEIROS, Hugo Andrade. **Qualidade do Ajuste da Função Logística Monofásica e Difásica com Estrutura de Erros Independentes e Autorregressivos**. Lavras: UFLA, 1999. 73p. (Dissertação - Mestrado em Agronomia / área de concentração em Estatística e Experimentação Agropecuária)*

Este trabalho teve por objetivo avaliar a qualidade e as características do ajuste da função logística monofásica e difásica com estrutura de erros independentes (EI) e autorregressivos de primeira e segunda ordens, AR(1) e AR(2). A função logística monofásica é expressa na forma: $y_t = a / (1 + \exp(-b(t-d))) + u_t$, sendo a forma difásica representada pela soma de duas destas funções, em que y_t é o peso à idade t , a é o valor assintótico, b é um parâmetro de crescimento e d é a idade no ponto de inflexão, em cada uma das fases. A estrutura de autocorrelação dos erros, no caso mais geral considerado neste estudo, AR(2), é da forma: $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$, na qual o último termo corresponde ao chamado ruído branco. Em uma primeira etapa elaborou-se um programa, usando os procedimentos do SAS, para simular 2000 amostras correspondentes a um intervalo de tempo de 0 a 60 meses, dividido em um número de partes igual a 10, 50, 100 e 1000. Fez-se o ajuste da função logística monofásica e difásica, através do procedimento MODEL do SAS, com valores de R^2 estabelecidos em 0,95 e 0,75 e estrutura de erros EI e AR(1) com valores do coeficiente de autocorrelação (ρ) estipulados em 0,80 e 0,30. Na segunda etapa, como exemplo de aplicação em dados reais, ajustou-se a função logística monofásica e difásica, com estrutura de erros EI, AR(1) e AR(2) a dados de crescimento de cinco vacas obtidos da literatura. Os resultados permitiram verificar que o número de iterações nas quais não se chegou a um ponto de convergência ou que convergiram para valores muito distantes do esperado, foi reduzido com o aumento de n , R^2 e a diminuição do valor de ρ . Verificou-se ainda que, aumentando-se n , os estimadores tendiam aos valores paramétricos e com maior precisão em erro quadrático médio. Quando a função logística foi ajustada aos dados da literatura, observou-se, pelos indicadores de qualidade do ajustamento e pela análise gráfica dos resíduos, que a introdução da estrutura de autocorrelação nos erros melhorou o ajuste, minimizando o problema das medidas

* Comitê Orientador: Luiz Henrique de Aquino - UFLA (Orientador), Eduardo Bearzoti - UFLA e Daniel Furtado Ferreira - UFLA

repetidas. Notou-se ainda uma sensível melhoria na qualidade do ajustamento quando foi utilizada a função difásica, em relação à monofásica, em todas as situações abordadas neste estudo. O uso da função logística monofásica e difásica, com estrutura de erros autorregressivos, é indicado para descrever o crescimento ponderal de vacas leiteiras com aproximadamente duas pesagens por mês ou uma por semana, em um intervalo de sessenta meses.

ABSTRACT

MEDEIROS, Hugo Andrade. **Quality of the fit of monophasic and diphasic logistic function with independent and autoregressive error structure.** Lavras: UFLA, 1999. 73p. (Dissertation - Master Program in Agronomy / major in Statistics and Agricultural Experimentation)*

This work aimed to evaluate both quality and characteristics of the fit for monophasic and diphasic function with independent errors (EI) and first and second order autorregressive errors structure, AR(1) and AR(2). The monophasic logistic function is expressed in the form: $y_t = a / (1 + \exp(-b(t-d))) + u_t$, being the diphasic form the sum of two of these functions, where y_t is the weight to the age t , a is the asymptotic value, b is a growth parameter and d is the age at the inflexion point, this in each of the phases. The autocorrelation structure of the errors, in the most general case considered in this study, AR(2), is the form: $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$, where ε_t is the white noise. In a first step, a program was designed by using the SAS procedures to simulate two thousand samples corresponding to a time interval from 0 to 60 months, split into a number of parts equal to 10, 50, 100 and 1,000. The fit of the monophasic and diphasic function was performed through MODEL procedure of SAS, with R^2 values set at 0.95 and 0.75 and EI and AR(1) error structure, with values of the correlation coefficients (ρ) stipulated in 0.80 and 0.30. In the second step, as an example of application in real data, the monophasic and diphasic logistic function was fitted with EI, AR(1) and AR(2) errors structures to a five cows growth data obtained from literature. The results allowed to verify that the number of iterations in which either a convergence point was not reached or that they converged to values beyond those expected, was decreased with the increase of n and R^2 and by decrease of the ρ value. It was verified that by increasing n , the parameter estimators tended toward the stipulated values and with a greater precision into average quadratic error. When logistic function was fitted to the data of the literature, it was found, by the goodness of fit indicators and by residues graphic analysis, that the introduction of autocorrelation structure into errors improved the adjustment, minimizing the problem of repeated measurements. In addition, a sensitive improvement in the goodness of fit was observed when diphasic function

* Guidance Committee: Luiz Henrique de Aquino - UFLA (Major Professor),
Eduardo Bearzoti - UFLA and Daniel Furtado Ferreira - UFLA

was employed as compared with monophasic function in every situation approached in this study. The use of monophasic and diphasic logistic function with autoregressive errors structure is pointed out to report dairy cows ponderal growth with about two weighings a month or once a week in a sixty-month interval.

1 INTRODUÇÃO

↘ A possibilidade de analisar o crescimento de seres vivos por meio de curvas de crescimento não é algo novo, mas as dificuldades envolvidas nos cálculos praticamente impediam sua implementação. Atualmente, com a difusão do uso de computadores e sendo estes cada vez mais rápidos, não apenas se tornou viável esta abordagem, como também fizeram dela algo cada vez mais popular no meio científico.

↘ As funções de crescimento são equações não lineares que tentam refletir as relações de crescimento inerente a várias partes do organismo animal, até a sua finalização com a maturidade, durante a sua vida, em situações ambientais distintas. Elas refletem as relações entre o impulso de crescimento inerente a um indivíduo e a maturidade em todas as partes do corpo e o ambiente no qual se expressam.

↘ O conhecimento das curvas de crescimento é importante para todos os pesquisadores da área animal que estejam interessados em recomendações sobre a eficiência da produção durante a vida do animal.

↘ De forma geral, as funções de crescimento condensam grande quantidade de informação inerente ao animal, durante toda a sua vida em, no máximo, quatro parâmetros não lineares, que podem ser interpretados biologicamente. Dentre os diversos modelos existentes, a função logística se destaca por sua flexibilidade e interpretabilidade. A curva logística é simétrica em relação ao seu ponto de inflexão, cuja ordenada corresponde à metade do valor assintótico. Ela possui

duas assíntotas, sendo uma delas representada pelo parâmetro a e outra pelo eixo das abcissas. A função logística difásica corresponde à soma de duas logísticas, conforme o enfoque proposto por Koops (1986).

Freqüentemente admite-se como verdadeira a hipótese da independência dos erros das observações repetidas num mesmo indivíduo. Esta é, sem dúvida, uma maneira de simplificar as análises e a solução do problema, pois permite usar o método de mínimos quadrados ordinário para o ajuste da regressão. No entanto, quando as observações são tomadas num mesmo indivíduo, o ideal seria levar em conta a existência de autocorrelação nos erros individuais. Em vista destas considerações, pretendeu-se, neste trabalho, avaliar a qualidade do ajuste da função logística monofásica e difásica, com estrutura de erros independentes e autorregressivos, com autocorrelação de primeira e de segunda ordens, AR(1) e AR(2), mediante dados de simulação e de dados reais, obtidos da literatura.

2 REFERENCIAL TEÓRICO

2.1 Curvas de Crescimento

Para se descrever relações tamanho-idade em animais, existe um grande número de modelos de crescimento disponíveis. Nomes como Brody, von Bertalanffy, Gompertz, Richards ou logística, associados à modelos assintóticos de crescimento com propriedades ligeiramente diferentes, surgem regularmente (Richards, 1969 In: Koops, 1986). Este autor afirma que os parâmetros dos modelos não lineares devem ter interpretação biológica, mas comenta que há dificuldade em se explicar o crescimento de um ser vivo por meio de equações contendo poucos parâmetros e que, no caso da curva de crescimento do peso corporal e altura para humanos, existe uma clara inadequabilidade de um simples modelo "sigmóide", pois com a idade de dois anos, por exemplo, já se atingiu a metade da altura adulta.

Robertson (1923) distinguiu três ciclos de crescimento para o homem e para várias espécies de animais, sendo que esta teoria só é completamente aceita para a curva de crescimento humano (Cheek, 1968). Para estudos de crescimento animal, todavia, curvas de crescimento de ciclo único são as geralmente usadas, conforme Koops (1986).

Peil e Helwin (1981) propuseram um modelo de crescimento capaz de ajustar muitos ciclos simultaneamente. Como Robertson (1923), eles usaram uma função de crescimento na qual cada ciclo era baseado em uma curva logística separada. Bock et al. (1973) combinaram duas funções logísticas em

um “modelo logístico duplo”, a fim de modelar o crescimento individual humano, em altura, do nascimento à idade adulta.

△ É sabido que diferentes partes ou componentes do corpo têm diferentes razões de crescimento (Palsson, 1955). Se estas diferenças são consideráveis, elas possivelmente podem conduzir a fases na curva de crescimento total do corpo. Este é o fator interno de maior importância no surgimento de fases no crescimento. Segundo este autor, o crescimento máximo ocorre primeiramente no sistema nervoso, depois o esqueleto ósseo, em seguida são os músculos e finalmente os depósitos de gordura e, para analisar a composição do corpo usando a função multifásica, deve-se assumir que diferentes componentes do corpo têm ganho máximo em diferentes idades.

A fim de se obter um melhor ajuste, a função multifásica pode ser modificada, ou os parâmetros podem ser agrupados ou substituídos por constantes, pois um decréscimo no número de parâmetros é freqüentemente desejável. Mas a escolha da modificação correta depende do conhecimento da função e de se perceber o significado dos parâmetros (Koops e Grossman, 1991). Ainda conforme estes autores, a escolha da função a ser usada para o ajuste dos dados é um tanto arbitrária e difícil, dependendo em grande parte da preferência e da experiência do pesquisador, que atualmente dispõe de maior facilidade no uso de recursos computacionais. A comparação das diferentes funções de crescimento é difícil, segundo eles, pois cada uma delas possui um conjunto único de propriedades. Por meio de um estudo comparativo sobre quatro funções, através de simulação, estes autores mostraram que uma função de múltiplas fases é uma alternativa razoável à função de fase única e que uma importante vantagem é não se ter que selecionar a “melhor” função de fase única. Eles concluíram ainda que a função multifásica pode ser usada em quase todos os casos, sendo que os

problemas de se ter que estimar um grande número de parâmetros para uma função de múltiplas fases, podem ser superados pelo fato de as estimativas destes parâmetros serem menos correlacionadas do que aquelas para uma função de fase única, as quais, embora forneçam uma impressão geral do crescimento, freqüentemente mostram desvios sistemáticos dos valores preditos em relação aos observados.

Conforme Fitzhugh (1976), o termo curva de crescimento sugere uma curva sigmóide descrevendo uma seqüência de medidas de tamanho, freqüentemente peso corporal. Segundo ele, uma terminologia mais geral seriam curvas tamanho-idade. Informa, ainda, que Cock (1966), segundo Tanner (1951), caracterizou os dados em três tipos básicos: estático, seccional-cruzado e longitudinal. No tipo estático, o tamanho é observado apenas uma vez para um grupo de indivíduos, todos à mesma idade ou estágio de desenvolvimento. No seccional-cruzado, assim como no do tipo estático, cada indivíduo é medido apenas uma vez, entretanto, as mesmas medidas são feitas em outras idades em outros indivíduos amostrados da mesma população. Em Estatística, a análise de curvas de crescimento aplica-se a dados longitudinais, consistindo de medidas repetidas, por exemplo no tempo, de alguma característica avaliada em cada membro de um grupo de indivíduos. Segundo este autor, os objetivos primários para o ajuste de curvas de crescimento são descritivos, pois a informação contida em uma seqüência de pontos tamanho-idade está consolidada em relativamente poucos parâmetros e preditivos, pois os parâmetros de curvas de crescimento são utilizados separadamente ou em conjunto para predizer taxas de crescimento, necessidades alimentares, respostas à seleção dentre outros fatores. Acrescenta, também, que os objetivos da análise determinam a escolha do método de ajustamento da curva de crescimento. Ele afirma que as bases primárias para a

comparação de métodos de ajuste de curvas de crescimento incluem a interpretabilidade biológica dos parâmetros, a qual geralmente depende do entendimento dos inter-relacionamentos de fatores genéticos e ambientais, os quais podem produzir um padrão particular; a qualidade do ajuste para os dados observados, que se refere à minimização dos desvios entre os valores observados e os preditos; e a dificuldade operacional, que varia de acordo com a escolha da função e as características de um conjunto específico de dados, pois a maioria das funções são sensíveis à frequência e regularidade dos dados. Ele menciona que os algoritmos envolvendo iterações são sensíveis à escolha dos valores iniciais e podem não convergirem para uma solução, além de ser possível encontrar estimativas dos parâmetros que sejam matematicamente corretas, mas absurdas biologicamente. No que diz respeito à qualidade do ajuste das funções de crescimento, o autor comenta que os testes usuais envolvendo variância residual, não são apropriados a dados longitudinais, devido à autocorrelação nos erros em medidas repetidas no tempo e acrescenta que os pesquisadores têm tendido a contar com uma comparação visual da curva ajustada aos dados observados e com uma inspeção dos parâmetros estimados em termos de sua coerência matemática e biológica.

Ainda no que se refere à qualidade do ajuste, DeNise e Brinks (1985) afirmam que o valor da soma de quadrados dos desvios entre os valores observados e os preditos é uma característica da qualidade do ajuste global da curva.

Modelos que não são lineares em seus parâmetros têm sido extensivamente usados para ajustar relações tamanho-idade durante a vida. Os casos particulares de uma função de quatro parâmetros (Richards, 1959; Nelder, 1961, 1962; Grosenbaugh, 1965), têm sido especialmente populares. A forma

3
1. Regimes med. 2, 1999

geral indeterminada, a qual pode ser referida como função de Richards, representa um número infinito de casos específicos determinados, incluindo a monomolecular (Brody, 1945), logística, Gompertz (Laird e Howard, 1967) e Bertalanffy (Bertalanffy, 1957).

A função de Richards é apresentada na Tabela 1, no caso mais geral e em quatro casos específicos. A inclusão do parâmetro M permite uma flexibilidade na localização do ponto de inflexão da curva de crescimento, visando a um melhor ajustamento.

TABELA 1. Funções de Richards, no caso geral e em quatro casos específicos, com valores do parâmetro M em cada caso.

Modelo	Equação para y_t	Parâmetro M
Richards :	$\alpha (1 \pm e^{-k(t-d)})^M$	variável
Brody :	$\alpha (1 - e^{-k(t-d)})$	1
Bertalanffy:	$\alpha (1 - e^{-k(t-d)})^3$	3
Logística :	$\alpha (1 + e^{-k(t-d)})^{-1}$	-1
Gompertz :	$\alpha \exp(e^{-k(t-d)})$	$M \rightarrow \infty$

FONTE: Fitzhugh (1976).

sendo:

α - valor assintótico da medida de tamanho quando $t \rightarrow \infty$. Geralmente é interpretado como o tamanho médio à maturidade, independente de flutuações de curta duração no tamanho devido a efeitos ambientais de clima e oferta de alimento;

- k - parâmetro de crescimento. Geralmente referido como um índice de maturidade, serve como medida da taxa de crescimento e para medir a taxa de mudança na taxa de crescimento. O valor $b = e^{kd}$ é um parâmetro de escala (constante de integração) o qual é estabelecido pelos valores iniciais de y_0 e t_0 . Este (b) é um parâmetro de ajuste para a situação em que $y_0 \neq 0$ e/ou $t_0 \neq 0$; por exemplo, quando apenas observações pós-natal são disponíveis e t_0 é tido como a data de nascimento;
- d - é a idade no ponto de inflexão;
- y_t - é o valor predito da medida para a idade t.

As interpretações biológicas relativas aos parâmetros da função de Richards e de seus casos particulares têm sido a principal razão para a sua popularidade. Geralmente, no entanto, as interpretações biológicas são elaboradas após a escolha empírica do modelo matemático (Fitzhugh, 1976).

Os parâmetros ^{ccc} destes modelos de crescimento são geralmente estimados usando o procedimento iterativo do método dos mínimos quadrados generalizado (Nelder, 1961; Marquardt, 1963; Fabens, 1965; Causton, 1969). Uma base de dados com mais que uns poucos indivíduos requer o uso de algoritmos eficientes e computadores de alta velocidade. Estimativas iniciais precariamente determinadas podem aumentar o número de iterações requeridas ou até impedir a convergência para uma solução razoável. A não ortogonalidade entre as colunas da matriz jacobiana pode levar a estimativas que, embora sejam matematicamente viáveis, são biologicamente impossíveis.

Severas flutuações nas relações peso-idade são bastante comuns; entretanto, os vários casos da função de Richards geralmente presumem um

aumento monotônico em tamanho, da origem até a assíntota. A curva ajustada obviamente suaviza as irregularidades dos dados reais, conforme Brown *et al.* (1972), mas ocorrem problemas de interpretação quando estas irregularidades afetam substancialmente os valores dos parâmetros estimados.

Kroll (1990) realizou um estudo do crescimento de vacas leiteiras, através de modelos monofásicos com estrutura de erros EI e AR(1). Conclui que se deve optar pelo processo autorregressivo, em relação ao usual (EI), devido à estrutura de autocorrelação dos erros. Dentre os modelos utilizados neste trabalho e com base nos valores dos coeficientes de determinação, este autor concluiu que o modelo logístico foi o que pior descreveu o crescimento dos animais.

Tornero (1996) trabalhou com modelos multifásicos de crescimento, com estrutura de erros EI e AR(1), aplicados a dados de dez vacas holandesas e de outros animais, obtidos da literatura. Através da análise de componentes principais, aplicados aos dados das vacas, a autora concluiu pelo modelo composto de duas fases e escolheu a função logística para descrever o crescimento dos animais. Comparando os modelos logísticos mono e difásicos, por meio das estimativas das variâncias dos erros, diz que o modelo difásico é mais eficiente que o monofásico. Ela observou que os pesos ao nascer são superestimados, em média, em torno de 40 kg no modelo monofásico com estrutura de erros EI e 3 kg no modelo difásico AR(1). A autora concluiu que, para se descrever o crescimento ponderal de vacas holandesas, do nascimento ao trigésimo mês de idade, o modelo logístico difásico com estrutura de erros AR(1) é suficiente em termos de ajuste e tem as qualidades necessárias para se fazer inferências de “parâmetros zootécnicos”.

2.2 Modelos Não-lineares

Stevens (1951) estudou o modelo de regressão:

$$E(y_i) = \alpha + \beta \rho^{x_i}, \quad (i = 1, 2, \dots, n)$$

em que α , β e ρ são parâmetros com $0 < \rho < 1$ e x_i é o tempo. Ele ressaltou o uso deste modelo em relações de crescimento, tanto animal como de plantas. Nesta equação, y_i tende assintoticamente para um certo limite, quando x_i tende para o infinito. O autor afirmou, ainda, que outros modelos assintóticos, como os de Gompertz e o logístico, por simples transformações matemáticas, podem ser escritos como o citado acima. Apresentou também tabelas para facilitar os cálculos da matriz de covariância, bem como exemplos de sua utilização.

Nelder (1961) descreveu o ajuste da família de curvas de quatro parâmetros, definidas pela equação diferencial:

$$\frac{dw}{dx} = kw \left(1 - \left(\frac{w}{A} \right)^{1/\theta} \right)$$

sendo w o peso, x o tempo e k , A e θ parâmetros. Supôs $\theta > 0$ e obteve a solução da equação:

$$w = \frac{A}{[1 + e^{-(\lambda+kx)/\theta}]^\theta}$$

que é chamada “Logística Generalizada”, em que λ é a constante de integração. Quando $\theta = 1$ obtém-se a função logística.

Nelder (1962) apresentou uma forma alternativa da função Logística Generalizada de maneira que θ pudesse apresentar valores positivos ou negativos. Substituiu na equação anterior:

$$\frac{1}{\theta} = \phi$$

$$\lambda' = -\ln \theta + \frac{\lambda}{\theta}$$

e

$$k' = \frac{k}{\theta}$$

obtendo:

$$w = \frac{A}{[1 + \phi e^{-(\lambda' + k'x)/\theta}]^{\frac{1}{\phi}}}$$

Nesta última função, se $\phi = 1$, tem-se a função logística; se $\phi = -1$, a função de Mitscherlich e se ϕ tender a zero, a função de Gompertz. Conforme este autor, o modelo logístico pode ser ajustado obtendo-se os valores iniciais para ϕ e A por método gráfico. A seguir calculam-se os valores iniciais de λ' e k' por meio de regressão ponderada de $\ln[w^\phi/(A^\phi - w^\phi)]$ contra x , com pesos $[|A^* - w^*|/A^*]^2$; a inclinação da reta de regressão fornece então k' e o intercepto estima $\lambda' - \ln|\phi|$. Finalmente, escreve-se $\alpha = \ln(A)$ e calculam-se as primeiras correções; o processo é então repetido.

Theil (1971) apresentou o método de mínimos quadrados generalizado. Considerou conhecida a matriz de transformação, discutiu e apresentou as propriedades dos estimadores. De acordo com o autor, ajusta-se uma regressão de Y_x contra Y_{ix} ($i= 0, 1, 2, \dots, k$; $x= 2, 3, \dots, n$) pelo método de mínimos quadrados ordinário. A partir dos desvios de regressão, indicados por e_x , calculou a estimativa de ρ_v , ($v= 1, 2, \dots$) como segue:

$$\rho_v = \frac{\sum_{x=z}^n e_{vx} \cdot e_{vz-1}}{(n-1)(QM \text{ Residuo})}$$

De posse desta estimativa, aplicou o método de mínimos quadrados generalizado, usando $\hat{\rho}_v$ em lugar de ρ_v .

Vieira & Mischan (1976), comparando curvas de crescimento, ressaltaram a importância das funções logística e de Gompertz em diversas

aplicações biológicas. Afirmaram que, embora a função de Gompertz apresente características interessantes, ela não é tão conhecida como a função logística.

A função logística pode ser ajustada pelo método proposto por Nelder (1961). Este método utiliza um processo iterativo para os cálculos, que neste caso, é facilitado pelo uso do computador.

Johnston (1977) apresentou uma teoria sobre autocorrelação dos resíduos, quando estes seguem um esquema autorregressivo de primeira ordem, cujo modelo é:

$$u_t = \rho u_{t-1} + \varepsilon_t ,$$

em que:

$$|\rho| < 1 ,$$

$$E(\varepsilon_t) = 0 ,$$

$$E(\varepsilon_t \varepsilon_{t-s}) = \sigma^2 , \text{ se } s = 0 ,$$

$$E(\varepsilon_t \varepsilon_{t-s}) = 0 , \text{ se } s \neq 0 ,$$

$t = 2, 3, \dots, n$; u_t é o erro autorregressivo de primeira ordem. A solução é conseguida pelo método dos mínimos quadrados generalizado (MQG), apresentando também uma metodologia para se estimar ρ . É notório que o MQG, convenientemente aplicado em um modelo linear:

$$Y = X\beta + u$$

no qual Y é o vetor das observações; X é a matriz de incidência; β é o vetor dos parâmetros e u é o vetor dos erros, gera o modelo que permite obter as estimativas dos parâmetros de forma não tentenciosa e de variância mínima.

Glasbey (1979) propôs um modelo para crescimento, que incorpora como termo estocástico um processo autorregressivo de primeira ordem. Isso contorna uma deficiência encontrada na literatura sobre regressão não linear com resíduos correlacionados. As estimativas de σ_0^2 e ρ_0 são dadas pelas fórmulas:

$$\sigma_0^2 = \frac{\sum_{i=1}^n e_i^2}{n} ,$$

e

$$\rho_0 = \frac{\sum_{i=2}^n e_i e_{i-1}}{(n-1)\sigma_0^2} ,$$

em que e_i é o desvio de regressão da função logística e σ_0^2 é a variância residual do modelo.

Glasbey (1980) afirmou que, em análise estatística, quando existe regressão entre duas variáveis, é comum supor que os erros em cada observação sejam independentes entre si. Entretanto, se num ajuste os resíduos se mostram correlacionados, a função em estudo não é apropriada. Quando os erros são correlacionados, as estimativas dos parâmetros de regressão, pelo método dos mínimos quadrados ordinário são ineficientes, segundo Berger (1981).

Machado (1980) usou médias de dados de crescimento ponderal e estimou os parâmetros da função logística, utilizando o método dos mínimos quadrados ordinário para chegar às equações normais e, em seguida, aplicando o método de Newton, obteve a solução do sistema por meio de processo iterativo.

Draper e Smith (1981) apresentaram um procedimento para estimar os parâmetros em modelos não lineares, por meio do método dos mínimos quadrados ordinário, supondo o erro aleatório com média zero e variância σ^2 . O sistema de equações normais é então resolvido pelo método da linearização, "steepest descent" (gradiente) ou "Marquardt's compromise". Os autores discutiram, ainda neste contexto, a obtenção do intervalo de confiança para os parâmetros, pelo método de mínimos quadrados aplicado aos modelos lineares, apresentando exemplos de aplicação.

Oliver (1986) discutiu a função de crescimento logística com três parâmetros

$$y = \frac{\lambda}{1 + \beta e^{-\alpha x}},$$

na qual λ , β e α são parâmetros e x é o tempo, que apresenta propriedades convenientes para estudos em Economia e Biologia. Um termo ε_i estocástico é

adicionado nesta função, possuindo média zero e variância σ^2 não correlacionadas, com valores de x predeterminados.

Godoi (1983) trabalhou com dados experimentais, provenientes de resultados de lactação em vacas, coletados repetidamente em um mesmo animal. Com estes dados, obteve as estimativas do parâmetro ρ_v (coeficiente de autocorrelação), com o emprego dos desvios de regressão representados por:

$$\hat{e}_{ijk} = y_{ijk} - \hat{y}_{ijk}$$

em que y_{ijk} é o valor da observação do animal i , na repetição j , no tempo k e \hat{y}_{ijk} é uma estimativa dessa mesma observação. Com os desvios de regressão e por meio da fórmula de Pearson, o autor estimou os coeficientes de correlação para cada período, chegando à estimativa final através da média aritmética. A análise dos dados levou-o a concluir que havia uma estrutura de autocorrelação, sendo recomendada então uma metodologia que levasse isto em conta.

Crocci (1984) trabalhou com uma seqüência de tratamentos em períodos sucessivos e apresentou um modelo autorregressivo simplificado, no qual os erros possuem uma distribuição probabilística que incorpora a correlação existente entre os dados de uma mesma observação. Afirmou que a hipótese de independência nos erros das parcelas facilita o desenvolvimento dos estudos. No entanto, para tornar essa afirmação aceitável, é necessário que o modelo inclua todos os parâmetros que possam, de alguma forma, afetar a estrutura de correlação das observações e admita a correlação na forma de um processo autorregressivo de primeira ordem. Adotou o modelo matemático com estrutura

de correlação e comparou com dois outros, sem essa estrutura. Afirmou, ainda, concordando com Godoi (1983), que o modelo autorregressivo é mais sensível em detectar os efeitos de tratamentos pelo teste F. Entretanto, essa sensibilidade só é maior nos casos em que o coeficiente de correlação linear entre os períodos é alto ($\rho_v \geq 0,70$). No caso de a matriz V ser desconhecida, ou seja, ρ_v desconhecido, pode-se adotar o procedimento recomendado por Theil (1971), que consiste na obtenção de uma estimativa $\hat{\rho}_v$, por meio dos desvios de regressão, conforme descrito em Hoffmann e Vieira (1987). Os autores também descreveram a presença da autocorrelação nos erros em medidas repetidas, tomadas em um mesmo indivíduo, apresentando também a solução do problema por transformações matriciais. Consideraram o modelo:

$$Y = X\beta + u$$

em que:

$$u_x = \rho u_{x-1} + \varepsilon_x$$

com $E(u) = 0$ e $E(u'u) = V\sigma^2$, sendo V uma matriz $n \times n$, simétrica e positiva definida. Então, existe também, V^{-1} simétrica e positiva definida e Λ tal que:

$$\Lambda' \Lambda = V^{-1}$$

e

$$\Lambda y = \Lambda X \beta + \Lambda u$$

com

$$\Lambda u = \varepsilon$$

3) Nota: Não é possível fazer a correção (1982)

Esta transformação matricial, convenientemente aplicada aos modelos que se pretende estudar, permite obter as estimativas dos parâmetros de forma não tendenciosa. (Estes autores citam, ainda, que a validade dos estimadores de mínimos quadrados (do vetor de parâmetros), da estatística F na análise de variância, da distribuição t para construção de intervalos de confiança, dentre outros, está vinculada a certas pressuposições sobre o vetor de erros (ε). A primeira delas é que a média dos erros é nula, isto é:

$$E(\varepsilon) = \emptyset$$

A seguir, tem-se que a variância do erro ε_i , $i = 1, 2, \dots, n$, é constante e igual a σ^2 , isto é:

$$E(\varepsilon' \varepsilon) = I\sigma^2$$

Nesse caso, diz-se que os erros são homocedásticos. Em terceiro lugar, o erro de uma observação é não correlacionado com o erro de outra observação, isto é:

$$E(\varepsilon_i \varepsilon_j) = 0, i \neq j$$

A quarta pressuposição é que os erros têm distribuição normal.

A primeira pressuposição exclui, por exemplo, a existência de erros sistemáticos de medida da variável Y. Se a segunda pressuposição (erros homocedásticos) não é atendida, pode-se usar o método dos mínimos quadrados ponderados. Se a terceira não é obedecida, diz-se que há autocorrelação nos resíduos e neste caso pode-se usar modelos generalizados ou então utilizar outros modelos. A quarta pressuposição é necessária para que, dentre outras estatísticas, seja possível usar o teste F da análise de variância e/ou construir intervalos de confiança para os parâmetros.

O método dos mínimos quadrados ordinário aplicado aos modelos não lineares, segue o seguinte procedimento:

seja o conjunto de variáveis (X ; Y) e o modelo: $Y = f(X ; \theta) + \varepsilon$. O objetivo do método é encontrar um vetor $\hat{\theta}$, para o qual a norma quadrática de ε seja mínima. Considerando que $f(X ; \theta)$ é diferenciável com respeito a θ , pode-se escrever:

$$Z = [z_{ij}] = \left[\frac{\partial f(X_i ; \theta)}{\partial \theta_j} \right] ; i = 1, \dots, n, j = 1, \dots, p$$

Sabe-se que uma condição para a existência do mínimo é:

$$\frac{\partial \varepsilon}{\partial \theta_j} \Big|_{\hat{\theta}} = 0 ; j = 1, \dots, p$$

isto é:

$$Z'[Y - f(X; \theta)] = 0$$

que é conhecido como Sistema de Equações Normais (SEN).

Conforme o guia do usuário SAS/ETS® (1993), a simulação diz respeito à determinação das variáveis endógenas ou dependentes, como uma função dos valores de entrada de outras variáveis, mesmo quando os dados reais (observados) para algumas das variáveis resposta estão disponíveis no conjunto de dados de entrada. A simulação é útil na verificação do ajuste dos parâmetros do modelo e pode ser executada para checar quão bem o modelo prediz os valores observados no período histórico, para investigar a sensibilidade da solução para mudanças nos valores de entrada ou dos parâmetros, para examinar as características dinâmicas do modelo, checar a estabilidade de soluções simultâneas, ou ainda para estimar a distribuição estatística dos valores preditos de um modelo não-linear usando métodos de Monte-Carlo.

Souza (1998), tratando da obtenção de estimadores de MQO não-lineares, sugere escrever o modelo de regressão não-linear na forma

$$Y = f(\theta^0) + \varepsilon$$

no qual Y tem componentes y_t , $f(\theta^0)$ tem componentes $f(x_t; \theta^0)$ e ε tem componentes ε_t . Sendo $F(\theta)$ a matriz jacobiana de $f(\theta)$ e $F = F(\theta^0)$. O autor diz que o estimador ($\hat{\theta}$) de mínimos quadrados de θ^0 é obtido mediante a pesquisa do mínimo (em Θ) da soma de quadrados residuais

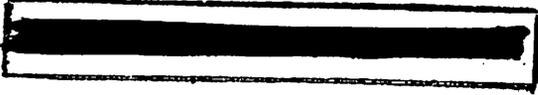
$$\begin{aligned} \text{SSE}(\theta) &= \sum_{t=1}^n (y_t - f(x_t, \theta))^2 = \\ &= (Y - f(\theta))'(Y - f(\theta)) \end{aligned}$$

Diz que, por analogia com o modelo linear, toma-se como estimador de σ^2 o valor obtido pela expressão:

$$\hat{\sigma}^2 = \frac{\text{SSE}(\hat{\theta})}{n - p}$$

Este autor ressalta ainda que outra alternativa seria

$$\hat{\sigma}^2 = \frac{\text{SSE}(\hat{\theta})}{n}$$



Neste caso, com erros normais, o par $(\hat{\theta}, \hat{\sigma}^2)$ define o estimador de máxima verossimilhança de (θ^0, σ^2) . O autor prossegue a discussão tratando da geometria de MQO não-lineares e de suas propriedades distribucionais. Ao descrever o cálculo de MQO não-lineares, ele comenta que existem vários métodos numéricos disponíveis, como a modificação do método de Gauss-Newton, conforme Hartley (1961). Comenta que esta técnica é popular e tem-se mostrado muito eficiente na maioria das aplicações. Tratando de resíduos em série de tempo, o autor expressa o modelo na forma:

$$y_t = f(x_t; \theta^0) + u_t$$

em que a série u_t pode ser aproximada por um processo AR(p) com p suficientemente grande, isto é,

$$u_t \phi_1 u_{t-1} + \dots + \phi_p u_{t-p} = \varepsilon_t \quad t = 0, \pm 1, \pm 2, \dots$$

sendo a seqüência ε_t iid normal com média zero e variância σ^2 . Nestas condições, $\text{Var}(u) = \Gamma = (\gamma_{ij})$, onde $\gamma_{ij} = \gamma(|i - j|)$ sendo $\gamma(h) = \text{Cov}(u_t, u_{t+h})$ a função de autocovariância do processo estocástico ε_t . A matriz Γ pode ser escrita na forma $\sigma^2 \Omega$, onde os elementos de Ω são funções suaves dos parâmetros ϕ_j do AR(1). Comenta que o procedimento estatístico adequado ao modelo não-linear subordinado à estrutura do AR(1) envolve a determinação de um estimador consistente $\hat{\Omega}$ de Ω e a procura de um estimador de θ^0 mais eficiente do que o de

mínimos quadrados ordinário. Diz que neste contexto, procura-se pelo mínimo da soma de quadrados residual “ponderada”

$$(Y - f(\theta))' \hat{\Omega}^{-1}(Y - f(\theta))$$

Ele diz que nenhuma idéia nova está envolvida aqui, sendo o valor de θ que minimiza esta soma de quadrados o estimador de mínimos quadrados generalizados $\hat{\theta}_G$. Diz ainda que este vetor é determinado pela fatoração $\hat{\Omega}^{-1} = P'P$ e da regressão não-linear

$$PY = Pf(\theta) + v$$

em que a matriz P , para o $AR(p)$, é determinada a partir dos resíduos de mínimos quadrados ordinários não-lineares e da solução do sistema de equações de Yule-Walker. Comenta que, após as primeiras p observações, a transformação P induz o modelo

$$(PY)_t = (Pf(\theta))_t + v_t, \text{ sendo:}$$

$$(PY)_t = y_t + \hat{\phi}_1 y_{t-1} + \hat{\phi}_2 y_{t-2} + \dots + \hat{\phi}_p y_{t-p}$$

e

$$(Pf(\theta))_i = f(x_i, \theta) + \hat{\phi}_1 f(x_{i-1}, \theta) + \hat{\phi}_2 f(x_{i-2}, \theta) + \dots + \hat{\phi}_p f(x_{i-p}, \theta)$$

Comenta ainda que com o SAS, o cálculo é levado a efeito com a construção de um conjunto de dados adequado e que a abordagem aparece em Gallant (1987, p.134).

3 MATERIAL E MÉTODOS

O presente trabalho foi desenvolvido em duas etapas, sendo que a primeira delas envolveu um estudo por meio de simulação. Na segunda etapa é apresentado um exemplo de aplicação, para o qual foram utilizados dados reais obtidos da literatura.

3.1 Simulação

Para levar a efeito a simulação, adotou-se o modelo logístico monofásico:

$$y_t = \frac{a}{1 + e^{(-b(t-d))}} + u_t$$

e o modelo logístico difásico:

$$y_t = \frac{a_1}{1 + e^{(-b_1(t-d_1))}} + \frac{a_2}{1 + e^{(-b_2(t-d_2))}} + u_t$$

em que u_t tem estrutura AR(1), $\rho u_{t-1} + \varepsilon_t$, sendo ρ o coeficiente de autocorrelação e ε_t o ruído branco.

Ambos os modelos foram analisados em duas situações quanto à estrutura dos erros: independentes e autorregressivos AR(1). Foram simuladas 2000 amostras correspondentes a um intervalo de tempo de 0 a 60 meses, o qual foi dividido em 10, 50, 100 e 1000 partes. Os parâmetros a serem utilizados nestes modelos foram tomados, tendo por referência, aqueles obtidos através do exemplo de aplicação com dados da literatura e estão apresentados na Tabela 2.

TABELA 2. Parâmetros usados na simulação com os modelos logísticos monofásico e difásico com estrutura de erros independentes (EI) e autorregressivos de primeira ordem AR(1).

Parâmetros	EI	AR(1)
1. Modelo monofásico		
a	370,00	370,00
b	0,10	0,11
c	15,00	14,00
ρ	-	0,80(0,30)
2. Modelo difásico		
a_1	133,00	133,00
b_1	0,60	0,63
d_1	4,50	4,43
a_2	232,00	232,00
b_2	0,17	0,17
d_2	22,00	22,00
ρ		0,80(0,30)

Dado que a variação de y (variável dependente) explicada pelo modelo foi pré fixada em função de R^2 , considerou-se o seguinte procedimento partindo-se do modelo:

$$y_t = f(a, b, d) + u_t,$$

no qual $f(a, b, d)$ é a parte determinística e u_t é o erro com estrutura AR(1).

O coeficiente de determinação, neste caso, é dado por:

$$R^2 = \frac{\sigma_M^2 + (\sigma_u^2 - \sigma_\varepsilon^2)}{\sigma_{\text{Total}}^2},$$

em que $(\sigma_u^2 - \sigma_\varepsilon^2)$ corresponde à variação devido à estrutura de autocorrelação nos erros. Isolando σ_ε^2 e substituindo σ_u^2 por sua expressão, conforme se encontra em Hoffmann e Vieira (1987),

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2},$$

determina-se a expressão da variância do erro:

$$\sigma_\varepsilon^2 = \frac{(1 - R^2)(1 - \rho^2)}{R^2 - \rho^2} \sigma_M^2, \quad R^2 > \rho^2,$$

na qual a variância do modelo (σ_M^2) é obtida através do procedimento UNIVARIATE do SAS, levando-se em consideração apenas a parte determinística do modelo.

Desta forma, foi possível determinar os demais parâmetros necessários à simulação, fixando-se R^2 em 0,95 e 0,75 e ρ (parâmetro de autocorrelação) em 0,80 e 0,30. Estes parâmetros são apresentados nas Tabelas 3, 4, 5 e 6. As Tabelas 3 e 5 apresentam os parâmetros utilizados no modelo logístico monofásico e difásico, respectivamente, com estrutura de erros independentes.

Nas Tabelas 4 e 6 se encontram os parâmetros utilizados nos modelos monofásico e difásico, com estrutura de autocorrelação de primeira ordem nos erros, AR(1).

TABELA 3. Valores de n, sementes, R^2 , σ_M^2 e σ_ε^2 usados na simulação do modelo logístico monofásico com estrutura de erros EI.

Tamanho da amostra (n)	Semente	Parâmetros		
		R^2	σ_M^2	σ_ε^2
10	1437583	0,95	12150,8700	639,5195
		0,75		4050,2900
50	4538579	0,95	9806,2800	516,1200
		0,75		2157,7600
100	3758725	0,95	9549,1750	502,5882
		0,75		3183,0583
1000	6983526	0,95	9323,7600	490,7242
		0,75		3107,9200

TABELA 4. Valores de n , sementes, R^2 , ρ , σ_M^2 e σ_ε^2 usados na simulação do modelo logístico monofásico com estrutura de erros AR(1).

Tamanho da amostra (n)		Parâmetros			
amostra (n)	Semente	R^2	ρ	σ_M^2	σ_ε^2
10	1437583	0,95	0,80	12150,8700	705,5344
		0,75	0,80		9941,6209
		0,95	0,30		642,8658
		0,75	0,30		4188,3681
50	4538579	0,95	0,80	9806,2800	569,3969
		0,75	0,80		8023,3200
		0,95	0,30		518,8206
		0,75	0,30		3380,1950
100	3758725	0,95	0,80	9549,1750	554,4682
		0,75	0,80		7812,9614
		0,95	0,30		505,2180
		0,75	0,30		3291,5717
1000	6983526	0,95	0,80	9323,7600	541,3796
		0,75	0,80		7628,5309
		0,95	0,30		493,2920
		0,75	0,30		3213,8718

TABELA 5. Valores de n , sementes, R^2 , σ_M^2 e σ_ε^2 usados na simulação do modelo logístico difásico com estrutura de erros EI.

Tamanho da amostra (n)		Parâmetros		
	Semente	R^2	σ_M^2	σ_ε^2
10	1437583	0,95	12102,9200	794,8905
		0,75		5034,3067
50	4538579	0,95	11730,7700	617,4089
		0,75		3910,2567
100	3758725	0,95	11345,0300	597,1067
		0,75		3781,6767
1000	6983526	0,95	11005,8700	579,2563
		0,75		3668,6233

TABELA 6. Valores de n, sementes, R^2 , ρ , σ_M^2 e σ_ε^2 usados na simulação do modelo logístico difásico com estrutura de erros AR(1).

Tamanho da amostra (n)	Semente	Parâmetros			
		R^2	ρ	σ_M^2	σ_ε^2
10	1437583	0,95	0,80	15102,9200	876,9437
		0,75	0,80		12356,9346
		0,95	0,30		799,0498
		0,75	0,30		5205,9308
50	4538579	0,95	0,80	11730,7700	681,1415
		0,75	0,80		9597,9027
		0,95	0,30		620,6369
		0,75	0,30		4043,5609
100	3758725	0,95	0,80	11345,0300	658,7437
		0,75	0,80		9282,2973
		0,95	0,30		600,2312
		0,75	0,30		3910,5975
1000	6983526	0,95	0,80	11005,8700	639,0505
		0,75	0,80		9004,8027
		0,95	0,30		582,2873
		0,75	0,30		3793,6900

A simulação foi conduzida através da utilização do Programa apresentado no Anexo 1A. Este programa, em linguagem SAS, foi elaborado tendo por

referência Souza (1998, p. 381) e o guia do usuário SAS/ETS® (1993, cap. 11). Na primeira parte do programa foram gerados os dados referentes à distribuição logística. Foram fornecidos os valores de σ_ε , n , semente, ρ e dos parâmetros do modelo conforme constam na Tabela 2, além do valor do incremento a ser dado à variável x a fim de se obter o número desejado de observações. Este incremento foi obtido dividindo-se 60 por $n-1$. Na segunda parte foram obtidos os estimadores dos parâmetros, através do procedimento MODEL, que solicita que se forneçam de novo os valores dos parâmetros do modelo. A próxima parte do programa consistiu na eliminação das iterações que não convergiram ou com parâmetros sob restrição. A seguir foram determinados o viés e o erro quadrático médio, por meio do procedimento IML. Nesta parte foram fornecidos os valores paramétricos, ρ e σ_ε^2 , além de se ter feito a substituição de N por seu valor atual, que foi obtido subtraindo-se de 2000 o número de iterações que foram eliminadas.

3.2 Exemplo de Aplicação

Os dados utilizados para exemplificar o ajuste da função logística encontram-se descritos em Kroll (1990), sendo que foram tomados aqueles referentes a cinco dos seis animais, os quais estão apresentados no Anexo B. Foram considerados os modelos logísticos monofásico e difásico, contendo estrutura de erros independentes (EI) e autocorrelacionados, com autocorrelação de primeira ordem AR(1) e de segunda ordem AR(2). A estrutura de autocorrelação nos erros, no modelo AR(2) é da forma:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t .$$

Os valores iniciais requeridos para reduzir o número de iterações necessárias para a solução do sistema de equações não lineares foram obtidos a partir dos gráficos referentes às pesagens dos cinco animais e pelo método dos quadrados mínimos ordinário para modelos de regressão não lineares com erros independentes, através do Programa apresentado no Anexo 2A, o qual faz uso do procedimento NLIN do SAS. Caso a convergência não fosse alcançada, alternava-se sucessivamente o método de Gauss pelo método DUD ou Marquardt, conforme opções do SAS. O processamento dos dados para o desenvolvimento do método dos quadrados mínimos generalizado, em modelos de regressão não linear com erros autocorrelacionados, foi feito através do Programa apresentado no Anexo 3A, o qual faz uso do procedimento MODEL do SAS.

A qualidade dos ajustamentos foi verificada por meio da comparação entre as curvas ajustadas, pelos valores dos coeficientes de determinação ajustados (R^2_{aj}), pela análise gráfica dos resíduos e pelos valores obtidos para os parâmetros em termos de sua interpretabilidade biológica.

4 RESULTADOS E DISCUSSÃO

As Tabelas 7 e 8 mostram os resultados obtidos quanto ao número de iterações utilizadas, às estimativas médias dos parâmetros, às raízes quadradas dos erros quadráticos médios e aos valores mínimos e máximos assumidos pelos estimadores dos parâmetros da função logística monofásica, com estrutura de erros independentes (EI) e valores do coeficiente de determinação (R^2) iguais a 0,95 e 0,75 respectivamente. O mesmo ocorre com as Tabelas 9 e 10, sendo que estas se referem à função logística difásica. Nas Tabelas 11 a 14 estão os resultados obtidos para o modelo logístico monofásico com estrutura de erros autorregressivos de ordem 1 (AR(1)), alterando-se em cada uma os valores de R^2 e do coeficiente de autocorrelação (ρ). Na Tabela 11 tem-se R^2 de 0,95 e ρ de 0,80. Nas Tabelas 12, 13 e 14 têm-se os pares de valores para R^2 e ρ como se segue: 0,75 e 0,80; 0,95 e 0,30; 0,75 e 0,30. O mesmo ocorre nas Tabelas 15 a 18, sendo que estas se referem aos resultados para o modelo logístico difásico.

No modelo monofásico com estrutura de erros independentes (Tabelas 7 e 8), praticamente não houve problemas com iterações e os valores médios dos estimadores dos parâmetros ficaram bem próximos daqueles estipulados para os mesmos, excetuando-se a situação em que $R^2 = 0,75$ e $n = 10$, na qual 93,6% das iterações puderam ser utilizadas no cálculo dos estimadores, embora ainda assim a média dos estimadores não tenha se distanciado muito dos seus respectivos valores paramétricos, o que indica um viés pequeno. Observou-se também um aumento na precisão dos estimadores, medida pela raiz quadrada do erro quadrático médio e pelos valores mínimos e máximos (L_1 e L_2), à medida que se aumentou n , sendo que esta observação se aplicou a todas as situações descritas nas demais tabelas. Já no modelo difásico com estrutura de erros independentes, Tabelas 9 e 10, ocorreu o inverso, ou seja, houve problemas em todas as

situações, exceto quando se teve valores elevados de R^2 e n , ou seja, 0,95 e 1000 respectivamente. O viés foi maior do que no modelo monofásico, principalmente para menores valores de n . Observou-se que houve redução na precisão dos estimadores, em relação ao modelo monofásico, o que pode ser devido ao maior número de parâmetros estimados. Para a situação em que $n=10$ e $R^2 = 0,75$, chegou-se a utilizar apenas 35,8% das iterações.

A introdução da estrutura de autocorrelação nos erros elevou sensivelmente o número de iterações eliminadas em ambos os casos, monofásico e difásico, conforme se observa pelas Tabelas 11 a 18. No modelo monofásico, em que $R^2 = 0,95$ e $\rho = 0,80$ (Tabela 11), houve apenas quatro iterações que foram eliminadas do processo, quando $n = 10$, sendo utilizadas nos demais casos todas as 2000 iterações. Os valores médios dos estimadores não se distanciaram muito daqueles estipulados, o que indica um pequeno viés, sendo que para o parâmetro ρ ele foi mais elevado. Observou-se, pela Tabela 12, que quando ρ foi mantido em 0,80 e R^2 reduzido para 0,75, os valores médios dos estimadores se distanciaram mais dos paramétricos, ou seja, houve um aumento no viés. Para n igual a 10 chegou-se a eliminar aproximadamente 25% das iterações. Comparando-se as Tabelas 11 e 13 com a Tabela 7 pode-se observar o efeito de ρ no modelo monofásico com R^2 igual a 0,95. Observou-se que, para baixos valores de ρ (0,30), a precisão dos estimadores não diferiu muito daquela obtida para o modelo com erros EI, sendo que houve aumento do viés para valores elevados de ρ (0,80). Para ρ igual a 0,30, observou-se que não houve problemas com iterações. Comparando as Tabelas 12 e 14 com a Tabela 8, nas quais se tem o modelo monofásico com R^2 igual a 0,75, verificou-se um comportamento similar ao observado quando se tinha R^2 igual a 0,95, exceto pelos valores mais elevados da raiz quadrada do erro quadrático médio e da maior amplitude do intervalo entre os valores máximo e mínimo (L_1 e L_2), indicando uma redução na precisão

do estimadores quando se reduz R^2 . Observou-se que, além do estimador do parâmetro d , o de b , para $n = 10$, também foi muito afetado neste caso.

Pela análise comparativa entre as Tabelas 9, 15 e 17, as quais apresentam os resultados obtidos para o modelo logístico difásico com R^2 igual a 0,95, observou-se um comportamento semelhante ao do modelo monofásico com o mesmo valor de R^2 . Notou-se, entretanto, que para o tamanho de amostra (n) igual a 10, o valor de N foi maior quando $\rho = 0,30$, o que significa menor número de iterações eliminadas. Nos demais casos, observou-se que o valor de N foi mais elevado para o modelo com erros EI. A convergência aos valores paramétricos, tanto no modelo monofásico quanto no difásico, foi mais rápida e com maior precisão no modelo com estrutura de erros EI, decrescendo no sentido daquele em que o valor de ρ era mais elevado.

Uma comparação entre os resultados das Tabelas 10, 16 e 18, nas quais se tem o modelo difásico com R^2 igual a 0,75, mostra um comportamento dos estimadores semelhante ao observado anteriormente, ou seja, no modelo difásico com R^2 igual a 0,95. A diferença está nos menores valores de N e na menor precisão dos estimadores, além de uma convergência mais lenta aos valores paramétricos. Novamente, observou-se que o número de iterações utilizadas (N) decresceu com o aumento de ρ , exceto para $n=10$, em que o valor de N para ρ igual a 0,30 foi superior ao do modelo com erros EI. Em geral, o valor de N foi maior para maiores valores de n , o que não se verificou para ρ igual a 0,80 (Tabela 16), em que foram obtidos valores muito baixos de N para tamanhos de amostra n iguais a 50 e 100.

De maneira geral, pode-se observar que valores de ρ mais elevados e menores valores de R^2 resultaram numa convergência mais lenta aos valores paramétricos, em uma menor precisão e em menores valores de N adotados. O modelo difásico apresentou maiores problemas com as iterações, ou seja, menores

valores de N do que o modelo monofásico, o que era de se esperar, conforme Koops e Grossman (1991).

Os estimadores tenderam assintoticamente aos valores paramétricos em todas as situações analisadas, o que indica serem eles consistentes. Observou-se também que os valores médios dos estimadores ainda tendiam aos paramétricos e com maior precisão, para valores de n superiores a 100.

Notou-se, pela observação dos valores médios dos estimadores e dos valores máximos e mínimos, que havia uma ligeira assimetria à direita para pequenos valores de n , que tendia à simetria à medida em que se aumentava n .

TABELA 7. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio (\sqrt{EQM}) e valores mínimo e máximo (L_1 L_2), obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros EI em amostras de 10, 50, 100 e 1000 observações ($R^2=0,95$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L_1	L_2
a (370,00)	2000	372,20	21,06	312,83	583,61
		370,51	8,04	345,04	422,93
		370,22	5,42	347,29	397,90
		370,01	1,74	364,24	376,06
b (0,10)	2000	0,10	0,02	0,03	0,17
		0,10	0,01	0,07	0,13
		0,10	0,01	0,08	0,12
		0,10	0,002	0,09	0,11
d (15,00)	2000	21,83	7,17	15,37	41,84
		16,25	1,48	13,76	19,88
		15,64	0,84	13,30	17,31
		15,06	0,18	14,52	15,60
Resíduo	2000	643,79	350,99	51,36	2401,16
		516,69	105,18	247,91	981,44
		503,26	70,71	299,21	766,18
		490,20	21,78	420,06	577,66

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 8. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros EI em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a (370,00)	1872	381,930	60,02	256,49	713,28
	2000	373,55	23,25	314,02	627,39
	2000	371,48	14,41	322,79	469,55
	2000	370,13	4,39	355,43	385,73
b (0,10)	1872	0,14	0,20	0,02	5,39
	2000	0,10	0,02	0,04	0,20
	2000	0,10	0,01	0,06	0,16
	2000	0,10	0,004	0,09	0,12
d (15,00)	1872	22,56	9,74	1,39	49,76
	2000	16,47	2,67	9,68	35,71
	2000	15,76	1,60	9,56	21,39
	2000	15,08	0,44	13,70	16,44
Resíduo	1872	3971,98	2180,41	345,81	15244,95
	2000	3270,74	666,34	1553,42	6214,36
	2000	3186,80	447,68	1894,76	4845,05
	2000	3105,36	137,93	2661,13	3658,82

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 9. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difásico com estrutura de erros EI em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,95$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a ₁ (133,00)	1327	177,19	68,95	20,00	358,40
	1755	143,24	41,92	44,41	303,31
	1922	134,84	31,90	38,88	258,14
	2000	132,32	10,54	89,95	175,79
b ₁ (0,60)	1327	1,70	1,89	0,03	9,89
	1755	0,85	0,81	0,13	9,59
	1922	0,86	0,88	0,20	9,76
	2000	0,61	0,09	0,36	1,17
d1 (4,50)	1327	14,42	12,83	4,21	59,96
	1755	6,10	2,01	2,81	13,24
	1922	5,30	1,09	2,64	9,78
	2000	4,57	0,21	3,84	5,65
a2 (232,00)	1327	190,08	70,82	36,16	457,83
	1755	221,60	44,69	61,32	341,39
	1922	230,18	34,34	104,85	337,71
	2000	232,73	11,41	186,73	277,95
b2 (0,17)	1327	0,91	1,49	0,04	9,94
	1755	0,21	0,21	0,09	5,96
	1922	0,18	0,04	0,10	0,50
	2000	0,17	0,01	0,14	0,23
d2 (22,00)	1327	30,48	10,90	7,39	59,68
	1755	23,88	3,10	17,05	36,24
	1922	22,76	2,07	16,53	29,20
	2000	22,02	0,64	19,71	24,57
Resíduo	1327	692,90	493,03	12,72	4536,21
	1755	614,04	129,94	283,97	1224,91
	1922	596,16	85,61	361,75	907,86
	2000	578,84	25,71	495,86	681,29

- (1) os valores entre parênteses são os estipulados na simulação;
- (2) n^o de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;
- (3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 10. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difásico com estrutura de erros EI em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a ₁ (133,00)	716	184,34	77,28	52,99	576,80
	981	175,92	65,82	24,99	352,78
	1412	164,70	59,59	32,39	353,39
	1971	132,11	26,64	43,50	279,58
b ₁ (0,60)	716	2,33	2,49	0,04	9,99
	981	0,94	1,01	0,04	6,80
	1412	1,04	1,36	0,06	9,5
	1971	0,77	0,68	0,19	9,03
d1 (4,50)	716	13,90	14,02	1,44	59,96
	981	6,91	3,55	0,49	17,95
	1412	6,09	2,71	0,39	24,97
	1971	4,67	0,58	2,93	10,21
a2 (232,00)	716	203,66	71,53	72,76	538,79
	981	192,99	67,80	42,55	602,73
	1412	200,26	62,10	30,76	356,11
	1971	233,02	28,77	84,97	331,85
b2 (0,17)	716	1,90	2,40	0,04	9,44
	981	0,44	0,66	0,03	7,13
	1412	0,37	0,68	0,04	9,96
	1971	0,18	0,03	0,11	0,49
d2 (22,00)	716	33,06	14,53	7,30	59,85
	981	26,56	6,77	14,82	59,80
	1412	24,70	4,81	15,35	56,56
	1971	22,03	1,62	16,71	31,87
Resíduo	716	4175,81	2962,53	80,78	18029,65
	981	3859,78	837,86	1737,56	7787,63
	1412	3759,14	529,84	2322,22	5738,19
	1971	3664,81	162,78	3140,14	4309,24

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 11. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio (\sqrt{EQM}) e valores mínimo e máximo (L_1 e L_2), obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,95$ $\rho = 0,80$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L_1	L_2
a (370,00)	1996	376,98	49,60	202,43	639,59
	2000	379,48	73,23	283,23	2703,26
	2000	374,72	30,49	301,74	790,10
	2000	370,56	9,16	339,09	406,77
b (0,11)	1996	0,11	0,03	0,03	0,36
	2000	0,11	0,04	0,03	0,61
	2000	0,10	0,03	0,03	0,24
	2000	0,10	0,01	0,07	0,13
d (14,00)	1996	22,29	8,82	7,57	49,65
	2000	16,95	5,74	2,98	114,66
	2000	16,08	3,31	5,79	54,11
	2000	15,12	0,92	12,23	18,09
ρ	1996	-0,15	1,01	-2,23	1,43
	2000	0,63	0,22	-0,04	0,94
	2000	0,72	0,11	0,41	0,92
	2000	0,79	0,02	0,72	0,85
Resíduo	1996	465,73	3597,71	12,07	1940,39
	2000	551,79	118,25	233,14	1071,12
	2000	550,27	78,04	319,15	833,95
	2000	540,93	24,09	463,44	638,25

(1) os valores entre parênteses são os estipulados na simulação;

(2) n^o de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 12. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio (\sqrt{EQM}) e valores mínimo e máximo (L_1 e L_2), obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$ $\rho = 0,80$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L_1	L_2
a (370,00)	1505	417,69	144,40	151,97	992,14
	1511	404,27	121,81	132,45	966,52
	1616	396,64	111,04	177,33	986,34
	1986	380,88	50,83	277,36	831,31
b (0,11)	1505	0,54	1,04	0,02	7,01
	1511	0,56	0,99	0,02	7,99
	1616	0,64	1,46	0,02	9,78
	1986	0,11	0,06	0,03	1,64
d (14,00)	1505	22,99	13,34	0,42	49,75
	1511	18,48	11,48	0,21	59,61
	1616	17,87	10,43	0,04	59,96
	1986	16,07	4,87	0,95	56,33
ρ	1505	-0,01	0,92	-1,27	1,32
	1511	0,03	0,22	0,19	0,94
	1616	0,71	0,12	0,39	0,92
	1986	0,79	0,02	0,72	0,85
Resíduo	1505	6693,83	5451,30	274,09	3178,25
	1511	7586,57	1670,65	3282,68	15310,22
	1616	7666,39	1117,68	4503,63	11544,64
	1986	7619,41	339,64	6530,02	8939,53

(1) os valores entre parênteses são os estipulados na simulação;

(2) n^o de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 13. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio (\sqrt{EQM}) e valores mínimo e máximo (L_1 e L_2), obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,95$ $\rho = 0,30$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L_1	L_2
a (370,00)	2000	373,76	29,60	272,83	827,34
	2000	370,98	11,49	337,41	453,73
	2000	370,44	7,75	339,59	411,98
	2000	370,03	2,49	361,74	378,68
b (0,11)	2000	0,10	0,03	0,03	0,27
	2000	0,10	0,01	0,06	0,14
	2000	0,10	0,01	0,07	0,13
	2000	0,10	0,002	0,09	0,11
d (14,00)	2000	21,96	7,65	12,86	67,10
	2000	16,28	1,72	12,69	22,21
	2000	15,66	1,02	12,26	18,02
	2000	15,07	0,25	14,29	15,83
ρ	2000	-0,27	0,64	-1,24	1,43
	2000	0,20	0,17	-0,32	0,62
	2000	0,25	0,11	-0,06	0,57
	2000	0,30	0,03	0,19	0,40
Resíduo	2000	556,11	355,51	26,67	2298,46
	2000	517,00	107,20	228,84	969,28
	2000	505,31	71,40	301,86	769,64
	2000	492,82	21,93	422,25	579,87

- (1) os valores entre parênteses são os estipulados na simulação;
 (2) n^o de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;
 (3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 14. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio (\sqrt{EQM}) e valores mínimo e máximo (L_1 e L_2), obtidos através de 2000 simulações do modelo logístico monofásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$ $\rho = 0,30$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L_1	L_2
a (370,00)	1852	381,57	69,49	236,56	780,01
	2000	378,88	47,04	298,32	1304,46
	2000	373,31	20,48	310,18	559,06
	2000	370,28	6,42	348,71	393,96
b (0,11)	1852	0,22	0,51	0,02	6,29
	2000	0,11	0,03	0,03	0,44
	2000	0,10	0,02	0,04	0,21
	2000	0,10	0,01	0,08	0,12
d (14,00)	1852	22,42	10,38	0,74	49,22
	2000	16,91	4,63	5,57	73,59
	2000	15,92	2,35	6,93	28,37
	2000	15,09	0,64	13,08	17,08
ρ	1852	-0,25	0,64	-1,30	1,39
	2000	0,20	0,17	-0,34	0,62
	2000	0,25	0,11	-0,06	0,54
	2000	0,30	0,03	0,19	0,40
Resíduo	1852	3512,74	2286,27	212,73	13183,87
	2000	3365,33	698,69	1492,41	6314,58
	2000	3290,46	464,87	1965,58	5000,17
	2000	3210,73	142,87	2751,08	3776,74

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 15. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,95$ $\rho = 0,80$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a1 (133,00)	1277	184,87	85,87	17,64	432,23
	1548	170,98	69,53	32,84	377,81
	1604	168,11	64,67	38,04	352,86
	1762	144,93	44,36	33,07	316,84
b1 (0,63)	1277	1,42	1,75	0,06	9,86
	1548	0,97	0,96	0,06	8,11
	1604	1,33	1,63	0,08	9,04
	1762	1,02	1,32	0,15	9,76
d1 (4,43)	1277	14,97	14,08	5,30	59,92
	1548	6,88	4,80	0,36	58,04
	1604	6,2	3,65	0,11	46,44
	1762	5,00	1,41	0,88	11,75
a2 (232,00)	1277	193,8	80,42	41,53	526,12
	1548	200,88	72,86	44,85	485,33
	1604	200,63	68,67	52,46	421,88
	1762	218,67	47,73	42,86	339,87
b2 (0,17)	1277	0,77	1,30	0,03	9,58
	1548	0,54	0,77	0,04	7,37
	1604	0,61	1,10	0,04	9,81
	1762	0,23	0,32	0,09	9,69
d2 (22,00)	1277	29,83	11,53	6,95	59,96
	1548	26,72	8,55	3,09	59,27
	1604	25,69	6,74	7,99	55,4
	1762	22,83	2,80	15,56	33,01
ρ	1277	-0,37	1,31	-2,90	1,52
	1548	0,51	0,33	-0,03	0,92
	1604	0,65	0,17	0,26	0,89
	1762	0,79	0,02	0,72	0,85

...continua...

TABELA 15, Cont.

Parâmetros ⁽¹⁾	N ⁽²⁾	Média ⁽³⁾	Estatísticas		
			\sqrt{EQM}	L ₁	L ₂
Resíduo	1277	516,63	609,61	3,17	3425,85
	1548	606,98	157,22	263,18	1293,32
	1604	631,02	96,71	390,04	968,16
	1762	638,01	28,37	547,68	750,87

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 16. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difusico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$ $\rho = 0,80$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Média ⁽³⁾	Estatísticas		
			\sqrt{EQM}	L ₁	L ₂
a1	426	227,03	137,51	40,31	580,55
	114	226,82	117,35	79,34	470,57
	133	226,59	118,16	83,10	590,90
	438	189,93	75,55	88,44	519,28
b1	426	2,22	2,43	0,04	9,84
	114	1,76	1,80	0,04	7,18
	133	3,12	3,49	0,04	9,77
	438	1,09	1,33	0,05	8,79
d1	426	15,73	16,34	1,47	59,96
	114	11,56	14,67	0,29	59,56
	133	9,5	12,63	0,23	54,34
	438	6,42	6,03	0,01	59,91

...continua...

TABELA 16, Cont.

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a2 (232,00)	426	258,68	108,71	45,94	796,77
	114	267,88	112,47	68,35	877,54
	133	246,62	100,62	81,45	682,17
	438	191,62	71,49	63,81	428,76
b2 (0,17)	426	2,11	2,57	0,01	9,52
	114	1,24	1,68	0,03	7,35
	133	2,08	2,92	0,03	9,19
	438	0,72	1,33	0,02	9,70
d2 (22,00)	426	30,17	15,75	2,12	59,92
	114	35,57	19,38	2,82	59,7
	133	33,56	17,76	1,05	59,77
	438	27,61	9,32	7,12	59,16
ρ	426	-0,23	1,20	-3,15	1,69
	114	0,48	0,35	0,07	0,82
	133	0,63	0,20	0,35	0,86
	438	0,78	0,03	0,72	0,85
Resíduo	426	8095,52	8271,80	105,01	46195,00
	114	8295,79	2252,68	5133,36	13363,89
	133	8635,83	1373,61	5575,78	11004,60
	438	8955,53	381,25	7718,05	10014,18

- (1) os valores entre parênteses são os estipulados na simulação;
- (2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;
- (3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 17. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,95$ $\rho = 0,30$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a1 (133,00)	1442	179,56	75,98	16,55	364,92
	1719	150,73	51,60	33,48	348,90
	1853	141,06	41,38	41,42	299,67
	2000	131,82	15,51	63,85	196,32
b1 (0,63)	1442	1,64	1,95	0,03	9,92
	1719	0,89	0,84	0,08	5,81
	1853	1,00	1,18	0,14	9,76
	2000	0,64	0,17	0,31	3,49
d1 (4,43)	1442	14,70	13,52	2,14	59,99
	1719	6,33	2,70	1,78	50,53
	1853	5,49	1,52	2,29	12,78
	2000	4,59	0,30	3,55	6,45
a2 (232,00)	1442	189,39	76,63	33,72	522,49
	1719	214,29	54,50	30,22	366,06
	1853	223,71	44,26	61,37	344,8
	2000	233,27	16,79	165,39	303,86
b2 (0,17)	1442	0,83	1,38	0,04	9,63
	1719	0,26	0,29	0,07	3,93
	1853	0,20	0,14	0,10	3,21
	2000	0,17	0,02	0,12	0,27
d2 (22,00)	1442	30,48	11,40	6,95	59,70
	1719	24,39	4,07	14,69	58,37
	1853	23,13	2,70	15,53	32,38
	2000	22,00	0,94	18,06	25,74
ρ	1442	-0,53	0,96	-3,89	1,47
	1719	0,11	0,24	-0,48	0,52
	1853	0,21	0,14	-0,12	0,50
	2000	0,29	0,03	0,18	0,40

...continua...

TABELA 17, Cont.

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
Resíduo	1442	543,95	542,67	6,74	3739,61
	1719	597,94	133,35	289,89	1186,64
	1853	593,34	86,62	355,72	918,52
	2000	581,65	25,88	498,46	682,85

(1) os valores entre parênteses são os estipulados na simulação;

(2) n° de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;

(3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

TABELA 18. Estimativas médias dos parâmetros com a raiz quadrada do erro quadrático médio e valores mínimo e máximo, obtidos através de 2000 simulações do modelo logístico difásico com estrutura de erros AR(1) em amostras de 10, 50, 100 e 1000 observações ($R^2 = 0,75$ $\rho = 0,30$).

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a1 (133,00)	816	188,76	85,38	13,80	485,41
	833	181,63	72,85	48,48	401,56
	1128	176,57	65,90	41,04	402,34
	1881	136,60	35,12	49,36	295,12
b1 (0,63)	816	2,20	2,38	0,03	9,57
	833	1,16	1,27	0,03	8,75
	1128	1,27	1,81	0,03	9,97
	1881	0,89	1,01	0,17	9,92
d1 (4,43)	816	14,02	14,01	0,16	59,87
	833	7,28	5,21	0,07	49,37
	1128	6,48	4,21	0,01	56,05
	1881	4,80	0,94	2,37	11,00

...continua...

TABELA 18, Cont.

Parâmetros ⁽¹⁾	N ⁽²⁾	Estatísticas			
		Média ⁽³⁾	\sqrt{EQM}	L ₁	L ₂
a2 (232,00)	816	204,83	80,19	51,55	573,14
	833	193,88	69,23	46,50	453,45
	1128	190,70	67,22	46,80	413,76
	1881	228,34	37,93	69,87	322,33
b2 (0,17)	816	2,02	2,46	0,04	9,82
	833	0,63	0,97	0,03	6,93
	1128	0,57	1,03	0,04	9,97
	1881	0,19	0,16	0,10	5,72
d2 (22,00)	816	32,79	14,84	6,77	59,95
	833	27,46	8,70	5,68	59,94
	1128	25,59	5,93	5,53	56,93
	1881	22,31	2,16	16,55	33,80
ρ	816	-0,43	0,87	-3,14	1,40
	833	0,10	0,25	-0,38	0,53
	1128	0,19	0,15	-0,13	0,49
	1881	0,29	0,03	0,18	0,40
Resíduo	816	3594,78	3543,68	63,53	28266,87
	833	3804,81	861,29	1755,42	7784,74
	1128	3848,07	562,24	2405,30	5943,63
	1881	3789,54	168,64	3247,70	4442,94

- (1) os valores entre parênteses são os estipulados na simulação;
- (2) n^o de simulações em que se obtiveram convergência e estimativas em regiões aceitáveis;
- (3) valores de n iguais a 10, 50, 100 e 1000, em cada situação.

5 EXEMPLO DE APLICAÇÃO

Como exemplo de aplicação, foram feitas análises para as quais se fez uso dos programas apresentados nos Anexos 2A e 3A. Os dados foram obtidos em Kroll (1990) e correspondem a pesagens mensais feitas em cinco vacas, durante um período de sessenta meses.

Os resultados obtidos estão contidos nas Tabelas 19, 20 e 21, nas quais constam os valores das estimativas das assíntotas, dos parâmetros de crescimento dos pontos de inflexão e dos coeficientes de determinação ajustados, para os modelos monofásico e difásico. Na Tabela 19 têm-se os modelos com estrutura de erros EI e nas Tabelas 20 e 21 os modelos com estrutura de erros AR(1) e AR(2), respectivamente. As Figuras 1 a 4 mostram as curvas ajustadas e os gráficos de resíduos, para a vaca 1. A Figura 1 mostra o ajuste da função logística monofásica com estrutura de erros independentes (EI), AR(1) e AR(2). Na Figura 2 é mostrado o ajuste da função logística difásica com estrutura de erros EI, AR(1) e AR(2). Na Figura 3 é apresentado o gráfico de resíduos do modelo monofásico com erros EI, AR(1) e AR(2). A Figura 4 mostra o gráfico de resíduos para o modelo difásico com erros EI, AR(1) e AR(2).

TABELA 34. Estimativas da assíntota (a, a_1, a_2)⁽¹⁾, do parâmetro de crescimento (b, b_1, b_2)⁽²⁾, do tempo de inflexão (d, d_1, d_2)⁽³⁾ e R^2_{aj} , dos modelos logístico monofásico e difásico, com estrutura de erros independentes. Dados ponderais de 60 observações em vacas.

Vaca	Monofásico				Difásico						
	a	b	d	R^2_{aj}	a_1	b_1	d_1	a_2	b_2	d_2	R^2_{aj}
1	369,2	0,12	14,6	0,96	133,1	0,63	4,4	232,0	0,17	21,8	0,97
2	493,3	0,10	17,5	0,97	121,3	0,82	3,8	362,7	0,13	22,9	0,98
3	384,2	0,10	15,4	0,96	77,6	0,89	5,1	309,8	0,11	19,7	0,97
4	383,6	0,11	13,7	0,94	124,8	0,82	4,4	258,3	0,13	20,5	0,96
5	448,4	0,09	16,9	0,97	72,4	3,44	4,2	377,5	0,09	20,9	0,98

(1) kg (2) mês⁻¹ (3) mês

TABELA 35. Estimativas da assíntota (a, a_1, a_2)⁽¹⁾, do parâmetro de crescimento (b, b_1, b_2)⁽²⁾, do tempo de inflexão (d, d_1, d_2)⁽³⁾ e R^2_{aj} , dos modelos logístico monofásico e difásico, com estrutura de erros AR(1). Dados ponderais de 60 observações em vacas.

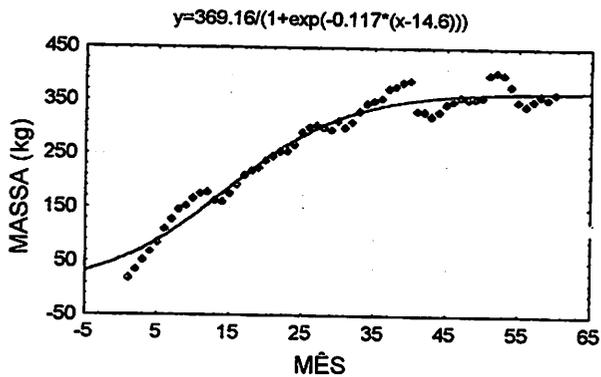
Vaca	Monofásico				Difásico						
	a	b	d	R^2_{aj}	a_1	b_1	d_1	a_2	b_2	d_2	R^2_{aj}
1	362,7	0,14	15,1	0,98	139,6	0,58	4,6	225,2	0,17	22,2	0,98
2	497,9	0,12	19,1	0,99	117,2	0,77	4,2	384,4	0,11	23,7	0,99
3	391,4	0,12	17,0	0,98	92,4	0,83	4,6	309,3	0,10	22,1	0,98
4	380,8	0,15	15,4	0,98	126,4	0,83	4,7	265,0	0,13	21,2	0,99
5	429,2	0,12	17,0	0,98	34,6	2,45	11,4	400,9	0,11	18,3	0,98

(1) kg (2) mês⁻¹ (3) mês

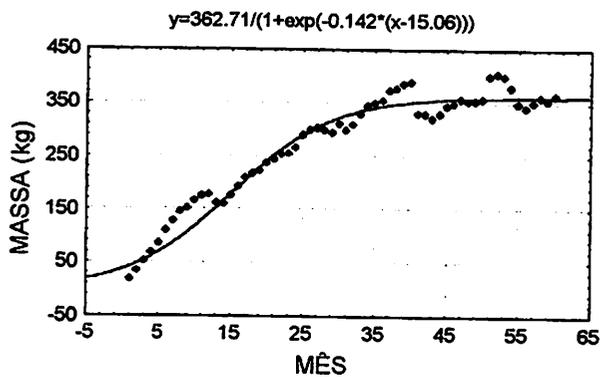
TABELA 36. Estimativas da assíntota (a, a_1, a_2)⁽¹⁾, do parâmetro de crescimento (b, b_1, b_2)⁽²⁾, do tempo de inflexão (d, d_1, d_2)⁽³⁾ e R^2_{aj} , dos modelos logístico monofásico e difásico, com estrutura de erros AR(2). Dados ponderais de 60 observações em vacas.

Vaca	Monofásico				Difásico						
	a	b	d	R^2_{aj}	a_1	b_1	d_1	a_2	b_2	d_2	R^2_{aj}
1	364,0	0,14	14,9	0,98	133,2	0,64	4,5	232,3	0,17	21,9	0,99
2	492,0	0,12	16,6	0,98	115,4	0,79	4,1	380,8	0,11	23,3	0,99
3	388,7	0,12	16,6	0,98	90,9	0,82	4,8	308,1	0,10	21,7	0,98
4	372,2	0,17	14,6	0,99	133,2	0,64	4,5	232,3	0,17	21,9	0,99
5	427,3	0,12	16,7	0,98	36,3	2,30	11,3	397,6	0,11	18,1	0,98

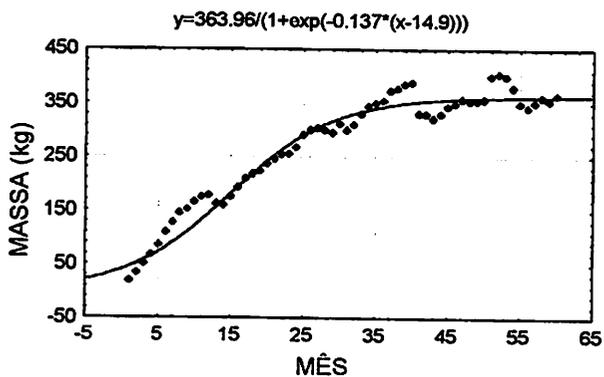
(1) kg (2) mês⁻¹ (3) mês



(a)



(b)



(c)

FIGURA 1. Ajuste da função logística monofásica com estrutura de erros EI (a), AR(1) (b) e AR(2) (c), para a vaca 1. (♦ observado — predito).

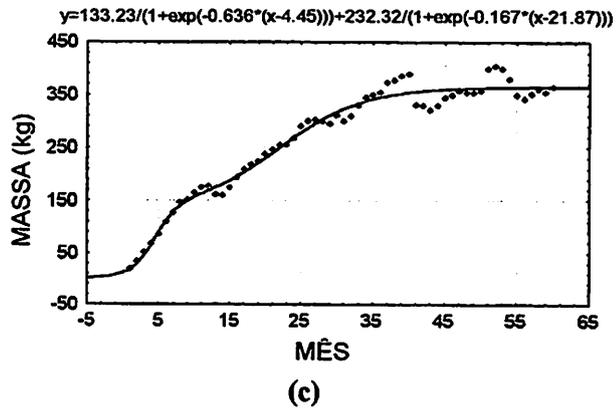
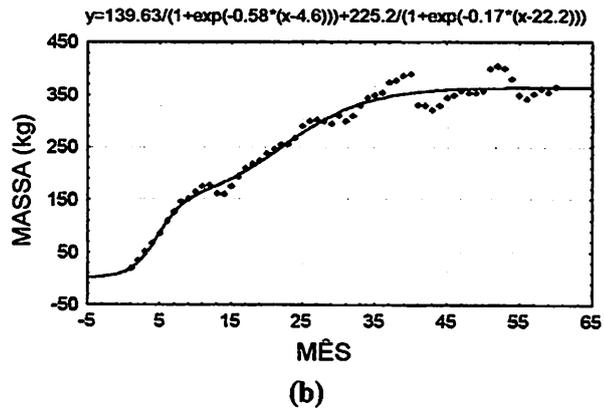
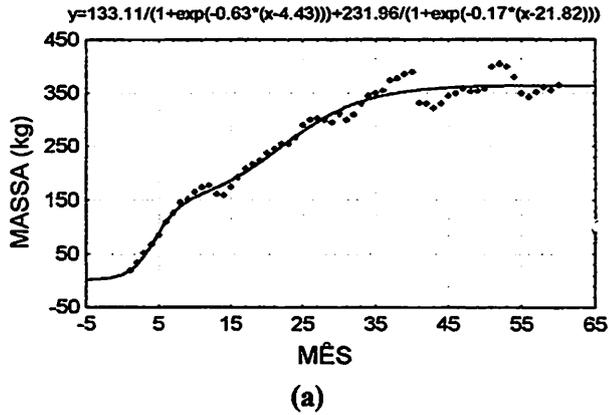


FIGURA 2. Ajuste da função logística difásica com estrutura de erros EI (a), AR(1) (b) e AR(2) (c), para a vaca 1. (♦ observado — predito).

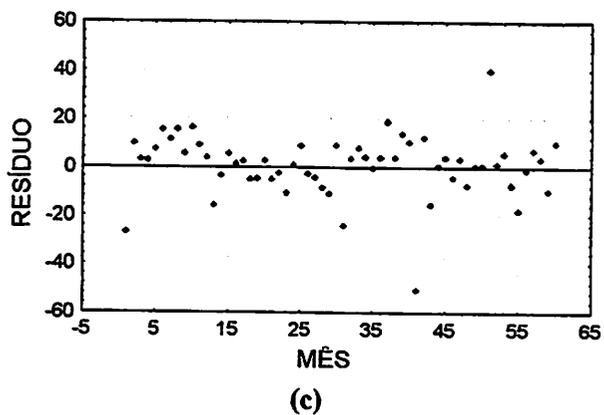
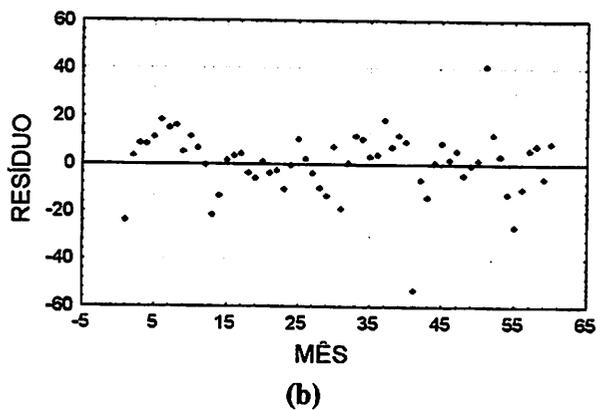
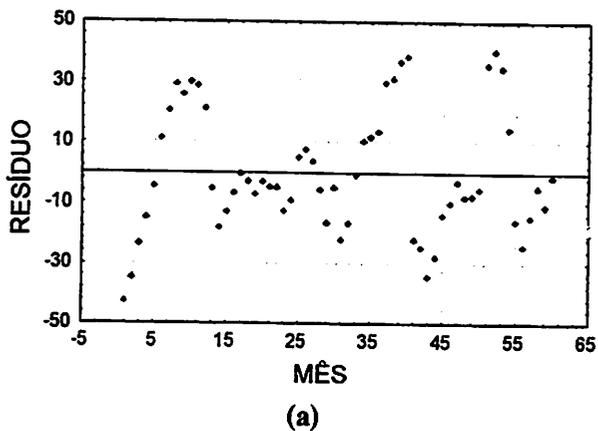
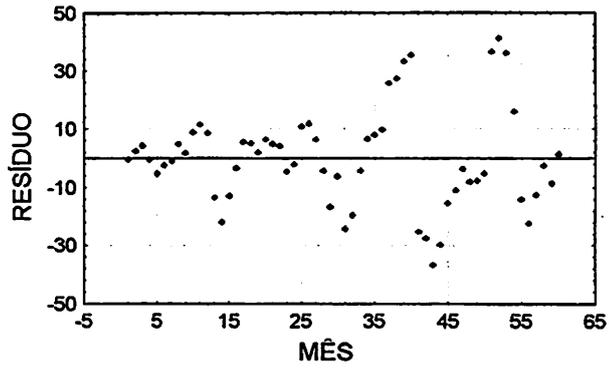
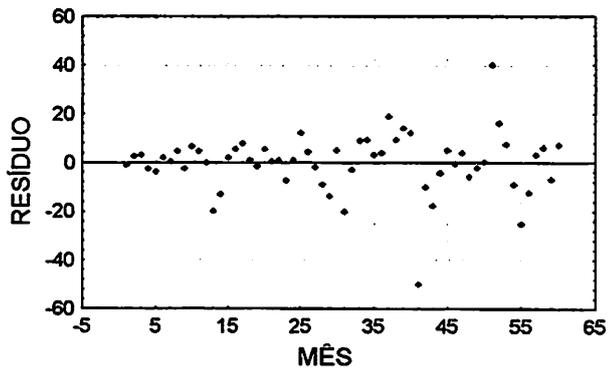


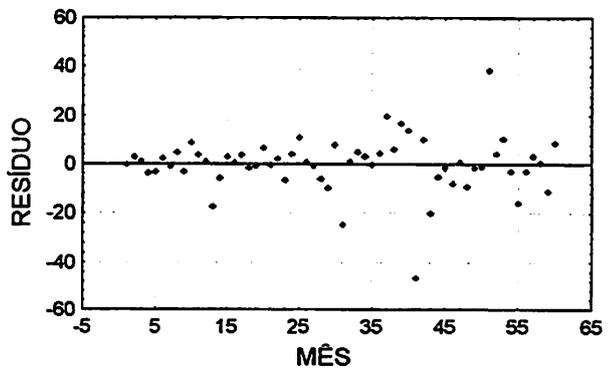
FIGURA 3. Resíduos do modelo logístico monofásico com estrutura de erros EI (a), AR(1) (b) e AR(2) (c), para a vaca 1.



(a)



(b)



(c)

FIGURA 4. Resíduos do modelo logístico difásico com estrutura de erros EI (a), AR(1) (b) e AR(2) (c), para a vaca 1.

Observa-se nas Tabelas 19, 20 e 21 que os valores de R^2 ajustados, mesmo para o modelo monofásico, foram superiores a 0,96 para os cinco animais analisados, o que indica um bom ajustamento da função logística a este tipo de dados de crescimento. Este resultado diverge dos obtidos por Kroll (1990), que obteve valores bem inferiores. Ele concluiu que o modelo logístico foi o que pior descreveu o crescimento dos animais, seja com estrutura de erros EI seja com autorregressivos, em relação aos demais modelos analisados em seu trabalho.

Pode-se observar pela Figura 1 que o ajuste da função monofásica foi bom após o 12^o mês, havendo uma estimação superior aos valores observados nos meses iniciais, seguida de uma estimação inferior nos meses finais, o que está de acordo com Fitzhugh (1976) e Tornero (1996). A qualidade do ajuste foi melhor com a introdução da estrutura AR(1) e pouco se alterou, em relação a esta, com o uso da estrutura AR(2).

O uso do modelo difásico, Figura 2, levou a uma expressiva melhoria na qualidade do ajuste em relação ao modelo monofásico, principalmente nas observações iniciais. A introdução da estrutura de autocorrelação nos erros, também produziu uma melhoria na qualidade do ajuste, em relação ao modelo com erros independentes, sendo que pouco se alterou ao se passar do modelo AR(1) para o modelo AR(2).

Quando se utilizou o modelo com estrutura de erros EI, percebeu-se claramente o problema causado pela autocorrelação, conforme se observa na Figura 3 (a) e da não homogeneidade das variâncias (Figura 4 (a)). Estes problemas foram praticamente contornados quando se levou em consideração os modelos com estruturas de erros AR(1) e AR(2), conforme mostram os gráficos nas Figuras 3 (b), 3 (c), 4 (b) e 4 (c).

6 CONCLUSÕES

- 6.1 O número de iterações, para as quais não se chegou a um ponto de convergência, ou que convergiram para valores muito distantes do esperado, foi reduzido com o aumento de n e R^2 e a diminuição do valor de ρ .
- 6.2 Os estimadores tenderam aos valores paramétricos e com maior precisão em erro quadrático médio, à medida que se aumentou o valor de n .
- 6.3 A introdução da estrutura de autocorrelação nos erros melhorou o ajuste, minimizando o problema das medidas repetidas.
- 6.4 Houve uma sensível melhoria na qualidade do ajustamento quando foi utilizada a função difásica, em relação à monofásica.
- 6.5 O uso da função logística monofásica e difásica, com estrutura de erros autorregressivos, é indicado para descrever o crescimento ponderal de vacas leiteiras com, aproximadamente, duas pesagens por mês ou uma por semana, em um intervalo de sessenta meses.

REFERÊNCIAS BIBLIOGRÁFICAS

BERGER, R. D. Comparison of the Gompertz and logistic equations to describe plant disease progress. **Phytopatology**, v. 71, p. 716-719, 1981.

BERTALANFFY, L. Von. **Quantitative laws in metabolism and growth**. **Quart. Rev. Biol.** v. 32, p. 217, 1957.

BOCK, R. D.; WAINER, H.; PETERSEN, A.; THISSEN, D.; MURRAY, J.; ROCHE, A. A parametrization for individual human growth curves. **Hum. Biol.** v. 45, p. 63-80, 1973.

BROWN, J. E., FITZ HUGH, H. A. Jr., CARTWRIGHT, T. C. A comparison of nonlinear models for describing weight-age relationships in cattle. **Journal of Animal Science**. v. 42, p. 810-818, 1976.

BRODY, S. **Bioenergetics and Growth**. Reinhold Publishing Corp., New York. 1945.

CAUSTON, D. R. A computer program for fitting the Richards function. **Biometrics**. v. 25, p. 401, 1969.

CHEECK, D. B. A new look at growth. Capítulo 1 em: **Human Growth. Body Composition, Cell Growth, Energy and Intelligence**. Lea & Febiger, Philadelphia. 1968.

CHIACCHIO, E. J. **Regressão Não-Linear: desenvolvimento de um sistema computacional e aplicações**, Piracicaba, 1993. 149p. Dissertação (Mestrado) - Escola Superior de Agricultura Luiz de Queiroz - USP - São Paulo.

COCK, A. G. Genetical aspects of metrical growth and form in animals. **Quart. Ver. of Biol.** v. 41, p. 131, 1966.

F3
F4
F5
F6
F7
F8
F9
F10
F11
F12
F13
F14
F15
F16
F17
F18
F19
F20
F21
F22
F23
F24
F25
F26
F27
F28
F29
F30
F31
F32
F33
F34
F35
F36
F37
F38
F39
F40
F41
F42
F43
F44
F45
F46
F47
F48
F49
F50
F51
F52
F53
F54
F55
F56
F57
F58
F59
F60
F61
F62
F63
F64
F65
F66
F67
F68
F69
F70
F71
F72
F73
F74
F75
F76
F77
F78
F79
F80
F81
F82
F83
F84
F85
F86
F87
F88
F89
F90
F91
F92
F93
F94
F95
F96
F97
F98
F99
F100

CROCCI, A. J. Modelo autorregressivo para análise de experimentos com vacas em lactação. Piracicaba, 1984. 86p. (Tese- Doutorado - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo).

DeNISE, R. S. K.; BRINKS, J. S. Genetic and environmental aspects of the growth curve parameters in beef cows. *Journal of Animal Science*, v. 61, p. 1431-1440, 1985.

DRAPER, R.; SMITH, H. Applied regression analysis, New York: Wiley, 1981.

FABENS, A. J. Properties and fitting of the von Bertalanffy growth curve. *Growth*, v. 29, p. 265, 1965.

FITZHUGH JR., H. A. Analysis of growth curves and strategies for altering their shape. *Journal of Animal Science*, v. 42, n. 6, p. 1036-1050, 1976.

GALLANT, A. R. Nonlinear statistical models. New York: Wiley, 1987.

GLASBEY, C. A. Correlated residual in non-linear regressions applied to growth data. *Appl. Statist.*, v 28, p. 251-259, 1979.

GLASBEY, C. A. Non-linear regression with autorregressive time series errors. *Biometrics*, v. 36, p. 135-140, 1980.

GODOI, C. R. M. Modelo linear com autocorrelação nos resíduos para análise de produção leiteira em experimento com dupla reversão. Piracicaba, 1983. 110p. (Tese - Livre Docência - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo)

GROSENBAUGH, L. R. Generalization and reparametrization of some sigmoid and other nonlinear functions. *Biometrics*. v. 21, p. 708 - 714, 1965.

HARTLEY, H. O. The modified Gauss-Newton method for the fitting of nonlinear regression function by least squares. *Tecnometrics*, Richmond, v. 3 p. 269-280, 1961.

HOFFMAN, R., VIEIRA, S. Análise De Regressão: uma introdução à econometria. 2 ed. Hucitec. São Paulo, 1987. 379 p.

- JOHNSTON, J. **Métodos econométricos**. São Paulo, Atlas, 1977. 318 p.
- X KOOPS, W. J. **Multiphasic Growth Curve Analysis**. *Growth* 50, p. 169-177, 1986.
- KOOPS, W. J.; GROSSMAN, M. **Multiphasic growth and allometry**. *Growth Dev. Aging*, v. 55, p. 203-212, 1991.
- KROLL, L. B. **Estudo do crescimento de vacas leiteiras através do modelo de correlação nos erros**. Botucatu, 1990. 97p. Tese (Doutorado) - Faculdade de Ciências Agrônômicas do Campus de Botucatu, Universidade Estadual Paulista.
- LAIRD, A. K., HOWARD, A. **Growth curves in inbred mice**. *Nature*. Basingstoke. v. 213, n. 5078. p. 786-788, 1967.
- MACHADO, M. A. F. T. **A função logística e sua aplicação a dados de crescimento**. In: **Seminários de Pós-graduação em Experimentação e Estatística**. Piracicaba, ESALQ, USP, 1980. 19p.
- MARQUARDT, D. W. **An algorithm for least squares estimation of nonlinear parameters**. *Journal Society for Industrial and Applied Mathematics*. Philadelphia. v. 11(2). p. 431-441, 1963.
- NELDER, J. A. **The fitting of a generalization of the logistic curve**. *Biometrics* 17, p. 89 - 110, 1961.
- NELDER, J. A. **182 Note: Na alternative form of a generalized logistic equation**. *Biometrics*, v. 18, p. 614-616, 1962.
- OLIVER, F. R. **Aspects of maximum likelihood of the logistics growth function**. *JASA*, v. 61, p. 697-705, 1986.
- PALSSON, H. **Conformation and body composition**. In: **Progress in the Physiology of Farm Animals**. (ed. J. Hammond) v. 2. Butterworths Scientific Publications, London, 1955.
- PEIL, J., HELWIN, H. **A phenomenologic-mathematical model of growth dynamics**. *Biometrics J.* v. 23, p. 41 - 54, 1981.

RICHARDS, F. J. The quantitative analysis of plant growth. Chapter 1. In: **Plant Physiology**, v. VA. F. C. Steward (ed.). Academic Press, New York, London, 1969.

ROBERTSON, T. B. **The Chemical Basis of Growth and Senescence.** Monographs of Experimental Biology, Philadelphia. 1923.

SAS Institute Inc., SAS/ETS® User's Guide. Version 6. Second Edition. Cary, NC: SAS Institute Inc., 1993.

SOUZA, Geraldo da Silva e. **Introdução aos modelos de regressão linear e não-linear,** Brasília: Embrapa-SPI, Embrapa-SEA, 1998. 505p.

TANNER, J. M. Some notes on the reporting of growth data. **Human Biol.**, v. 23, p. 93, 1951.

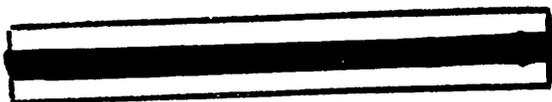
↖ **TORNERO, M. T. T.** **Utilização de modelos assintótico-sigmóides de crescimento multifásico com sazonalidade em estudos zootécnicos.** Botucatu, 1996. 131p. Tese (Doutorado) - Universidade Estadual Paulista, Campus de Botucatu.

THEIL, H. **Principles of econometrics.** New York. John Willey & Sons, 1971. 736 p.

↖ **VIEIRA, S.; MISCHAN, M. M.** A Logística e a Gompertz: duas funções alternativas no estudo de dados de crescimento. **Ciência e Cultura**, São Paulo, v. 28, p. 950-952, 1976.



ANEXOS



ANEXOS

Página

ANEXO 1A	Programa SAS utilizado para a simulação dos dados do modelo logístico.....	67
ANEXO 2A	Programa SAS utilizado para a determinação dos estimadores dos parâmetros do modelo logístico difásico com erros independentes.....	70
ANEXO 3A	Programa SAS utilizado para a determinação dos estimadores dos parâmetros do modelo logístico difásico com estrutura de erros AR(2).....	71
ANEXO B	Dados de pesagens mensais de cinco vacas	72

ANEXO 1A

```
/*PROGRAMA PARA SIMULAÇÃO */
/*Geração de dados -modelo difásico n=50 AR(1)*/

options nodate nonumber ps=60 ls=75;
data crown14h;
  N=2000;
  keep x d u1 y ii;
  do ii=1 to N;
    x=0;
    u=0;
    sigma=97.97;
    do i=1 to 50;
      z1=exp(-0.60*(x-4.50));
      z2=exp(-0.17*(x-22.00));
      d=133/(1+z1)+232/(1+z2);
      x=x+1.224490;
      seed=4538579;
      u1=rannor(seed)*sigma+0.80*u;
      y=d+u1;
      output;
      u=u1;
    end;
  end;

/* Estimativas dos parâmetros do modelo logístico difásico AR(1) */

proc model maxiter=100 noprint;
  z1=exp(-b1*(x-d1));
  z2=exp(-b2*(x-d2));
  y=a1/(1+z1)+a2/(1+z2);
  parms a1=133.00 b1=0.60 d1=4.50 a2=232.00 b2=0.17 d2=22.00;
  %ar(y,1);
  fit y/outest=est outs=resi;
  by ii;
run;
quit;
```

```
/* Elimina iterações que não convergiram (c = 0,001) e aquelas com parâmetros
sob restrição */
```

```
data estresi;
  merge est resi;
  by ii;
run;
data est1;
  set estresi;
  if (a1 le 0 or a1 ge 1000) or (b1 le 0 or b1 ge 10) or
     (d1 le 0 or d1 ge 60) or (a2 le 0 or a2 ge 1000) or
     (b2 le 0 or b2 ge 10) or (d2 le 0 or d2 ge 60) or
     ii= 3 or ii= 6 or ii= 10 or ii= 12 or ii= 15 or ii= 18 or
     ..
     ii=1977 or ii=1979 or ii=1981 or ii=1982 or ii=1993 or ii=1996
  then delete;
run;
```

```
/* Cálculo do viés e raiz quadrada do erro quadrático médio */
```

```
proc iml;
  N=1505 ;
  use est1;
  read all var{a1} into X1;
  read all var{b1} into X2;
  read all var{d1} into X3;
  read all var{a2} into X4;
  read all var{b2} into X5;
  read all var{d2} into X6;
  read all var{Y_L1} into X7;
  read all var{Y} into Y1;
  a1=133.00;
  X11=a1-X1;
  one=j(1,N,1);
  biasa1=(one*X11)/N;
  b1=0.60;
  X21=b1-X2;
  biasb1=(one*X21)/N;
  d1=4.50;
  X31=d1-X3;
```

```

biasd1=(one*X31)/N;
a2=232.00;
X41=a2-X4;
biasa2=(one*X41)/N;
b2=0.17;
X51=b2-X5;
biasb2=(one*X51)/N;
d2=22.00;
X61=d2-X6;
biasd2=(one*X61)/N;
Y_L1=0.80;
X71=Y_L1-X7;
biasY_L1=(one*X71)/N;
y=9597.9027;
X81=Y-Y1;
bias1=(one*X81)/N;
eqma1=(X11*X11)/N;
eqmb1=(X21*X21)/N;
eqmd1=(X31*X31)/N;
eqma2=(X41*X41)/N;
eqmb2=(X51*X51)/N;
eqmd2=(X61*X61)/N;
eqmY_L1=(X71*X71)/N;
eqm=(X81*X81)/N;
eqma1=sqrt(eqma1);
eqmb1=sqrt(eqmb1);
eqmd1=sqrt(eqmd1);
eqma2=sqrt(eqma2);
eqmb2=sqrt(eqmb2);
eqmd2=sqrt(eqmd2);
eqmY_L1=sqrt(eqmaY_L1);
eqmr=sqrt(eqmr);
print eqma1, eqmb1, eqmd1, eqma2, eqmb2, eqmd2, eqmY_L1, eqmr ;
print eqma1, eqmb1, eqmd1, eqma2, eqmb2, eqmd2, eqmY_L1, eqmr ;
quit;

```

ANEXO 2A

/* DETERMINAÇÃO DOS ESTIMADORES DOS PARÂMETROS DO MODELO DIFÁSICO */

```
options nodate nonumber ps=60 ls=75;
data vaca1;
input x y1 y2 y3 y4 y5 ;
cards;
1 20 36 30 24 32
2 34 54 30 38 52
...
59 356 530 420 406 495
60 366 540 450 414 434
;
/*Mínimos quadrados ordinário*/
proc nlin method=gauss data=vaca1 maxiter=50;
model y6=a1/(1+exp(-b1*(x-d1)))+a2/(1+exp(-b2*(x-d2)));
parms a1=120 b1=0.8 d1=4 a2=300 b2=0.13 d2=23;
run;
quit;
```

ANEXO 3A

/*PROGRAMA PARA OBTENÇÃO DOS ESTIMADORES DOS PARÂMETROS DO MODELO LOGÍSTICO DIFÁSICO AR(2) */

```
options nodate nonumber ps=60 ls=75;
data vaca;
input x y1 y2 y3 y4 y5 y6;
cards;
1 20 36 30 24 32
2 34 54 30 38 52
...
59 356 530 420 406 495
60 366 540 450 414 434
;
proc model data=vaca;
/*Modelo logístico difásico */
y6=a1/(1+exp(-b1*(x-d1)))+a2/(1+exp(-b2*(x-d2)));
/*Estimativas iniciais dos parâmetros */
parms a1=100 b1=2 d1=12 a2=350 b2=0.1 d2=20;
/*Estrutura ar(2) */
%ar(y6,2);
fit y6/outall out=all printall itall itdetaills itprint list listcode;
run;
quit;
proc print data=all;
run;
quit;
```

ANEXO B

Dados de pesagens mensais, em kg, de cinco vacas do Posto Zootécnico Luiz de Queiroz de Piracicaba - SP.

vaca 1	vaca 2	vaca 3	vaca 4	vaca 5
20	36	30	24	32
34	54	30	38	52
52	66	56	60	68
68	96	74	76	86
86	121	90	98	140
110	138	116	140	133
128	158	128	150	148
146	180	132	166	158
152	190	150	170	163
166	200	146	171	156
175	180	146	166	174
178	181	164	184	198
162	178	191	206	218
160	202	220	219	214
176	210	190	216	220
193	230	200	220	230
210	245	208	225	220
218	256	218	216	218
224	242	220	222	238
238	270	224	241	256
246	276	250	256	268
255	302	262	276	270
256	311	274	282	288
268	313	274	295	281
290	336	286	290	290
300	346	287	293	300
303	345	278	290	300
300	358	290	290	306
295	364	290	298	332

...continua...

ANEXO B, Cont.

vaca 1	vaca 2	vaca 3	vaca 4	vaca 5
312	354	312	325	352
300	386	320	348	350
310	395	330	363	353
330	410	335	377	384
345	418	352	396	390
350	444	356	406	397
355	440	351	392	392
374	448	352	344	378
378	448	317	336	390
386	430	325	330	391
390	436	370	314	405
331	456	370	320	418
330	462	372	334	406
322	476	372	350	385
330	494	384	358	384
345	486	338	368	390
350	420	340	374	392
358	428	344	380	396
354	434	358	390	390
355	438	360	392	402
358	450	342	400	410
400	448	350	390	412
405	450	351	370	414
400	448	355	365	423
380	454	370	360	434
350	480	386	376	440
342	484	402	382	450
352	496	402	380	475
362	509	410	394	478
356	530	420	406	495
366	540	450	414	434

FONTE: Kroll (1990).