



FERNANDO RIBEIRO CASSIANO

**MODELOS DE CRESCIMENTO ANIMAL PARA
TEMPOS IRREGULARES**

LAVRAS - MG

2014

FERNANDO RIBEIRO CASSIANO

**MODELOS DE CRESCIMENTO ANIMAL PARA TEMPOS
IRREGULARES**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientadora

Dra. Thelma Sáfyadi

LAVRAS - MG

2014

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e Serviços
da Biblioteca Universitária da UFLA**

Cassiano, Fernando Ribeiro

Modelos de crescimento animal para dados irregulares /
Fernando Ribeiro Cassiano. – Lavras: UFLA, 2014.
52 p. : il.

Dissertação (mestrado) - Universidade Federal de Lavras, 2014.

Orientador: Thelma Sáfyadi.

Bibliografia.

1. Modelos de crescimento. 2. Erros irregulares. 3. Modelos
autorregressivos. I. Universidade Federal de Lavras. II. Título.

CDD - 519.536

FERNANDO RIBEIRO CASSIANO

**MODELOS DE CRESCIMENTO ANIMAL PARA TEMPOS
IRREGULARES**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 19 de fevereiro de 2014.

Dr. Augusto Ramalho de Moraes	UFLA
Dr. Daniel Furtado Ferreira	UFLA
Dr. Denismar Alves Nogueira	UNIFAL-MG

Dra. Thelma Sáfyadi
Orientadora

LAVRAS - MG
2014

Ao Senhor Jesus pois sem Ele nada é possível;

*À mamãe (in memoriam) que, mesmo tendo ficado pouco tempo, me
ensinou lições que jamais esquecerei;*

*À minha tia Nívia que sempre cuidou de mim, me apoia e torce pelo meu
sucesso;*

Às minhas irmãs; Raimara e Nayara, que amo.

DEDICO

AGRADECIMENTOS

A Jesus por me proporcionar esta conquista;
À tia Delci e ao tio Paulo por me ajudarem a iniciar esta caminhada;
À minha família por me estimular a perseverar na busca pelos meus objetivos e apoiar sempre durante este curso;
Ao professor Tiãozinho por ter me indicado a UFLA para cursar este mestrado;
Aos professores Rômulo e Santos pelo grande apoio no início deste curso;
À professora Dra. Thelma Sáfydi pela orientação, pela simpatia e atenção em me atender quando eu preciso e pela paciência em me aguentar neste tempo;
Aos professores do DEX pela humildade e pelo conhecimento transmitido;
À secretária do DEX, Josi, pela prestatividade e simpatia;
À CAPES e à FAPEMIG pela concessão da bolsa de estudos;
Aos meus colegas e amigos de sala que ajudaram a diminuir, ou aumentar, o estresse em vários momentos deste curso;
Aos amigos de república pela ótima convivência e pelas horas de conversa fiada;
A todos os amigos que cooperaram de forma direta ou indireta para a conclusão deste curso.

Sê

Se não puderes ser um pinheiro, no topo de uma colina,

Sê um arbusto no vale, mas sê

O melhor arbusto à margem do regato.

Sê um ramo, se não puderes ser uma árvore.

Se não puderes ser um ramo, sê um pouco de relva

E dá alegria a algum caminho.

Se não puderes ser uma estrada,

Sê apenas uma senda,

Se não puderes ser o Sol, sê uma estrela.

Não é pelo tamanho que terás êxito ou fracasso...

Mas sê o melhor no que quer que sejas.

(Pablo Neruda)

RESUMO

O estudo do crescimento animal ajuda a definir pontos importantes ao longo da sua vida e também a traçar estratégias de manejo de maneira que as épocas de venda ou abate sejam definidas com mais precisão. A utilização de curvas não lineares para modelar o crescimento tem sido muito frequente na literatura. Tais funções geralmente fornecem um bom ajuste aos dados, seus parâmetros são interpretáveis biologicamente e normalmente têm menos parâmetros que as funções lineares no que se refere à modelagem do crescimento. A estrutura dos resíduos, no entanto, em grande parte dos casos não é considerada de maneira adequada, já que em algumas situações os dados são obtidos em tempos irregulares, ou seja, em tempos que não são igualmente espaçados. Ajustou-se modelos para o peso médio de 160 animais da raça hereford considerando resíduos autorregressivos de primeira ordem regulares (AR(1)) e irregulares (IS-AR(1)). Os modelos de crescimento utilizados neste estudo foram o Logístico, o Gompertz e o Von Bertalanffy. Uma nova parametrização do modelo IS-AR(1) foi apresentada; após a estimação dos parâmetros em cada um dos casos (regular e irregular) os modelos foram comparados por meio do critério de informação de Akaike e também da soma de quadrados dos resíduos. Os valores dos parâmetros dos modelos de melhor ajuste foram utilizados para fazer simulação de dados. Da simulação foram retiradas, sucessivamente, amostras de tamanho 100 e então os modelos foram novamente ajustados e comparados. Verificou-se que considerar a estrutura autorregressiva irregular melhorou a qualidade do ajuste dos modelos em praticamente todos os casos.

Palavras-chave: Modelos de crescimento. Erros irregulares. Modelos autorregressivos.

ABSTRACT

Studying animal growth aids in defining important points over its life, as well as to trace management strategies in a way in which the sale and slaughter periods be defined with more precision. The use of non-linear curves to model growth has been very frequent in literature. Such functions generally provide good data adjustment, its parameters are biologically interpretable and, normally, presents less parameters than the linear functions in regard to growth modeling. The residue structure, however, in a large portion of the cases, is not adequately considered, since, in some situations, the data are obtained at irregular periods, that is, in periods which are not equally spaced. The models were adjusted for the average weight of 160 animals of the Hereford breed, considering regular (AR(1)) and irregular (IS-AR(1)) first order autoregressive residues. The growth models used in this study were the Logistic, Gompertz and Von Bertalanffy. A new patterning of the IS-AR(1) model was presented; after the estimation of the parameters in each case (regular and irregular) the models were compared through the Akaike information criteria and through the square sum of the residues. The parameter values of the models with the best adjustments were used in data simulation. From the simulation, size 100 samples were successively removed, and the models were newly adjusted and compared. It was verified that to consider the irregular autoregressive structure improved the quality of the model adjustment in practically all cases.

Keywords: Growth models. Irregular errors. Autorregressive models.

LISTA DE FIGURAS

Figura 1	Valores médios reais de peso em função da idade de 160 animais da raça Hereford	33
Figura 2	Simulação de dados de peso animal, do 1 ^o ao 730 ^o dia após o nascimento, para o modelo Logístico	39
Figura 3	Simulação de dados de peso animal, do 1 ^o ao 730 ^o dia após o nascimento, para o modelo Gompertz	39
Figura 4	Simulação de dados de peso animal, do 1 ^o ao 730 ^o dia após o nascimento, para o modelo Von Bertalanffy	40

LISTA DE TABELAS

Tabela 1	Expressões de alguns modelos não lineares.	20
Tabela 2	Valores médios dos pesos observados de 160 animais em função da idade.	32
Tabela 3	Valores estimados dos parâmetros para o ajuste dos modelos aos dados reais considerando as estruturas regular e irregular, e as somas de quadrados dos resíduos em ambos o casos. . .	37
Tabela 4	Valores do AIC (dados reais) quando foram ajustados os diferentes tipos de modelos considerando as estruturas regular e irregular dos erros.	38
Tabela 5	Valores estimados para os parâmetros do modelo Logístico à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.	41
Tabela 6	Valores estimados para os parâmetros do modelo Gompertz à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.	42
Tabela 7	Valores estimados para os parâmetros do modelo Von Bertalanffy à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.	42
Tabela 8	Valores do AIC para o modelos Logístico à medida que foram sendo retiradas as amostras.	43
Tabela 9	Valores do AIC para o modelo Gompertz à medida que foram sendo retiradas as amostras.	43
Tabela 10	Valores do AIC para o modelo Von Bertalanffy à medida que foram sendo retiradas as amostras.	43

SUMÁRIO

1	INTRODUÇÃO	12
2	REFERENCIAL TEÓRICO	14
2.1	Revisão bibliográfica	14
2.2	Modelos de crescimento	18
2.2.1	Métodos de estimação para regressão não linear	22
2.2.1.1	O método Gauss-Newton	23
2.3	Séries Temporais	25
2.3.1	Processos estacionários	26
2.3.1.1	Operadores	27
2.3.2	Modelos Autorregressivos	28
2.3.3	Dados Irregulares	29
2.3.4	Análise de dados irregulares	29
2.4	Avaliadores de qualidade	31
3	MATERIAL E MÉTODOS	32
3.1	Reparametrização do modelo IS-AR(1)	34
4	RESULTADOS E DISCUSSÃO	37
5	CONCLUSÃO	44
	REFERÊNCIAS	45
	APÊNDICE	49

1 INTRODUÇÃO

A internacionalização do nosso país é algo indiscutível. A cada dia que se passa estamos conquistando uma visibilidade em escala mundial. Mais especificamente, o setor alimentício tem ótimas projeções de crescimento para os próximos anos, segundo estudo do Ministério da Agricultura, Pecuária e Abastecimento (BRASIL,2014). Desta forma estudos para se conseguir a otimização dos processos de produção, melhor qualidade dos produtos, redução de custos e outros são de fundamental importância para que o setor continue em ascensão neste cenário. Dentro desta ascensão, a pecuária é um setor de destaque.

Avaliar os processos de desenvolvimento animal permite conhecer pontos importantes ao longo do seu crescimento. Este conhecimento faz com que estratégias de manejo sejam realizadas, com o objetivo de maximizar o potencial do animal, seja para produção, venda ou abate. A modelagem auxilia na compreensão das operações de um sistema, sendo usada como mecanismo de predição, previsão de crescimento, época de abate, controle de alimentação, dentre outros.

Na literatura atual é possível encontrar muitos exemplos de estudos feitos na tentativa de modelar o desenvolvimento animal. Na maioria dos casos o estudo do crescimento é feito utilizando-se curvas não lineares, uma vez que estas curvas possuem um número pequeno de parâmetros que são em sua maioria biologicamente interpretáveis (MAZZINI et al., 2005). Cada modelagem, no entanto, é acompanhada de erros, que podem ser causados por fatores conhecidos ou aleatórios. A estrutura desses erros, ou seja, independência ou não, homocedasticidade ou não, regularidade ou não é de fundamental importância para que os resultados sejam mais precisos.

Percebe-se que em muitos estudos a estrutura dos erros (dependência e variabilidade) não é considerada de maneira adequada, assim como, a irregularidade

na coleta dos dados, que normalmente é desprezada. A coleta de dados muitas vezes é feita em intervalos de tempo diferentes, resultando assim em um conjunto irregular de valores. O que se tem feito, na maioria dos casos, é considerar erros independentes e regulares (observações coletadas em intervalos de tempo iguais). Isso pode introduzir vários vieses como, por exemplo, pode acontecer a suavização dos dados (ERDOGAN, 2005), também diminuir flutuações sazonais assim como reduz as informações contidas no conjunto de dados (ECKNER, 2013a). Observando estes aspectos surge a seguinte questão: Analisar curvas de crescimento considerando uma estrutura mais razoável referente aos erros irá ajudar a melhorar o ajuste dos modelos? Com o objetivo de analisar tal fato utilizando ferramentas de séries temporais e assim contribuir para a melhoria na qualidade dos ajustes e, conseqüentemente, as previsões é que se apresenta este trabalho.

Sendo mais específico, neste estudo ajustam-se as curvas de crescimento Logístico, Gompertz e de Von Bertalanffy para dados de peso corporal de animais bovinos da raça Hereford em diferentes idades entre o nascimento e, aproximadamente, dois anos considerando erros autorregressivos de primeira ordem e observações coletadas em espaços de tempo iguais e diferentes. Pretende-se verificar se o fato de considerar a irregularidade na coleta dos dados melhora a qualidade do ajuste dos modelos. Após o estudo dos dados reais são feitas simulações que objetivam avaliar e confirmar a eficiência do método utilizado neste trabalho. A análise será feita utilizando o método proposto por Erdogan et al. (2005) no entanto, uma nova parametrização do modelo é proposta com o objetivo de resolver algum problema de limitação computacional que possa ocorrer. Também se pretende obter todas as demonstrações omitidas no artigo deste mesmo autor.

2 REFERENCIAL TEÓRICO

Neste capítulo apresentamos alguns artigos relacionados a modelos de crescimento e os principais conceitos utilizados para a realização deste trabalho.

2.1 Revisão bibliográfica

Vários modelos de regressão não linear estão disponíveis na literatura e são utilizados para a modelar dados de crescimento nas mais diversas áreas do conhecimento. Muitos trabalhos consideram a estrutura de erros independentes ou autorregressivos, não levando em consideração, também, os períodos de tempo em que os dados foram coletados. A seguir são apresentados alguns artigos que usaram tais estruturas de erros e não observaram a irregularidade, quando havia, nos dados. Nos artigos em que não é destacada a estrutura dos erros, estes foram considerados independentes.

Tedeschi et al. (2000) analisaram o crescimento de animais da raça Guzera e seus cruzamentos alimentados a pasto, com e sem suplementação. Para este estudo foram utilizados dados de 119 animais, entre machos e fêmeas, nascidos em 1977 e 1978. Vários modelos não lineares foram ajustados aos dados com o uso do software SAS e, em seguida, foram comparados. Os autores concluíram que o modelo de Richards apresentou maior dificuldade de convergência, as funções Brody tenderam a superestimar o peso adulto enquanto as funções Gompertz modificada e Logística tendem a subestimá-lo. A função que foi selecionada pelos índices de comparação foi a Gompertz.

Oliveira, Lôbo e Pereira (2000), utilizando dados de crescimento e reprodução de 573 vacas da raça guzera entre 1961 e 1985, na fazenda Canoas, em Curvelo-MG, compararam modelos não lineares para descrever o crescimento dos

animais com o objetivo de estabelecer um padrão médio de crescimento. Ajustaram os modelos Brody, Von Bertalanffy, Logístico, Gompertz e Richards e avaliaram a qualidade de ajustamento. O modelo Richards apresentou dificuldades computacionais para a convergência e os outros modelos se mostraram adequados para descrever o crescimento sendo que o escolhido para representar a curva média de crescimento dos animais foi o Von Bertalanffy por apresentar um ajustamento superior no conjunto de critérios.

Mazzini et al. (2003) estudaram o crescimento de bovinos da raça hereford, nascidos de 1992 a 1994 na Agropecuária Recreio no município de Bajé-RS. Neste estudo eles ajustaram os modelos não lineares Brody, Gompertz, Logístico, Richards e Von Bertalanffy com o objetivo de comparar as cinco funções de crescimento considerando-se os modelos não ponderados e ponderados pelo inverso da variância dos pesos. Eles utilizaram o software SAS para a análise dos dados e concluíram que de forma geral, baseados no seu estudo, a ponderação pelo inverso das variâncias dos pesos melhorou a qualidade do ajuste. As funções de Gompertz e Von Bertalanffy foram as que melhor se ajustaram aos dados, a função logística embora tenha mostrado um bom ajuste aos dados tendeu a subestimar o peso adulto e as demais tiveram um pior desempenho. Os métodos de comparação utilizados foram: a interpretação biológica dos parâmetros, o coeficiente de determinação ajustado, o quadrado médio do resíduo, o erro de predição médio, o teste t de Student para as estimativas dos parâmetros "A" e "k" e o percentual de convergência.

Lopes et al. (2011) usaram registros de animais, bovinos da raça nelore, criados em sistema extensivo de criação, nascidos entre 1987 e 2007, para determinar qual o modelo não linear é mais adequado para descrever o crescimento dos animais. Por meio do procedimento NLIN do SAS os modelos Brody, Logístico,

Gompertz e Von Bertalanffy foram ajustados aos dados e comparados utilizando os critérios: quadrado médio do resíduo, coeficiente de determinação ajustado, erro de predição médio e desvio médio absoluto dos resíduos.

Souza et al. (2010) analisaram dados de crescimento de animais da raça Indubrasil no estado do Sergipe. Foram 1201 bovinos com, no mínimo, 8 pesagens coletadas em intervalos de aproximadamente três meses do nascimento até aproximadamente dois anos. Utilizando o procedimento NLIN do SAS ajustaram os modelos Brody, Von Bertalanffy, Richards, Logístico e Gompertz aos dados de peso-idade e comparando-os por meio do quadrado médio dos resíduos, do estimador de máxima verossimilhança da variância residual, do coeficiente de determinação dentre outros. Após a definição do modelo de melhor ajuste, calcularam a taxa de crescimento absoluto e avaliaram a influência de fatores (sexo, fazenda, mês e ano de nascimento) sobre os parâmetros da curva.

Silva (2009) utilizou modelos não lineares para estudar o crescimento de vacas da raça nelore e cruzadas. Os modelos Brody, Gompertz, Logístico, Von Bertalanffy e Richards foram ajustados aos dados de peso de animais de quatro cruzamentos criados em sistema de produção intensiva que foram submetidos a três níveis de suplementação pós-desmama. O objetivo foi comparar os modelos e avaliar a influência de efeitos de ambientes e grupos genéticos sobre os parâmetros estimados. O ajuste dos modelos foi feito utilizando o procedimento NLIN do SAS pelo método de mínimos quadrados ordinários e ponderados pelo inverso da variância dos pesos.

McManus, Louvandini e Campos (2010) estudaram dados de peso e altura de quatro grupos genéticos de cavalos criados pelo Exército brasileiro. Estimaram curvas de crescimento gerais usando os modelos de Gompertz, Brody, Logístico, Weibull e Richards, segundo o procedimento PROC NLIN do programa SAS. A

comparação do ajuste foi feita observando o quadrado médio dos resíduos e o coeficiente de determinação.

Mazucheli, Souza e Philippsen (2011), utilizando dados de crescimento de codornas, fizeram uma análise de regressão não linear usando o modelo Gompertz, sob a suposição de homogeneidade e heterocedasticidade multiplicativa dos resíduos. Para isso, utilizaram o peso de 414 animais, dentre machos e fêmeas, e ajustaram o modelo utilizando a maximização da função log-verossimilhança e, sob o enfoque bayesiano, via métodos MCMC. Aqui a maximização da log-verossimilhança foi feita por meio de procedimentos do SAS/NLMIXED, com erros independentes.

Mansano (2012) avaliou as características do crescimento morfométrico e composição corporal de girinos de Rã-touro em cativeiro. O experimento foi realizado utilizando um total de 3240 girinos de rã-touro com peso médio inicial de 0,044 g e comprimento total médio de 12,79 mm procedentes de uma mesma desova, que foram alojados em 18 caixas de amianto de 100 litros de capacidade. Os dados de peso vivo e comprimento total foram obtidos no início do experimento e com 13, 23, 33, 42, 55 e 64 dias e em seguida os modelos Gompertz, Von Bertalanffy, Logístico e Brody são ajustados. Aqui, a conclusão obtida foi que os modelos Gompertz e Logístico são adequados para descrever o crescimento em peso e comprimento de girinos de rã-touro em cativeiro.

Mazzini et al. (2005) estudaram o crescimento de bovinos da raça hereford, porém considerando heterocedasticidade e resíduos autorregressivos. Neste trabalho foram utilizados os modelos Brody, Gompertz, Logístico e Von Bertalanffy que, para cada animal, foram ajustados de acordo com os seguintes métodos: quadrados mínimos ordinários (QMO), quadrados mínimos ponderados (QMP), quadrados mínimos ponderados generalizados com erros autorregressi-

vos de primeira (QMPG-AR1) e segunda (QMPG-AR2) ordem. O software SAS foi utilizado para o ajuste dos modelos e a comparação destes foi feita utilizando o percentual de convergência, quadrado médio do resíduo, coeficiente de determinação ajustado e erro de predição médio. Chegaram a conclusão de que os modelos Gompertz e Von Bertalanffy ajustados pelos métodos QMP, (QMPG-AR1) e (QMPG-AR2) foram indicados para descrever o crescimento de novilhos hereford até os dois anos de idade.

Prado (2011) estudou o crescimento dos diâmetros transversal e longitudinal de frutos de coqueiro da variedade anã verde utilizando os modelos Logístico e Gompertz. Os modelos foram ajustados aos dados pelo método de mínimos quadrados e por meio do método iterativo de Monte Carlo via cadeias de Markov considerando uma estrutura de erros independentes e autorregressivos de primeira ordem (AR(1)). A qualidade do ajuste foi feita por meio da avaliação do coeficiente de determinação ajustado, desvio padrão residual e critério de informação de Akaike. A conclusão obtida foi de que o modelo Logístico proporcionou melhor qualidade de ajuste em ambos os métodos de estimação dos parâmetros. Ainda é feita uma análise usando dados simulados, que são comparados aos dados reais, e pelo resultado concluiu-se que o método de simulação foi eficiente uma vez que apresentou estimativas próximas aos valores paramétricos.

2.2 Modelos de crescimento

O primeiro modelo de crescimento foi proposto por Thomas Malthus em 1798 (MALTHUS, 1798). Essa atitude foi resposta a diversas conjecturas de outros estudiosos da época que enxergavam um mundo perfeito, no qual a população e os recursos para a sobrevivência caminhavam em perfeita harmonia. Malthus então sugere um modelo em que a população, se não controlada, cresceria de maneira

exponencial e a capacidade de produção aumentaria apenas linearmente, causando em certo momento um colapso.

Com o passar dos anos a ideia de se modelar o crescimento foi utilizada também por diversos outros autores que, baseados na ideia original, propuseram modelos matemáticos para descrever o crescimento não apenas de populações, mas também o crescimento animal, vegetal, etc. Esses modelos são, normalmente, não lineares e, na atualidade, têm sido amplamente utilizados por pesquisadores das diversas áreas do conhecimento. Por exemplo, Veloso et al. (2012) avaliam genótipos de frangos caipira, Costa et al. (2009) avaliam o crescimento de tilápias de diferentes linhagens utilizando modelos não lineares. Silva et al. (2012) propõem uma análise genética de curvas de crescimento de bovinos, McManus, Louvandini e Campos (2010) usaram curvas de crescimento em quatro espécies de cavalos. Já Prado (2011) avaliou o crescimento de frutos de coqueiro e Fernandes (2012) avaliou o crescimento do fruto do cafeeiro em diferentes alinhamentos de plantio, dentre outros.

Os modelos não lineares, geralmente fornecem um bom ajuste com menos parâmetros do que modelos lineares em relação ao crescimento. Além disso, apresentam parâmetros ou funções de parâmetros com interpretação biológica. Essa mesma ideia é defendida por Loaiza-Echeverri (2011, p. 10) que diz que

os benefícios de se utilizar modelos não lineares, em detrimento aos lineares, reside no fato de que os modelos não lineares sintetizam um grande número de medidas em alguns parâmetros e têm um significado biológico, uma vez que utilizam mensurações subsequentes no tempo sobre o mesmo indivíduo e são facilmente comparadas entre diferentes cenários de produção.

A Tabela 1 apresenta alguns dos principais modelos não lineares que são utilizados para modelar o crescimento animal.

TABELA 1 Expressões de alguns modelos não lineares.

Modelo de curva de crescimento	Referência
$y_i = \beta_1 e^{\left(-e^{(\beta_2 - \beta_3 x_i)}\right)} + e_i$	Gompertz Gompertz e Philos (1825)
$y_i = \beta_1 (1 + e^{\beta_2 - \beta_3 x_i})^{-1} + e_i$	Logístico Ratkowski (1983)
$y_i = \beta_1 x_i (x_i + \beta_2)^{-1} + e_i$	Michaelis-Menten Michaelis e Menten (1913)
$y_i = \beta_1 (1 - \beta_2 e^{-\beta_3 x_i})^3 + e_i$	Von Bertalanffy (1957)
$y_i = \frac{\beta_2 \beta_3^{\beta_4} + \beta_1 x_i^{\beta_4}}{\beta_3^{\beta_4} + x_i^{\beta_4}} + e_i$	Michaelis-Menten Modificado Lopez et al. (2000)

Fonte: Silveira (2010).

Para todos os modelos apresentados na Tabela 1, o parâmetro β_1 representa o peso adulto máximo ou assintótico do animal e o parâmetro β_3 a taxa de maturidade ou velocidade de crescimento. Os modelos que apresentam o parâmetro β_4 possuem ponto de inflexão variável, cuja localização é determinada pelo parâmetro em questão. Os demais modelos ou apresentam ponto de inflexão fixo ou não possuem, como é o caso do modelo Michaelis-Menten. De forma geral, não há uma interpretação prática para o parâmetro β_2 . Uma exceção é o modelo Michaelis-Menten no qual este parâmetro representa a idade na qual se atinge a metade do peso adulto (SILVEIRA, 2010). Em cada modelo y_i , x_i e e_i representam o peso observado do animal, o tempo da medição do peso do animal e o erro respectivamente, no tempo i .

Nas etapas iniciais da vida, o crescimento é mais acelerado e o aumento de peso é maior do que o que ocorre próximo à idade adulta, observando-se uma curva de evolução sigmóide ascendente. À medida que o indivíduo desenvolve-se, a velocidade de crescimento se altera e apresenta mudança na curvatura, identificando o ponto de maior velocidade de crescimento corporal, o que ocorre no chamado ponto de inflexão. Depois deste ponto de inflexão, o crescimento diminui paulatinamente e a taxa de crescimento é cada vez mais lenta. Essa tendência

continua até que o crescimento se estabiliza, fato que matematicamente corresponde com a assíntota horizontal (GOMEZ; MUÑOZ; BETANCUR, 2008). No que segue apresenta-se alguns comentários sobre os modelos Logístico, Gompertz e Von Bertalanffy que serão utilizados neste trabalho.

O modelo logístico, ou modelo de Verhulst, foi proposto por Pierre François Verhulst em 1838. O objetivo inicial foi a tentativa de descrever o comportamento do crescimento da população humana. O modelo em questão foi proposto com a expressão (1):

$$p_t = \frac{mp'e^{mt}}{np'e^{mt} + m - np'} \quad (1)$$

em que p' representa o número da população inicial e m e n são constantes arbitrárias. Com o passar dos anos outros estudiosos reescreveram este modelo, o que é conhecido como reparametrização, como se observa na Tabela 1 e também em Prado (2011). A utilização deste modelo nas diversas áreas do conhecimento gerou um leque de aplicações em que a modelagem se mostrou eficiente. Novas propostas surgiram ao longo do tempo para o modelo de Verhulst como, por exemplo, para modelar o crescimento de frutos (PRADO, 2011), modelar o crescimento animal (LOPES et al., 2011), dentre outros. Neste trabalho será utilizada a parametrização encontrada em Silveira (2010), como apresentado na Tabela 1.

Proposto em 1920 por Benjamin Gompertz, o modelo Gompertz também sofreu reparametrizações ao longo dos anos. Sua proposta inicial foi modelar o crescimento da população. Muitas outras aplicações foram feitas utilizando tal modelo. Zeviane (2013) sugere uma parametrização para o modelo, de modo que os parâmetros são mais facilmente interpretáveis.

Ludwig Von Bertalanffy foi um biólogo e estudioso que nasceu em Atzgerdorf no ano de 1901. O modelo proposto por Von Bertalanffy teve como objetivo estabelecer uma relação entre o metabolismo e o peso corporal dos animais (BER-

TALANFFY, 1957). Em sua pesquisa o autor objetiva comparar indivíduos de uma mesma espécie e de diferentes tamanhos. Com o passar dos anos este modelo foi utilizado para muitas outras modelagens. Alguns exemplos desta diversificação no seu uso podem ser encontradas nos trabalhos de Henriques e Casarini (2009) que avaliam o crescimento de mexilhões em banco natural na Ilha de Palmas-SP, e de Lopes et al. (2011) que estudaram o crescimento de bovinos da raça Nelore da região norte do Brasil, dentre outros.

2.2.1 Métodos de estimação para regressão não linear

Gallant (1987) estima os parâmetros de uma regressão não linear adotando convenções análogas à estimação paramétrica para uma regressão linear. A utilização do método de mínimos quadrados se mostra eficiente também para uma regressão não linear. Seja então a equação de uma regressão não linear dada por:

$$y_i = f(x_i, \boldsymbol{\theta}) + e_i, \quad i = 1, \dots, n$$

em que y_i representa cada variável observada, $f(x_i, \boldsymbol{\theta})$ representa o modelo que foi adotado para descrever a variável, x_i e $\boldsymbol{\theta}$ representam a variável independente e o vetor de parâmetros respectivamente, e e_i é a componente aleatória ou erro que é i.i.d. (independente e identicamente distribuído) e $e_i \sim \mathcal{N}(0, \sigma^2)$. No caso em que não ocorre o i.i.d. tem-se dados correlacionados e torna-se apropriado utilizar os modelos da família autorregressivos integrados e médias móveis (ARIMA), propostos por Box e Jenkins (1976), para trabalhar com os dados. Vetorialmente

esta equação pode ser escrita como:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f(x_1, \boldsymbol{\theta}) \\ f(x_2, \boldsymbol{\theta}) \\ \vdots \\ f(x_n, \boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}.$$

A soma de quadrados dos erros neste caso é dada por:

$$SQ = \sum_{i=1}^n [y_i - f(x_i, \boldsymbol{\theta})]^2.$$

Para obtermos o valor que minimiza a soma de quadrados é necessário fazer a derivação de SQ em relação a cada componente do parâmetro $\boldsymbol{\theta}$ e, em seguida, igualar a zero, resultando assim num sistema de equações normais não lineares. A solução deste sistema não linear não pode ser encontrada de forma explícita, como ocorre no caso de um modelo linear; necessitando de métodos iterativos para a sua resolução. Há vários métodos iterativos na literatura que podem ser utilizados para a solução de um sistema de equações não lineares, dentre eles podemos citar Método do gradiente, Método da secante, Gauss-Newton, Newton Raphson. A função *gnls* do software R, que será utilizada neste trabalho para o ajuste dos modelos tem como método padrão o Gauss-Newton e, por esse motivo, segue-se uma breve descrição do método.

2.2.1.1 O método Gauss-Newton

Assim como a maioria dos métodos iterativos, neste método é necessário um valor inicial para os parâmetros, o que comumente é chamado de “valor inicial arbitrário”. A partir daí, pelo método de mínimos quadrados, utilizando uma expansão de primeira ordem em série de Taylor como aproximação da função

de regressão não linear a raiz da equação ou solução do sistema de equações é aproximada (GALLANT, 1987).

Seja então o conjunto de valores iniciais dos parâmetros dados pelos valores $\theta^0 = [\theta_1^0, \theta_2^0, \dots, \theta_p^0]$ em que p é o número de parâmetros do modelo. Obtidos os valores iniciais e fazendo a expansão em série de Taylor até a primeira ordem centrada neste ponto tem-se:

$$f(\theta) = f(\theta^0) + \frac{\partial f(\theta^0)}{\partial \theta} (\theta - \theta^0)$$

em que $\frac{\partial f(\theta^0)}{\partial \theta}$ é a matriz de derivadas parciais em relação a cada parâmetro θ . Para simplificar, trate $\frac{\partial f(\theta^0)}{\partial \theta}$ por $F(\theta^0)$. Desta forma o modelo não linear pode ser reescrito matricialmente por:

$$Y = f(\theta) + e,$$

$$Y = f(\theta^0) + \frac{\partial f(\theta^0)}{\partial \theta} (\theta - \theta^0) + e,$$

$$Y = f(\theta^0) + F(\theta^0) (\theta - \theta^0) + e.$$

A aproximação da soma de quadrados é

$$SQE = \|Y - f(\theta^0) - F(\theta^0) (\theta - \theta^0)\|^2.$$

O valor que minimiza a soma de quadrados, θ^* , pode ser encontrado usando-se a expressão

$$\theta^* = [F'(\theta^0)F(\theta^0)]^{-1} F'(\theta^0) [Y - f(\theta^0)].$$

Terminado este primeiro processo, se o critério de convergência, que é adotado pelo pesquisador, não for satisfeito inicia-se um novo ciclo e o novo valor de θ para o passo seguinte será $\theta^1 = \theta^0 + \theta^*$, e assim sucessivamente até que o critério

seja satisfeito.

A Tabela 1 apresenta alguns modelos não lineares cujos erros podem ser uma série temporal, uma vez que as observações formam um conjunto ordenado no tempo. Estes são, geralmente, correlacionados.

2.3 Séries Temporais

Tentar descrever o futuro tendo por base dados e informações passadas é um dos principais objetivos da análise de séries temporais. Estes objetivos podem ser alcançados utilizando-se, basicamente, dois enfoques: análise no domínio temporal ou de frequências. Nos dois casos objetiva-se construir modelos para a série temporal. No domínio temporal são propostos modelos paramétricos e no domínio de frequências os modelos propostos são não paramétricos. Neste estudo utiliza-se apenas o enfoque no domínio temporal.

Morettin e Toloi (2006) conceituam uma série temporal como sendo qualquer conjunto de observações ordenadas no tempo e citam quatro interesses principais para sua análise: investigar o mecanismo gerador da série temporal; fazer previsões de valores futuros da série (a curto ou em longo prazo); descrever apenas o comportamento da série e procurar periodicidades relevantes nos dados. Uma série temporal Y_t geralmente pode ser decomposta das seguintes maneiras:

$$Y_t = T_t + S_t + e_t \quad (2)$$

$$Y_t = T_t S_t e_t \quad (3)$$

em que T_t é a componente de tendência, S_t é a componente de sazonalidade e e_t é a componente aleatória com média zero e variância finita. Em (2) diz-se que os efeitos da série são aditivos e em (3) diz-se que a série possui efeitos multiplica-

tivos. De maneira mais formal, uma série temporal pode ser entendida como uma realização de um processo estocástico, ou seja, um processo governado por leis probabilísticas.

Um processo estocástico é normalmente entendido como um conjunto de variáveis aleatórias que dependem de outras variáveis como, por exemplo, o tempo. Em outras palavras, um processo estocástico é o conjunto de todas as possíveis trajetórias de um fenômeno ou processo físico.

2.3.1 Processos estacionários

Morettin e Toloi (2006) definem um processo estocástico $e = \{e_t, t \in T\}$, em que T é normalmente tomado como o conjunto dos números inteiros ou dos números reais, como fracamente estacionário ou estacionário de segunda ordem (ou em sentido amplo) se, e somente se:

1. $E[Y_t] = \mu_t = \mu$, constante, para todo $t \in T$;
2. $E[Y_t^2] < \infty$, para todo $t \in T$;
3. $\gamma(t_1, t_2) = cov(Y_{t_1}, Y_{t_2})$ é uma função de $|t_1 - t_2|$.

Essa classe de processos é de suma importância na análise de séries temporais, pois o ajuste de modelos para a correlação é feito considerando estas características.

Observando-se as características de um processo estacionário e considerando as decomposições (2) e (3) percebe-se que uma série temporal que possui as componentes de tendência e sazonalidade é claramente não estacionária, pois ao menos uma das condições não é satisfeita. Então, faz-se necessária a identificação e a eliminação de tais componentes para que se tenha um processo estacionário. A inspeção visual da série temporal assim como testes específicos que foram criados para a detecção destas componentes ajudam pesquisadores no trato das séries

temporais. Esquivel (2012) cita alguns testes que são utilizados para testar a tendência numa série temporal, dentre eles tem-se o teste de Wald-Wolfowitz, o teste de Cox-Stuart (ou Teste do Sinal), etc. Além de testes para a tendência, a descrição de alguns testes para a sazonalidade podem ser encontrados em Morettin e Tolo (2006) dentre os quais estão o teste de Kruskal-Wallis, o teste de Friedman, e o teste da análise de variância. Outro teste bastante utilizado para detectar a sazonalidade, assim como o período em que ela ocorre é o teste de Fisher.

2.3.1.1 Operadores

Para a construção dos modelos matemáticos para as séries temporais e também facilitar a manipulação destas Box e Jenkins (1976) utilizam alguns operadores. São eles:

1. Operador translação para o passado B . Aplicado a um elemento da série temporal num instante t retorna o seu valor imediatamente anterior. Aplicando-se sucessivamente este operador, n vezes, por exemplo, tem-se o n -ésimo valor anterior ao instante t . Desta forma, tem-se:

$$BY_t = Y_{t-1}; B^2Y_t = Y_{t-2}; \dots; B^nY_t = Y_{t-n}.$$

2. Operador diferença Δ . É definido como a diferença entre dois pontos consecutivos da série. Assim tem-se:

$$\Delta Y_t = Y_t - Y_{t-1}.$$

Utilizando o operador B dentro desta relação obtem-se:

$$\Delta = 1 - B.$$

2.3.2 Modelos Autorregressivos

Box e Jenkins (1976) propuseram modelos para o ajuste da correlação em uma série temporal. A metodologia proposta por eles tem sido bastante utilizada na literatura e foi também a técnica aplicada neste trabalho. A ideia aqui é representar um elemento da série temporal num instante t como uma combinação linear de alguns valores passados até um tempo p .

Desta forma para a série estacionária e_t , o modelo autorregressivo de ordem p , $AR(p)$, é dado por:

$$e_t = \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \dots + \varphi_p e_{t-p} + a_t, \quad (4)$$

em que $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_p]$ é o conjunto dos pesos e a_t é uma variável aleatória com média zero e variância constante σ^2 . A componente a_t é chamada *ruído branco*. Aplicando-se o operador translação para o passado, B , em (4) consegue-se resumi-la em $\varphi(B)e_t = a_t$, em que $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$ é o polinômio autorregressivo de ordem p . De acordo com Morretin e Tolo (2006) este processo é estacionário quando as raízes de $\varphi(B) = 0$ estiverem fora do círculo unitário. Os processos de primeira e segunda ordem, $AR(1)$ e $AR(2)$,

$$e_t = \varphi e_{t-1} + a_t$$

$$e_t = \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + a_t$$

são de uma importância prática considerável (BOX; JENKINS, 1976). Segundo Morettin e Tolo (2006) os processos $AR(1)$ e $AR(2)$ são estacionários quando $-1 < \varphi < 1$ e $\varphi_1 + \varphi_2 < 1, \varphi_2 - \varphi_1 < 1, -1 < \varphi_2 < 1$, respectivamente.

2.3.3 Dados Irregulares

A coleta dos dados nem sempre é fácil de se fazer, o que pode resultar em uma série não igualmente espaçada aqui definida para a situação em que os dados não são coletados em intervalos iguais de tempo. De acordo com Erdogan et al. (2005) este tipo de série é comum em muitas aplicações da vida real quando as medições são restritas por condições práticas. Os autores destacam ainda, que é uma prática comum ignorar os momentos (irregularidades) e tratar os dados como se fossem regulares e que este fato pode claramente introduzir um viés significativo nas estimativas dos parâmetros, levando a predições incorretas.

Eckner (2013b) também reforça esta ideia dizendo que

talvez a maioria das abordagens comuns seja transformar dados não igualmente espaçados em dados igualmente espaçados usando alguma forma de interpolação - a maioria frequentemente linear - e então aplicar métodos existentes para séries temporais igualmente espaçadas. Porém, transformar dados de séries temporais usando tal caminho introduz vários vieses.

Daí a preocupação em se considerar a estrutura original da série temporal.

2.3.4 Análise de dados irregulares

Erdogan et al. (2005) propõem um método para analisar uma série temporal irregular (dados não igualmente espaçados). Em sua proposta os autores consideram uma série temporal Y_t que não possui a componente sazonal e pode, então, ser decomposta como

$$Y_t = b(t) + e_t$$

em que $b(t)$ é uma função determinística que representa a componente de tendência e e_t é a componente de ruído aleatória. Geralmente, apenas os valores de Y_t podem ser observados. Então, o primeiro passo é estimar a componente $b(t)$ e

extrair a componente aleatória.

Seja $\{y_{t_i}, t_i\}$ uma amostra de Y_t . Assumindo que $b(t)$ é um polinômio de grau p em t e, portanto,

$$b_p(t) = \rho_0 + \rho_1 t + \rho_2 t^2 + \cdots + \rho_p t^p$$

em que p é um inteiro não negativo e $\boldsymbol{\rho} = [\rho_0; \rho_1; \cdots; \rho_p]$ é o vetor de coeficientes. Este vetor pode ser estimado usando o método de mínimos quadrados, ou seja, deve-se minimizar a soma de quadrados dos resíduos dada por $\sum_{i=0}^n (y_{t_i} - b_p(t_i))^2$. Define-se então $e_{t_i} = y_{t_i} - b_p(t_i)$ e trata-se $\{e_{t_i}, t_i\}$ como uma amostra de e_t .

Após ajustado o polinômio os autores definem um processo estacionário autorregressivo de primeira ordem para dados irregulares como:

Definição. (Processo IS-AR(1)) Uma série temporal $\{e_{t_i}, t_i\}$ é (IS-AR(1)) estacionária amostrada irregularmente se $\{e_t\}$ é estacionária e se para cada t e $\Delta_i > 0$,

$$e_{t+\Delta_i} = \varphi^{\Delta_i} e_t + \sigma_{\Delta_i} \epsilon_{t+\Delta_i} \quad (5)$$

em que $\epsilon_{t+\Delta_i} \sim \mathcal{N}(0,1)$ e $cov(\epsilon_t, \epsilon_s) = 0$ para cada $t \neq s$ e $\sigma_{\Delta_i}^2 = \sigma^2 \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2} \right)$ para algum $\sigma > 0$.

Estimação Paramétrica: Dado um conjunto de observações $\{e_{t_0}, e_{t_1}, \dots, e_{t_n}\}$ de e_t , define-se $\Delta_i = t_{i+1} - t_i$ para todo $i = 0, 1, 2, \dots, n-1$.

Uma vez que $E[\epsilon_t] = 0$ e $cov(\epsilon_t, \epsilon_s) = 0$ para todo $t \neq s$, pode-se estimar φ pelo método de mínimos quadrados. Em particular precisa-se encontrar $\varphi \in (-1, 1)$ minimizando $\sum_{i=0}^{n-1} (e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i})^2$. Uma vez que $\Delta_i \geq 1$ para todo i esta soma é uma função convexa de φ , e pode ser minimizada usando otimização convexa.

Para estimar σ , define-se $\epsilon_i = e_{t_{i+1}} - (\hat{\varphi})^{\Delta_i} e_{t_i}$. Na equação (5) tem-se

que $\epsilon_i \sim \mathcal{N}(0, \sigma_{\Delta_i}^2)$. Assim, usando o estimador de máxima verossimilhança dos resíduos $\epsilon_{t_0}, \epsilon_{t_1}, \dots, \epsilon_{t_{n-1}}$ tem-se que o estimador de σ é

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} [e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}.$$

Uma demonstração de como se obtém o estimador $\hat{\sigma}$ analiticamente é disponibilizada no Apêndice.

Neste trabalho, para a representação da componente de tendência, utilizamos algumas curvas não lineares que são conhecidas como modelos de crescimento. Desta forma ao invés de estimar-se os coeficientes do polinômio, obtivemos os valores das estimativas dos parâmetros das curvas não lineares.

2.4 Avaliadores de qualidade

Há vários métodos para se medir ou avaliar a qualidade de ajuste de um modelo e então, se for o caso, fazer comparações entre modelos. Neste trabalho os métodos utilizados serão o critério de informação de Akaike (AIC) e pela soma de quadrado dos resíduos (SQ).

O critério de informação de Akaike, AIC, é dado por

$$AIC = \ln(\text{QME}) + \frac{2(p+1)}{n},$$

em que QME é o quadrado médios dos resíduos, p o número de parâmetros do modelo e n o tamanho da amostra.

A soma de quadrados dos resíduos, SSE, é dada pela expressão

$$SQ = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

em que y_i é o valor observado e \hat{y}_i é o valor ajustado.

3 MATERIAL E MÉTODOS

Neste trabalho serão considerados dados reais e dados simulados. Para os dados simulados serão utilizados os valores paramétricos obtidos do ajuste dos modelos com os dados reais. Os dados reais são informações relativas ao peso corporal de animais, machos não castrados, da raça Hereford, obtidos por meio do trabalho de Emiliano (2013). Foram consideradas as dezesseis primeiras pesagens de 160 animais, desde o nascimento até aproximadamente dois anos de idade. Os pesos médios dos 160 animais nas 16 pesagens estão apresentados na Tabela 2 e também na Figura 1.

TABELA 2 Valores médios dos pesos observados de 160 animais em função da idade.

Idade (em dias)	Peso (em kg)
0	35,00
168	189,97
214	213,76
243	243,64
275	281,12
311	324,39
344	363,90
377	404,22
414	443,92
472	489,51
508	518,54
542	541,34
594	559,21
624	583,81
660	611,76
692	639,24

Fonte: Emiliano (2013)

Os modelos Gompertz, Logístico e Von Bertalanffy serão ajustados aos dados considerando uma estrutura de erros autorregressivos de primeira ordem

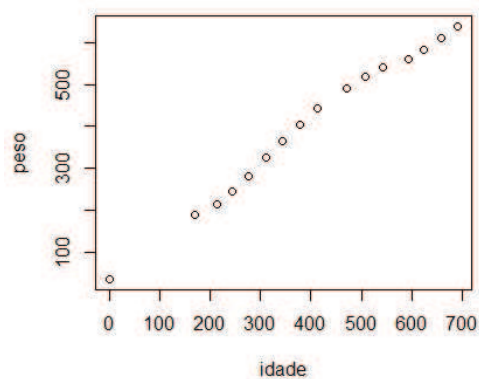


Figura 1 Valores médios reais de peso em função da idade de 160 animais da raça Hereford

regular e irregular (observações não igualmente espaçadas). Considerando, então, as parametrizações apresentadas na Tabela 1, tem-se as seguinte estruturas dos resíduos:

$$e_i = \varphi e_{i-1} + a_i$$

$$e_i = \varphi^{\Delta_i} e_{i-1} + a_i$$

em que Δ_i representa o intervalo de tempo entre duas observações subsequentes e a_i é uma variável aleatória com média 0 e variância constante σ^2 , chamada ruído branco.

Inicialmente cada um dos modelos, Logístico, Gompertz e Von Bertalanffy será ajustado aos dados reais utilizando-se a função *gnls* do software R considerando-se uma estrutura regular e independente dos resíduos. Em seguida os resíduos serão tomados por meio da diferença entre o valor observado e o valor ajustado. Obtidos os resíduos, será então, considerada a estrutura regular e irregular destes para o ajuste dos modelos AR(1) e IS-AR(1), respectivamente.

As somas de quadrados para o caso regular (SQR) e irregular (SQIR), nesta sequência, são dadas pelas equações (6) e (7), nesta ordem.

$$\text{SQR} = \sum_{i=0}^{n-1} (e_{t_{i+1}} - \varphi e_{t_i})^2 \quad (6)$$

$$\text{SQIR} = \sum_{i=0}^{n-1} (e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i})^2. \quad (7)$$

No caso de resíduos irregulares (equação (7)) utilizando o método de discretização do espaço paramétrico, minimiza-se (7) para os resíduos de cada um dos modelos. Este método consiste em testar “todas” as possibilidades para φ dentro do espaço paramétrico e verificar qual proporciona a solução desejada. Para isso, será feita a discretização do espaço paramétrico, a saber o intervalo real $] - 1, 1[$, em números numa escala de 0,0001. Dessa forma tem-se o conjunto $\{-0,9999, -0,9998, \dots, 0,9999\}$ como espaço paramétrico.

3.1 Reparametrização do modelo IS-AR(1)

Ao considerar a estrutura irregular dos resíduos, insere-se no modelo estatístico potências cujos expoentes são números muito grandes e as bases são números muito pequenos. Isso torna o processo computacional de cálculo inviável ou muitas vezes impossível, uma vez que computacionalmente há um limite máximo de casas decimais suportado. Após esse limite os valores são reconhecidos como sendo zero. Com o objetivo de evitar tal fato uma modificação no modelo é proposta.

Seja então o modelo original IS-AR(1) dado pela equação (5) e o processo de minimização dado pela equação (7). Ao invés de trabalhar com as distâncias reais entre os pontos, trabalhar-se-á com uma transformação linear das distâncias

dada por

$$\delta_i = \frac{\Delta_i}{\min(\Delta_i)}, \quad (8)$$

em que $\min(\Delta_i)$ refere-se à menor distância entre observações consecutivas. As novas expressões serão respectivamente:

$$e_{t+\delta_i} = \varphi^{\delta_i} e_t + \sigma_{\delta_i} \epsilon_{t+\delta_i}, \quad (9)$$

$$\sum_{i=0}^{n-1} (e_{t_{i+1}} - \varphi^{\delta_i} e_{t_i})^2. \quad (10)$$

Essa transformação reduzirá, em muitos casos, o espaço paramétrico de φ para $0 < \varphi < 1$ uma vez que se tem expoentes fracionários. Para encontrar o valor que minimiza a soma dada pela equação (10) utiliza-se também o método de discretização do espaço paramétrico, sendo que a escala utilizada para tal é a mesma usada para o caso regular.

Após ajustados, os modelos serão comparados por meio do critério de Akaike para identificação do melhor ajuste. As estimativas dos parâmetros do modelo melhor ajustado serão utilizadas para fazer simulação de dados para cada um dos modelos, a saber Logístico, Gompertz e Von Betalanffy. Para cada um deles será gerada uma simulação de peso animal do 1^o ao 730^o dia após o nascimento ($n = 730$). Em cada simulação os erros serão gerados como autorregressivos de primeira ordem.

A cada simulação ($n = 730$) ajustamos o respectivo modelo; por exemplo, para a simulação feita com o modelo Logístico, a curva Logística será ajustada considerando as diferentes estruturas de erros (regular e irregular). Em seguida serão retiradas, sucessivamente, amostras aleatórias de tamanho cem das simulações e ajustado o modelo para as n observações restantes. A cada retirada de uma

amostra os modelos serão novamente ajustados e comparados. Isso será feito com o objetivo de verificar a eficiência do método considerado neste trabalho.

4 RESULTADOS E DISCUSSÃO

Para ajuste dos modelos e estimação de seus parâmetros foi utilizada a função *gnls* do software R. Após ajustados os modelos, foram obtidos os resíduos de cada ajuste e, então, tratados como autorregressivos de primeira ordem (AR(1)). As estruturas regular e irregular dos resíduos foram consideradas para a estimação do parâmetro autorregressivo.

A Tabela 3 apresenta os valores estimados dos parâmetros que minimizam as somas de quadrados dos resíduos para cada um dos modelos,

TABELA 3 Valores estimados dos parâmetros para o ajuste dos modelos aos dados reais considerando as estruturas regular e irregular, e as somas de quadrados dos resíduos em ambos os casos.

Parâmetros	Modelos		
	Logístico	Gompertz	Von Bertalanffy
β_1	660,854	745,4478	819,245
β_2	9,270	1,069079	0,668134
β_3	0,00702	0,004071	0,00301
φ	-0,0758	0,1442	0,4236
SQR	1465,192	1142,094	1059,316
$\varphi(\text{IR})$	0,4477	0,3632	0,4695
SQIR	1295,272	1051,974	1050,843

em que $\varphi(\text{IR})$ é a estimativa do parâmetro autorregressivo quando é considerada a estrutura irregular dos resíduos, SQIR é a soma de mínimos quadrados quando considera-se essa mesma estrutura, φ é o parâmetro autorregressivo quando considerada uma estrutura regular dos resíduos e SQR é a soma de mínimos quadrados para esta estrutura. Em cada ajuste considerando as diferentes estruturas dos resíduos os valores de β_1 , β_2 , e β_3 são os mesmos, alterando apenas os valores dos parâmetros autorregressivos.

Observou-se em cada modelo qual a estrutura de erros proporcionou melhor ajuste observando-se a soma de quadrado de resíduos e também o critério de informação de Akaike (AIC).

A Tabela 4 apresenta os valores do AIC em cada um dos casos citados anteriormente para os dados reais. Percebeu-se que quando se considera a estrutura irregular dos resíduos há uma melhora no ajuste dos modelos, no caso do modelo Logístico a soma de quadrados de resíduos chegou a diminuir 12 %. As estimativas dos parâmetros autorregressivos também mudaram quando foi considerada a estrutura irregular chegando, no caso do modelo Logístico, a mudar de sinal (Tabela (3)). Em todos os ajustes o valor do AIC foi menor para os modelos com erros irregulares, seguindo a mesma tendência das somas de quadrados, indicando assim que houve melhor ajuste quando se considera esta estrutura.

Fez-se então uma simulação de dados de crescimento utilizando-se as estimativas obtidas para os parâmetros dos modelos que melhor se ajustaram aos dados. Nas figuras 2, 3 e 4 estão apresentadas as simulações para os modelos Logístico, Gompertz e Von Bertalanffy, respectivamente.

TABELA 4 Valores do AIC (dados reais) quando foram ajustados os diferentes tipos de modelos considerando as estruturas regular e irregular dos erros.

Tipos de erros	Modelos		
	Logístico	Gompertz	Von Bertalanffy
regulares	5,653823	5,4047	5,32946
irregulares	5,530558	5,3225	5,32143

Cada simulação contém pesos que variam do primeiro até o septuagésimo trigésimo dia após o nascimento, totalizando 730 valores.

Os dados simulados foram ajustados a cada um dos seus respectivos modelos considerando-se as diferentes estruturas. Neste caso os ajustes foram coin-

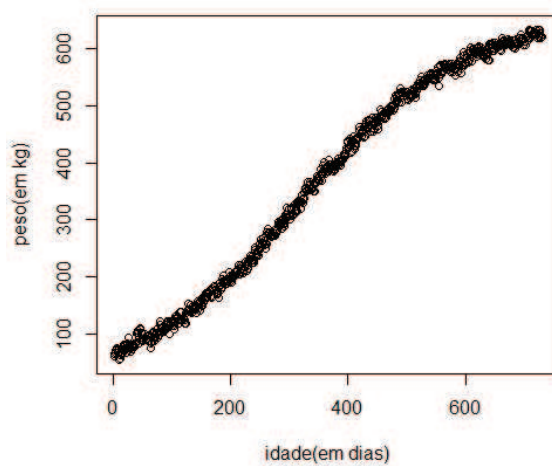


Figura 2 Simulação de dados de peso animal, do 1^o ao 730^o dia após o nascimento, para o modelo Logístico

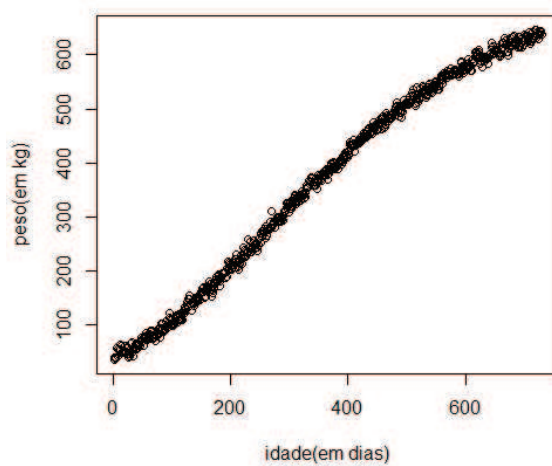


Figura 3 Simulação de dados de peso animal, do 1^o ao 730^o dia após o nascimento, para o modelo Gompertz

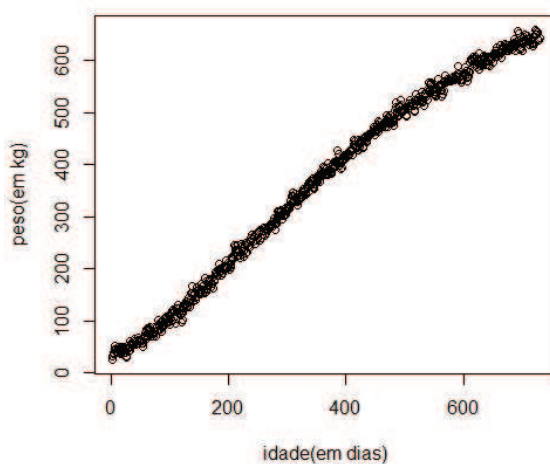


Figura 4 Simulação de dados de peso animal, do 1^o ao 730^o dia após o nascimento, para o modelo Von Bertalanffy

cientes uma vez que as observações são igualmente espaçadas. Isso mostra que o modelo IS-AR(1) pode ser visto como uma extensão do modelo AR(1) usual. Foram retiradas, sucessivamente, sete amostras aleatórias de tamanho 100 de cada uma das simulações. A cada retirada, para amostras de tamanho n (730, 630, 530, 430, 330, 230, 130, 30), ajustou-se os respectivos modelos considerando-se as diferentes estruturas dos resíduos. Objetivou-se com isso verificar se a observância de tais estruturas melhora a qualidade do ajuste dos modelos. As Tabelas 5, 6 e 7 apresentam os resultados de cada ajuste à medida em que se retirou as amostras.

Percebe-se nas Tabelas 5, 6 e 7 que quando o ajuste foi feito com todos os dados simulados as estimativas dos parâmetros foram próximas das estimativas reais. À medida que foram sendo tiradas as amostras as estimativas foram se distanciando e oscilando em torno das estimativas reais; porém em todos os casos, ou quase todos, as estimativas do modelo irregular foram mais precisas. Apenas para o modelo Gompertz (Tabela (6)), quando $n = 30$ ou 630, as somas de quadrados

TABELA 5 Valores estimados para os parâmetros do modelo Logístico à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.

Valores reais							
$\beta_1=660,854$	$\beta_2=9,27$	$\beta_3=7,02e^{-3}$	$\varphi=0,4477$				
Valores estimados							
n	β_1	β_2	β_3	φ (IR)	SQIR	φ	SQR
30	657,45	9,23	$7,104 e^{-3}$	0,5572	1343,83	0,0664	1367,37
130	655,83	9,48	$7,134 e^{-3}$	0,6006	7221,61	0,1879	7401,05
230	659,69	9,42	$7,063 e^{-3}$	0,4852	12883,51	0,2248	13148,06
330	658,79	9,52	$7,113 e^{-3}$	0,4608	18661,40	0,3094	19348,22
430	659,53	9,41	$7,077 e^{-3}$	0,4935	23437,34	0,3950	23834,60
530	659,54	9,41	$7,084 e^{-3}$	0,4599	28402,12	0,3929	28927,41
630	660,09	9,42	$7,079 e^{-3}$	0,4430	32306,11	0,4078	32899,46
730	660,23	9,40	$7,072 e^{-3}$	0,4348	36300,44	0,4348	36300,44

foram menores para o caso regular.

As Tabelas 8, 9 e 10 apresentam os valores do AIC para os modelos com dados simulados.

Observa-se que em praticamente todos os casos a observância da irregularidade nos dados proporcionou a redução da soma de quadrados dos resíduos levando, assim, a um melhor ajuste do modelo.

TABELA 6 Valores estimados para os parâmetros do modelo Gompertz à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.

Valores reais							
$\beta_1=745,4478$	$\beta_2=1,069079$	$\beta_3=4,071e^{-3}$	$\varphi=0,3632$				
Valores estimados							
n	β_1	β_2	β_3	φ (IR)	SQIR	φ	SQR
30	740,05	1,0863	$4,164e^{-3}$	0,0001	1595,90	-0,3170	1432,01
130	737,90	1,0778	$4,148e^{-3}$	0,3576	6115,47	0,0642	6322,66
230	740,35	1,0786	$4,125e^{-3}$	0,3558	10849,65	0,0821	11436,29
330	741,93	1,0769	$4,112e^{-3}$	0,4124	15683,02	0,2368	16462,85
430	743,28	1,0737	$4,098e^{-3}$	0,3812	20172,34	0,2499	21359,56
530	743,88	1,0714	$4,088e^{-3}$	0,4037	25205,42	0,3346	25769,74
630	744,02	1,0713	$4,087e^{-3}$	0,3879	29699,87	0,3636	29688,30
730	743,242	1,0690	$4,088e^{-3}$	0,3896	34410,14	0,3896	34410,14

TABELA 7 Valores estimados para os parâmetros do modelo Von Bertalanffy à medida que foram sendo retiradas as amostras e as somas de quadrados de resíduos de cada ajuste.

Valores reais							
$\beta_1=819,245$	$\beta_2=0,668134$	$\beta_3=3,01e^{-3}$	$\varphi=0,4695$				
Valores estimados							
n	β_1	β_2	β_3	φ (IR)	SQIR	φ	SQR
30	793,61	0,6671	$3,138e^{-3}$	0,4279	1789,91	-0,0624	1824,14
130	814,92	0,6699	$3,045e^{-3}$	0,5836	7901,76	0,0847	8282,98
230	822,22	0,6663	$2,990e^{-3}$	0,4086	14403,25	0,2161	14456,50
330	821,02	0,6668	$2,997e^{-3}$	0,5140	20156,07	0,3308	20950,85
430	819,24	0,6681	$3,008e^{-3}$	0,4327	26192,30	0,3083	27193,22
530	820,55	0,6671	$2,998e^{-3}$	0,4458	32004,60	0,3671	32639,09
630	820,82	0,6674	$2,997e^{-3}$	0,4827	35930,24	0,4479	36059,00
730	820,89	0,6677	$2,999e^{-3}$	0,4647	39551,49	0,4647	39551,49

TABELA 8 Valores do AIC para o modelos Logístico à medida que foram sendo retiradas as amostras.

n	erros regulares	erros irregulares
30	4,387418	4,370053
130	4,166615	4,142071
230	4,116052	4,095726
330	4,120007	4,083864
430	4,05247	4,035662
530	4,029957	4,011631
630	3,980959	3,96276
730	3,928511	3,928511

TABELA 9 Valores do AIC para o modelo Gompertz à medida que foram sendo retiradas as amostras.

n	erros regulares	erros irregulares
30	4,433608	4,541967
130	4,009133	3,975815
230	3,976569	3,92391
330	3,958513	3,909985
430	3,942831	3,885644
530	3,914368	3,892226
630	3,878256	3,878646
730	3,875033	3,875033

TABELA 10 Valores do AIC para o modelo Von Bertalanffy à medida que foram sendo retiradas as amostras.

n	erros regulares	erros irregulares
30	4,675638	4,656694
130	4,279196	4,232079
230	4,210921	4,207231
330	4,199586	4,160912
430	4,1843	4,146797
530	4,150678	4,131047
630	4,07266	4,069082
730	4,014285	4,014285

5 CONCLUSÃO

Com as análises realizadas verifica-se que, apesar da baixa escala de melhora, considerar a estrutura irregular dos resíduos leva a menores valores de AIC e a menores somas de quadrados de resíduos. No caso real o modelo Von Bertalanffy foi o que proporcionou a menor soma de quadrado dos resíduos. Nos dados simulados os modelos se ajustaram bem aos dados. À medida que foram sendo retiradas as amostras as estimativas foram se distanciando das reais. Mesmo assim, as estimativas obtidas quando se leva em consideração a irregularidade foram mais próximas das reais. O que nos leva a constatar mais uma vez a eficiência do método utilizado. De acordo com o Critério de Informação de Akaike, nos dados simulados o modelo Gompertz foi o que obteve melhor desempenho em praticamente todos os casos. Isso mostra que cada situação deve ser considerada de maneira individual. Nem sempre um modelo será o melhor no ajuste de diferentes tipos de dados ou em um mesmo tipo de dados com diferentes comportamentos.

REFERÊNCIAS

BERTALANFFY, L. V. Quantitative laws in metabolism and growth. **Biological Research**, Los Angeles, v. 32, n. 3, p. 217-231, Sept. 1957.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. Ockland: Holden-Day, 1976. 575p.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Assessoria de Gestão Estratégica. **Brasil projeções do agronegócio 2011/2012 a 2021/2022**. Disponível em: <[http://www.agricultura.gov.br/arq_editor/file/Ministerio/gestao/projecao/ProjecoeddoAgronegocioBrasil2011-20012a2021-2022\(2\)\(1\).pdf](http://www.agricultura.gov.br/arq_editor/file/Ministerio/gestao/projecao/ProjecoeddoAgronegocioBrasil2011-20012a2021-2022(2)(1).pdf)>. Acesso em: 28 jan. 2014.

COSTA, A. C. et al. Avaliação do crescimento de tilápias de diferentes linhagens através de modelos não lineares. **Archivos de Zootecnia**, Cordoba, v. 58, n. 1, p. 561-564, 2009. Suplemento.

ECKNER, A. **A Framework for the Analysis of unevenly spaced time series data**. Disponível em: <<http://www.eckner.com/research.html>>. Acesso em: 22 abr. 2013a.

ECKNER, A. **A Note on Trend and Seasonality Estimation for Unevenly-Spaced Time Series**. Disponível em: <<http://www.eckner.com/research.html>>. Acesso em: 22 abr. 2013b.

EMILIANO, P. C. **Critérios de informação: como eles se comportam em diferentes modelos?** 2013. 179p. Tese(Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2013.

ERDOGAN, E. et al. Statistical models for unequally spaced time series. In: **SIAM INTERNATIONAL CONFERENCE DATA MINING, 2005**, Auckland. **Proceedings...** Auckland: University of Auckland, 2005. p. 626-630.

ESQUIVEL, R. M. **Análise espectral singular**: modelagem de séries temporais através de estudos comparativos usando diferentes estratégias de previsão. 2012. 160 p. Dissertação (Mestrado em Modelagem Computacional e Tecnologia Industrial) - Serviço Nacional de Aprendizagem Industrial, Salvador, 2012.

FERNANDES, T. J. **Curvas de crescimento do fruto do cafeeiro em diferentes alinhamentos de plantio utilizando modelos não lineares**. 2012. 80 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2012.

GALLANT, A. R. **Nonlinear statistical models**. Ottawa: Willey, 1987, 610 p.

GÓMEZ, D. A. A.; MUÑOZ, M. F. C.; BETANCUR, L. F. R. Modelación de las funciones de crecimiento aplicadas a la producción animal. **Revista Colombiana de Ciências Pecuárias**, Medellín, v. 21, n. 1, p. 39-58, mar. 2008.

HENRIQUES, M. B.; CASARINI, L. M. Avaliação do mexilhão *Perna perna* e da espécie invasora *Isognomon bicolor* em banco natural da Ilha das Palmas, Baía de Santos, Estado de São Paulo, Brasil. **Boletim do Instituto de Pesca**, São Paulo, v. 35, n. 4, p. 577-586, 2009.

LOAIZA-ECHEVERRI, A. M., Utilização de modelos não lineares para descrever o crescimento do perímetro escrotal em ruminantes. **Veterinaria y Zootecnia**, Manizales, v. 5, n. 2, p. 9-23, 2011.

LOPES, F. B. et al. Ajustes de curvas de crescimento em bovinos Nelore da região norte do Brasil. **Revista Brasileira de Saúde e Produção Animal**, Salvador, v. 12, n.3, p. 607-617, jul./set. 2011.

MALTHUS, T. R. **An Essay on the Principle of Population**. London: Church-yard, 1798. Disponível em: <<http://www.econlib.org/library/Malthus/malPop.html>>. Acesso em: 9 maio 2014.

MANSANO, C. F. M. **Características do crescimento morfométrico e composição corporal de girinos de Rã-touro em cativeiro**. 2012. 87 p. Dissertação (Mestrado em Agricultura) - Universidade Estadual Paulista, Jaboticabal, 2012.

MAZUCHELI, J.; SOUZA, R. M. de; PHILIPPSEN, A. S. Modelo de crescimento de Gompertz na presença de erros normais heterocedásticos: um estudo de caso. **Revista Brasileira de Biometria**, São Paulo, v. 29, n. 1, p. 91-101, 2011.

MAZZINI, A. R. de A. et al. Análise da curva de crescimento de machos Hereford. **Ciência e Agrotecnologia**, Lavras. v. 27, n. 5, p. 1105-1112, set./out. 2003.

MAZZINI, A. R. de A. et al. Curva de crescimento de novilhos Hereford: heterocedasticidade e resíduos autorregressivos. **Ciência Rural**, Santa Maria, v. 35, n. 2, p. 422-427, mar./abr. 2005.

MCMANUS, C. M.; LOUVANDINI, H.; CAMPOS, V. A. L. Non linear growth curves for weight and height in four genetic groups of horses. **Ciência Animal Brasileira**, Goiânia, v. 11, n. 1, p. 80-89, jan./mar. 2010.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. rev. e ampl. São Paulo: E. Blucher, 2006. 560 p.

OLIVEIRA, H. N. de; LÔBO, R. B.; PEREIRA, C. S. Comparação de modelos não lineares para descrever o crescimento de fêmeas da raça Guzera. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 9, p. 1843-1851, set. 2000.

PRADO, T. K. L. do. **Modelos não lineares no crescimento de frutos de coqueiro**. 2011. 91 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2011.

SILVA, F. de L. **Curvas de crescimento e produtividade de vacas Nelore e cruzadas, de diferentes tipos biológicos, em sistemas de produção intensiva**. 2009. 88 p. Dissertação (Mestrado em Ciências. Área de concentração: Ciência animal e pastagens) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2009.

SILVA, N. A. M. et al. Proposta de análise genética de curvas de crescimento de bovinos por meio do algoritmo SAEM. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 64, n. 5, p. 1256-1264, 2012.

SILVEIRA, F. G. da. **Classificação multivariada de modelos de crescimento para grupos genéticos de ovinos de corte**. 2010. 61f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, MG, 2010.

SOUZA, L. de A. et al. Curvas de crescimento em bovinos da raça Indubrasil criados no estado do Sergipe. **Ciência Agrônômica**, Fortaleza, v. 41, n. 4, p. 671-676, out./dez. 2010.

TEDESCHI, L. O. et al. Estudo da curva de crescimento de animais da raça Guzera e seus cruzamentos alimentados a pasto, com e sem suplementação. 1. Análise e seleção das funções não-lineares. **Revista brasileira de zootecnia**, Brasília, v. 29, n. 2, p. 630-637, 2000.

VELOSO, R. de C. et al. Curvas de Crescimento de Genótipos de Frangos Caipiras. In: SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 9., 2012, João Pessoa. **Anais...** João Pessoa: SBM, 2012. Disponível em: <<http://sbmaonline.org.br/anais/ix/trabalhos/>>. Acesso em: 28 jan 2014.

ZEVIANE, W. M. **Parametrizações interpretáveis em modelos não lineares**. 2013. 146 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2013.

APÊNDICE

Nesta seção são apresentadas as demonstrações para a obtenção do estimador de máxima verossimilhança (EMV) de σ , da variância de X_{t+h} e da $cov(X_t, X_{t+h})$, respectivamente. Demonstração do procedimento para encontrar o estimador de σ utilizando o estimador de máxima verossimilhança.

A função de verossimilhança para o modelo (5) é dada por

$$L(\sigma|e, \varphi) = \prod_{i=0}^{n-1} \left(\frac{1}{\sqrt{2\pi\sigma^2 \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)}} \right) \exp \left(-\frac{1}{2} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\sigma^2 \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \right)$$

Tomando o logaritmo natural da função de verossimilhança obtém-se

$$\begin{aligned} l(L(\sigma|e, \varphi)) &= \sum_{i=0}^{n-1} \ln \left(\frac{1}{\sqrt{2\pi\sigma^2 \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)}} \right) - \frac{1}{2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\sigma^2 \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \\ &= \sum_{i=0}^{n-1} \ln \left(\frac{1}{\sqrt{2\pi \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right) \sigma}} \right) - \frac{1}{2\sigma^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \\ &= \sum_{i=0}^{n-1} -\ln \left(\sqrt{2\pi \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right) \sigma} \right) - \frac{1}{2\sigma^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \\ &= \sum_{i=0}^{n-1} - \left\{ \ln \left(\sqrt{2\pi \left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \right) + \ln(\sigma^2)^{1/2} \right\} - \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} \end{aligned}$$

$$= - \sum_{i=0}^{n-1} \ln \left(\sqrt{2\pi \left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)} \right) - n \ln (\sigma^2)^{1/2} - \\ - \frac{1}{2\sigma^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

Para obter-se o valor de σ que maximiza a função $l(L(\sigma|e, \varphi))$ tem-se a derivada

$$\frac{dl(L(\sigma|e, \varphi))}{d\sigma^2} = -n \frac{1}{(\sigma^2)^{1/2}} \frac{1}{2} (\sigma^2)^{-1/2} - \frac{1}{2} (-1) \frac{1}{(\sigma^2)^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

Igualando a 0 obter-se o valor crítico $\hat{\sigma}$

$$\frac{dl(L(\sigma|e, \varphi))}{d\sigma^2} = 0$$

$$0 = -n \frac{1}{(\sigma^2)^{1/2}} \frac{1}{2} (\sigma^2)^{-1/2} - \\ - \frac{1}{2} (-1) \frac{1}{(\sigma^2)^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

$$n \frac{1}{(\sigma^2)^{1/2}} \frac{1}{2} (\sigma^2)^{-1/2} = \frac{1}{2} \frac{1}{(\sigma^2)^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

$$n \frac{1}{2\sigma^2} = \frac{1}{2} \frac{1}{(\sigma^2)^2} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

$$n \frac{2(\sigma^2)^2}{2\sigma^2} = \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1 - \varphi^{2\Delta_i}}{1 - \varphi^2} \right)}$$

$$\sigma^2 = \frac{1}{n} \sum_{i=0}^{n-1} \frac{[e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)}$$

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{\left(\frac{1-\varphi^{2\Delta_i}}{1-\varphi^2}\right)} [e_{t_{i+1}} - \varphi^{\Delta_i} e_{t_i}]^2}$$

Demonstração da fórmula da variância da variável aleatória e_{t+h} escrita em função da variável aleatória e_t .

$$\begin{aligned} \text{var}(e_{t+h}) &= \text{var}\left[\varphi^h e_t + \sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}\right] \\ &= \text{var}\left[\varphi^h e_t\right] + \text{var}\left[\sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}\right] + \\ &\quad + 2\text{cov}\left[\varphi^h e_t, \sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}\right] \xrightarrow{0} \\ &= \left(\varphi^h\right)^2 \text{var}[e_t] + \sigma^2 \text{var}\left[\sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}\right] \\ &= \left(\varphi^h\right)^2 \text{var}[e_t] + \sigma^2 \left[\sum_{j=0}^{h-1} \text{var}\left[\varphi^{h-1-j} \epsilon_{t+1+j}\right] + \right. \\ &\quad \left. + 2 \sum_{j \neq k} \text{cov}\left[\varphi^{h-1-j} \epsilon_{t+1+j}, \varphi^{h-1-k} \epsilon_{t+1+k}\right]\right] \xrightarrow{0} \\ &= \left(\varphi^h\right)^2 \text{var}[e_t] + \sigma^2 \sum_{j=0}^{h-1} \left(\varphi^{h-1-j}\right)^2 \text{var}[\epsilon_{t+1+j}] \end{aligned}$$

$$\begin{aligned}
&= (\varphi^h)^2 \text{var}[e_t] + \sigma^2 \sum_{j=0}^{h-1} (\varphi^{h-1-j})^2 \\
&= (\varphi^h)^2 \text{var}[e_t] + \sigma^2 \left(\frac{1 - \varphi^{2h}}{1 - \varphi^2} \right)
\end{aligned}$$

Demonstração da fórmula da covariância entre e_t e e_{t+h} .

$$\begin{aligned}
\text{cov}(e_t, e_{t+h}) &= E[(e_t - E[e_t])(e_{t+h} - E[e_{t+h}])] \\
&= E[e_t e_{t+h} - e_t E[e_{t+h}] - E[e_t] e_{t+h} + \\
&\quad + E[e_t] E[e_{t+h}]] \\
&= E[e_t e_{t+h}] \\
&= E[e_t (\varphi^h e_t + \sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j})] \\
&= E[\varphi^h e_t^2 + e_t \sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}] \\
&= E[\varphi^h e_t^2] + E[e_t \sigma \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}] \\
&= \varphi^h E[e_t^2] + \sigma E[e_t \sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j}] \\
&= \varphi^h E[e_t^2] + \sigma E[e_t] E \left[\sum_{j=0}^{h-1} \varphi^{h-1-j} \epsilon_{t+1+j} \right] \\
&= \varphi^h \text{var}[e_t]
\end{aligned}$$