



IÁBITA FABIANA SOUSA

**AJUSTE DE MODELOS NÃO LINEARES NA
DESCRIÇÃO DE GERMINAÇÃO DE SEMENTES DE
CAFÉ (*Coffea arabica* L.) CV. CATUAÍ**

LAVRAS - MG

2012

IÁBITA FABIANA SOUSA

**AJUSTE DE MODELOS NÃO LINEARES NA DESCRIÇÃO DE
GERMINAÇÃO DE SEMENTES DE CAFÉ (*Coffea arabica* L.) CV.
CATUAÍ**

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.

Orientador

Dr. Joel Augusto Muniz

Coorientadora

Dra. Taciana Villela Savian

LAVRAS - MG

2012

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

Sousa, Iábida Fabiana.

Ajuste de modelos não lineares na descrição de germinação de sementes de café (*coffea arábica* L.) cv. Catuaí / Iábida Fabiana Sousa.–
Lavras : UFLA, 2012.

72 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2012.

Orientador: Joel Augusto Muniz.

Bibliografia.

1. Curva de crescimento. 2. Erros correlacionados. 3. Análise de regressão. 4. Regressão não linear. I. Universidade Federal de Lavras.
II. Título.

CDD - 519.536

IÁBITA FABIANA SOUSA

**AJUSTE DE MODELOS NÃO LINEARES NA DESCRIÇÃO DE
GERMINAÇÃO DE SEMENTES DE CAFÉ (*Coffea arabica* L.) CV.
CATUAÍ**

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 13 de fevereiro de 2012.

Dr. Augusto Ramalho de Morais

UFLA

Dr. Virgílio Anastácio da Silva

UFLA

Dr. Joel Augusto Muniz
Orientador

**LAVRAS - MG
2012**

*Aos meus familiares, em especial minha mãe, Maria José e meus avós José Raimundo e Raimunda, pelo amor, apoio e força em todos os momentos ;
Aos meus irmãos, Ana Cristina, Antonio Alberto e Albertino Neto pelo apoio e amizade;
A minha bisavó, Francisca (in memoriam) pelo amor, pelos momentos compartilhados.*

DEDICO.

AGRADECIMENTOS

Acima de tudo, agradeço a Jesus Cristo por sua presença em minha vida, mostrando-me, muitas vezes, que durante a angústia e o cansaço, ainda vale a pena. Obrigada por seu amor, força e sabedoria ao longo dessa caminhada.

Ao meu orientador, Dr. Joel Augusto Muniz, pelo acolhimento, apoio, estímulo, confiança e disponibilidade durante todo o desenvolvimento da orientação.

A professora Dr^a. Taciana Villela Savian, por estar sempre disponível a esclarecer dúvidas.

À minha família maravilhosa: minha mãe, Maria José; meus irmãos Ana Cristina, Antonio Alberto e Albertino Neto; meus avós, tios, primos e afilhados pelo apoio, carinho e incentivo que mesmo de longe estavam sempre presentes.

À Universidade Federal de Lavras, que me proporcionou cursar o mestrado em Estatística e Experimentação Agropecuária.

Aos amigos da UFLA, Carolina, Juliano, Mariele, Adriele, Adriana, André, Elayne, Tábata, Marcelo, Leandro, Juracy, Larissa, Daniele, Lourenço, Juscélia, Siomara, Jair Rocha, Diogo, Jair Wyzykowsky. Obrigada pelo aprendizado e trocas de informações, pelas reuniões, viagens, passeios. Agradecimento especial às irmãs de orientação Adriana Dias, Thais e Diana ao irmão de orientação Walmes Zeviani pelas dúvidas tiradas e por sempre me socorrer com o R. Guardarei em meu coração todos os momentos de alegria compartilhados com todos vocês.

Aos professores da Pós-Graduação pelos conhecimentos transmitidos e pela disposição em ajudar. Só eu sei o quanto aprendi com vocês.

A todos os funcionários do DEX, em especial, as "Josis" e Selminha. Exemplos de cordialidade e profissionalismo.

Ao professor Dr. Renato Mendes, Johan e aos funcionários do Laboratório de Análise de Sementes por contribuir na realização do experimento desse estudo.

À SEDUC-PI, pelo apoio financeiro e institucional.

À CAPES, pela concessão de bolsa de estudos.

Aos amigos e irmãos da Primeira Igreja Batista de Parnaíba e Primeira Igreja Presbiteriana de Lavras, pelas palavras de encorajamento, e pelas muitas orações.

Aos amigos mais chegados que irmãos, Ana Sales, Ana Patricia, Eugênia, Marina, Patricia Rufino, Thalita, Rutinha, Barbara, Daniel, Letícia, Amanda, Juliany, Delva, Aracy, Rosângela, Maria, Rejane, Davi, Lucas, Theia, Ana Maria, Célio, Alba Sena, Jerfeson, Elica, Samara e Lucas. Em especial Fábio Azevedo por compartilhar tantos momentos especiais. Muito obrigada mesmo pelas vossas amizades.

Às irmãs de república Carol, Thaís e Ana Flávia. Obrigada por tudo.

A todos aqueles que de alguma maneira contribuíram, ou estiveram na torcida pela realização deste trabalho.

E que venha o futuro!

O conhecimento dá ocasião à arrogância, mas o amor edifica. Se alguém supõe conhecer alguma coisa, ainda não conhece até o ponto em que é necessário conhecer.

1 Coríntios 8, 1b-2

RESUMO

Sementes de café com qualidade fisiológica, de procedência conhecida e com alto desempenho germinativo é fundamental para a obtenção de mudas vigorosas. Na avaliação de sementes, o estudo da curva de germinação pode contribuir para melhor entendimento do processo de germinação. O objetivo desse estudo foi avaliar a qualidade do ajuste dos modelos Logístico e Gompertz, com estrutura de erros independentes e autorregressivos, com autocorrelação de primeira ordem, AR(1), na descrição de germinação de sementes de café (*Coffea arabica*L.) linhagem Catuaí vermelho IAC 99, em cinco diferentes percentuais de germinação. A estimação dos parâmetros para os modelos foi feita pelos métodos de mínimos quadrados e pelo processo iterativo de Gauss-Newton, utilizando-se a função gnls do pacote nlme do programa R versão 2.13.1. A seleção do melhor modelo, para descrever o processo germinativo, teve como base a precisão dos ajustes baseados no máximo da função de verossimilhança (MFV) através do teste de razão de máxima verossimilhança utilizado para modelos encaixados, critério de informação de Akaike (AIC) e critério Bayesiano de Schwarz (BIC), além dos avaliadores de qualidade de ajuste (coeficiente de determinação ajustado e desvio padrão residual). Os dados utilizados foram provenientes de um experimento conduzido no ano de 2011 no Laboratório de Análises de Sementes da Universidade Federal de Lavras. Os modelos não lineares Logístico e Gompertz apresentaram-se adequados para ajuste da porcentagem de germinação. O modelo Gompertz com estrutura de erros autorregressivo de ordem 1 apresentou-se como o melhor para descrever o processo germinativo ao longo do tempo.

Palavras-chave: Modelos de crescimento. Erros autocorrelacionados. Regressão não linear. Germinação de sementes de café.

ABSTRACT

Seeds of coffee with physiological quality and known origin with high germinative performance becomes mandatory to achieve homogeneous germination of vigorous seedlings. The germination curve may contribute for better understanding of the process in seed analysis studies. The objective of this study was to evaluate the fit of Gompertz and Logistic models, with independent structure and autocorrelated errors, with first-order autocorrelation, AR (1), in description of the germination of coffee seeds *Coffea arabica* red line Catuaí IAC 99 in five different percentages of germination. The estimation of parameters for the models was made by Least Squares method and the iterative Gauss-Newton, using the function `gnls` Package `nlme` of the R program version 2.13.1. The selection of the best model to describe the germination process was based on the accuracy of the adjustments based on the maximum likelihood function (MLF) by testing maximum likelihood ratio used for nested models, Akaike information criterion (AIC) and Schwarz Bayesian criterion (BIC), and evaluators goodness of fit (adjusted coefficient of determination and residual standard deviation). The data used were taken from an experiment that was carried out in 2011 in Seeds Laboratory Analysis of the Federal University of Lavras. The nonlinear Logistic and Gompertz models were adequate to adjust the percentage of germination. The Gompertz model with order 1 autoregressive error structure seemed to be the most suitable to describe the germination process over time.

Key words: Growth models. Autocorrelated errors. Non linear regression. Germination of coffee seeds.

LISTA DE FIGURAS

Figura 1	Instalação do processo de germinação	32
Figura 2	Instalação do processo de germinação	32
Figura 3	Sementes de café germinadas aos 15 dias - Lote III	32
Figura 4	Sementes de café germinadas aos 21 dias - Lote V	32
Figura 5	Sementes de café germinadas aos 30 dias - Lote I	32
Figura 6	Sementes de café germinadas aos 30 dias - Lote IV	32
Figura 7	Análise gráfica dos resíduos do modelo Logístico para o lote I . .	41
Figura 8	Análise gráfica dos resíduos do modelo Logístico para o lote II .	41
Figura 9	Análise gráfica dos resíduos do modelo Logístico para o lote III .	42
Figura 10	Análise gráfica dos resíduos do modelo Logístico para o lote IV .	42
Figura 11	Análise gráfica dos resíduos do modelo Logístico para o lote V .	43
Figura 12	Análise gráfica dos resíduos do modelo Gompertz para o lote I .	43
Figura 13	Análise gráfica dos resíduos do modelo Gompertz para o lote II .	44
Figura 14	Análise gráfica dos resíduos do modelo Gompertz para o lote III	44
Figura 15	Análise gráfica dos resíduos do modelo Gompertz para o lote IV	45
Figura 16	Análise gráfica dos resíduos do modelo Gompertz para o lote V .	45
Figura 17	Modelo Logístico na descrição de germinação de sementes de café - Lote I, em função do tempo com estruturas de erros inde- pendentes	48
Figura 18	Modelo Logístico na descrição de germinação de sementes de café - Lote II, em função do tempo com estruturas de erros inde- pendentes	48
Figura 19	Modelo Logístico na descrição de germinação de sementes de café - Lote III, em função do tempo com estruturas de erros in- dependentes	49
Figura 20	Modelo Logístico na descrição de germinação de sementes de café - Lote IV, em função do tempo com estruturas de erros in- dependentes	49
Figura 21	Modelo Logístico na descrição de germinação de sementes de café - Lote V, em função do tempo com estruturas de erros inde- pendentes	50
Figura 22	Histograma representativo da taxa de germinação absoluta (TGA) de sementes de café estimada pela função Logística nos cinco lotes	52
Figura 23	Modelo Gompertz na descrição de germinação de sementes de café- Lote I, em função do tempo com estruturas de erros inde- pendentes	53

Figura 24	Modelo Gompertz na descrição de germinação de sementes de café - Lote II, em função do tempo com estruturas de erros independentes	54
Figura 25	Modelo Gompertz na descrição de germinação de sementes de café - Lote III, em função do tempo com estruturas de erros independentes	54
Figura 26	Modelo Gompertz na descrição de germinação de sementes de café - Lote IV, em função do tempo com estruturas de erros independentes	55
Figura 27	Modelo Gompertz na descrição de germinação de sementes de café - Lote V, em função do tempo com estruturas de erros independentes	55
Figura 28	Histograma representativo da taxa de germinação absoluta (TGA) de sementes de café estimada pela função Gompertz nos cinco lotes	57
Figura 29	Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico e Gompertz, com estruturas de erros AR (1) para o lote I	60
Figura 30	Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote II	61
Figura 31	Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote III	61
Figura 32	Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote IV	62
Figura 33	Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote V	62

LISTA DE TABELAS

Tabela 1	Estimativas das estatísticas dos testes de Shapiro-Wilk, Breusch-Pagan e Durbin-Watson e os respectivos valor-p dos resíduos do modelo Logístico para os cinco lotes de semente de café.	46
Tabela 2	Estimativa das estatísticas dos testes de Shapiro-Wilk, Breusch-Pagan e Durbin-Watson e os respectivos valor-p dos resíduos do modelo Gompertz para os cinco lotes de semente de café.	46
Tabela 3	Estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, aos valores de porcentagem de germinação de sementes para os cinco lotes	47
Tabela 4	Estimativas da estatística do teste de Shapiro-Wilk dos resíduos do modelo Logístico e valor - p para os cinco lotes de sementes de café.	50
Tabela 5	Estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, aos valores de porcentagem de germinação de sementes para os cinco lotes	52
Tabela 6	Estimativas da estatística do teste de Shapiro-Wilk dos resíduos e valor - p do modelo Gompertz para os cinco Lotes.	56
Tabela 7	Estatísticas do teste de razão de máxima verossimilhança (TRV) para os modelos estudados.	57
Tabela 8	Estimativas dos parâmetros para o ajuste do modelo Logístico e Gompertz, com estrutura de erros autoregressivos de ordem 1, aos valores de porcentagem de germinação de sementes para os cinco lotes.	58
Tabela 9	Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_{aj}^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC) e critério bayesiano de Schwarz (BIC) para os modelos ajustados, na descrição de germinação de sementes para os cinco lotes	59

SUMÁRIO

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	15
2.1	Caracterização e Germinação das sementes de café (<i>Coffea arabica</i> L.)	15
2.2	Modelos de regressão não linear	18
2.3	Estimação dos parâmetros de modelos não lineares	19
2.3.1	Método dos mínimos quadrados ordinários	20
2.3.2	Método dos mínimos quadrados ponderados	23
2.3.3	Método dos mínimos quadrados generalizados	24
2.3.4	Processos iterativos	26
2.4	Modelos de crescimento	27
3	MATERIAI E MÉTODOS	31
3.1	Material	31
3.2	Métodos	33
3.2.1	Modelos estudados	33
3.2.2	Crítérios para seleção do modelo	34
3.2.2.1	Coefficiente de determinação ajustado (R_{aj}^2)	35
3.2.2.2	Desvio padrão residual (DPR)	35
3.2.2.3	Crítérios de informação	36
3.3	Análise dos resíduos	37
3.4	Recursos computacionais	39
4	RESULTADOS E DISCUSSÃO	40
4.1	Ajuste dos modelos não lineares	40
4.2	Seleção dos modelos ajustados	58
5	CONCLUSÃO	63
	REFERÊNCIAS	64
	ANEXOS	71

1 INTRODUÇÃO

Num mercado ainda em franca expansão, o café contribui em torno de 7,5% do valor total das exportações brasileiras e em comparação com os demais países, representou mais de 30% da produção mundial previstas para as safras 2010/2011. Segundo o Ministério da Agricultura, Pecuária e Abastecimento/MAPA, atualmente o Brasil é considerado o maior produtor e exportador mundial de café em grãos, além de ser o responsável por mais de 8 milhões de empregos diretos e indiretos, onde as principais áreas de cultivo localizam-se nos estados de Minas Gerais, Espírito Santo, São Paulo, Bahia, Paraná, Rondônia e Rio de Janeiro (Brasil, 2011).

Entre as inúmeras espécies de café cultivadas no mundo, as mais conhecidas são: *Coffea arabica* e o *Coffea canephora*, sendo que ambas possuem um grande número de variedades e linhagens. O *coffea arabica* possui maior importância econômica, representando 70% do café comercializado no mundo. Originário do Oriente, produz cafés de melhor qualidade, mais finos e requintados, tem grãos de cor esverdeada e é cultivado em regiões com altitude média de 600m. A sua produção representa 73,9% (32,18 milhões de sacas) da produção do País, e tem como maior produtor o estado de Minas Gerais, com 67,9% (21,85 milhões de sacas) de café beneficiado de acordo com a Companhia Nacional de Abastecimento (CONAB, 2011).

Dentre as cultivares do café arábica, as linhagens do Catuaí Vermelho têm ampla capacidade de adaptação, apresentando produtividade elevada na maioria das nossas regiões cafeeiras. De baixa estatura, permitem maior densidade de plantio, tornam mais fácil a colheita e mais eficientes os tratamentos fitossanitários (GUERREIRO *et al.*, 2006).

A propagação usual do cafeeiro ocorre por meio de mudas provenientes

de sementes, nas quais exibem germinação lenta e desuniforme. A germinação é uma sequência de eventos fisiológicos influenciada por fatores externos e internos às sementes, em que cada fator pode atuar por si ou em interação com os demais. Em geral, a germinação acumulada de sementes, constitui-se em uma curva de formato sigmoideal que pode ser avaliada por meio de modelos não lineares.

Em estatística, a regressão não linear é uma forma de análise em que os dados são modelados por uma função que é uma combinação não linear de parâmetros do modelo e depende de uma ou mais variáveis independentes.

O presente trabalho teve como objetivo avaliar a qualidade do ajuste dos modelos não lineares Logístico e Gompertz, com estrutura de erros independentes e autorregressivos, com autocorrelação de primeira ordem, AR(1), aos dados de germinação de sementes de café da linhagem Catuaí vermelho IAC 99, com diferentes índices de germinação, visando identificar a melhor estrutura para a descrição do processo germinativo.

2 REFERENCIAL TEÓRICO

2.1 Caracterização e Germinação das sementes de café (*Coffea arabica* L.)

Elemento principal no estabelecimento, expansão, diversificação e desenvolvimento da agricultura, as sementes têm a função de perpetuação e multiplicação das espécies. Sob o aspecto ecológico, representam a interligação natural entre as espécies vegetais e o ambiente (MARCOS FILHO, 2005).

Detentoras de reservas nutritivas necessárias ao desenvolvimento da planta durante o início da germinação, as sementes, do ponto de vista funcional, são constituídas por um tecido de reserva endospermático ou cotiledonar e eixo embrioná-

rio, cujo crescimento está temporariamente suspenso, ambos protegidos por uma cobertura protetora constituída de tecidos vivos e mortos (POPINIGIS, 1985).

As mudas das variedades de café arabica podem ser produzidas a partir de sementes, já que a fecundação se dá em sua maioria (90 a 95%). Portanto, a obtenção de sementes de café de qualidade fisiológica de procedência conhecida e com alto desempenho germinativo tem sido considerada como o principal fator responsável pela obtenção de mudas vigorosas em condições de campo. As sementes devem ser obtidas em Fazendas Experimentais de Órgãos de Pesquisa, federais ou estaduais, ou de produtores registrados para a produção de sementes, tendo - se certeza quanto à boa origem do material, devendo ser conhecida variedade/linhagem (MATIELLO *et al.*, 2005).

Os frutos da espécie *Coffea arabica* produzem sementes viáveis quando estão fisiologicamente maduros, tendo sua descrição morfológica confundida, em muitas vezes, com a da semente. As sementes de café são elípticas ou em forma de ovo, plano convexa, possuindo um sulco longitudinal na superfície plana, constituídas por embrião, endosperma e tegumento (DEDECCA, 1957), ou seja, o fruto exceto o epicarpo (casca) e o mesocarpo (mucilagem). O tegumento da semente é um envoltório (película prateada), que reveste o endosperma. O endosperma garante as reservas energéticas durante os processos de germinação e emergência, composto basicamente de carboidratos. O embrião, parte viva da semente, é formado por radícula, hipocótilo e cotilédones. O endocarpo, também chamado de pergaminho, protege as sementes contra microorganismos e danos físicos. Segundo Black, Bewley & Halmer (2006), o grão de café apresenta teores aproximados de 6% de água, 12% de proteínas, 11% de óleos, 70% de carboidratos e 4% de minerais.

Considerada pelos botânicos como a retomada do crescimento do embrião

(NASSIF *et.al*, 1998), a fase inicial de germinação consiste na ativação dos processos fisiológicos do crescimento pela ativação do aumento do teor de umidade e da atividade respiratória da semente.

Para que a germinação ocorra determinadas condições devem ser satisfeitas, a saber (POPINIGIS, 1985):

- a) A semente deve ser viável;
- b) Condições internas das semente devem ser favoráveis à germinação;
- c) As condições ambientais devem ser favoráveis (água, temperatura, oxigênio e luz);
- d) Condições satisfatórias de sanidade (ausência de agentes patogênicos).

A germinação de sementes em teste de laboratório é a emergência e desenvolvimento das estruturas essenciais do embrião, demonstrando sua aptidão para produzir uma planta normal sob condições favoráveis de campo, no caso do café em teste padrão que leva em média 30 dias (BRASIL, 2009). O método tradicionalmente usado para avaliar a qualidade fisiológica de sementes se baseia na realização do teste de germinação em que os resultados são de grande valia para comparação entre lotes de sementes para fins de comercialização e para fins de semeadura. Os testes de germinação são altamente padronizados, permitindo, assim, a obtenção de resultados semelhantes quando executados sobre um mesmo lote, em laboratórios diferentes. (POPINIGIS, 1985)

As sementes de *Coffea arabica* apresentam germinação lenta e desuniforme. Assim, torna-se importante conhecer melhor o processo germinativo, visando reduzir o tempo de germinação que é essencial para a melhoria das práticas agrícolas e desenvolvimento da produção de café.

2.2 Modelos de regressão não linear

Em estatística, regressão é um método comumente usado que permite explorar e inferir a relação de uma variável resposta com variáveis preditoras específicas. Pode-se atingir este objetivo por meio dos bem conhecidos modelos de regressão, Drapper & Smith (1998) classificam os modelos de regressão como: lineares em relação aos parâmetros (as derivadas parciais em relação aos parâmetros não depende dos parâmetros); modelos linearizáveis, aqueles que, por meio de alguma transformação, tornam-se lineares e modelos inerentemente não lineares são modelos em que pelo menos uma das derivadas parciais depende de algum parâmetro do modelo.

Em qualquer tipo de modelagem, além do ajuste realizado, é necessário fazer inferências sobre os parâmetros em estudo. A inferência nos modelos não lineares ocorre por aproximação em Série de Taylor na região próxima às estimativas, e essa aproximação pode ser considerada boa ou ruim dependendo do modelo a ser estudado, delineamento experimental e conjunto de dados (SOUZA *et al.*, 2010). Em conformidade com Bates & Watts (1988), uma das melhores alternativas que se pode fazer para assegurar uma análise não linear bem sucedida é obter bons valores iniciais para os parâmetros, a partir dos quais a convergência poderá ser obtida rapidamente.

A classe de modelos lineares é bastante flexível, uma vez que muitos modelos podem ser formulados. Estes modelos podem ser uma boa opção quando o objetivo é descrever as observações durante um período limitado do processo de crescimento sem incorporar informações sobre os dados (DRAPER & SMITH, 1998; SEBER, 1977). Já no caso não linear, na maioria das vezes, as formulações de possíveis modelos são baseadas em considerações teóricas inerentes ao fenômeno que se tem interesse modelar. Modelos formulados desta forma são

chamados de modelos mecanísticos (MAZUCHELI & ACHCAR, 2002; SEBER & WILD, 1989; BATES & WATTS, 1988). Em geral, os parâmetros na função esperança têm interpretação prática para os cientistas ou pesquisadores.

O ajuste de modelos não lineares é bastante difundido e investigado do ponto de vista metodológico, na descrição de curvas de crescimento animal (MAZZINI *et al.*, 2003, 2005; MENDES *et al.*, 2008), na descrição de cinética de digestão de animais ruminantes (SAVIAN *et al.*, 2007a, 2007b), utilizando a abordagem clássica e (SAVIAN *et al.*, 2009) a abordagem bayesiana. Na área de produção vegetal, é utilizado em situações experimentais como no estudo de crioconservação de sementes de aroeira (*Astronium urundeuva* Engl.) e baraúna (*Schinopsis brasiliensis* Engl.) (GONZAGA *et al.*, 2003), estudo da dinâmica e disponibilização de nutrientes no sistema solo-planta (PEREIRA *et al.*, 2005), estudo do crescimento de espécies vegetais (MARTINS FILHO *et al.*, 2008), em estudos de velocidade de reações químicas (MACHADO, 2008), sendo também normalmente utilizados na descrição de curvas de lactação, quando ajustados a dados decorrentes de produção de leite em intervalos consecutivos de tempo (SILVA *et al.*, 2005). Um consenso dos autores que trabalham com ajuste de modelos não lineares na descrição dos fenômenos supracitados é sobre a possibilidade de se condensar informações de uma série de dados, tomados ao longo do tempo, em um pequeno conjunto de parâmetros biologicamente interpretáveis (TERRA *et al.*, 2010).

2.3 Estimação dos parâmetros de modelos não lineares

Utilizado na estimação dos parâmetros em modelos não lineares, da mesma maneira que em modelos lineares (GALLANT, 1987), o método dos mínimos quadrados consiste em encontrar o melhor ajustamento para um conjunto de dados tentando minimizar a soma de quadrados dos resíduos entre a curva ajustada e os

valores observados. O método apresenta algumas variações para estrutura da matriz de covariâncias, estruturas estas que diferem dependendo do conhecimento do fenômeno em estudo. A escolha de uma estrutura adequada pode tornar o modelo mais realista em relação aos dados como também melhorar a eficiência das estimativas.

Savian (2005) apresentou a caracterização da regressão em função do vetor de erros da seguinte maneira: a) modelos ordinários - aqueles cuja estrutura dos erros não viola nenhuma das pressuposições, ou seja, $\varepsilon \sim N(0, I\sigma^2)$. A estrutura da matriz de covariância pode ser representada da seguinte forma: $\Sigma = I\sigma^2$, em que σ^2 é a variância das observações e I é a matriz identidade de dimensão p. b) modelos ponderados - são aqueles cuja estrutura dos erros viola a pressuposição de homogeneidade de variâncias. Nesse caso, diz-se que os erros são heteroscedásticos, e representa-se por $\varepsilon \sim N(0, D\sigma^2)$ em que D é uma matriz diagonal, positiva definida, que pondera a variância σ^2 . c) modelos generalizados - são aqueles cuja estrutura dos erros viola a pressuposição de independência dos erros e/ou de homogeneidade de variâncias. Os erros são correlacionados e possivelmente heteroscedásticos, $\varepsilon \sim N(0, W\sigma^2)$, em que W é uma matriz simétrica, positiva definida, que representa as variâncias e covariâncias dos erros.

2.3.1 Método dos mínimos quadrados ordinários

Gallant (1987) apresenta o modelo de regressão linear da seguinte forma:

$$y_t = f(x_t, \theta^0) + \varepsilon_t; \quad t=1, \dots, n$$

no qual y_t representa o valor observado, $f(x_t, \theta^0)$ é a relação funcional conhecida, x_t é um vetor de observações em k variáveis regressoras, $\theta^0 = [\theta_1^0, \theta_2^0, \dots, \theta_p^0]'$ é um

vetor de parâmetros p dimensional desconhecido e ε_i um vetor de erros aleatórios supostos normalmente e identicamente distribuídos com média zero e variância desconhecida σ^2 , ou seja, $\varepsilon_i \sim N(0, \sigma^2)$.

Na forma matricial tem - se:

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

onde y tem componentes de y_t , $f(\theta^0)$ tem componentes de $f(x_t, \theta^0)$ e ε tem componentes de ε_t .

O estimador $\hat{\theta}$ de mínimos quadrados de θ^0 , pode ser obtido via minimização da soma de quadrados residuais, $SSE(\theta)$, por meio da pesquisa do mínimo em todos os valores do espaço paramétrico de $(\theta^0 \in \Theta)$,

$$SSE(\theta) = \sum_{t=1}^n [y_t - f(x_t, \theta)]^2$$

pode-se ocorrer ao contrário da soma de mínimo quadrados linear $SSE(\theta)$, e de se ter vários mínimos relativos, além do mínimo absoluto.

Na forma matricial, a soma de quadrados dos erros é:

$$SSE(\theta) = \varepsilon' \varepsilon = [y - f(\theta)]' [y - f(\theta)]$$

Considere $F(\theta) = \frac{\partial f(\theta)}{\partial \theta'}$ a matriz jacobiana de $f(\theta)$ em que $f(\theta)$ é uma função vetor coluna $n \times 1$ de um argumento p dimensional de (θ) .

O estimador de mínimos quadrados de $\hat{\theta}$ satisfaz a equação:

$$\frac{\partial SSE(\theta)}{\partial \theta'} \Big|_{\hat{\theta}=\theta} = 0$$

Sendo:

$$\begin{aligned}\frac{\partial SSE(\theta)}{\partial \theta'} &= \frac{\partial}{\partial \theta'} [y - f(\theta)]' [y - f(\theta)] \\ &= -2 [y - f(\theta)]' F(\theta)\end{aligned}$$

Igualando-se a equação acima a zero, obtém-se:

$$\begin{aligned}-2 [y - f(\hat{\theta})]' F(\hat{\theta}) &= 0 \\ -2F'(\hat{\theta}) [y - f(\hat{\theta})] &= 0 \\ F'(\hat{\theta}) [y - f(\hat{\theta})] &= 0\end{aligned}$$

$F'(\hat{\theta}) [y - f(\hat{\theta})] = 0$, é o sistema de Equações Normais (SEN) não linear.

A matriz $F(\hat{\theta})$ desempenha o mesmo papel que a matriz X na regressão linear $y = X\beta + \epsilon$.

Substituindo-se $F(\hat{\theta})$ por X , tem - se:

$$\begin{aligned}X' [y - f(\hat{\theta})] &= 0 \\ X'Y - X' f(\hat{\theta}) &= 0 \\ X' f(\hat{\theta}) &= X'Y \text{ (SEN) não linear.}\end{aligned}$$

Fazendo-se no SEN $[y - f(\hat{\theta})] = \hat{\epsilon}$, tem-se $X'\hat{\epsilon} = 0$. Resulta que o vetor residual $\hat{\epsilon}$ satisfaz $F'(\hat{\theta})\hat{\epsilon} = 0$ e é, portanto ortogonal às colunas da matriz Jacobiana $F(\theta)$ calculada em $\theta = \hat{\theta}$, assim como em regressão linear. A identificação entre X e $F(\hat{\theta})$, vale em geral, isto é, todas as expressões que se obtém no estudo da inferência do modelo linear, assumindo a pressuposição de normalidade dos erros tem uma contrapartida no caso não linear que se obtém por intermédio da substituição da matriz X por $F(\hat{\theta})$.

Porém, no SEN não-linear, as equações são não lineares em relação aos parâmetros, não possuindo em geral uma solução explícita, na qual as mesmas

devem ser obtida através de processos iterativos, usando um software estatístico.

2.3.2 Método dos mínimos quadrados ponderados

Com a finalidade de estabilização das variâncias, o método dos mínimos quadrados ponderados na estimação dos parâmetros é mais adequado por fornecer estimadores não tendenciosos e de variância mínima (HOFFMAN & VIEIRA, 1998).

Seja o modelo linear,

$$Y = X\beta + u$$

supondo $u \sim N(0, V\sigma^2)$, $E(u) = 0$ e $Var(u) = V\sigma^2$, em que V é uma matriz diagonal positiva definida, que representa as variâncias associadas a cada u_i . Os elementos nulos, fora da diagonal da matriz V significa válida a pressuposição de independência das observações, ou seja, $E(u_i, u_j) = 0, i \neq j$.

Seja um novo modelo $Z = Q\beta + \phi$, e uma matriz simétrica não-singular P tal que $V = PP = P'P$. Desta forma, tem-se: $V^{-1} = P^{-1}P^{-1}$.

Escrevendo $\phi = P^{-1}\epsilon$ tem-se: $E(\phi) = E(P^{-1}u) = P^{-1}E(u) = 0$ e ainda, $Var(\phi) = E(\phi\phi') - [E(\phi)]^2 = E(P^{-1}uu'P^{-1}) = P^{-1}V\sigma^2P^{-1} = \sigma^2P^{-1}PPP^{-1} = \sigma^2I$.

Multiplicando cada um dos termos de $Y = X\beta + u$ por P^{-1} , obtém-se o novo modelo:

$$P^{-1}Y = P^{-1}X\beta + P^{-1}u \Leftrightarrow Z = Q\beta + \phi,$$

em que o vetor de erros é dado por $\phi = P^{-1}u$.

Aplicando-se o método do MQO ao modelo novo, tem-se que o MQP fornece o SEN:

$$X'V^{-1}X\hat{\beta} = X'V^{-1}Y$$

cuja solução:

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y.$$

2.3.3 Método dos mínimos quadrados generalizados

Hoffman & Vieira (1998) consideram que, em presença de heterogeneidade de variâncias e/ou autocorrelação dos resíduos, o método dos mínimos quadrados generalizado é mais eficiente do que o método dos mínimos quadrados ponderados e ordinários. Se esses elementos estão incluídos como componentes do modelo, então as hipóteses terão mais probabilidade de estarem satisfeitas, os estimadores do parâmetro terão melhores propriedades estatísticas, e o modelo pode ser mais informativo.

Embora algumas transformações de dados sejam conhecidas para algumas formas de relação entre variância e média, como por exemplo, a transformação arco-seno para dados oriundos de proporção, em que nessas situações podem ocorrer devido a variabilidade da média que já é esperada, excesso de zero, correlação entre indivíduos, a busca de novos procedimentos de análises se faz necessário.

Segundo Mazzini *et al.*, (2005) se fosse utilizado o estimador de MQO, não levando em consideração a não-homoscedasticidade e autocorrelação dos erros, os testes de significância das estimativas seriam viesados e ocorreria a subestimação das variâncias das estimativas dos parâmetros respectivamente.

Seja o modelo linear

$$Y = X\beta + u$$

supondo $u \sim N(0, W\sigma^2)$, $E(u) = 0$ e $Var(u) = V\sigma^2$, em que W é uma matriz simétrica, positiva definida, que representa as variâncias e covariâncias dos erros. Admitindo-se que os erros são autocorrelacionados na forma de um processo autorregressivo estacionário de primeira ordem AR(1).

$$u_t = \phi_1 u_{t-1} + \varepsilon_t$$

em que $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$, $E(\varepsilon_t \varepsilon_{t-h}) = 0$ se $h \neq 0$.

O modelo u_t será estacionário se

$$-1 \leq \phi \leq 1; t = 2, \dots, n.$$

Nessas condições, $\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\phi_1^2}$ e $cov_u = \frac{\sigma_\varepsilon^2}{1-\phi_1^2} \phi_1^h$. De maneira análoga ao método dos mínimos quadrados ponderados encontra-se:

$$\hat{\beta} = (X'W^{-1}X)^{-1}X'W^{-1}Y.$$

Os métodos dos mínimos quadrados, ordinários e ponderados, são casos particulares dos MQG.

2.3.4 Processos iterativos

Vários métodos iterativos são propostos para obtenção das estimativas de mínimos quadrados dos parâmetros de um modelo de regressão não linear. Os mais utilizados são o método de Gauss-Newton ou método da linearização, o método de Newton, o método Steepest-Descent ou método gradiente e o método de Marquardt (BATES & WATTS, 1988), esses métodos diferem na forma como $\theta - \theta^0$ é calculado para propiciar as atualizações no vetor de parâmetros.

De uma forma geral, os critérios básicos são:

$$\text{Gradiente: } \theta - \theta^0 = X' \varepsilon$$

$$\text{Gauss - Newton: } \theta - \theta^0 = (X' X) X' \varepsilon$$

$$\text{Newton: } \theta - \theta^0 = G^{-1} X' \varepsilon$$

$$\text{Maquardt: } \theta - \theta^0 = [X' X + \delta \text{diag}(X' X)]^{-1} X' \varepsilon$$

em que $(X' X)^{-}$ é uma inversa generalizada.

O método Gradiente é baseado no gradiente ou grau de variação dos resíduos $\varepsilon' \varepsilon$. Os métodos Gauss-Newton e Marquardt realizam a regressão dos resíduos em relação as primeiras derivadas do modelo não-linear em relação aos parâmetros, até que haja a convergência. O método de Newton faz a regressão desses resíduos em relação a uma função das segundas derivadas do modelo não linear com relação aos parâmetros.

Segundo Souza (1998), o sucesso na convergência de um algoritmo para um método iterativo no processo de estimação não linear está diretamente associado ao uso de uma função resposta apropriada e de valores iniciais adequados ao procedimento numérico. Uma má escolha dos valores iniciais pode resultar num número muito grande de interações até convergir; convergir ao um mínimo local, ou mesmo, não convergir. Entretanto, bons valores iniciais podem levar a

um mínimo global, mesmo com a existência de vários mínimos locais.

No programa computacional R, os quatros métodos citados estão implementados, sendo que os dois mais utilizados são os de Gauss Newton e Maquardt.

O método Gauss-Newton usa a expansão em série de Taylor do vetor de funções: $f(\theta) = f(\theta^0) + F(\theta^0)(\theta - \theta^0) + \dots$ em que $F(\theta^0)$ é a matriz de primeiras derivadas de X , avaliada no ponto θ^0 .

Substituindo os termos dessa expansão no SEN, obtemos:

$$\begin{aligned} X' f(\theta) &= X' Y \\ X' [f(\theta^0) + F(\theta^0)(\theta - \theta^0)] &= X' Y \\ X' [f(\theta^0) + X(\theta - \theta^0)] &= X' Y \\ X' X(\theta - \theta^0) &= X' Y - X' f(\theta^0) \\ (\theta - \theta^0) &= (X' X)^{-1} - X' \varepsilon \end{aligned}$$

Portanto, a fórmula iterativa conhecida como método de Gauss-Newton é:

$$\theta^1 = \theta^0 + (X' X)^{-1} - X' \varepsilon$$

O processo é repetido colocando-se θ^1 no lugar de θ^0 (vetor de valores iniciais que podem ser obtidos de estudos anteriores, conhecimentos teóricos ou por uma grade de valores que minimize a soma de quadrados do resíduos), repetindo um determinado número de vezes até que o vetor de estimativas não se altere mais dentro de uma precisão pré estipulada.

2.4 Modelos de crescimento

Curvas de crescimento (sigmoidais) possuem uma larga aplicação em áreas como biologia, agronomia, economia, dentre outras. Exemplos típicos ocorrem

em estudos de crescimento humano, animal e vegetal. Essas curvas começam em algum ponto fixo e aumentam a sua taxa de crescimento monotônico até um ponto de inflexão, após este ponto esta taxa de crescimento começa a diminuir até a curva se aproximar de um valor final, chamado de assíntota (RATKOWSKY, 1983). Geralmente estudam-se as curvas de crescimento por meio de funções não lineares em que apresentam algumas peculiaridades quanto ao procedimento de ajuste (DRAPER & SMITH, 1998). Os modelos são constituídos a partir de uma base teórica em que os parâmetros fornecem um maior conhecimento sobre o fenômeno estudado, o que gera um bom ajuste com menos parâmetros.

A escolha do modelo que melhor descreve a curva de crescimento usualmente é baseada na qualidade de seu ajuste aos dados, na interpretação biológica dos parâmetros. Vários tipos de modelos estatísticos podem ser usados, de acordo com suas habilidades, para facilitar a interpretação dos processos envolvidos no sistema de produção vegetal.

Os modelos empíricos usados frequentemente para estimar o crescimento vegetal incluem a função Logística e de Gompertz (SEBER & WILD, 1989; RATKOWSKY, 1983).

O modelo Logístico como em Seber & Wild (1989) é dado pela equação:

$$y = \frac{\alpha}{1 + \exp[-\gamma(x - \beta)]}$$

em que:

α corresponde ao valor assintótico, indicando o valor de estabilização da variável dependente em relação ao tempo;

β corresponde ao parâmetro de localização;

γ medida relativa a taxa de crescimento da curva;

A curva logística é simétrica em relação ao ponto de inflexão $x = \beta$ quando $y = \frac{\alpha}{2}$. Neste ponto, a taxa de crescimento atinge seu valor máximo, ou seja, $W_{max} = \gamma \frac{\alpha}{4}$.

Quando os resultados inferenciais obtidos não são muito precisos, outras parametrizações podem ser consideradas de maneira que se ajustem melhor aos dados, como apresentado por Ratkowsky (1983).

O modelo Gompertz (SEBER & WILD, 1989), definido pela equação,

$$y = \alpha \exp[-\exp(\gamma(x - \beta))]$$

em que:

α corresponde ao valor assintótico, indicando o valor de estabilização da variável dependente em relação ao tempo;

β corresponde ao parâmetro de localização;

γ medida relativa à taxa de crescimento da curva;

é frequentemente usado em situações em que o crescimento não é simétrico em relação ao ponto de inflexão que ocorre em $x = \beta$ quando $y = \frac{\alpha}{e}$, sendo $W_{max} = \gamma \frac{\alpha}{e}$.

Na análise de germinação de sementes, Guillén *et al.*, (2009) estudando a germinação de sementes cactus colunares em diferentes populações no vale do Tehuacán-Cuicatlán, México, utilizaram o modelo logístico no ajuste dos dados e observaram que em todas as situações o modelo ajustado foi significativo para explicar a porcentagem máxima acumulada em função do tempo. Ikeda *et al.*, (2008) em estudo relacionado ao efeito da Luz e KNO_3 na germinação de sementes de *Tridax Procumbens* sob temperatura constante e alternada, identificaram o modelo logístico na assincronia do percentual de germinação acumulada. Gazola *et al.*, (2011) em estudos de modelagem não linear, envolvendo curva de crescimento do desempenho germinativo de sementes de milho híbrido, verificaram que o mo-

delo logístico apresentou ajuste adequado aos dados de percentuais germinativos. Em estudo apresentado por Duarte *et al.*, (2010), em germinação de sementes de *Dyckia goehringii* Gross e Rauh (Bromeliaceae), sob diferentes temperaturas encontraram um comportamento sigmoidal na descrição de germinação. Tomaz *et al.*, (2010) com o objetivo de determinar o tempo necessário para a condução do teste de germinação de *Panicum maximum* cv. Tanzânia ajustaram o modelo logístico para a avaliação da germinação, percebendo que a curva esta próxima aos valores observados, indicando que modelo se ajustou bem aos dados. Vidigal *et al.*, (2011), verificando a ocorrência de mudanças nas sementes, durante a maturação de frutos de pimentão, utilizaram o modelo logístico no ajuste dos dados de germinação das sementes ao longo do tempo.

3 MATERIAI E MÉTODOS

3.1 Material

A base de dados utilizada neste trabalho provém de um experimento realizado no período de 06 de setembro a 06 de outubro de 2011, no Laboratório de Análise de Sementes (LAS) do Departamento de Agricultura (DAG) da Universidade Federal de Lavras (UFLA), no município de Lavras, MG. Foram utilizados cinco lotes, com respectivos potenciais mínimos de germinação 73%, 88%, 90%, 90% e 97% de sementes de café *Coffea arabica* L. da cultivar "Catuaí Vermelho IAC 99", associando as porcentagens acima citadas aos lotes I, II, III, IV, V, respectivamente. Cada tratamento constou de quatro repetições com 100 sementes, utilizando como substrato rolos de papel toalha (germitest) umedecidos com água destilada na proporção de 2,5 vezes o peso do papel seco, colocados em câmara de germinação, regulados à temperatura constante de 30°C, de acordo com as regras para análise de sementes (BRASIL, 2009). A germinação das sementes foi monitorada durante 30 dias, sendo anotado diariamente o número de sementes germinadas que tiveram como critério de germinação a protusão radicular.

Nas figuras 1 a 6, encontram-se algumas fotos ilustrando a instalação do experimento e diferentes etapas do processo de germinação das sementes.



Figura 1 Instalação do processo de germinação



Figura 2 Instalação do processo de germinação



Figura 3 Sementes de café germinadas aos 15 dias - Lote III



Figura 4 Sementes de café germinadas aos 21 dias - Lote V



Figura 5 Sementes de café germinadas aos 30 dias - Lote I



Figura 6 Sementes de café germinadas aos 30 dias - Lote IV

3.2 Métodos

3.2.1 Modelos estudados

Foram ajustados os modelos de crescimento não linear Logístico e Gompertz de acordo com a parametrização apresentada por Seber & Wild (1989), sendo descritos pelas seguintes expressões:

1. Modelo Logístico com erros independentes e autoregressivos de primeira ordem AR(1), respectivamente:

$$Y_t = \frac{\alpha}{1 + \exp[-\gamma(X_t - \beta)]} + u_t$$

$$Y_t = \frac{\alpha}{1 + \exp[-\gamma(X_t - \beta)]} + \phi u_{t-1} + \xi_t$$

2. Modelo Gompertz com erros independentes e autoregressivos de primeira ordem AR(1), respectivamente:

$$Y_t = \alpha \exp[-\exp(\gamma(X_t - \beta))] + u_t$$

$$Y_t = \alpha \exp[-\exp(\gamma(X_t - \beta))] + \phi u_{t-1} + \xi_t$$

em que

Y_t - expressa a porcentagem de germinação acumulada no tempo t ;

α - corresponde ao valor assintótico, ou seja, representa a porcentagem máxima de germinação acumulada, valor de estabilização da variável dependente em relação ao tempo;

β - corresponde ao tempo no qual Y alcança metade do seu valor máximo

no modelo Logístico e no modelo Gompertz Y alcança proporção $1/e$ do seu valor máximo;

γ - está vinculado à taxa de crescimento, que representa a velocidade de germinação para atingir o valor assintótico;

exp - refere-se à base do logaritmo neperiano;

X_t - refere-se ao valor da variável independente, ou seja, tempo necessário para germinação de sementes, dado em dias, $t = 1, 2, 3...30$;

u_t - corresponde ao resíduo no tempo t ;

u_{t-1} - corresponde ao resíduo no tempo $t-1$;

ϕ - corresponde ao parâmetro autoregressivo de ordem 1;

ξ_t - é o ruído branco.

3.2.2 Critérios para seleção do modelo

Para a verificação da qualidade do ajuste de um modelo não linear (MOTULUSKY & CHRISTOPOULOS, 2003; BOZSDOGAN, 1987), vários aspectos podem ser observados, busca-se o modelo mais parcimonioso, isto é, o modelo que envolva o mínimo de parâmetros possíveis a serem estimados e que explique bem o comportamento da variável resposta. Para isso, ferramentas analíticas, tais como testes hipóteses e critérios de informação podem ser utilizados, sendo o seu o cálculo e interpretação, fáceis de serem realizados. Logo a seleção do modelo mais adequado para descrever a germinação de sementes foi realizada com base na precisão dos ajustes obedecendo aos seguintes critérios:

3.2.2.1 Coeficiente de determinação ajustado (R_{aj}^2)

$$R_{aj}^2 = 1 - \left[\frac{(1 - R^2)(n - i)}{(n - p)} \right]$$

em que:

$$R^2 = 1 - \frac{SQR}{SQT}$$

SQR = soma de quadrado do resíduo;

SQT = soma de quadrado total;

n = número de observações utilizadas para ajustar a curva;

p = número de parâmetros na função incluindo o intercepto;

i = 1 se tiver intercepto e 0 se não tiver intercepto.

Tem-se que quanto maior for o valor de R_{aj}^2 , melhor será o modelo.

3.2.2.2 Desvio padrão residual (DPR)

A estimativa do DPR é obtida pela seguinte expressão:

$$DPR = \sqrt{\frac{QME}{(n - p)}}$$

em que:

QME = quadrado médio do erro;

n = número de observações;

p = número de parâmetros do modelo.

Quanto menor for o desvio padrão residual, melhor será o modelo ajustado.

3.2.2.3 Critérios de informação

Os critérios de Informação de Akaike (AIC)(AKAIKE, 1974), o Critério de Informação Bayesiano (BIC) (SCHWARZ, 1978) e o Teste da Razão de Verossimilhança (TRV) são outras estatísticas que servem como medidas de comparação da qualidade de ajuste do modelo baseados no máximo da função de verossimilhança (MFV) sendo dependentes do número de observações e parâmetros do modelo em estudo.

A estimativa do AIC é dada pela seguinte fórmula:

$$AIC = -2\log L(\hat{\theta}) + 2k$$

O valor do critério BIC para um determinado modelo é:

$$BIC = -2\log L(\hat{\theta}) + k\log(n)$$

onde:

$L(\hat{\theta})$ é o máximo da função de verossimilhança;

k: o número de parâmetros no modelo;

n: é o número de observações usadas na estimação do modelo em estudo.

O Critério de Informação de Akaike (AIC) admite a existência de um modelo real que descreve os dados que é desconhecido, e tenta escolher dentre um grupo de modelos avaliados, o que minimiza a divergência de Kullback-Leibler (K-L). Esta divergência está relacionada à informação perdida por se usar um modelo aproximado e não o "real". O critério BIC é definido como a estatística que maximiza a probabilidade de se identificar o verdadeiro modelo dentre os avaliados. Apesar de usarem o mesmo critério estatístico, utilizam valores críticos

diferentes. Tanto o AIC e o BIC são usados para comparar modelos não aninhados e o melhor modelo, considerando o ajuste, será aquele que apresentar menor valor.

O teste da razão de máxima verossimilhança (TRV) é apropriado para testar dois modelos hierarquicamente aninhados, baseado no ln da razão entre duas verossimilhanças. Sob as hipóteses:

H_0 : O modelo mais simples é o correto;

H_a : O modelo mais parametrizado é o correto.

O TRV usa a estatística $L = 2\log\frac{L_1}{L_2} = 2[\log(L_2) - \log(L_1)]$ sendo L_2 o máximo do logaritmo natural da função de verossimilhança para o modelo mais parametrizado e L_1 para modelo mais simples. Sob a hipótese de H_0 verdadeira $L \sim \chi^2_{(\delta, v)}$ com δ graus de liberdade (gl), sendo v é a diferença entre o número de parâmetros do modelo mais parametrizado e o modelo mais simples. Rejeita-se H_0 quando $L > \chi^2_{(\delta, v)}$.

3.3 Análise dos resíduos

Em regressão não linear, assim como em regressão linear, a análise dos resíduos de um modelo é feita para verificar a plausibilidade das pressuposições envolvidas. Para a verificação dos pressupostos, pode-se utilizar análise gráfica dos resíduos, sendo este um método informal de análise que envolvem os gráficos de resíduos em relação às variáveis independentes e aos valores preditos, ou por meio de testes estatísticos, que é uma maneira mais objetiva de se analisar os resíduos por fornecer uma medida numérica para algumas das discrepâncias previamente descritas.

Para a verificação do pressuposto de normalidade, várias formas gráficas têm sido propostas como histogramas dos resíduos, em que se podem visualizar desvios grosseiros da normalidade ou "Q-Q plot" que é um gráfico que testa a con-

formidade entre a distribuição empírica e uma dada distribuição teórica, apresentando como regra de decisão se a observação gráfica entre os quantis da variável aparecem alinhados com os quantis da distribuição, se isso ocorre é porque os dados estão normalmente distribuídos. O teste estatístico de Shapiro - Wilk (SHAPIRO & WILK, 1965) é um dos mais utilizados para a verificação do pressuposto em que rejeita a hipótese de normalidade quando o valor de p é inferior ou igual a 0,05.

Um gráfico dos resíduos contra os valores estimados pode ser utilizado para examinar se as variâncias dos erros são constantes, observando se a dispersão dos resíduos ocorre aleatoriamente em torno da média zero. O teste estatístico frequentemente usado para heterocedasticidade é chamado de teste de Breusch-Pagan (BREUSCH & PAGAN, 1979), em que a hipótese nula é que os resíduos são homocedásticos, logo com valor de p inferior ou igual a 0,05 a hipótese nula é rejeitada.

Resíduos relacionados com o tempo são chamados autocorrelacionados e para verificar a existência de correlação entre os resíduos, pode-se plotar os resíduos contra o tempo ou os resíduos contra qualquer outra variável de interesse. Quando os erros são independentes, espera-se que os mesmos flutuem aleatoriamente em torno da média zero. O teste de Durbin-Watson (Durbin & Watson, 1950) testa a existência de autocorrelação de primeira ordem e é calculado por:

$$DW = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n (u_t)^2}$$

em que:

u_t : é o resíduo no tempo t;

u_{t-1} : é o resíduo no tempo t-1.

Tem - se que a um valor de p inferior ou igual a 0,05 rejeita-se a hipótese de dependência nos resíduos.

3.4 Recursos computacionais

Os procedimentos de ajuste para os modelos com estruturas de erros independentes e correlacionados foram implementados por meio da função *gnls* (Adequação do modelo linear utilizando mínimos quadrados generalizados) do pacote *nlme* (Modelos lineares e não lineares de efeitos mistos) (PINHEIRO *et al.*, 2011) do programa R versão 2.13.1 (2011-07-08), conforme anexos, os quais possibilitaram obter as estimativas dos parâmetros e os avaliadores de qualidade de ajuste.

Para análise dos resíduos dos modelos de regressão não linear, foram utilizados métodos que podem ser ajustados no R como a matriz gradiente (derivadas da função em relação ao vetor de parâmetros) do modelo ajustado dentro da função *lm* e, assim, obter testes estatísticos *bptest* e *dubirnWatsonTest* implementados no pacote *lmtest* como também o pacote *nlstools* (ferramentas de diagnóstico para modelos de regressão não linear)(Florent e Delignette-Muller, 2011) para a obtenção dos gráficos residuais.

As estimativas iniciais, necessárias para a obtenção das estimativas dos parâmetros, foram obtidas por meio do método gráfico iterativo para valores iniciais em regressão não linear, utilizando-se o pacote *manipulate*, função *manipulate()* disponível no RStudio (editor para script do R), para os modelos Logístico e Gompertz respectivamente.

4 RESULTADOS E DISCUSSÃO

4.1 Ajuste dos modelos não lineares

Após o ajuste dos modelos Logístico e Gompertz, escolhidos com base na observação gráfica dos dados, considerando erros independentes e variância constante, verificou-se por meio da análise gráfica dos resíduos, Figuras 7 a 16 respectivamente, pelo quadro resíduos ou resíduos studentizados versus os valores ajustados, que em todos os lotes analisados, os resíduos se mostram concentrados nas extremidades, ou seja, não há uma distribuição aleatória em torno da média zero, indicando, assim, a não homogeneidade das variâncias que pode ser confirmado pelas estimativas da estatística do teste de Breusch-Pagan apresentadas nas Tabelas 1 e 2. A verificação da pressuposição de normalidade mostrada pelo gráfico "Q-Q Plot" dos resíduos studentizados, nota-se que em todas as situações os valores extremos se afastaram da reta, mostrando a falta de normalidade que pode ter sido afetada pela variabilidade dos dados como mostra as Figuras 7 a 16 e as Tabelas 1 e 2. A suposição de dependência examinada pelo gráfico de autocorrelação, apresenta com exceção dos gráficos observados nas Figuras 8, 11 e 16, onde os resíduos se comportaram mais aleatórios, um agrupamento dos resíduos na linha de base zero, mostrando que as interferências em períodos separados são sistematicamente relacionados. O que por meio do teste de Durbin-Watson isso não se confirma em todas as situações, ressaltando a importância da utilização de testes estatísticos na validação das pressuposições.

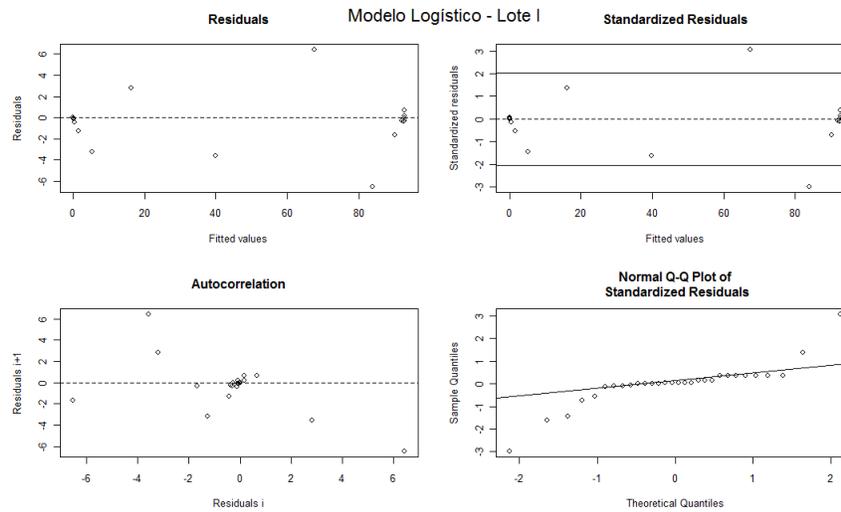


Figura 7 Análise gráfica dos resíduos do modelo Logístico para o lote I

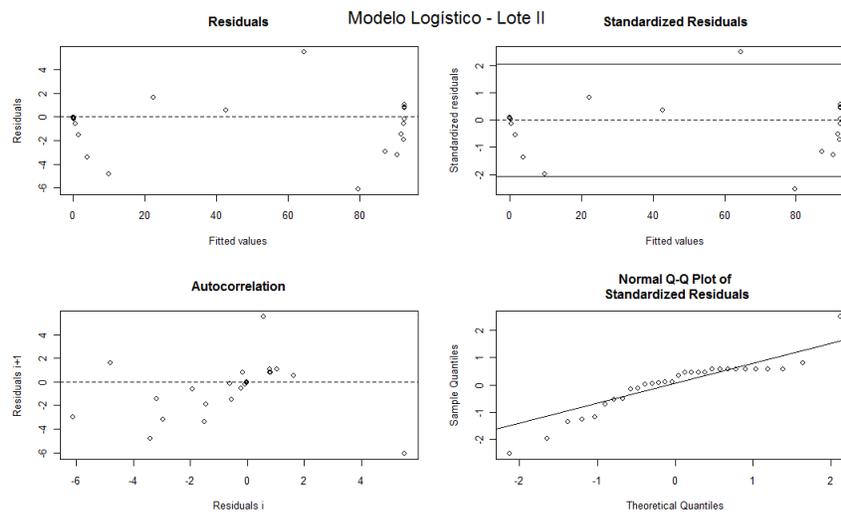


Figura 8 Análise gráfica dos resíduos do modelo Logístico para o lote II

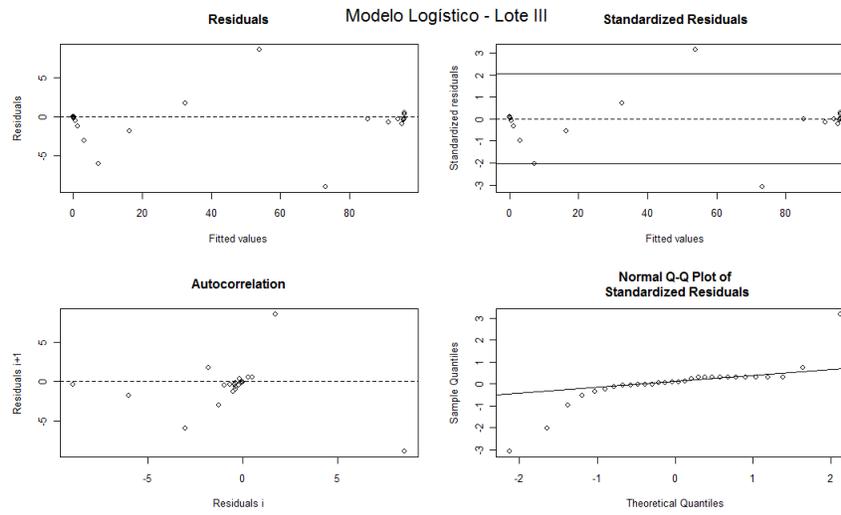


Figura 9 Análise gráfica dos resíduos do modelo Logístico para o lote III

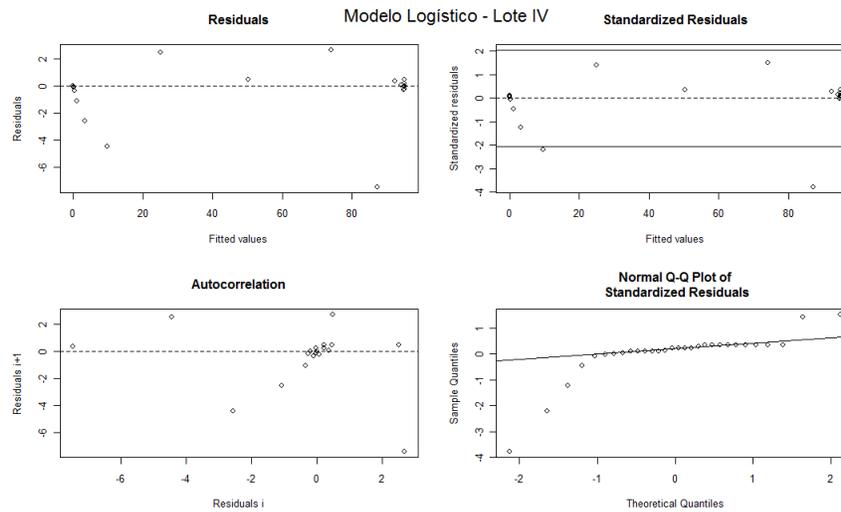


Figura 10 Análise gráfica dos resíduos do modelo Logístico para o lote IV

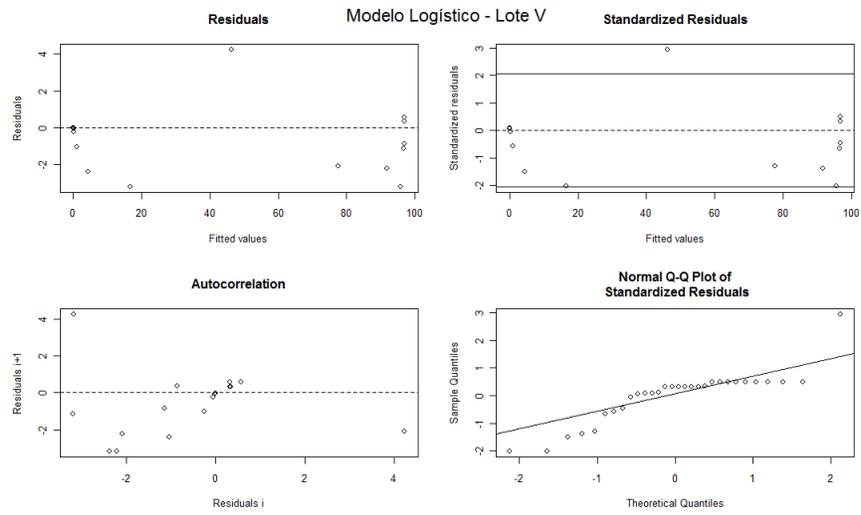


Figura 11 Análise gráfica dos resíduos do modelo Logístico para o lote V

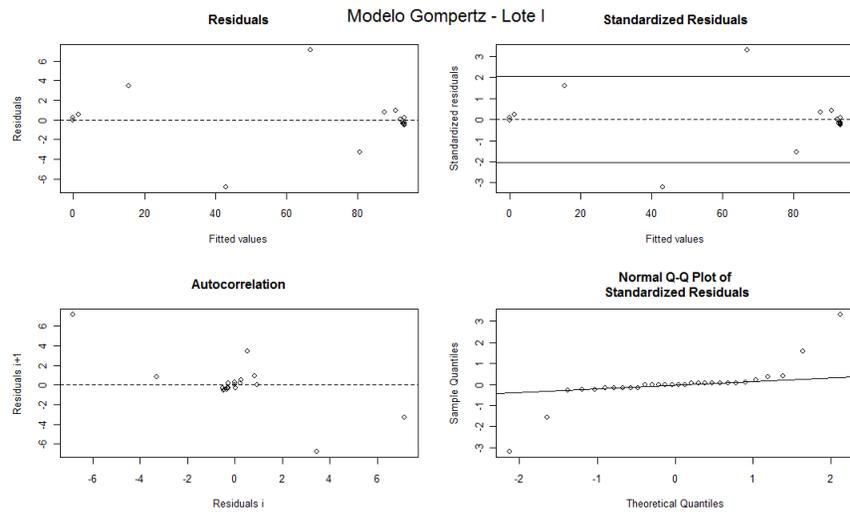


Figura 12 Análise gráfica dos resíduos do modelo Gompertz para o lote I

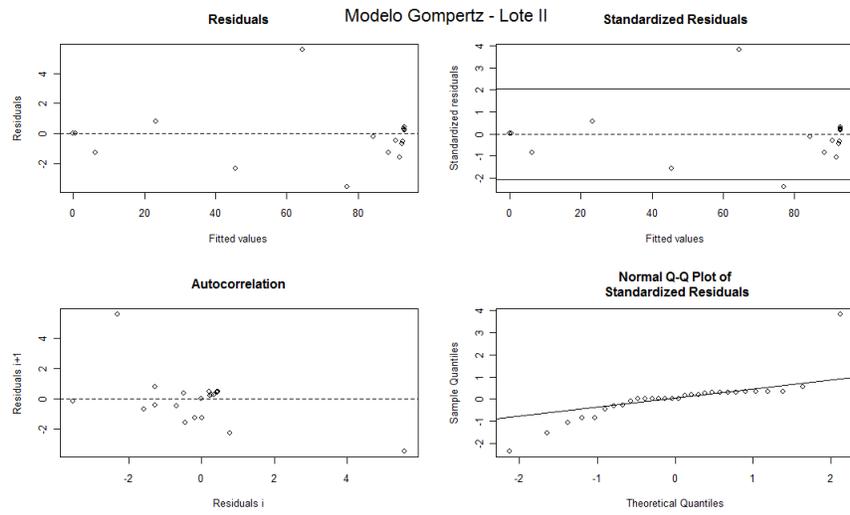


Figura 13 Análise gráfica dos resíduos do modelo Gompertz para o lote II

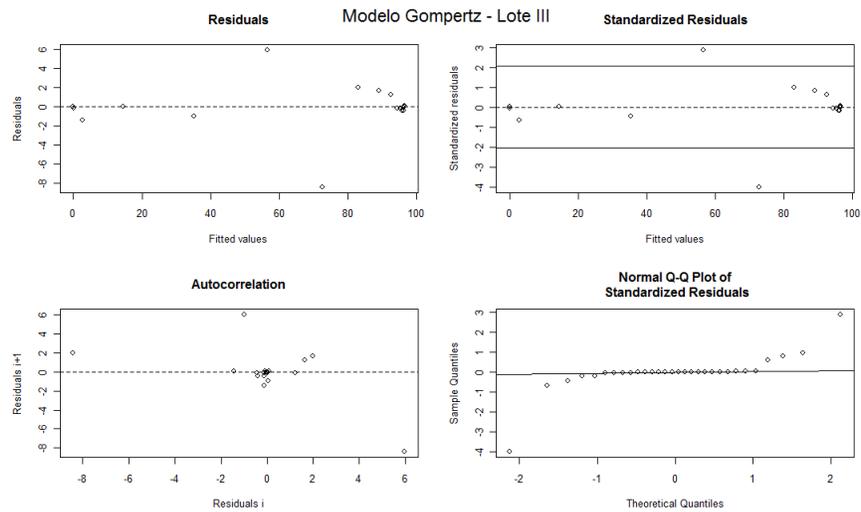


Figura 14 Análise gráfica dos resíduos do modelo Gompertz para o lote III

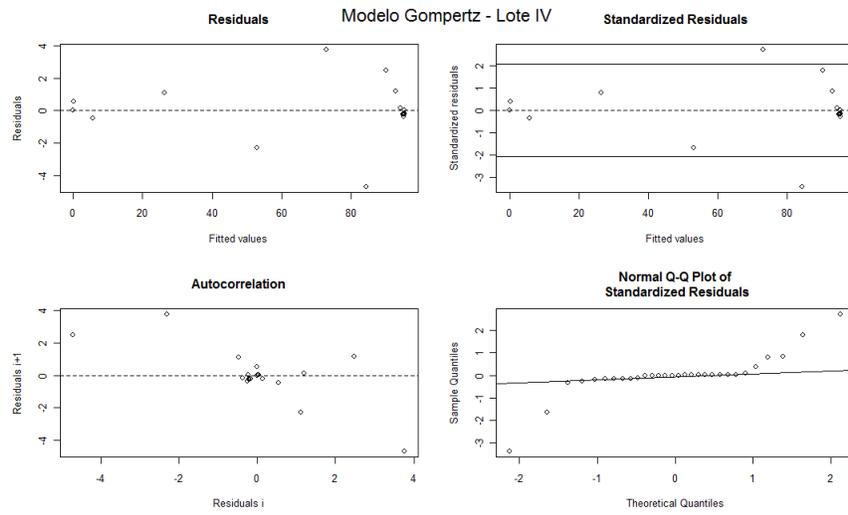


Figura 15 Análise gráfica dos resíduos do modelo Gompertz para o lote IV

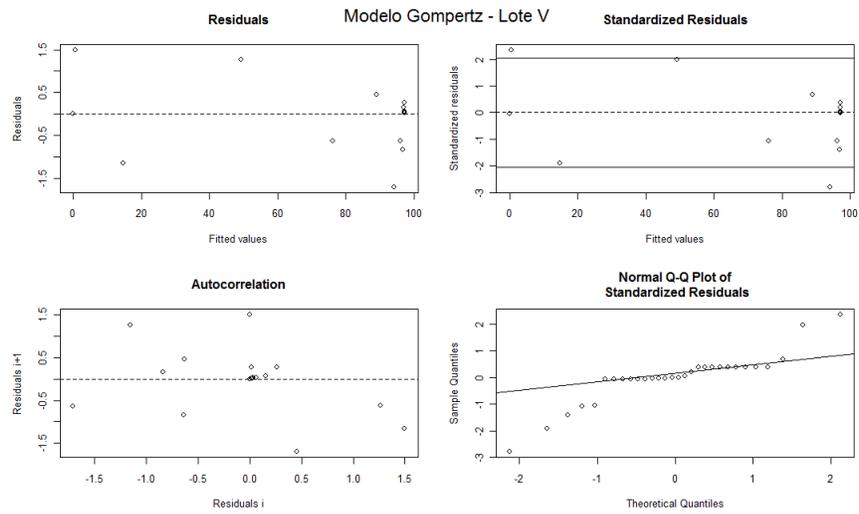


Figura 16 Análise gráfica dos resíduos do modelo Gompertz para o lote V

Na Tabela 1 e 2, estão apresentadas as estimativas das estatísticas dos testes de Shapiro-Wilk, Breusch-Pagan e Durbin-Watson e os respectivos valor-P para os modelos Logísticos e Gompertz dos cinco lotes analisados. Observa-se que tanto no ajuste do modelo Logístico como no modelo Gompertz, os erros relativos aos lotes analisados não apresentaram normalidade, homocedasticidade e algumas situações independência.

Tabela 1 Estimativas das estatísticas dos testes de Shapiro-Wilk, Breusch-Pagan e Durbin-Watson e os respectivos valor-p dos resíduos do modelo Logístico para os cinco lotes de semente de café.

Lotes	Shapiro-Wilk	Valor-p	Breusch-Pagan	Valor-p	Durbin-Watson	Valor-p
I	0,7793	0,0002	22,9718	0,0001	3,0506	0,0001
II	0,8825	0,0032	13,6698	0,0011	1,5574	0,1300
III	0,7137	0,0001	16,9820	0,0002	2,2253	0,7760
IV	0,6540	0,0001	4,4529	0,0011	2,2990	0,6620
V	0,8150	0,0001	21,3829	0,0002	1,7447	0,3500

Tabela 2 Estimativa das estatísticas dos testes de Shapiro-Wilk, Breusch-Pagan e Durbin-Watson e os respectivos valor-p dos resíduos do modelo Gompertz para os cinco lotes de semente de café.

Lotes	Shapiro-Wilk	Valor-p	Breusch-Pagan	Valor-p	Durbin-Watson	Valor-p
I	0,6613	0,0000	24,2076	0,0005	3,5277	0,0000
II	0,7298	0,0001	16,5231	0,0002	3,0417	0,0040
III	0,5719	0,0001	15,6517	0,0003	3,1387	0,0020
IV	0,706	0,0001	20,4560	0,0001	3,4783	0,0000
V	0,8362	0,0003	6,6433	0,0360	2,5639	0,1580

Diante do não atendimento das pressuposições, realizou-se o ajuste pelo método dos mínimos quadrados generalizados.

Na Tabela 3, estão apresentadas as estimativas dos parâmetros do modelo Logístico, com erros independentes, para cada lote de semente ao longo do tempo.

Tabela 3 Estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, aos valores de porcentagem de germinação de sementes para os cinco lotes

Lotes	Parâmetros		
	α	β	γ
I	92,82	10,22	1,27
II	92,46	10,16	0,99
III	95,97	10,73	0,92
IV	95,04	9,90	1,15
V	96,92	9,06	1,48

Todos os parâmetros foram significativos em 1% pelo teste t.

Os lotes de sementes apresentaram comportamentos semelhantes quanto ao percentual de germinação acumulada (α), todos acima de 90%. Com exceção do lote V, todos os percentuais obtidos mostraram-se superiores aos percentuais mínimos classificados de acordo com o lote, sendo o que mais se diferenciou com um aumento de 26% em relação aos 73% apresentado foi o lote I. O parâmetro (β) está relacionado ao dia em que germinação acumulada atinge 50%, sendo assim, nota-se que a qualidade V obteve o menor tempo para atingir metade do potencial de germinação, o que pode estar relacionado com alto poder de germinabilidade das sementes. As demais qualidades obtiveram metade do percentual de germinação em torno do 10º dia. O parâmetro (γ), representando a velocidade de germinação, teve na qualidade V o maior valor, o que indica que a porcentagem de germinação acumulada foi atingida em um menor tempo que as demais qualidades estudadas.

Observa-se por meio das Figuras 17 a 21, que as curvas dos dados expressam um comportamento sigmoidal.

Em estudo para caracterizar as fases de germinação de sementes de maracujá-doce, Ferrari *et al.*, (2008), ajustaram as curvas de germinação pela função Logística.

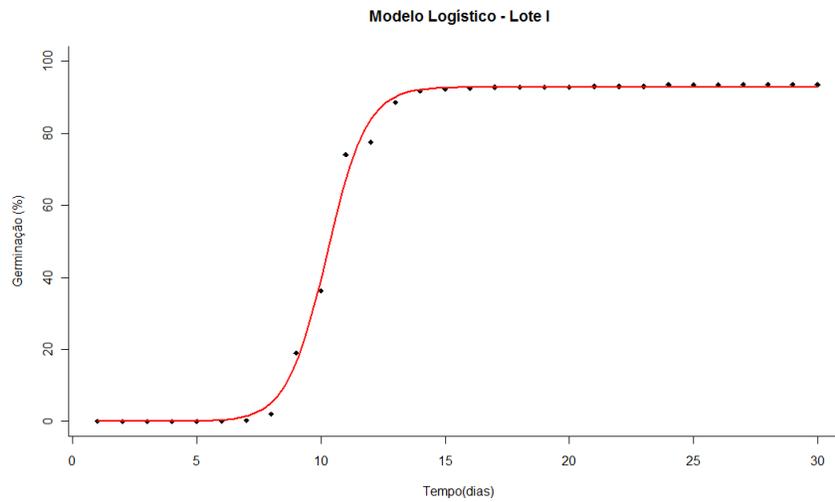


Figura 17 Modelo Logístico na descrição de germinação de sementes de café - Lote I, em função do tempo com estruturas de erros independentes

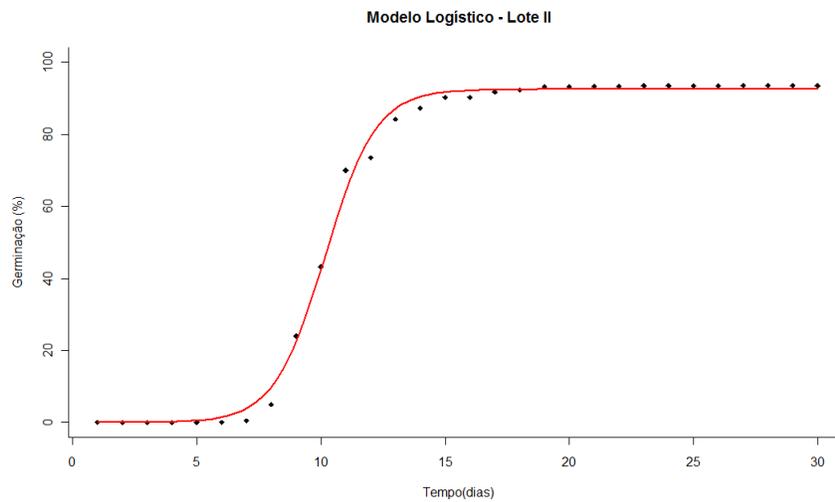


Figura 18 Modelo Logístico na descrição de germinação de sementes de café - Lote II, em função do tempo com estruturas de erros independentes

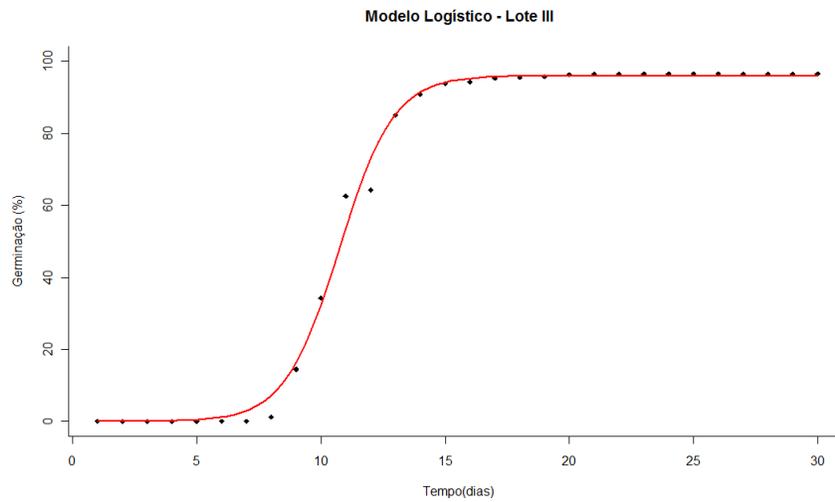


Figura 19 Modelo Logístico na descrição de germinação de sementes de café - Lote III, em função do tempo com estruturas de erros independentes

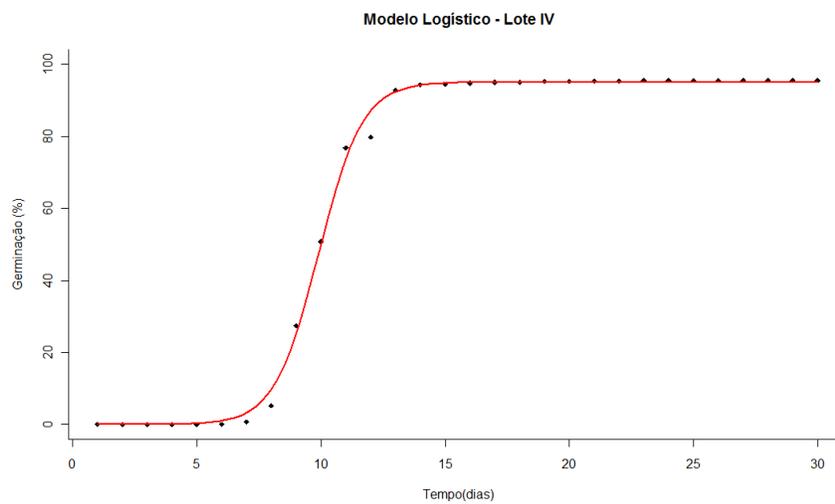


Figura 20 Modelo Logístico na descrição de germinação de sementes de café - Lote IV, em função do tempo com estruturas de erros independentes

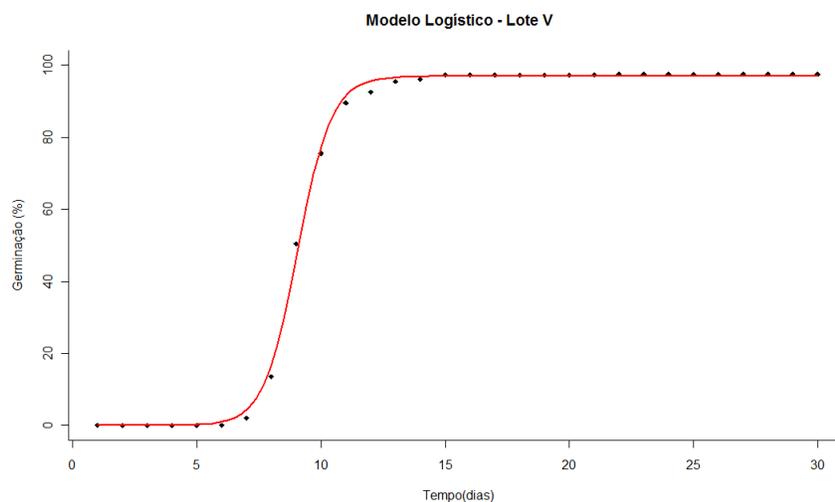


Figura 21 Modelo Logístico na descrição de germinação de sementes de café - Lote V, em função do tempo com estruturas de erros independentes

Após o ajuste do modelo Logístico, verifica-se pelo teste de Shapiro-Wilk, conforme Tabela 4, que os erros relativos a todos os lotes analisados apresentaram distribuição normal, indicando que o ajuste pelo modelo se mostrou adequado.

Tabela 4 Estimativas da estatística do teste de Shapiro-Wilk dos resíduos do modelo Logístico e valor - p para os cinco lotes de sementes de café.

Qualidades	Estatística do Teste Shapiro Wilk	Valor - p
I	0,9266	0,5736
II	0,9732	0,8951
III	0,9174	0,5131
IV	0,7928	0,0706
V	0,9570	0,7868

Através da primeira derivada do modelo Logístico em função do tempo, obtém-se a taxa de germinação absoluta (TGA), apresentada na Figura 22. As TGA foram crescentes até atingirem os máximos de 28%/dia para lote I, em torno de 22%/dia para os lotes II e III, 27%/dia para lote IV e 35%/dia para lote V, nessa

fase, a taxa de germinação se caracterizou elevada e positiva, chegando ao máximo no ponto de inflexão da curva, que ocorreu próximo ao 10º dia. A partir deste ponto, a concavidade da curva muda e a taxa de germinação começa a diminuir em virtude de fatores que inibem progressivamente a germinação. Observa-se que o lote V, caracterizado como o de melhor potencial germinativo, teve um menor tempo para o processo de germinação, já que no décimo terceiro dia praticamente todas as sementes já haviam germinado. Os resultados apresentados são coerentes com o que foi exposto por Eira *et al.*, (2006) em estudo sobre a fisiologia e germinação da semente de café, em que citaram que após o décimo dia, 50% da população de sementes de cafeeiro apresentaram emissão da radícula e, no décimo quinto dia, a maioria das sementes já germinaram.

Caron *et al.*, (2007), ao analisarem o crescimento de plantas de aroeira vermelha, utilizaram a taxa de crescimento absoluto ao longo do tempo na identificação de fitomassa acumulada. Verificaram o tempo de ocorrência da maior taxa de crescimento absoluto que ocorreu no intervalo de 22 a 32 dias, considerando um acúmulo de 34,2% fitomassa de 0,46g.planta⁻¹ e que a Taxa de Crescimento Absoluto (TCA), variou de acordo com o crescimento das mudas, segundo as respostas fisiológicas as condições ambientais.

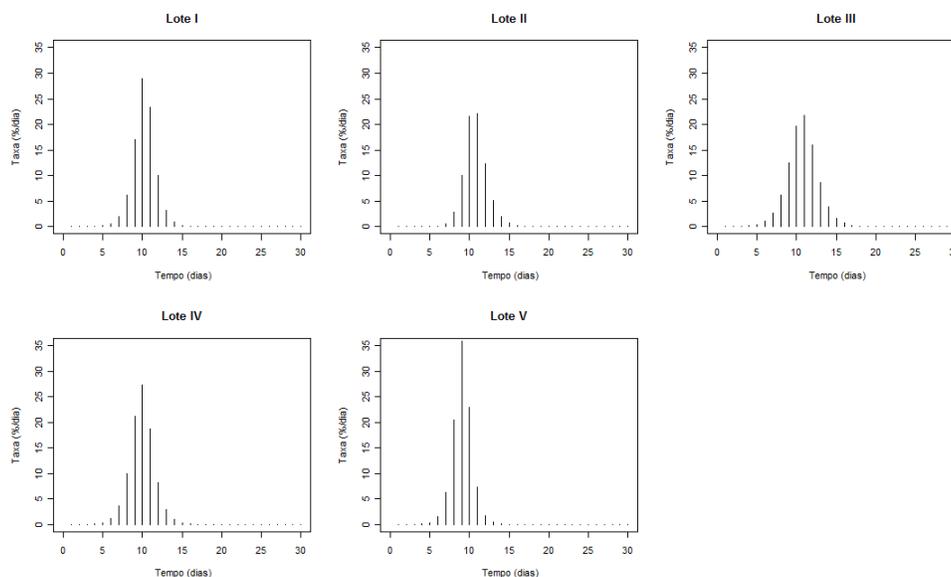


Figura 22 Histograma representativo da taxa de germinação absoluta (TGA) de sementes de café estimada pela função Logística nos cinco lotes

As estimativas dos parâmetros, para o modelo Gompertz, com erros independentes, são apresentados na Tabela 5.

Tabela 5 Estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, aos valores de porcentagem de germinação de sementes para os cinco lotes

Lotes	Parâmetros		
	α	β	γ
I	93,28	9,69	0,84
II	93,06	9,49	0,66
III	96,52	10,01	0,63
IV	95,47	9,32	0,78
V	97,23	8,62	1,02

Todos os parâmetros foram significativos em 1% pelo teste t.

Ao comparar as estimativas do parâmetro (α), obtidas pelo modelo Gompertz, percebe-se que os valores encontrados foram maiores que os obtidos pelo

modelo Logístico, como também diferenciando - se superiormente dos percentuais mínimos de germinação apresentados, exceto o lote V. As estimativas do parâmetro (β) apresentaram menores valores que os obtidos pelo modelo Logístico devido à sua interpretação que corresponde ao dia em que o potencial de germinação acumulada das sementes atinge 36,80%. Os valores referentes à velocidade de germinação, parâmetro (γ), mesmo que em menores valores não diferencia da explanação apresentada pelo modelo Logístico.

Nas Figuras 23 a 27, tem-se a representação gráfica dos dados e o ajuste do modelo Gompertz, com erros independentes.

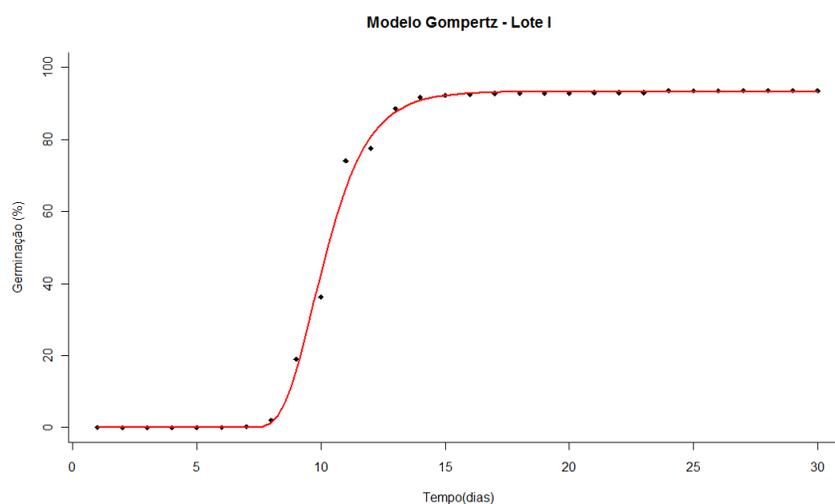


Figura 23 Modelo Gompertz na descrição de germinação de sementes de café-Lote I, em função do tempo com estruturas de erros independentes

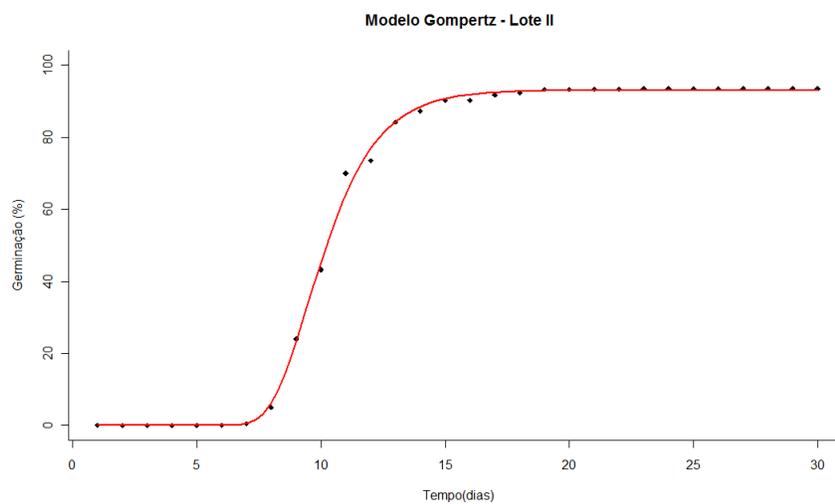


Figura 24 Modelo Gompertz na descrição de germinação de sementes de café - Lote II, em função do tempo com estruturas de erros independentes

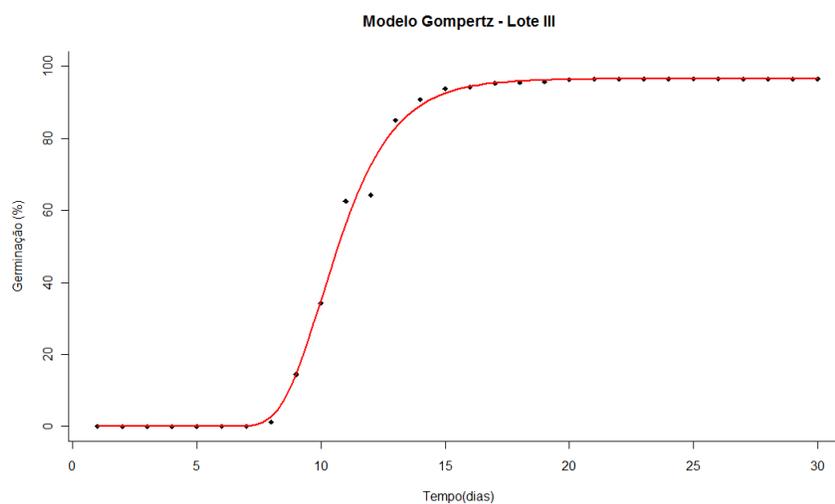


Figura 25 Modelo Gompertz na descrição de germinação de sementes de café - Lote III, em função do tempo com estruturas de erros independentes

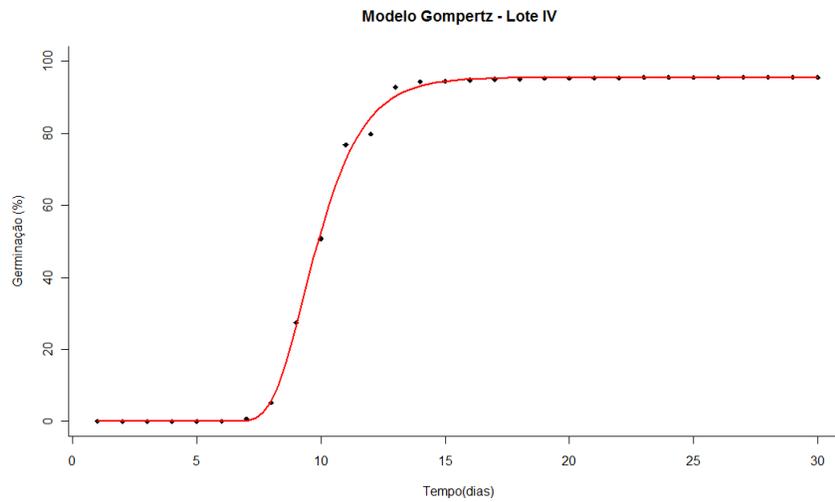


Figura 26 Modelo Gompertz na descrição de germinação de sementes de café - Lote IV, em função do tempo com estruturas de erros independentes

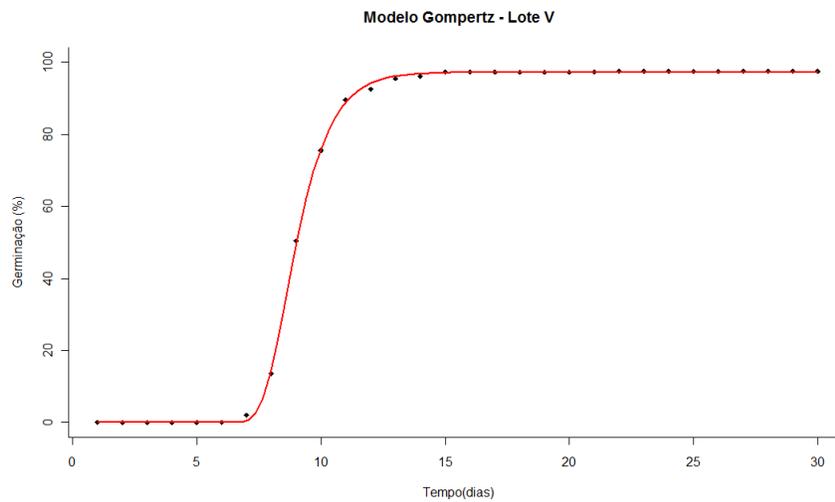


Figura 27 Modelo Gompertz na descrição de germinação de sementes de café - Lote V, em função do tempo com estruturas de erros independentes

O modelo Gompertz, de acordo com Gaspar-Oliveira *et al.*, (2008), proporcionou ajuste adequado aos dados de percentuais de germinação em função do tempo (em dias) para métodos de superação de dormência.

Como no ajuste do modelo Logístico, observa-se que pelo teste de Shapiro-Wilk, os erros relativos a todos os lotes analisados apresentaram distribuição normal para o modelo Gompertz, conforme Tabela 6.

Tabela 6 Estimativas da estatística do teste de Shapiro-Wilk dos resíduos e valor - p do modelo Gompertz para os cinco Lotes.

Lotes	Estatística do Teste Shapiro Wilk	Valor - p
I	0,9050	0,4382
II	0,9077	0,4541
III	0,8743	0,2842
IV	0,8958	0,3871
V	0,9209	0,5361

As TGA obtidas pelo modelo Gompertz têm sua interpretação semelhante ao do modelo Logístico, apresentando-se crescentes até atingirem os máximos de 28%/dia para Lote I, em torno de 21%/dia para o Lote II e 23%/dia para Lote III, 26%/dia para Lote IV e 34%/dia para Lote V, chegando ao ponto de inflexão da curva. A partir deste a taxa de germinação começa a diminuir exponencialmente, onde é possível por meio dos gráficos apresentados na Figura 28, verificar a assimetria do modelo Gompertz em relação ao ponto de inflexão.

Na Tabela 5, tem-se a estatística do teste da razão de máxima verossimilhança para os modelos Logístico e Gompertz, verificando-se para quais modelos e qualidades o ajuste com os resíduos correlacionados se mostrou mais plausível aos dados.

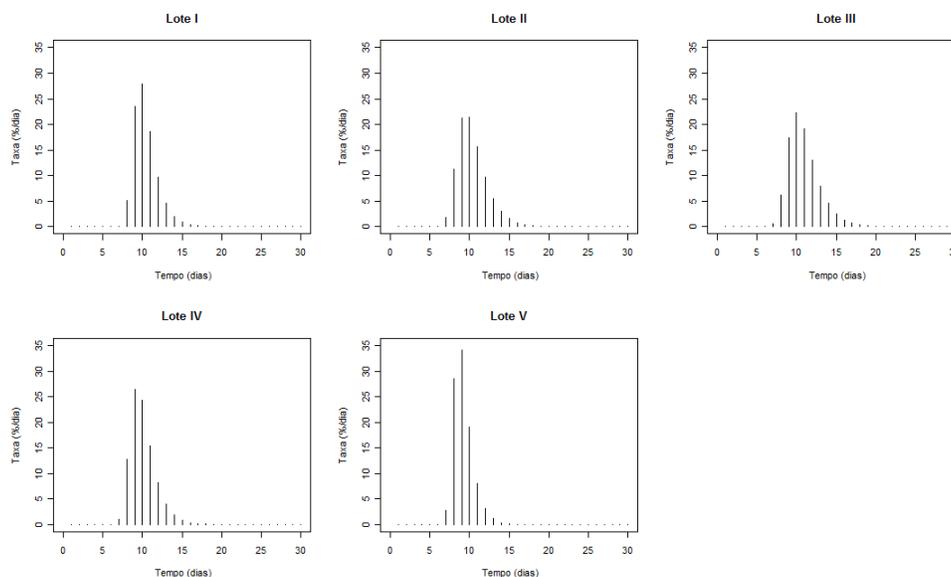


Figura 28 Histograma representativo da taxa de germinação absoluta (TGA) de sementes de café estimada pela função Gompertz nos cinco lotes

Tabela 7 Estatísticas do teste de razão de máxima verossimilhança (TRV) para os modelos estudados.

Qualidades	Modelos	Modelos	
		Logístico AR(1)	Gompertz AR(1)
I	Logístico	9,8417***	
	Gompertz		25,8619***
II	Logístico	1,5556 ^{NS}	
	Gompertz		9,3801***
III	Logístico	0,3875 ^{NS}	
	Gompertz		11,391***
IV	Logístico	1,1529 ^{NS}	
	Gompertz		22,9725***
V	Logístico	1,0140 ^{NS}	
	Gompertz		3,9992**

*** valores significativos a 1% ($P < 0,01$).

** valores significativos a 5% ($P < 0,05$).

Para o modelo Logístico, apenas para o lote I, o TRV, identificou um me-

lhor ajuste aos dados considerando aos resíduos correlacionados AR(1), já o ajuste com o modelo Gompertz em todos os lotes analisados, o melhor ajuste apresentou a estrutura de erros autocorrelacionados AR(1).

Na Tabela 8, observa-se que todas as estimativas dos parâmetros dos modelos Logístico e Gompertz com estrutura de erros autoregressivo de ordem 1 convergiu para valores próximos aos encontrados no ajuste com erros independentes. Verifica-se que o erro, associado ao índice de germinação das sementes, numa data, está negativamente correlacionado com o erro da data anterior.

Tabela 8 Estimativas dos parâmetros para o ajuste do modelo Logístico e Gompertz, com estrutura de erros autoregressivos de ordem 1, aos valores de porcentagem de germinação de sementes para os cinco lotes.

Lotes	Modelos	Parâmetros			
		α	β	γ	ϕ_1
I	Logístico	92,87	10,23	1,25	-0,52
	Gompertz	93,23	9,69	0,85	-0,74
II	Gompertz	93,08	9,49	0,66	-0,51
III	Gompertz	96,52	10,01	0,63	-0,55
IV	Gompertz	95,47	9,32	0,78	-0,72
V	Gompertz	97,26	8,61	1,01	-0,39

todos os valores significativos a 1% pelo teste t.

4.2 Seleção dos modelos ajustados

Através dos avaliadores de qualidade de ajuste apresentados na Tabela 9, observa-se o modelo Gompertz com estrutura de erros autoregressivos de ordem 1 apresentando um melhor ajuste aos dados, devido aos menores valores dos desvios padrão (DPR), os critérios de informação de Akaike (AIC) e Schwarz (BIC). Carvalho *et al.*, (2011), em estudo de modelos não lineares generalizados aplicados na predição da área basal e volume de *Eucalyptus* clonal, utilizaram os critérios acima

propostos para a seleção dos modelos devido à sua extrema relevância na análise de regressão. Já os valores obtidos no coeficiente de determinação, (R_{aj}^2) mostraram resultados semelhantes não podendo, assim, verificar a escolha do melhor modelo por essa estatística.

Tabela 9 Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_{aj}^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC) e critério bayesiano de Schwarz (BIC) para os modelos ajustados, na descrição de germinação de sementes para os cinco lotes

Lotes	Modelo	Critérios de seleção			
		R_{aj}^2	DPR	AIC	BIC
I	Logístico	0,99995	2,1329	135,4249	141,0297
	Gompertz	0,99998	2,1422	135,6855	141,2903
	Logístico AR(1)	0,99995	2,1122	127,5831	134,5891
	Gompertz AR(1)	0,99999	2,0609	111,8236	118,8296
II	Logístico	0,99975	2,3135	140,3023	145,9071
	Gompertz	0,99999	1,4631	112,8095	118,4143
	Logístico AR(1)	0,99968	2,3187	140,7467	145,7527
	Gompertz AR(1)	0,99999	1,4477	105,4295	112,4355
III	Logístico	0,99974	2,7985	151,7209	157,3257
	Gompertz	0,99999	2,0910	134,2351	139,8399
	Logístico AR(1)	0,99978	2,7977	153,3334	160,3394
	Gompertz AR(1)	0,99999	2,0613	124,8441	131,8500
IV	Logístico	0,99986	1,9117	128,8573	134,4651
	Gompertz	0,99999	1,3780	109,2151	114,8199
	Logístico AR(1)	0,99987	1,9111	130,1595	137,1655
	Gompertz AR(1)	0,99999	1,3119	88,2426	95,2486
V	Logístico	0,99991	1,4915	113,9644	119,5692
	Gompertz	0,99999	0,6221	61,5024	67,10771
	Logístico AR(1)	0,99965	2,0397	114,9503	121,9563
	Gompertz AR(1)	0,99999	0,6306	59,5031	66,5091

Nas Figuras 29 a 33, observa-se a representação gráfica do ajuste dos dois modelos não lineares que, segundo os avaliadores da qualidade de ajuste, obtiveram um melhor ajustamento aos dados. Nota-se que ambos os modelos descrevem de modo adequado a germinação das sementes, com curvas muito próximas aos valores observados, mas com melhor destaque para o modelo considerando a estrutura de erros AR(1). Cunha (2011), identificou o modelo Gompertz com erros autocorrelacionados de ordem 1 como o melhor modelo ao estudar os diâmetros dos frutos de ameixeira ao longo do tempo.

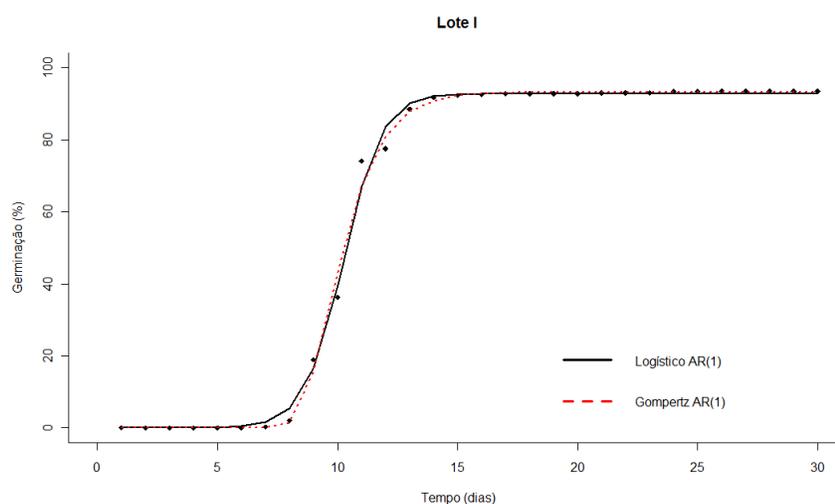


Figura 29 Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico e Gompertz, com estruturas de erros AR (1) para o lote I

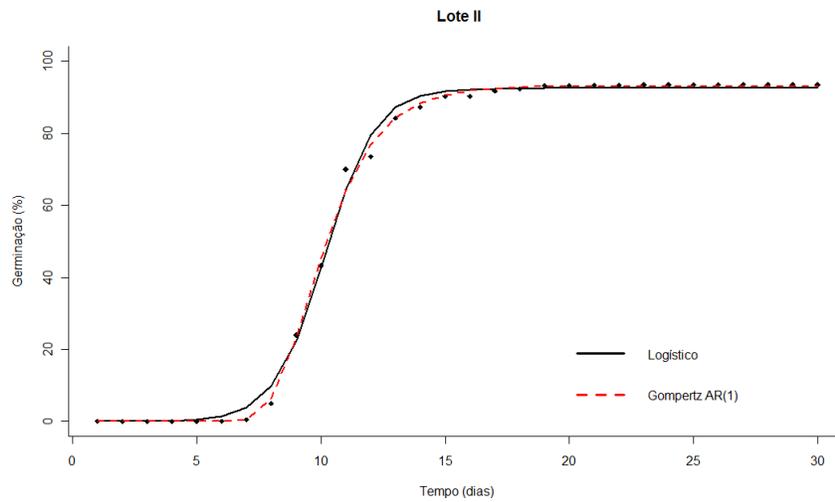


Figura 30 Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote II

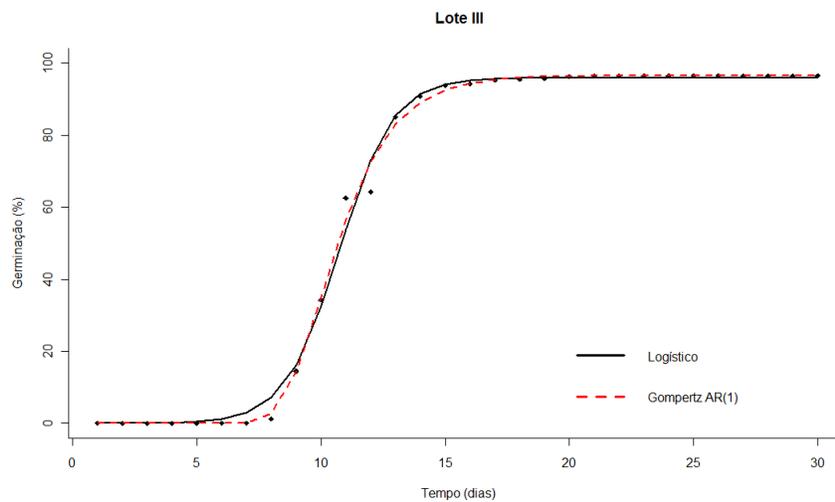


Figura 31 Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote III

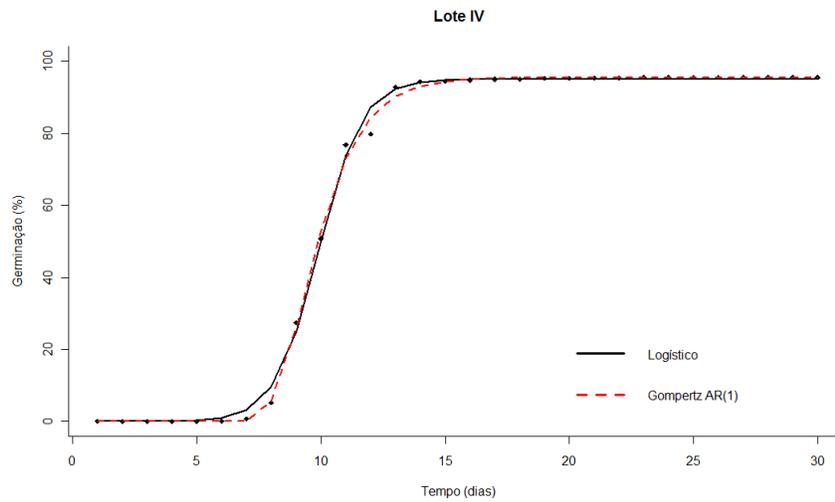


Figura 32 Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote IV

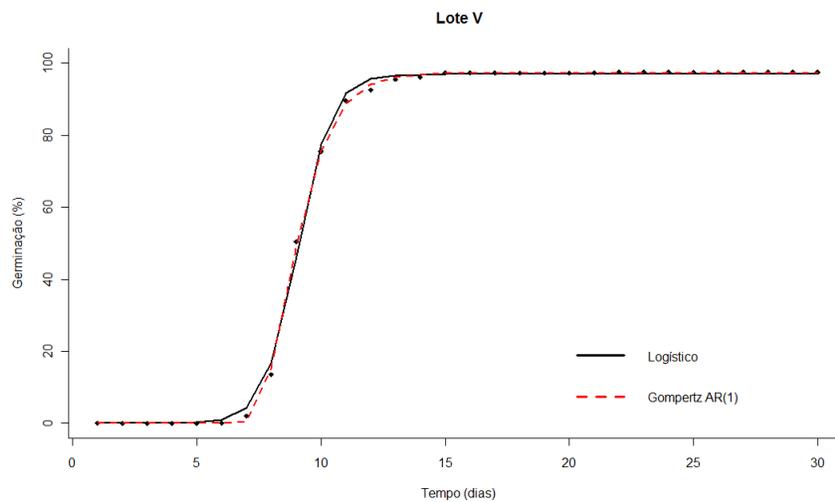


Figura 33 Valores médios da porcentagem de germinação acumulada, da curva de ajuste dos modelos Logístico com estrutura de erros independentes e Gompertz com estrutura de erro AR(1) para o lote V

5 CONCLUSÃO

Conclui-se que:

(i) O modelo de Gompertz, com estrutura de erros autoregressivos de primeira ordem, proporciona melhor qualidade de ajuste ao percentual de germinação acumulada que o modelo Logístico;

(ii) A germinação das sementes de café é descrita pelos modelos não lineares Logístico e Gompertz, apresentando adequada interpretação biológica dos parâmetros;

REFERÊNCIAS

- AKAIKE, H. A. New Look at the Statistical Model Identification. **IEEE Transactions on automatic control**, Minato-ku, Tokyo, Japan. v.19, n.6, p.716-723, Dec. 1974.
- BATES, D.M.; WATTS, D.G. **Nonlinear Regression Analysis and its Applications**. New York: John Wiley, 1988. 365p.
- BLACK M.; BEWLEY D.J.; HALMER. P. **The Encyclopedia of Seeds: Science Technology and Uses**. Wallingford, UK: CABI 2006. 900p.
- BOZSDOGAN, H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, v. 52, p.345-370. 1987.
- BRASIL. Ministério da Agricultura Pecuária e Abastecimento/MAPA. Café. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cafe/saiba-mais>>. Acesso em 20 Nov.2011.
- BRASIL. Ministério da Agricultura Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. **Regras para análise de sementes**. Brasília: MAPA, p.147-220. 2009.
- BREUSCH, T ; PAGAN, A. Teste simples para heterocedasticidade e coeficiente de variação aleatória Econométrica. **Sociedade Econométrica** v.47, p.1287-1294. 1979.
- CARON, B. O.; MEIRA, W. R.; SCHIMDT, D.; SANTOS FILHO, B. G.; MEDEIROS, S. L. P.; MANFRON, P. A.; MULLER, L.; Análise de crescimento de plantas de aroeira vermelha no município de JI-Paraná, RO. **Revista da FZVA**. Uruguaiana, v.14, n.1, p.1-13, 2007.
- CARVALHO, S. de P. C. e; CALEGARIO, N.; SILVA, F. F. e; BORGES, L. A. C.; MENDONÇA, A. R. de; LIMA, M. P. de; Modelos não lineares generalizados

aplicados na predição da área basal e volume de *Eucalyptus* clonal **Cerne**, Lavras, v. 17, n. 4, p. 541-548, out./dez. 2011.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira Café, Safra 2011 segunda estimativa, maio/2011**. Brasília: Conab, 2011. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/11_05_10_09_04_16_boletim_ccafe_portugues_-_maio_-_2011_2o_lev.>. Acesso em: 1 set. 2011.

CUNHA, L. S. **Modelos de regressão para a descrição do crescimento de frutos da ameixeira**. Dissertação (Mestrado em Agronomia/Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2011. 79p.

DEDECCA, D. M. Anatomia e desenvolvimento ontogenético de *Coffea arabica* L. var. *typica* Cramer. **Bragantia**, Campinas, v.16, p.315-366, 1957.

DRAPER, N.R.; SMITH, H. **Applied regression analysis**. 3.ed. New York.:Jonh Wiley, 1998. 706p.

DUARTE, E.F.; CARNEIRO, I.F.; SILVA, N.F.; GUIMARÃES, N.N.R. Características físicas e germinação de sementes de *Dyckia goehringii* Gross e Rauh (Bromeliaceae) sob diferentes temperaturas. **Pesquisa Agropecuária Tropical**, Goiânia, v.40, n.4, p.422-429, dez.2010.

DURBIN, J.; WATSON, G.S. Teste de Correlação Serial em Mínimos Quadrados de Regressão I. **Biometrika** v.37 , p.409-428, 1950.

EIRA, M.T. S.; SILVA, E. A. A. da.; CASTRO, R. D.; DUSSERT, S.; WALTERS, C.; BEWLEY, J. D.; HILHORST, H. W. M. Coffee seed physiology. **Brazilian Journal of Plant Physiology**, Londrina, v.18, n.1,p.149-163, Mar.2006.

FERRARI, T. B.; FERREIRA. G.; MISCHAN, M. M.; PINHO, S. Z. Germinação de sementes de maracujá-doce (*Passiflora alata* Curtis) : Fases e efeito de reguladores vegetais. **Biotemas**, v.21, n.3, p.65-74, set. 2008.

FLORENT, B.; DELIGNETTE-MULLER, M. L. nlstools: tools for nonlinear regression diagnostics. **R Development Core Team**, 2011.

GALLANT, A.R. **Nonlinear Statistical Models**. New York.:Jonh Wiley. p.1-146. 1987. 624p.

GASPAR-OLIVEIRA, C. M.; MARTINS, C. C; NAKAGAWA, J.; CAVARIANI, C. Duração do teste de germinação de *Brachiaria brizantha* cv. marandu (Hochst. ex A. Rich.) Stapf. **Revista Brasileira de Sementes**, Londrina, vol.30, n.3, p.30-38, 2008.

GAZOLA, S.; SCAPIM, C. A.; GUEDES, T. A.; BRACCINI, A. de L. e. Proposta de modelagem não-linear do desempenho germinativo de sementes de milho híbrido. **Ciência Rural**, Santa Maria, v.41, n.4, p.551-556, abr. 2011.

GONZAGA, T.W.C.; MATA, M.; SILVA, H.; DUARTE, M.E.M. Crioconservação de sementes de aroeira (*Astronium urundeuva* Engl.) e baraúna (*Schinopsis brasiliensis* Engl). **Revista Brasileira de Produtos Agroindustriais**, Campina Grande, v.5, n.2, p.145-154, out. 2003.

GUERREIRO FILHO, O.; FAZUOLI, L. C.; EIRA AGUIAR, A. T. Cultivares de *Coffea arabica* selecionadas pelo IAC: características botânicas, tecnológicas, agrônômicas e descritores mínimos. 2006. Disponível em: <http://www.infobibos.com/Artigos/2006_2/Cultivares_cafe/Index.htm>. Acesso em: 20 nov. 2011.

GUILLÉN, S., BENÍTEZ, J., MARTÍNEZ-RAMOS., M., CASAS, A. Seed germination of wild, in situ-managed, and cultivated populations of columnar cacti in the Tehuacán Cuicatlán Valley, México. **Journal of Arid Environments**, v.73, p.407-413, 2009.

HOFFMANN, R.; VIEIRA, S. **Análise de regressão: uma introdução à econometria**. São Paulo: HUCITEC, 379p. 1998.

IKEDA, F.S.; CARMONA, R.; MITJA, D.; GUIMARÃES, R.M. Luz e KNO₃ na germinação de sementes de *Ageratum conyzoides* L. sob temperaturas constantes e alternadas **Revista Brasileira de Sementes**, Londrina, v.30, n.2, p.193-199,

2008.

MACHADO, E. J. **Uso do 'Bootstrap' na estimação de parâmetros em modelos não-lineares- uma aplicação em mecanismos cinéticos de Michaelis-Menten.** Dissertação (Mestrado em Estatística e Experimentação Agropecuária.) - Universidade Federal de Lavras, Lavras, 2008. 110 p.

MARCOS FILHO, J. **Fisiologia de sementes de plantas cultivadas.** Piracicaba: Fealq, 2005. 495p.

MARTINS FILHO, S.; SILVA, F. F.; CARNEIRO, A. P. S.; MUNIZ, J. A. Abordagem bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, Santa Maria, v.38, n.6, p. 1516-1521, 2008.

MATIELLO, J.B.; SANTINATO, R.; GARCIA, A.W.R.; ALMEIDA, S.R.; FERNANDES, D.R. **Cultura de café no Brasil: novo manual de recomendações.** Rio de Janeiro, MAPA/PROCAFE, 2005. 438p.

MAZUCHELI, J.; ACHCAR, J.A. Algumas considerações importantes em regressão não linear. **Acta Scientiarum**, Maringá, v.24, n.6, p.1761-1770. 2002.

MAZZINI, A.R.A.; MUNIZ, J.A.; Aquino, L. H; SILVA, F. F. Análise da curva de crescimento de machos Hereford. **Ciência e Agrotecnologia**, Lavras, v.27, n.5, p.1105-1112. 2003.

MAZZINI, A.R.A.; MUNIZ, J.A.; SILVA, F. F.; Aquino, L. H. Curva de crescimento de novilhos Hereford: heterocedasticidade e resíduos autoregressivos. **Ciência Rural**, Santa Maria, v.35, n.2, p. 422-427, 2005.

MENDES, P.N., MUNIZ, J.A., SILVA, F. F., MAZZINI, A. R. de A. Modelo logístico difásico no estudo de crescimento de fêmeas Hereford. **Ciência Rural**, Santa Maria, v. 38, p. 1984-1990. 2008.

MOTULSKY, H.; CHRISTOPOULOS, A. **Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting.** 4.ed.

San Diego, CA:GraphPad Software, 2003. 351p.

NASSIF, S.M.L.; VIEIRA, I. G.; FERNANDES, G. D. Fatores Externos (ambientais) que Influenciam na Germinação de Sementes. Informativo Sementes IPEF - Abril 1998. Disponível em : <<http://www.ipef.br/tecsementes/germinacao.asp>>. Acesso em: 22 Nov. 2011.

PEREIRA, J.M.; MUNIZ, J.A.; SILVA, C.A. Non linear models to predict nitrogen mineralization in na oxisol. **Scientia Agricola**, Santa Maria, v.62, p.395-400, 2005.

PINHEIRO, J. C., BATES, D. M., DEBROY, S., SARKAR, D. nlme: Linear and Nonlinear Mixed Effects Models. **R Development Core Team**, 2011.

POPINIGIS, F. **Fisiologia da Semente**. 2.ed. Brasília, 1985. 298p.

RATKOWSKY, D.A. **Nonlinear Regression Modeling**. New York, Marcel Dekker, INC., p. 61 - 73. 1983.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, versão 2.13.1. 2011. Disponível em <<http://www.r-project.org>>. Acesso em: 5 out. 2011.

SAVIAN, T. V. **Estimação dos parâmetros no modelo para degradabilidade in situ de Mertens e Loften**. Dissertação (Mestrado em Agronomia/Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2005. 55p.

SAVIAN, T.V.; MUNIZ, J.A.; AQUINO, L.H.; BANYNS, V.L.; FERREIRA, D.F. Degradability study of neutral detergent fiber of coast cross (*Cynodon dactylon* x *Cynodon nlemfunensis*). **Ciência e Agrotecnologia**, Lavras, v.31, p.1184-1190, 2007a.

SAVIAN, T. V.; MUNIZ, J.A. A study of in situ degradability: heterogeneity of variances and correlated errors. **Scientia Agricola**, Piracicaba, v.64, p.548-554,

2007b.

SAVIAN, T. V.; MUNIZ, J. A.; SÁFADI, T.; SILVA, F. F. Análise bayesiana para modelos de degradabilidade ruminal. **Ciência Rural**, Santa Maria, v.39, p.2169-2177, 2009.

SCHWARZ, G. Estimating the dimensional of a model. **Annals of Statistics**, Hayward, v.6, n.2, p.461-464, Mar. 1978.

SEBER, G.A.F. **Linear Regression Analysis**. New York.:Jonh Wiley. 1977. 465p.

SEBER, G.A.F.; WILD, C.J. **Nonlinear regression**. New York: John Wiley, 1989. 792p.

SHAPIRO,S.S.; WILK, M. B. An Analysis of variance test for normality. **Biometrika**, Cambridge, v. 52, n. 3/4, p.591-611, dec. 1965.

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

SOUZA, E.M.; MUNIZ, J.A.; MARCHI, G.; GUILHERME, L. R. G. Modelagem não-linear da extração de zinco de um lodo de esgoto. **Acta Scientiarum. Technology**, Maringá, v.32, n.2, p.193-199, 2010.

SOUZA, G. da S. **Introdução aos modelos de regressão linear e não-linear**. Brasília: Embrapa-SPI/Embrapa-SEA. 1998. 489p.

TERRA, M.P.; MUNIZ, J. A.; SAVIAN, T.V. Ajuste dos modelos Logístico e Gompertz aos dados de crescimento de frutos da tamareira anã (*Phoenix roebelenii* O'BRIEN). **Magistra**, Cruz das Almas, v.22, n.1, p.1-7, 2010.

TOMAZ, C. A; C. M.; MARTINS. C. C; CARVALHO,L.R.; NAKAGAWA, J. Duração do teste de germinação do Capim-Tanzânia. **Revista Brasileira de**

Sementes, v.32, n.4, p.080-087, 2010.

VIDIGAL, D. S.; DIAS, D. C. F. dos S.; DIAS, L. A. dos S.; FINGER, F. L.
Changes in seed quality during fruit maturation of sweet pepper. **Scientia
Agricola**. Piracicaba, v.68, n.5, out. 2011.

ANEXO A - Rotina no R

```

-----
                        Rotina no R utilizada para a
obtenção das estimativas dos parâmetros dos modelos estudados.
#-----

#pacotes necessários para a análise.
require(manipulate)
require(nlstools)
require(car)
require(nlme)
require(lmtest)

# modelos estudados
# Logístico:  $b1/(1+\exp(-b3*(x-b2)))$ 
# Gompertz:  $b1*\exp(-\exp(-b3*(x-b2)))$ 

rm(list=ls(all=TRUE))
#--# função para o cálculo do R2
R2 <- function(residuals, observed){
  1-(sum(residuals)^2)/sum((observed-mean(observed))^2)
}
#-----
germinaçãocafe<-read.table("cafe medias.txt", header=TRUE,)
germinaçãocafe
germinaçãocafe$var <- factor(germinaçãocafe$var)
germinaçãocafe$var
#-----
#função que representa o modelo
logi <- function(x, b1, b2, b3){
  b1/(1+exp(-b3*(x-b2)))
}
#-----
# diagnose gráfica e primeiro chute para variedade I
plot(y~x, data=germinaçãocafe, col=as.numeric(germinaçãocafe$var=="I"),
xlab="Dias", ylab="Porcentagem de Germinação",bty="n",pch=22)
b1<- 94; b2 <- 10; b3 <- 0.9 # primeiro chute
curve(logi(x, b1, b2, b3), add=TRUE, col=2)
#-----
# procedimento gráfico para obter chutes e conhecer a função dos parâmetros
require(manipulate) # carrega pacote que permite manipulação gráfica

start <- list()      # cria uma lista vazia para receber os valores finais

#-----
manipulate({plot(y~x, data=subset(germinaçãocafe, var=="I"))
  curve(logi(x, b1=b1, b2=b2, b3=b3), add=TRUE)
  start <<- list(b1=b1, b2=b2, b3=b3)},
b1=slider(90, 96, initial=90), # chute para o intervalo de assintota
b2=slider(8, 13, initial=9), # chute para o parâmetro de localização
b3=slider(0.7, 1.8, initial=0.7) # chute para a taxa de crescimento
)
#-----
start # valores salvos do último movimento
#-----
rm(list=ls(all=TRUE))
# Ajuste do modelo logístico aos dados com os chutes do procedimento gráfico
##ajuste de do modelo logistico sem correlção para as variáveis

```

```

nI <- nls(y~b1/(1+exp(-b3*(x-b2))),
data=subset(germinaçãocafe, var=="I"), start=c(b1=93,b2=10,b3=1.2632))
summary(nI) # quadro de estimativas
R2(residuals(nI), germinaçãocafe$y)

# procedimentos para obter - testes estatísticos na análise de resíduos
residuosnI<-summary(nI)$resid
residuosnI
shapiro.test(residuosnI) # teste normalidade
#extraíndo a matriz gradiente avaliada nas estimativas dos parâmetros
FIc <- attr(nI$m$fitted(), "gradient")
# matriz gradiente para a lm(),
mlc <- lm(y~-1+FIc, data=subset(germinaçãocafe, var=="I"))
FIIn<-summary(nI)
bptest(mlc) # teste para homogeneidade de variâncias
durbin.watson(mlc)# teste para independência dos resíduos
# gráfico dos resíduos
nr <- nlsResiduals (nI)
plot (nr, que = 0)
  mtext("Modelo Logístico - Lote I",
        outer=TRUE, line=-2, cex=1.4)

# Para o modelo Gompertz e os demais lotes
o procedimento foi o mesmo acima descrito

#-----
# Ajuste dos modelos considerando os erros independentes e correlacionados.
#-----

#Ajuste do modelo Logístico -lote I considerando os erros independentes

nI <- gnls(y~b1/(1+exp(-b3*(x-b2))),
data=subset(germinaçãocafe, var=="I"), start=c(b1=93,b2=10,b3=1.0272))
summary(nI) # quadro de estimativas
R2(residuals(nI), germinaçãocafe$y)
shapiro.test(residuosnI)
residuosnI<-summary(nI)$resid
residuosnI
shapiro.test(residuosnI)
#-----
# gráfico dos dados com a curva estimada
lis <- c(list(x=NULL), as.list(coef(nI)), body(logi))
plot(y~x,data=subset(germinaçãocafe, var=="I"),xlab=" Tempo(dias)",
ylab="Germinação (%)",bty="l",pch=18,ylim=c(0,100), main= " Lote I" )
tmp <- as.function(lis)
curve(tmp, add=TRUE, col=2, lwd=2)
# -----
# Ajuste do modelo Logístico considerando - AR(1)

n1I <- gnls(y~b1/(1+exp(-b3*(x-b2))),
data=subset(germinaçãocafe, var=="I"),correlation=corAR1(form=~x),
start=c(b1=93,b2=10,b3=1.2632))
summary(n1I)
confint.default(n1I)
R2(residuals(n1I), germinaçãocafe$y)
residuosclI<-summary(n1I)$resid
shapiro.test(residuosclI)
#-----
# gráfico dos dados com a curva estimada

lis <- c(list(x=NULL), as.list(coef(n1I)), body(logi))
plot(y~x,data=subset(germinaçãocafe, var=="I"),xlab="Dias",
ylab="Porcentagem de Germinação",bty="l",pch=16, main="Lote I")
tmp <- as.function(lis)

```

```

curve(tmp, add=TRUE, col=2, lwd=2.0)
#-----
# Ajuste modelo - Gompertz - Erros independentes e correlacionados
#-----
#Ajuste do modelo Gompertz sem correlação
gI <- gnls(y~b1*exp(-exp(-b3*(x-b2))),
data=subset(germinaçãocafe, var=="I"), start=c(b1=94,b2=10,b3=0.9368),
control=nls.control(maxiter=500, minFactor=1/10000))

summary(gI) # quadro de estimativas
confint.default(gI)
R2(residuals(gI), germinaçãocafe$y)
residuoscgI<-summary(gI)$resid
shapiro.test(residuoscgI)
#-----
# gráfico dos dados com a curva estimada

lis <- c(list(x=NULL), as.list(coef(gI)), body(gomp))
plot(y~x,data=subset(germinaçãocafe, var=="I"),xlab=" Tempo (dias)",
ylab="Germinação (%)",bty="l",ylim=c(0,100),
main=" Modelo Gompertz - Lote I" )
tmp <- as.function(lis)
curve(tmp, add=TRUE, col=2, lwd=2.0)
#-----
#Ajuste do modelo Gompertz considerando - AR(1)

glI <- gnls(y~b1*exp(-exp(-b3*(x-b2))),
data=subset(germinaçãocafe, var=="I"),correlation=corAR1(form=~x),
start=c(b1=94,b2=10,b3=0.9368),
control=nls.control(maxiter=500, minFactor=1/10000))
summary(glI)
confint.default(glI)
R2(residuals(glI), germinaçãocafe$y)
residuosg1I<-summary(glI)$resid
shapiro.test(residuosg1I)
#-----
# gráfico dos dados com a curva estimada

lis <- c(list(x=NULL), as.list(coef(glI)), body(gomp))
plot(y~x,data=subset(germinaçãocafe, var=="I"),
xlab=" Tempo (dias)",ylab="Germinação (%)",bty="l",pch=16,
main=" Modelo Gompertz AR(1) - Lote I")
tmp <- as.function(lis)
curve(tmp, add=TRUE, col=2, lwd=2.0)
#-----
#Obtenção dos avaliadores de qualidade de ajuste AIC BIC e TRV
#-----
anova(nI,n1I)
anova(gI,glI)
#-----
# Obtenção da Taxa de germinação absoluta
#-----Derivadas da função-----
logis <- expression(b1/(1+exp(-b3*(x-b2))))
D(slogis,"x")
# gráfico
txcal<-read.table("txcrescabslogi.txt", header=TRUE)
txcal
plot(ym~x,type="h",data=subset(txcal, var=="I"),xlab=" Tempo (dias)",
ylab="Taxa (%/dia)",xlim=c(0,30),ylim=c(0,35),pch=17, main= "Lote I")

# Para o modelo gompertz e os demais lotes
o procedimento foi o mesmo acima descrito
#-----#-----

```