



LUCAS SANTANA DA CUNHA

**MODELOS DE REGRESSÃO PARA A
DESCRIÇÃO DO CRESCIMENTO DE FRUTOS
DA AMEIXEIRA**

**LAVRAS - MG
2011**

LUCAS SANTANA DA CUNHA

**MODELOS DE REGRESSÃO PARA A DESCRIÇÃO DO
CRESCIMENTO DE FRUTOS DA AMEIXEIRA**

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. Augusto Ramalho de Moraes

**LAVRAS - MG
2011**

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

Cunha, Lucas Santana da.

Modelos de regressão para a descrição do crescimento de frutos da ameixeira / Lucas Santana da Cunha. – Lavras : UFLA, 2011.
79 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2011.
Orientador: Augusto Ramalho de Moraes.
Bibliografia.

1. Análise de regressão. 2. Curva de crescimento. 3. Modelos não-lineares. 4. Regressão polinomial. 5. Diâmetro de frutos. I. Universidade Federal de Lavras. II. Título.

CDD – 519.536

LUCAS SANTANA DA CUNHA

**MODELOS DE REGRESSÃO PARA A DESCRIÇÃO DO
CRESCIMENTO DE FRUTOS DA AMEIXEIRA**

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 08 de fevereiro de 2011.

Dra. Taciana Villela Savian ESALQ

Dra. Ester Alice Ferreira UFLA

Dr. João Domingos Scalon UFLA

Dr. Augusto Ramalho de Moraes

Orientador

**LAVRAS - MG
2011**

AGRADECIMENTOS

Primeiramente a Deus, por me ter dado sabedoria e discernimento para que eu concluísse o curso de forma adequada.

Ao professor Augusto Ramalho de Moraes, pela atenção, dedicação e auxílio no desenvolvimento deste trabalho.

A todos os demais professores do Departamento de Ciências Exatas (DEX) da UFLA, em especial aos professores Renato, Scalon, Lucas, Daniel, Marcelo, Delly.

A todos os funcionários do DEX e da UFLA pelos serviços prestados, em especial à secretária Josi, sempre competente e dedicada ao trabalho.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

A todos os amigos da Pós-Graduação do DEX da UFLA. Especialmente, o Jair, Diogo, Thalita e Adriana.

Aos amigos de república Renato, Ricardo, Matheus, Wanda, Elma, Amanda e Giceli, pelo acolhimento na hora em que mais precisei.

Aos meus pais, Celso e Eli, pelo apoio, carinho, amor e dedicação dado desde o início de minha vida.

Ao meu irmão Tiago e aos meus avós; Armindo, Enelzira, Osvaldo e Arlete, pela compreensão nas vezes em que estive ausente.

À minha namorada Elienai, que apesar da distância sempre esteve presente nas horas boas e ruins desde o primeiro instante em que vim para Lavras.

Muito obrigado a todos!

RESUMO

Durante o período de crescimento e maturação, o fruto da ameixeira passa por mudanças físicas e químicas, sugerindo que esses indicadores de maturação determinam o momento adequado da colheita. A avaliação da tendência de desenvolvimento dos frutos e a modelagem de seu tipo de crescimento podem contribuir para melhor conhecimento e aplicação eficiente das práticas agrícolas. O objetivo desse estudo foi o de avaliar e comparar diferentes tipos de modelos de regressão lineares e não lineares utilizados na descrição do desenvolvimento de frutos da ameixeira “Gulfblase”, bem como verificar as pressuposições de validação inerentes a estes modelos. Foram utilizados dados provenientes de um experimento conduzido no ano de 2008, no município de Delfim Moreira, o qual está localizado no extremo sul do estado de Minas Gerais. Foi utilizado o delineamento experimental inteiramente ao acaso. Avaliou-se o desenvolvimento do fruto da ameixeira “Gulfblase” periodicamente ao longo de 50 dias após 31 dias do início da frutificação. Observou-se que o diâmetro do fruto durante sua fase de crescimento foi ajustado pelo modelo de regressão polinomial quadrático e os modelos não lineares. Os modelos não lineares, com estrutura de erros autorregressivos de ordem 1, descrevem melhor o crescimento dos frutos da ameixeira “Gulfblase” ao longo do tempo de avaliação.

Palavras-chave: Análise de regressão. Curva de crescimento. Modelo não-linear.

ABSTRACT

During the period of growth and maturation, the fruit of plum passes through physical and chemical changes, suggesting that these indicators determine the appropriate time of harvesting. The trend evaluation of fruit development and modeling its type of growing can contribute to a better understanding and effective implementation of agriculture practices. The aim of this study was to evaluate and compare different models of linear and nonlinear regression used to describe the development of plum fruits "Gulfbalse" and check the assumptions inherent in validating these models. It were applied data from an experiment conducted in 2008 in the town of Delfim Moreira, which is located in southern Minas Gerais state. It was used the randomized experimental design, evaluating the development of the fruit of plum tree "Gulfbalse" periodically over 50 days after 31 days after fructification. It was observed that the diameter of the fruit during its growth phase was fitted by quadratic polynomial regression model and nonlinear models. Nonlinear models with autoregressive error structure of order one better describe the fruit growth of plum "Gulfbalse" over time of evaluation.

Keywords: Regression Analysis. Growth curves. Nonlinear models.

LISTA DE GRÁFICOS

Gráfico 1	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo linear polinomial com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R_a^2) para as quatro plantas.....	40
Gráfico 2	Representação gráfica da relação entre os quantis observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo linear polinomial para as quatro plantas; histograma dos resíduos; estatística do teste de normalidade de Shapiro-Wilk.....	41
Gráfico 3	Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo polinomial ao longo do tempo para as quatro plantas de ameixeira avaliadas.....	43
Gráfico 4	Representação gráfica dos intervalos de confiança das estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	46
Gráfico 5	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R_a^2) para as quatro plantas.....	47
Gráfico 6	Representação gráfica da relação entre os quantis	

	observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo Gompertz com erros independentes, para as quatro plantas; estatística do teste de normalidade de Shapiro-Wilk.....	48
Gráfico 7	Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo Gompertz ao longo do tempo para as quatro plantas de ameixeira avaliadas.....	49
Gráfico 8	Representação gráfica dos intervalos de confiança das estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	51
Gráfico 9	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Logístico com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R_a^2) para as quatro plantas.....	52
Gráfico 10	Representação gráfica da relação entre os quantis observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo Logístico, com erros independentes, para as quatro plantas; estatística do teste de normalidade de Shapiro-Wilk.....	53
Gráfico 11	Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo Logístico ao longo do tempo para as quatro plantas de ameixeira avaliadas.....	54
Gráfico 12	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de	

	avaliação, da curva de ajuste do modelo Gompertz com estrutura de erros autorregressiva de ordem 1 e estimativa do coeficiente de determinação ajustado (R_a^2) para as quatro plantas.....	57
Gráfico 13	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Logístico com estrutura de erros autorregressiva de ordem 1 e estimativa do coeficiente de determinação ajustado (R_a^2) para as quatro plantas.....	59
Gráfico 14	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz e Logístico com estrutura de erros autorregressiva de ordem 1 para a planta “8”.....	62
Gráfico 15	Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz e Logístico com estrutura de erros autorregressiva de ordem 1 para as plantas “14”, “29” e “36”.....	64

LISTA DE TABELAS

Tabela 1	Análise de Variância para o diâmetro dos frutos da ameixeira “Gulfblase”.....	36
Tabela 2	Estimativas dos parâmetros para o ajuste do modelo polinomial, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	38
Tabela 3	Estatísticas do teste de Durbin Watson dos resíduos do modelo linear polinomial quadrático para as quatro plantas.....	44
Tabela 4	Estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	44
Tabela 5	Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Gompertz para as quatro plantas.....	50
Tabela 6	Estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	50
Tabela 7	Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Logístico para as quatro plantas.....	55
Tabela 8	Estimativas dos parâmetros para o ajuste do modelo Gompertz, com estrutura de erros autorregressiva de ordem 1, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	55
Tabela 9	Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Gompertz AR(1) para as quatro plantas.	58
Tabela 10	Estimativas dos parâmetros para o ajuste do modelo Logístico, com estrutura de erros autorregressiva de ordem 1,	

	ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas.....	58
Tabela 11	Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Logístico AR(1) para as quatro plantas	60
Tabela 12	Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_a^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC), teste de Durbin Watson (DW), para os modelos ajustados para a planta “8” da cultivar “Gulfblase” de ameixa.....	61
Tabela 13	Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_a^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC), teste de Durbin Watson (DW), para os modelos ajustados para as plantas “14”, “29” e “36” da cultivar “Gulfblase” de ameixa.....	63

SUMÁRIO

1	INTRODUÇÃO.....	13
2	REFERENCIAL TEÓRICO.....	15
2.1	A cultura da ameixeira.....	15
2.2	A modelagem do crescimento de frutos.....	16
2.3	Modelos de regressão linear.....	17
2.3.1	O Ajuste pelo método de mínimos quadrados ordinários.....	19
2.4	Modelos de regressão não linear.....	20
2.4.1	Estimação dos parâmetros de modelos não lineares.....	21
2.4.2	Processos iterativos.....	23
2.4.2.1	Método Gauss Newton.....	24
2.4.2.2	Método de Marquardt.....	25
2.5	Análise dos resíduos.....	26
3	MATERIAL E MÉTODOS.....	30
3.1	Realização do experimento.....	30
3.2	Modelos testados.....	30
3.2.1	Modelo linear polinomial.....	31
3.2.2	Modelos não lineares.....	31
3.3	Seleção dos modelos.....	32
3.3.1	Coefficiente de determinação ajustado.....	33
3.3.2	Desvio padrão residual.....	33
3.3.3	Critério de informação de Akaike (AIC).....	34
3.4	Recursos computacionais.....	34
4	RESULTADOS E DISCUSSÃO.....	36
4.1	Análise de variância.....	36
4.2	Ajuste do modelo linear polinomial.....	37
4.3	Ajuste dos modelos não lineares.....	44
4.4	Seleção dos modelos ajustados.....	60
5	CONCLUSÕES.....	66
	REFERÊNCIAS.....	67
	APÊNDICES.....	72

1 INTRODUÇÃO

A produção de ameixas com elevado padrão de qualidade, com tamanho, aparência, sabor e comercialmente aceitável, é fundamental para o êxito comercial da cultura.

A ameixa, originária das regiões frias da Europa e da China, tem sido cultivada na região Sul do Brasil, em especial nas áreas montanhosas, de clima ameno. Em relação as suas características nutricionais, o fruto fornece ao organismo humano sais minerais e vitaminas (em especial a A). A ameixa é muito utilizada em preparo de doces, conservas, geleia e licores. Ela possui também propriedades medicinais, que ajuda a tratar de afecções hepáticas (cálculos da vesícula biliar, hepatite, icterícia, fígado preguiçoso), digestão difícil, intestino preso e tosse.

Durante a maturação, o fruto da ameixeira sofre mudanças de cor, aroma, sabor e textura. Os indicadores de maturação determinam o momento adequado da colheita. Os dois fatores mais importantes referentes à colheita implicam em cuidado durante a mesma e a maturação adequada.

Diante desses fatos, o conhecimento do crescimento do fruto da ameixeira é fundamental para um manejo adequado e detecção de problemas no desenvolvimento da cultura. Assim, o desenvolvimento e aperfeiçoamento de métodos que permitam estudar o crescimento dos frutos constituem importante ferramenta no processo de descrever e prever relações entre as características envolvidas no processo de maturação das frutas. Uma das formas é encontrar modelos que relacionem variáveis que descrevem esse processo; e esses modelos envolvem as curvas de crescimento, as quais envolvem o ajuste de modelos de regressão linear e não linear, que podem ser usados para interpretação do processo envolvido no sistema de produção.

A análise do crescimento e do desenvolvimento dos frutos é fundamental para auxiliar o processo de conhecimento dos fatores que afetam as mudanças que ocorrem nos constituintes físicos e químicos de um fruto.

Tendo em vista que há poucos relatos, na literatura consultada, sobre curvas de crescimento que tenham usados os modelos não lineares para descrição do crescimento de frutos da ameixeira, este trabalho teve por objetivo estudar as curvas de crescimento referentes ao diâmetro longitudinal de frutos de ameixeira utilizando um modelo de regressão polinomial e os modelos não lineares Logístico e Gompertz. Incluiu-se nos objetivos, a verificação das pressuposições de validação de normalidade e independência dos erros.

2 REFERENCIAL TEÓRICO

2.1 A cultura da ameixeira

A ameixeira pertence à família *Rosaceae*, à subfamília *Prunoidae* e ao gênero *Prunus*, que compreende mais de 20 espécies (WEINBERGUER, 1975). Dentre as espécies que compõe as principais cultivares tem-se a ameixeira-japonesa (*Prunus salicina* Lindl.), originária do extremo oriente da China, que há mais de 400 anos foi levada para o Japão, onde foi cultivada pela primeira vez com fins alimentares (GRUMBERG, 1944), daí a origem do seu nome.

A ameixeira é uma das espécies frutíferas que mais se difundiu pelo mundo e seu cultivo adaptou-se a uma grande variedade de situações climáticas. Essa adaptabilidade da ameixeira às diferentes condições climáticas se deve basicamente a muitas espécies existentes e do resultado de hibridações ocorridas ao longo do desenvolvimento da cultura (CASTRO et al., 2003).

No Brasil, a ameixeira é uma das frutíferas de cultivo mais antigo, não tendo conhecimento certo de quando foi introduzida (CASTRO et al., 2003), sendo cultivada com maior destaque principalmente nos Estados do Rio Grande do Sul, Santa Catarina, Paraná, São Paulo e Sul de Minas Gerais.

As ameixeiras produzem frutos de diversos tamanhos e formas, com película fina, adstringente e com pouca pruína, apresentando várias colorações entre amarelo e vermelho, mas nunca azulada; a polpa é firme, de cor amarela, vermelha ou roxa, fibrosa, doce e aromática (CASTRO, 2005).

A colheita demasiadamente precoce resulta em perda de peso e tamanho, além de predispor as frutas à desidratação, amadurecimento anormal e distúrbios fisiológicos durante o armazenamento refrigerado (MUÑOZ, 1982; RESZCZYNSKY, 1977).

Por outro lado, a colheita demasiadamente tarde encurta o período de armazenamento da fruta por aproximá-la da senescência, além de favorecer a desidratação, o amaciamento demasiado, os distúrbios fisiológicos e o aparecimento de sabores estranhos ou sobrematuração (TORRELLARDONA, 1983).

Assim, a determinação do ponto ótimo da colheita do fruto da ameixeira é de extrema importância. Isso permite assegurar uma boa conservação, adequada resistência ao transporte e a manutenção das condições necessárias para que a fruta chegue até o consumidor com qualidade.

2.2 A modelagem do crescimento de frutos

O estudo das curvas de crescimento é usado largamente em biologia e agricultura, entre outras áreas. Assim, muitas funções e formas de análise foram propostas para modelar o crescimento de algumas plantas e de animais.

O estudo do crescimento de frutos e de vegetais tem grande aplicação na pesquisa agrícola, pois pode permitir que se compare o comportamento de uma planta e suas partes. Jauer et al. (2003) estudaram as variações no comportamento de algumas características fisiológicas no feijoeiro comum por meio da análise de crescimento.

Álvarez e Castro (1999) utilizaram o modelo linear polinomial de quarto grau para ajustar o diâmetro da cana; para a descrição do desenvolvimento de frutos de cirigueira, Martins et al. (2003) verificaram que o diâmetro dos frutos (*Spondias purpurea* L.) foram descritos por um modelo de regressão polinomial quadrático durante a fase de crescimento e maturação. Entretanto, cresce cada vez mais o interesse por modelos não lineares como os modelos sigmoidais de Gompertz, Logístico e Richards, descritos por Graybill e Iyer (2004) e Ratkowsky (1983), entre outros.

Devido à maior facilidade de cálculos, os modelos de regressão linear, como o polinomial, ainda são bem utilizados na modelagem do crescimento de plantas em geral. Seus princípios básicos estão descritos em vários textos de estatística, como em Drapper e Smith (1998) e Graybill e Iyer (2004).

Com o avanço da capacidade de operação dos computadores e o aprimoramento dos programas computacionais estatísticos nos últimos anos, o uso de regressão não linear na descrição de curvas resposta tem aumentado significativamente. Castro Neto e Reinhardt (2003) verificaram que a curva de crescimento de frutos de manga cv. Haden segue o padrão sigmoidal; Bruna (2007) verificou que os frutos das variedades de ciclo curto e médio de pêssego crescem exponencialmente;

2.3 Modelos de regressão linear

A análise de regressão é um método comumente usado para a obtenção de uma função de predição para prever os valores de uma variável resposta Y usando uma ou mais variáveis preditoras X_1, \dots, X_k .

Entretanto, a simples obtenção da função estimada não responde ao pesquisador se a variação da(s) variável(eis) preditor(a)s influencia significativamente na variação da variável resposta.

Assim, é necessária a utilização de testes estatísticos para validar as estimativas dos coeficientes da equação de regressão estimada. Um destes testes é o teste F da análise de variância proposto por Fisher e Snedecor (BATES; WATTS, 1988).

Um modelo estatístico consiste de no mínimo os componentes: resposta, erro, parâmetros e a parte sistemática. Segundo Schabenberger e Pierce (2002) a maioria dos modelos pode ser representada pela decomposição:

$$\textit{Resposta} = \textit{estrutura} + \textit{erro}$$

em que “Resposta” é a medida de interesse a ser mensurada, geralmente representada por “Y”; “estrutura” é a parte sistemática que contém os parâmetros do modelo, chamada de função média; o “erro” é a diferença entre a função média e as observações.

O modelo de regressão linear é uma das classes de modelos utilizada pelos pesquisadores na compreensão dos problemas de suas pesquisas desempenhando um papel importante em diferentes áreas do conhecimento.

Um modelo de regressão é dito linear se as derivadas parciais de primeira ordem da variável dependente, em relação a cada parâmetro, não forem funções dos próprios parâmetros. Assim, um modelo linear pode ser escrito da forma:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im} + e_i \quad (1)$$

em que Y_i é a i -ésima observação da variável resposta; X_{ik} é a i -ésima observação da k -ésima variável; β_k são os parâmetros do modelo; e_i é o erro experimental associado à i -ésima observação Y_i ; $k=0,1,2,\dots,m$ e $i=0,1,2,\dots,n$; X_{i0} é constante com todos os valores iguais a 1; m representa o número de variáveis e n o tamanho da amostra.

Utilizando a notação matricial, tem-se o modelo expresso da seguinte forma:

$$\underline{Y} = X\underline{\beta} + \underline{e} \quad (2)$$

em que \underline{Y} é o vetor de observações, de dimensões $n \times 1$; X é a matriz do modelo, de dimensões $n \times (m+1)$ das derivadas parciais de Y_i em relação aos parâmetros; $\underline{\beta}$ é o vetor de parâmetros, de dimensões $[(m+1) \times 1]$; e \underline{e} é o vetor de erros experimentais, de dimensões $(n \times 1)$, considerados

independentes e normalmente distribuídos com média zero e variância constante,

$$\underline{e} \sim N(\underline{\phi}; I\sigma^2)$$

em que $\underline{\phi}$ é o vetor nulo, de dimensões $n \times 1$, I é a matriz identidade, de dimensões $n \times n$, e N denota a distribuição normal.

2.3.1 O Ajuste pelo método de mínimos quadrados ordinários

O método dos mínimos quadrados ordinários (MQO) sob certas premissas tem propriedades estatísticas interessantes, o que o torna um dos métodos mais poderosos e difundidos na literatura (GUJARATI, 2006).

Esse método tem por objetivo minimizar a soma dos quadrados dos desvios entre os valores observados e os valores preditos com a finalidade de estimar os parâmetros.

Assim, com base na equação (2) temos:

$$Q = \underline{e}'\underline{e} = (\underline{Y} - X\underline{\beta})'(\underline{Y} - X\underline{\beta}) \quad (3)$$

$$Q = \underline{e}'\underline{e} = (\underline{Y}'\underline{Y} - 2\underline{\beta}'X'\underline{Y} + \underline{\beta}'X'X\underline{\beta})$$

Para estudar o comportamento da função Q , obtêm-se a derivada de Q com relação à $\underline{\beta}$, assim:

$$\frac{\partial Q}{\partial \underline{\beta}} = -2X'\underline{Y} + 2X'X\underline{\beta} \quad (4)$$

Igualando a zero, obtêm-se o sistema de equações normais (SEN),

$$-2X'\underline{Y} + 2X'X\hat{\underline{\beta}} = 0$$

$$X'X\hat{\underline{\beta}} = X'\underline{Y} \quad (5)$$

Como a matriz X tem posto coluna completo e igual a $r(X) = m + 1$, a matriz $X'X$ é positiva definida, de posto completo; logo $X'X$ é uma matriz simétrica de posto completo, logo, existe a matriz inversa de $X'X$. Assim, a solução única do sistema (5) é

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

O vetor $\hat{\beta}$, assim determinado, minimiza a soma de quadrados dos erros em (3), pois a segunda derivada de (4) em relação a β é

$$\frac{\partial^2 \hat{e}'\hat{e}}{\partial \hat{\beta}^2} = 2X'X$$

que é uma função positiva, implicando que o modelo de (3) tem mínimo. Assim, $\hat{\beta}$ é chamado de estimador de mínimos quadrados (RENCHE; SCHAALJE, 2008). Se a matriz X for singular, então uma alternativa é a utilização da inversa generalizada (SEARLE, 1971).

2.4 Modelos de regressão não linear

Segundo Bates e Watts (1988), Draper e Smith (1998) e Ratkowsky (1983), entre outros, a definição para que um modelo seja designado não linear é de que pelo menos uma das derivadas parciais da variável dependente com relação a qualquer parâmetro presente no modelo, dependa de algum parâmetro.

Um crescente número de pesquisadores na área de ciências biológicas acredita que as relações entre as variáveis biológicas são melhores descritas por funções não lineares. Assim, é interessante sempre levar em consideração esta possibilidade de análise.

O uso da análise de regressão não linear requer também algumas

pressuposições e alguns cuidados, sendo popularizada com o avanço da informática fornecendo muitas vezes resultados mais aplicáveis do que a própria regressão linear. A regressão não linear é preferível porque os parâmetros da curva ajustada têm um significado biológico (CALEGARIO et al., 2005).

Alguns modelos podem parecer não lineares nos parâmetros, mas são inerente ou intrinsecamente lineares porque, com devidas transformações, podem se tornar modelos de regressão lineares nos parâmetros (GUJARATI, 2006). Entretanto, existe um inconveniente de que além do parâmetro perder sua interpretação intrínseca, pode-se alterar a estrutura e distribuição do erro e assim, as suposições de normalidade, independência e homogeneidade da variância, podem não ser satisfeitas (BATES; WATTS, 1988). Além disso, as devidas transformações podem levar a resultados insatisfatórios para situações práticas.

Entre várias aplicações, Mazzini et al. (2003) utilizaram as funções não lineares de Brody, Gompertz, Logística, Richards e Von Bertalanffy para ajustar o crescimento de bovinos Hereford; Gurjão et al. (2006) afirmam que o crescimento dos frutos de tamarindo pode ser descrito por um modelo sigmoidal simples.

2.4.1 Estimação dos parâmetros de modelos não lineares

Um modelo não linear F é definido de forma geral para o vetor de parâmetros $\beta = [\beta_1 \ \beta_2 \ \dots \ \beta_m]'$ e para o vetor de variáveis regressoras da j -ésima unidade amostral $Z_j' = [Z_{1j} \ Z_{2j} \ \dots \ Z_{pj}]$ por

$$Y_j = F_j(\beta, Z_j) + e_j.$$

Em notação matricial tem-se:

$$\underline{Y} = \underline{F}(\underline{\beta}) + \underline{\varepsilon}.$$

em que pode-se expressar o vetor do modelo $\underline{F}(\underline{\beta})$, simplesmente por \underline{F} .

Gallant (1987) afirma que o método dos quadrados mínimos é utilizado na estimação dos parâmetros em modelos não lineares, da mesma maneira que em modelos lineares. Assim, para

$$\underline{Y} = \underline{F}(\underline{\beta}) + \underline{\varepsilon}$$

tem-se que a soma de quadrados dos erros é,

$$S(\underline{\beta}) = \underline{\varepsilon}'\underline{\varepsilon} = (\underline{Y} - \underline{F})'(\underline{Y} - \underline{F}) = \underline{Y}'\underline{Y} - 2\underline{Y}'\underline{F} + \underline{F}'\underline{F}$$

$$\frac{\partial S}{\partial \underline{\beta}} = \frac{-\partial 2\underline{Y}'\underline{F}}{\partial \underline{\beta}} + \frac{\partial \underline{F}'\underline{F}}{\partial \underline{\beta}} \quad (6)$$

Desenvolvendo o primeiro termo após a igualdade, tem-se:

$$\frac{-\partial 2\underline{Y}'\underline{F}}{\partial \underline{\beta}} = \frac{-\partial 2\underline{Y}'\underline{F}}{\partial \underline{F}} \times \frac{\partial \underline{F}}{\partial \underline{\beta}} = -2\underline{Y}'X$$

e que $X = \partial \underline{F} / \partial \underline{\beta}$ é a matriz de derivadas parciais.

Tem-se também que do segundo termo da igualdade (6),

$$\frac{\partial \underline{F}'\underline{F}}{\partial \underline{\beta}} = \frac{\partial \underline{F}'\underline{F}}{\partial \underline{F}} \times \frac{\partial \underline{F}}{\partial \underline{\beta}} = 2\underline{F}'X$$

Logo,

$$\frac{\partial S}{\partial \underline{\beta}} = -2\underline{Y}'X + 2\underline{F}'X$$

Igualando a zero a primeira derivada, tem-se que

$$X'\underline{F} = X'\underline{Y} \quad (7)$$

é a equação normal para os modelos não lineares.

Como \underline{F} e X são funções de $\underline{\beta}$ então não existe uma forma fechada para a solução. Portanto, deve-se utilizar um processo iterativo para resolução

desse sistema de equações. Em geral este processo de estimação inicia-se com base em um valor inicial para os parâmetros, o qual deve ser melhorado continuamente até que se tenha atingido um critério de convergência; este critério tem como referência que a soma de quadrados dos resíduos $\underline{e}'\underline{e}$ seja minimizada. Assim, o valor dos parâmetros na $(k+1)$ -ésima iteração é

$$\underline{\beta}_{k+1} = \underline{\beta}_k + \lambda \underline{\Delta} \quad (8)$$

em que o escalar λ é escolhido na k -ésima iteração de tal forma que

$$SQE(\underline{\beta}_{k+1}) < SQE(\underline{\beta}_k) \quad (9)$$

em que SQE é a soma de quadrados dos resíduos.

2.4.2 Processos iterativos

Existem vários métodos iterativos para a estimação de parâmetros de modelos não lineares. Dentre estes métodos iterativos, os mais utilizados são o de Gauss-Newton (HARTLEY, 1964) ou método de linearização, o método de Newton, o método gradiente (Steepest-Descent) e o método de Marquardt (BATES; WATTS, 1988). Esses métodos se diferem na forma como o $\underline{\Delta}$ é calculado para propiciar as atualizações dos valores no vetor de parâmetros.

De uma forma geral, os critérios básicos são:

$$\textit{Gradiente:} \quad \underline{\Delta} = X' \underline{e}$$

$$\textit{Gauss - Newton:} \quad \underline{\Delta} = (X'X)^{-} X' \underline{e}$$

$$\textit{Newton:} \quad \underline{\Delta} = G^{-} X' \underline{e}$$

$$\textit{Marquardt:} \quad \underline{\Delta} = \left[X'X + \delta \textit{diag} (X'X)^{-} X' \underline{e} \right]$$

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

O método Gradiente, como o próprio nome diz, é baseado no gradiente ou grau de variação dos resíduos $\underline{e}'\underline{e}$. Os métodos Gauss-Newton e Marquardt

realizam a regressão dos resíduos em relação às primeiras derivadas do modelo não linear em relação aos parâmetros, até que haja a convergência. O método de Newton faz a regressão destes resíduos em relação a uma função das segundas derivadas do modelo não linear com relação aos parâmetros (G^-).

Na utilização destes métodos, o sucesso na convergência de um algoritmo está diretamente associado ao uso de uma função resposta apropriado e de valores iniciais adequados ao procedimento numérico, não sendo garantida por qualquer método (GALLANT, 1987).

Então, uma maneira de verificar a precisão da solução numérica é provocando distúrbios nas condições iniciais e verificando se os novos valores iniciais conduzem à mesma solução.

Dentre os vários procedimentos para estimar os valores iniciais, destacam-se: regressão linear via método dos mínimos quadrados ordinários; transformação da variável resposta e ou da variável tempo, tentando transformar a função original em função mais simples; redução do vetor de parâmetros por imposição de valores para alguns deles e estimação dos demais; análise da função esperada e de suas derivadas com respeito aos parâmetros.

No programa computacional R, os quatro métodos citados estão implementados, sendo que os dois mais utilizados são o de *Gauss Newton* e o *Marquardt*.

2.4.2.1 Método *Gauss Newton*

O método de *Gauss-Newton* usa a expansão em série de Taylor do vetor de funções

$$\tilde{F}(\tilde{\beta}) = \tilde{F}(\tilde{\beta}_0) + X(\tilde{\beta} - \tilde{\beta}_0) + \dots$$

em que a matriz de primeiras derivadas X é avaliada no ponto $\tilde{\beta}_0$.

Se substituir os dois termos desta expansão nas equações normais obtém-se

$$\begin{aligned} X'F(\underline{\beta}) &= X'Y \\ X' \left[F(\underline{\beta}_0) + X(\underline{\beta} - \underline{\beta}_0) \right] &= X'Y \\ X'X(\underline{\beta} - \underline{\beta}_0) &= X'Y - X'F(\underline{\beta}_0) \\ \underline{\Delta} &= X'e \\ X'X\underline{\Delta} &= X'e \end{aligned}$$

e, portanto,

$$\underline{\Delta} = (X'X)^{-} X'e$$

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

Estimado o valor de $\underline{\Delta}$ para o vetor $\underline{\beta}_0$, aplicam-se a equação (8) e a inequação (9) para se obter o vetor de estimativas do passo 1. O processo é repetido um determinado número de vezes até que o vetor de estimativas não se altere mais dentro de uma precisão pré-estipulada.

2.4.2.2 Método de *Marquardt*

O método de *Marquardt* mantém um compromisso entre o método de *Gauss-Newton* e o método do gradiente. A fórmula de atualização do vetor de parâmetros é dada por:

$$\underline{\Delta} = \left[(X'X) + \delta \text{diag}(X'X) \right]^{-} X'e$$

Se $\delta \rightarrow 0$, há uma aproximação ao método de Gauss-Newton e se $\delta \rightarrow \infty$, há uma aproximação ao método do gradiente. Assim, se a soma de quadrados do resíduo decresce a cada iteração, estar-se-á utilizando essencialmente o método de Gauss-Newton; se ocorrer o contrário o valor de δ é aumentado em cada

iteração, este método aproxima-se ao método gradiente.

2.5 Análise dos resíduos

A análise de resíduos desempenha papel fundamental na avaliação do ajuste de um modelo de regressão linear, oferecendo recursos para investigar a adequação do modelo no que se refere às suposições básicas (CHARNET et al., 2008). Assim, o preditor de resíduos é dado por:

$$\underline{e} = \underline{Y} - X \hat{\underline{\beta}} \quad (10)$$

Segundo Schabenberger e Pierce (2002), a verificação de propriedades dos resíduos brutos da equação (10) não é a melhor forma de análise, dado que, os mesmos podem ser correlacionados e heteroscedásticos. Por isso deve dar preferência em trabalhar com os resíduos padronizados, pois minimizam o problema da heterogeneidade.

Desde que os resíduos ou erros do modelo são variáveis aleatórias não observáveis parece natural, utilizar os resíduos obtidos com relação aos valores ajustados. Assim, os resíduos $\underline{e} = \underline{Y} - X \hat{\underline{\beta}}$, em que $\hat{\underline{\beta}}$ é um estimador não viesado de $\underline{\beta}$, tem a propriedade de que a média seja nula, logo $E[\underline{e}] = 0$; mas os resíduos estimados podem ser correlacionados e heterocedásticos. Os resíduos podem ser definidos pela expressão $\underline{e} = \underline{Y} - \hat{\underline{Y}} = Y - HY$ em que $H = X(X'X)^{-1}X'$ é a matriz chamada “hat” dado que ela produz os valores de $\hat{\underline{Y}}$ quando multiplicada por \underline{Y} e, representa o projetor ortogonal das observações sobre o espaço paramétrico da matriz de delineamento X . A matriz H possui algumas propriedades importantes, entre elas, a matriz H é simétrica ($H' = H$) e idempotente ($HH = H$), assim $(I - H)(I - H) = (I - H)$.

A variância dos erros é dada por:

$$\begin{aligned} \text{Var}[\underline{e}] &= \text{Var}[\underline{Y} - H\underline{Y}] \\ &= \sigma^2 I(I - H)(I - H) = \sigma^2(I - H) \end{aligned}$$

em que H não é uma matriz diagonal e os elementos na sua diagonal nem sempre são iguais. O i -ésimo valor da diagonal de H é chamado de “leverage” do i -ésimo dado. Considerando que h_{ii} é o elemento da diagonal da matriz H , a variância de um erro é

$$\text{Var}[e_i] = \sigma^2(1 - h_{ii})$$

E, os resíduos padronizados com média zero e variância 1 são obtidos por:

$$e_i^* = \frac{e_i}{\left(\sigma\sqrt{(1-h_{ii})}\right)}$$

Como σ^2 é desconhecida, σ é substituído por $\hat{\sigma}$, raiz quadrada do quadrado médio do erro do modelo. O resíduo

$$r_i = \frac{e_i}{\hat{\sigma}\sqrt{(1-h_{ii})}}$$

é chamado de resíduo padronizado.

Este tipo de resíduo é mais interessante que o anterior, devido ao fato de considerar a variância individual de cada resíduo ordinário. Entretanto, se a i -ésima- observação for um outlier pode ocorrer que a estimativa da variância estará afetada por este valor.

Os resíduos studentizados eliminam a i -ésima observação e obtêm-se um estimador de variância, digamos $S_{(i)}^2$. Assim, esse estimador obtido a partir da análise original é dado por:

$$S_{(i)}^2 = \frac{(n-m-1)S^2}{n-m-2} - \frac{e_i^2}{(n-m-2)(1-h_i)}$$

em que:

m é o número de variáveis;

n é o tamanho da amostra.

Logo, tem-se que:

$$s_i = \frac{e_i}{\sqrt{(1-h_i)S_{(i)}^2}}$$

é chamado de resíduo studentizado e é mais apropriado para as medidas de diagnóstico.

Para verificar o pressuposto de normalidade, podem-se utilizar métodos gráficos como o que relaciona os quantis teóricos com os quantis esperados (qq-plot) ou testes de normalidade tradicionais como o de Shapiro Wilk e Kolmogorov-Smirnov cujas aplicações com mais detalhes podem ser vistas em Shapiro e Wilk (1965) e Stephens (1974).

O qq-plot ou gráfico quantil-quantil é um gráfico que testa a conformidade entre a distribuição empírica e uma dada distribuição teórica. A construção dos qq-plots é feita colocando-se os resíduos em ordem crescente e associando cada dado com o quantil da distribuição normal padrão. Como regra de decisão observa-se se a representação gráfica entre os quantis da variável aparecem alinhados com os quantis da distribuição. Se isso ocorre é porque os dados são normalmente distribuídos.

Geralmente, nos ajustes dos modelos, assume-se que os erros são independentes. Em curvas de crescimento de frutos, em que o mesmo fruto é medido em vários tempos diferentes ou até mesmo frutos diferentes que são medidos ao longo do tempo, mas, provenientes da mesma planta, a hipótese de independência dos erros pode não ser apropriada. Essa dependência é caracterizada pela correlação entre o erro relativo a um período e o erro da observação anterior.

Autocorrelação é um caso especial de correlação, e não se refere à relação entre duas ou mais variáveis, mas a relação entre os valores sucessivos de uma mesma variável.

A violação do pressuposto de erros independentes tem três consequências importantes para a análise de regressão. A primeira é que os testes estatísticos de significância dos parâmetros e limites de confiança para os valores previstos não são corretos (perda da eficiência dos estimadores de mínimos quadrados). A segunda, as estimativas dos coeficientes de regressão não são tão eficientes quanto eles seriam se a autocorrelação tivesse sido considerada (os estimadores dos mínimos quadrados não são os estimadores com variância mínima). Por fim, a terceira consequência, é que uma vez que os resíduos da análise de regressão não são independentes, eles contêm informações que podem ser usadas para melhorar a previsão de valores futuros (os estimadores dos mínimos quadrados serão viesados e não vão possuir mínima variância, ou seja, os estimadores não serão BLUE).

O teste de Durbin-Watson o qual indica a auto-correlação de primeira ordem é calculado por:

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

sendo:

e_t : o resíduo no tempo t;

e_{t-1} : o resíduo no tempo t-1.

Tem-se que seu valor está entre 0 e 4. Assim, um valor próximo de zero indica uma auto-correlação positiva, os valores de cerca de 2 revelam a inexistência de auto-correlação, e se você se aproxima 4, existe uma auto-correlação negativa.

Para verificar o pressuposto de homogeneidade dos resíduos, podem-se utilizar métodos gráficos como, por exemplo, a representação gráfica da tendência do comportamento dos resíduos do ajuste dos modelos.

3 MATERIAL E MÉTODOS

3.1 Realização do experimento

Os dados utilizados nesse trabalho foram provenientes de um experimento conduzido no ano de 2008, no município de Delfim Moreira, o qual está localizado no extremo sul do estado de Minas Gerais. A região em que o pomar está situado fica à altitude de aproximadamente 1380 m, e nas coordenadas 22°52'30" de latitude sul e 45°30'70" de longitude oeste.

Para as avaliações do desenvolvimento dos frutos da ameixeira, variedade "Gulfblase", foi utilizado o delineamento experimental inteiramente ao acaso, sendo avaliadas quatro plantas, que foram analisadas utilizando-se parcelas subdivididas no tempo. Cada planta foi subdividida nos diferentes quadrantes (norte, sul, leste e oeste) onde foram amostrados (10) frutos por quadrante. Assim, os fatores considerados na análise de variância foram: planta, quadrante em que o fruto se encontra (posição) e tempo (dias após frutificação).

Foram registrados o início e término do período de floração e de frutificação. Foi avaliado semanalmente o diâmetro transversal dos frutos. Segundo Oliveira (2009), o início do período de floração ocorreu em 01/07 e o início do período de frutificação em 31/07 e as anotações do diâmetro dos frutos foram feitas a partir do dia 31/08 e se prolongaram até a colheita realizada em 20/10. Assim, foram feitas oito medições, sendo a primeira, aos trinta e um dias após o início de frutificação, ou seja, os tempos de avaliação foram: $t_1 = 31$, $t_2 = 38$, $t_3 = 45$, $t_4 = 52$, $t_5 = 59$, $t_6 = 66$, $t_7 = 73$, $t_8 = 80$ dias após frutificação.

3.2 Modelos testados

Vários modelos podem ser usados para descrever o crescimento do fruto de plantas, entre eles os modelos lineares e os modelos não-lineares.

Nesse trabalho foram utilizados o modelo linear polinomial e os modelos não-lineares Logístico e de Gompertz, devido ao fato de serem os modelos mais utilizados na descrição do crescimento e desenvolvimento de frutos existentes na literatura, apresentados por Graybill e Iyer (2004). Para os não-lineares foram considerados os modelos com estrutura de erros independentes e com estrutura de erros auto-correlacionados de primeira ordem.

3.2.1 Modelo linear polinomial

Para verificar a tendência de comportamento do desenvolvimento dos frutos realizou-se uma visualização inicial dos dados observados em função do tempo, por meio de uma análise gráfica.

Para cada planta ajustou-se os modelos de regressão linear polinomial de segundo grau com estrutura de erros independentes:

$$y_k = A + Bt_k + Ct_k^2 + e_k$$

em que y_k é a resposta obtida no k-ésimo tempo; A , B e C são os parâmetros a serem estimados, t_k é a variável tempo e e_k é o erro aleatório associado à resposta obtida para o k-ésimo tempo.

3.2.2 Modelos não lineares

Para cada planta ajustou-se os modelos de regressão não linear de Gompertz com estrutura de erros independentes:

$$y_k = Ae^{-e^{-(B+Ct_k)}} + e_k$$

e, estrutura de erros auto-correlacionados (AR) de ordem um:

$$y_k = Ae^{-e^{-(B+Ct_k)}} + e_k + \phi e_{k-1}$$

em que y_k é a variável resposta obtida no k -ésimo tempo, A representa a assíntota de crescimento do diâmetro do fruto, representando o valor de estabilização do diâmetro do fruto; B é o parâmetro numérico relacionado ao valor da função quando o tempo tende a zero $t = 0$, C é o parâmetro numérico para a escala no eixo do tempo t e e_k, e_{k-1} são os erros aleatórios no tempo k e $(k-1)$ respectivamente, considerados como tendo distribuição normal e variância constante.

De modo análogo, ajustou-se os modelos de regressão não linear Logístico com estrutura de erros independentes:

$$y_k = \frac{A}{1 + e^{-(B+Ct_k)}} + e_k$$

e, estrutura de erros auto-correlacionados (AR) de ordem um:

$$y_k = \frac{A}{1 + e^{-(B+Ct_k)}} + e_k + \phi e_{k-1}$$

em que y_k é a variável resposta obtida no k -ésimo tempo, A representa a assíntota de crescimento do diâmetro do fruto, tendendo para o diâmetro máximo quando o tempo tende para o infinito, o parâmetro B está relacionado com o valor da função quando o tempo tende a zero $t = 0$, C é o parâmetro que está relacionado com a taxa média de crescimento da curva e e_k, e_{k-1} são os termos erros aleatórios no tempo k e $(k-1)$ respectivamente.

3.3 Seleção dos modelos

A seleção do modelo mais adequado para explicar o crescimento dos frutos foi realizada com base na significância do modelo de regressão e na precisão dos ajustes obedecendo aos seguintes critérios: coeficiente de determinação ajustado, desvio padrão residual (DPR) e critério de informação de Akaike (AIC) (MOTULSKY; CHRISTOPOULOS, 2003).

3.3.1 Coeficiente de determinação ajustado

O coeficiente de determinação R_a^2 ajustado para o número de parâmetros do modelo de regressão foi calculado por:

$$R_a^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - p}$$

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.

Tem-se que quanto maior for o coeficiente de determinação ajustado (R_a^2) melhor será o modelo ajustado.

3.3.2 Desvio padrão residual

A estimativa do desvio padrão residual foi obtida pela expressão:

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

em que:

QME = quadrado médio do erro;

n = número de observações;

p = número de parâmetros do modelo.

Tem-se que quanto menor for o desvio padrão residual (DPR) melhor será o modelo ajustado.

3.3.3 Critério de informação de Akaike (AIC)

O critério de informação de Akaike foi calculado pela seguinte expressão:

$$AIC(p) = n \ln(SQR/n) + 2p$$

sendo:

p: o número de parâmetros;

n: o número total de observações;

SQR: a soma de quadrados dos resíduos.

O AIC é usado para comparar modelos não-aninhados ou quando estão sendo comparados três ou mais modelos. Quanto menor o valor AIC melhor é o modelo ajustado.

3.4 Recursos computacionais

Os procedimentos de ajuste e dos diagnósticos para os modelos com estrutura de erros independentes foram implementados por meio do programa computacional R (R DEVELOPMENT CORE TEAM, 2010), conforme Anexos A, B e C, os quais possibilitaram obter as estimativas dos parâmetros, os intervalos de confiança para os parâmetros, além da análise dos resíduos, sendo verificados os pressupostos de normalidade utilizando-se o teste de Shapiro e Wilk (1965), a homogeneidade de variância dos resíduos padronizados e plotagem dos resultados, conforme sugestão de Drapper e Smith (1998) e Schabenberger e Pierce (2002).

As estimativas iniciais, necessárias para obtenção das estimativas dos

parâmetros por meio do processo iterativo, foram obtidas, com devidas parametrizações, pelas funções *SSlogis* e *SSGompertz* do programa computacional R para o modelo Logístico e de Gompertz, respectivamente.

Os procedimentos de ajuste dos modelos com estrutura de erros autorregressivos foram realizados por meio do programa computacional para análise estatística Statistical Analysis System Institute - SAS Institute (1999), apresentado no Anexo D, pelo procedimento (proc model) sendo que a macro utilizada é dada por:

% AR(y,p), em que p indica a ordem do modelo.

4 RESULTADOS E DISCUSSÃO

4.1 Análise de variância

A análise da variância para o diâmetro do fruto (Tabela 1) indica que houve efeito significativo ($p = 0,05$) de planta, do tempo e da interação tempo x planta. A interação tripla tempo x planta x posição não foi significativa, no nível de 5% de probabilidade, indicando que nessa significância, esses três fatores em conjunto são independentes uns dos outros. Assim, pode-se inferir que o diâmetro do fruto da ameixeira “Gulfblase” para essa amostra possui o mesmo comportamento nas quatro posições de coleta.

Tabela 1 Análise de Variância para o diâmetro dos frutos da ameixeira “Gulfblase”

FV	GL	SQ	QM	F	p-valor
Posição	3	1466,4	488,81	2,5987	0,1167
Planta	3	2566,9	855,65	4,5489	0,0333
Erro (1)	9	1692,9	188,1	-	-
(Parcela)	(15)	(5726,2)	-	-	-
Tempo	7	125187,0	17883,9	541,7646	0,0000
Tempo*Planta	21	1372,0	65,4	1,9798	0,0053
Tempo*Posição	21	799,0	38,0	1,1519	0,2865
Tempo*Planta*Posição	63	2633,0	41,8	1,2660	0,0824
Erro	1152	38028	33,0	-	-
TOTAL	1279	173745,3	-	-	-

Por outro lado, observa-se que as plantas têm um comportamento diferenciado em relação ao tempo de avaliação. A interação tempo x planta significativa, indica que há pelo menos para uma planta em que o comportamento do diâmetro dos frutos da ameixeira “Gulfblase”, difere das outras, ao longo do tempo.

4.2 Ajuste do modelo linear polinomial

Na Tabela 2 são apresentadas as estimativas dos coeficientes de regressão ajustados pelo método dos mínimos quadrados ordinários para cada planta ao longo do tempo. Tem-se que os coeficientes (*A*) relativos ao intercepto não são importantes para o presente estudo, pois eles estimam o valor do diâmetro do fruto quando o tempo é igual a zero, já que o intervalo de tempo em que foram feitas as avaliações é de 31 a 80 dias após a frutificação. Veja que se essa inferência fosse viável, seria estimado um diâmetro de fruto negativo como se observa para as plantas “14”, “29” e “36”, o que não tem sentido prático. Este fato ressalta o perigo de se fazer extrapolação, fazendo-se inferência fora do domínio dos dados em estudo. Para as estimativas dos coeficientes (*B*), significativas a 1% para todas as plantas, observa-se que a tendência de crescimento foi menor para a planta “8” e, maior para a planta “36”. Já as estimativas dos coeficientes (*C*), apesar de serem baixas, é possível observar que apenas a planta “8” não obteve o coeficiente significativo.

Tabela 2 Estimativas dos parâmetros para o ajuste do modelo polinomial, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Plantas	Coeficientes Estimados		
	A	B	C
8	2,2948	0,6236 *	-0,0005 ^{ns}
14	-13,9594	1,0891 *	-0,0041 *
29	-11,0184	1,0973 *	-0,0048 *
36	-24,3070	1,4201 *	-0,0070 *

* Significativo no nível de 1% e ^{ns} Não significativos pelo teste t.

Percebe-se que os coeficientes estimados das diferentes plantas possuem estimativas parecidas, com exceção do coeficiente (C) para a planta “8”, que não foi significativo e está bem abaixo em relação às outras plantas, sugerindo que o crescimento dos frutos nessa planta seja linear no intervalo de tempo considerado.

O Gráfico 1 mostra os ajustes do modelo polinomial para cada planta. Observa-se que todas as plantas tiveram um bom ajuste, levando-se em conta o reduzido número de instantes no tempo (dias após a frutificação), dado pelo coeficiente de determinação acima de 0,71, sugerindo que os modelos explicam acima de 71% das variações ocorridas no diâmetro do fruto da ameixeira devido às variações no tempo. Nota-se claramente que a tendência do crescimento do diâmetro dos frutos foi sempre crescente.

Martins et al. (2003) verificaram que o diâmetro dos frutos de ciriguela (Spondias purpurea L.) se ajustou bem por um modelo de regressão quadrático durante a fase de crescimento e maturação. Godoy e Cardoso (2003) ajustaram o modelo polinomial ao diâmetro longitudinal do fruto do melão, cultivar “Louis”, obtendo um coeficiente de determinação igual a 0,77.

Os frutos de plantas como pessegueiro e ameixeira apresentam três estádios de desenvolvimento do fruto. Segundo Dejong e Goudriaan (1989), no Estádio I, ocorre a multiplicação celular, e o fruto passa por um período de rápido crescimento do pericarpo e da semente. O Estádio II é um período de baixo crescimento e geralmente é dominado pelo endurecimento e lignificação do endocarpo (caroço). O Estádio III é período de rápida expansão das células e maturação do fruto. Em geral, esses frutos seguem o padrão de crescimento, representado graficamente por uma curva sigmoideal dupla. Devido a esse comportamento, verificam-se nos frutos duas fases de crescimento exponencial (estádios I e III) e uma de crescimento reduzido (estádio II) (BARBOSA et al., 1993). Este tipo de comportamento parece ser claro para as plantas “14”, “29” e “36”, mas, por ser um cultivar com menor ciclo e com um intervalo de apenas sete dias entre as avaliações, não foi possível ajustar o modelo sigmoideal duplo que contém o dobro de parâmetros dos modelos sigmoideais simples. Bruna (2007) verificou que as variedades de ciclo médio e de ciclo curto de pessegueiros não apresentaram o estágio II, com pouco crescimento, passando do estágio I diretamente para o estágio III.

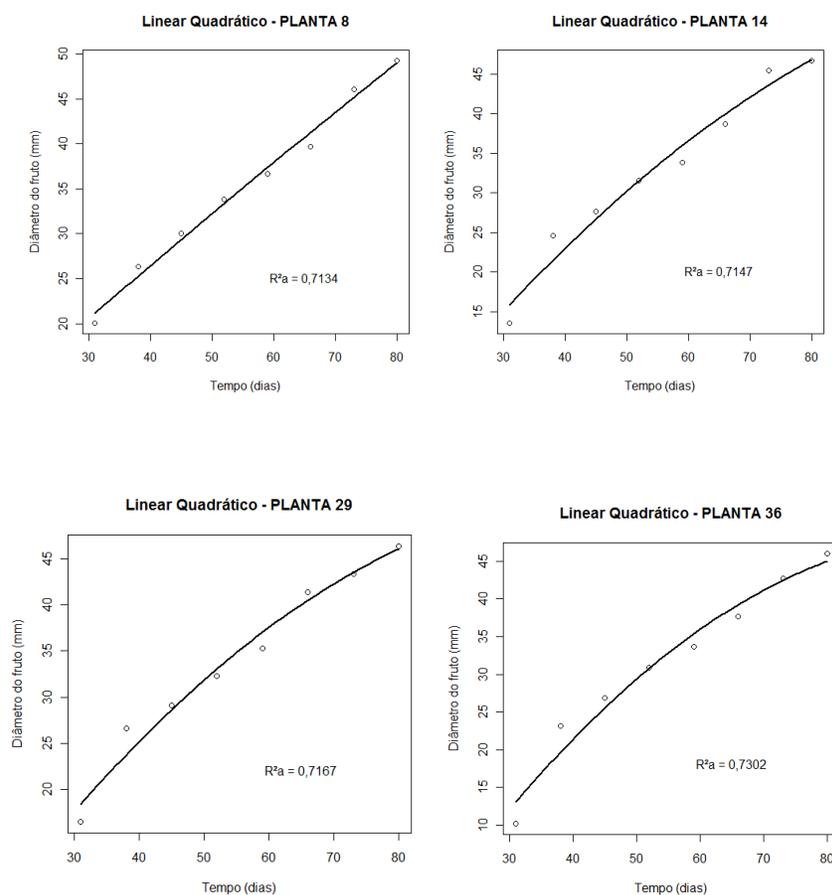


Gráfico 1 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo linear polinomial com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R^2_a) para as quatro plantas

O Gráfico 2 mostra os quantis observados, em função dos quantis esperados para os resíduos relativos ao ajuste do modelo polinomial referente a cada planta. Nota-se que alguns valores estão distantes da reta para as plantas “8” e “29”,

sugerindo a existência de problema de normalidade dos erros. Isto é confirmado pelo teste de Shapiro-Wilk que o pressuposto de normalidade da distribuição dos resíduos padronizados não foi satisfeito para essas plantas.

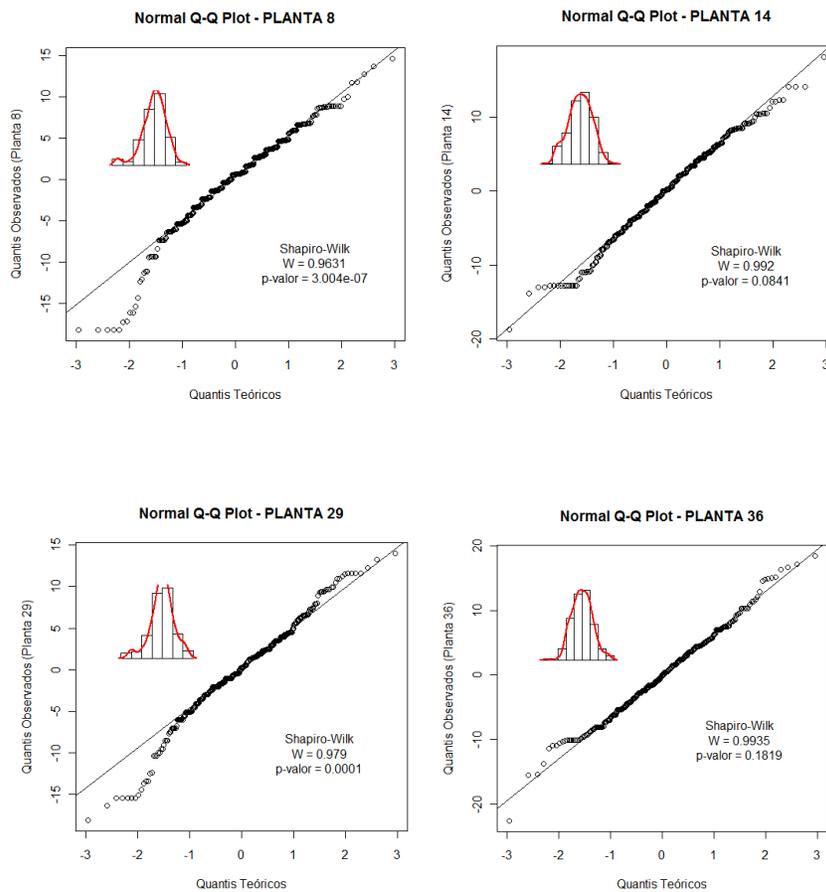


Gráfico 2 Representação gráfica da relação entre os quantis observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo linear polinomial para as quatro plantas; histograma dos resíduos; estatística do teste de normalidade de Shapiro-Wilk

No Gráfico 3, tem-se o comportamento dos resíduos do ajuste do modelo polinomial para cada planta ao longo do tempo. Observa-se que há uma variabilidade muito grande dos resíduos dentro de cada tempo de avaliação. Isto pode implicar na ocorrência de frutos de diferentes tamanhos, o que pode prejudicar a obtenção de um produto uniforme. Agora, em média, tem-se que a variabilidade dos resíduos é baixa ao longo do tempo (ilustrada pela linha contínua). No entanto, não estão de maneira aleatória em torno do zero; tem-se que para as plantas “8” e “29” parecem descrever uma curva descendente, e para as plantas “14” e “36” apresentam um comportamento cíclico. Talvez esse problema possa ser causado por uma provável auto-correlação existente entre as medidas de avaliação feitas nos diferentes tempos. Isto se comprova para todas as plantas, pelo teste de independência dos resíduos de Durbin Watson, conforme a Tabela 3.

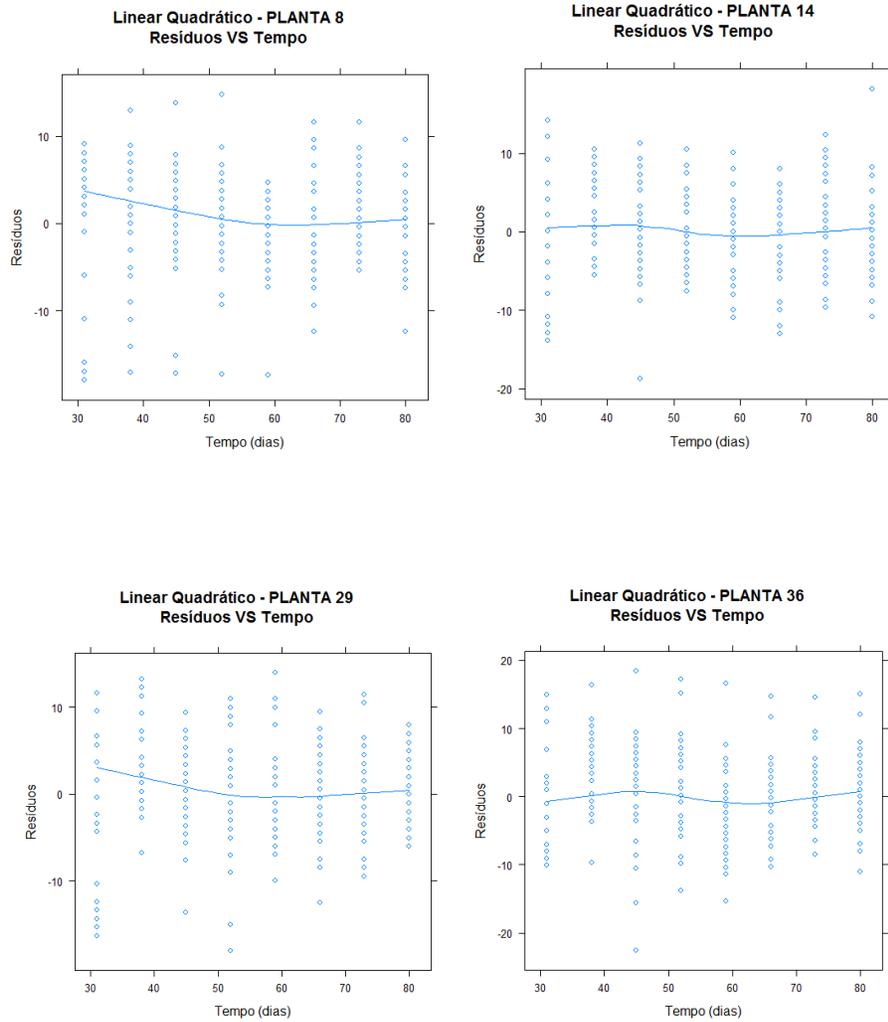


Gráfico 3 Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo polinomial ao longo do tempo para as quatro plantas de ameixeira avaliadas

Tabela 3 Estatísticas do teste de Durbin Watson dos resíduos do modelo linear polinomial quadrático para as quatro plantas

Planta	Estatística do Teste Durbin Watson (Dw)
8	1,3091
14	1,4050
29	1,4987
36	1,3080

Verifica-se que as estatísticas não estão próximas de 2, ou seja, há evidências de auto-correlação entre os resíduos, ou seja, os resíduos são dependentes.

4.3 Ajuste dos modelos não lineares

As estimativas dos parâmetros do modelo Gompertz (Tabela 4), com erros independentes, foram significativas em 1% pelo teste *t*.

Tabela 4 Estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Plantas	Coeficientes de regressão		
	<i>A</i>	<i>B</i>	<i>C</i>
8	90,1601	-0,9071	0,0175
14	60,1548	-1,3040	0,0334
29	55,6458	-1,1936	0,0355
36	50,8868	-1,7411	0,0470

Todos os parâmetros foram significativos em 1%.

Percebe-se que as plantas apresentam comportamento semelhante, com exceção da planta “8”, pois todas as estimativas estão próximas, dando indícios de não haver diferença significativa entre elas. Estes valores estão bem maiores do que os obtidos por Chagas (2008), que encontrou nas condições de Jundiaí, SP, frutos da ameixeira com diâmetro máximo de 48,3 milímetros. Isto se deve provavelmente aos fatos de as condições climáticas das regiões serem diferentes, como também, a cultivar. Assim, para a planta “8”, estima-se um diâmetro máximo de 90,1 mm bem acima dos valores observados, sugerindo que esse valor do parâmetro (A) pode estar superestimado. Isto se comprova quando se apresentam graficamente os intervalos de confiança dos parâmetros estimados, conforme o Gráfico 4, e também pela representação gráfica (Gráfico 5), na qual se tem um comportamento quase linear crescente para o diâmetro em função do tempo.

Vale ressaltar que o parâmetro (B) é um parâmetro de escala, que está relacionado com o valor do diâmetro do fruto quando o tempo tende a zero, sendo assim, não tem interpretação prática direta em nosso estudo. Pelos intervalos de confiança, a planta “36” difere da planta “8”, o mesmo ocorrendo com relação ao parâmetro (C). Observando o Gráfico 5, nota-se que o comportamento do diâmetro da planta “36” apresenta certa tendência de estabilizar o seu desenvolvimento, pela suave curvatura do modelo, enquanto que para a planta “8”, o desenvolvimento é linear, não apresentando tendência de estabilização.

Do mais, tem-se que as outras plantas possuem comportamento semelhante, dado que os intervalos de confiança para os parâmetros estão sobrepostos.

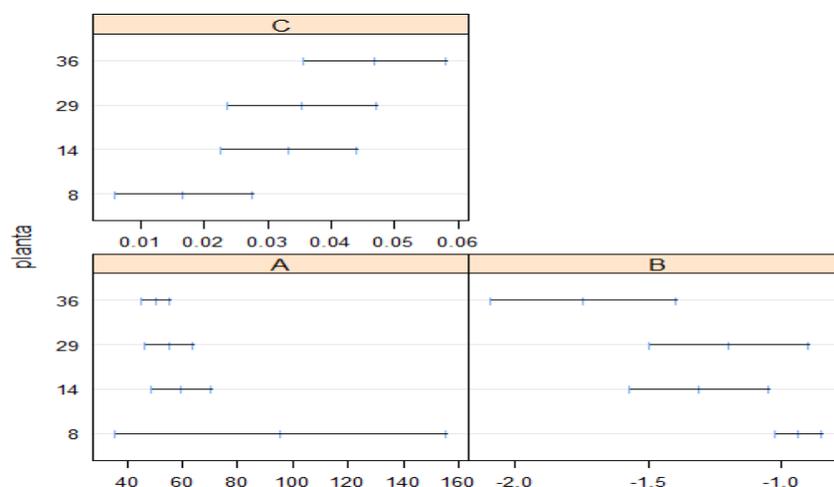


Gráfico 4 Representação gráfica dos intervalos de confiança das estimativas dos parâmetros para o ajuste do modelo Gompertz, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Pelo Gráfico 5 tem-se que as curvas tiveram um ajuste relativamente bom, verificadas pelo coeficiente de determinação ajustado (R_a^2) acima de 0,71 para todas as plantas; sugerindo que os modelos explicam acima de 71% das variações ocorridas no diâmetro do fruto da ameixeira devido às variações no tempo.

Moura et al. (2004) encontraram um comportamento sigmoidal no crescimento do diâmetro transversal e longitudinal de frutos de tomate, cv. Santa Clara.

Após o ajuste do modelo de Gompertz, verifica-se pelo teste de Shapiro-Wilk que os resíduos das plantas “8” e “29” não apresentaram uma distribuição normal (Gráfico 6). Isto implica que o ajuste por este modelo não foi adequado ou que os dados do diâmetro do fruto tem algum problema de variabilidade.

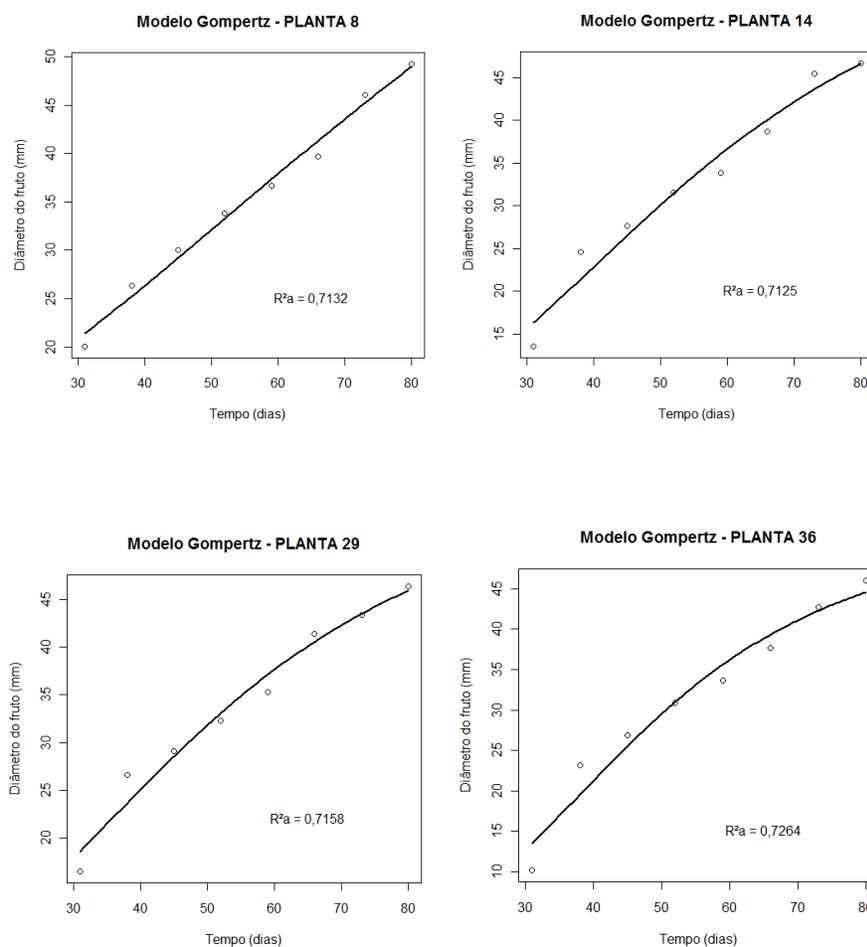


Gráfico 5 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R^2_a) para as quatro plantas

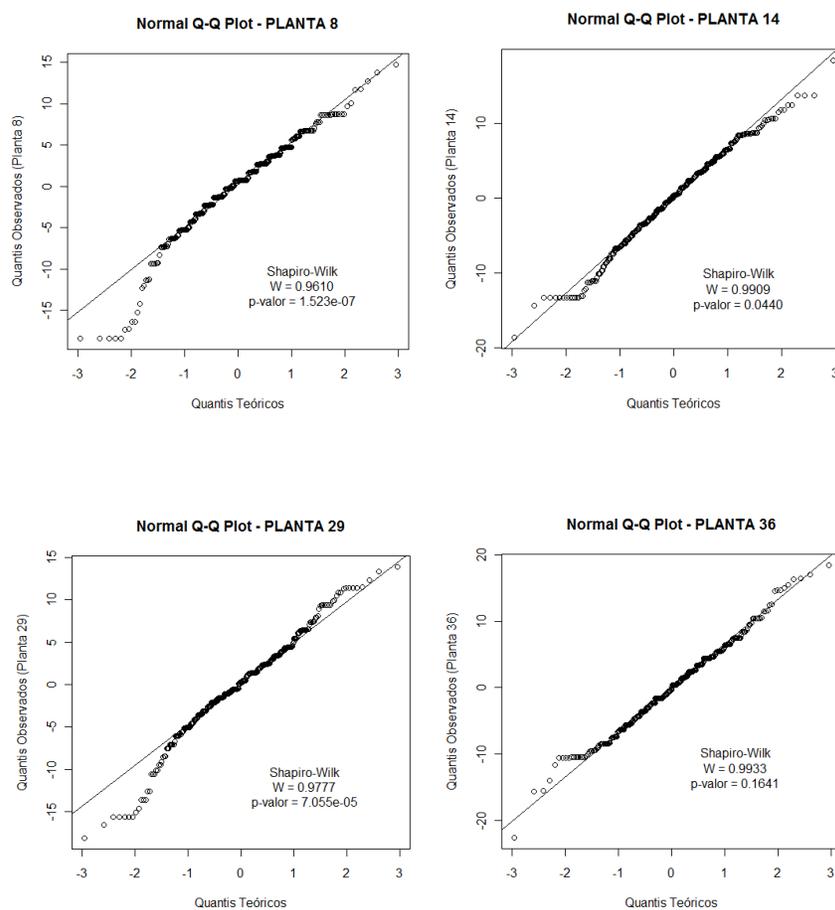


Gráfico 6 Representação gráfica da relação entre os quantis observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo Gompertz, com erros independentes, para as quatro plantas; estatística do teste de normalidade de Shapiro-Wilk

No Gráfico 7 é possível visualizar o comportamento dos resíduos do ajuste do modelo não linear Gompertz com estrutura de erros independentes ao longo do tempo para todas as plantas. Observa-se que o comportamento dos resíduos para o ajuste desse modelo é muito semelhante ao comportamento dos

resíduos para o ajuste do modelo polinomial, talvez devido ao tipo de crescimento dos frutos, que foi sempre ascendente e sem atingir um diâmetro máximo. Na Tabela 5 têm-se as estatísticas do teste de Durbin Watson que testa a independência dos resíduos.

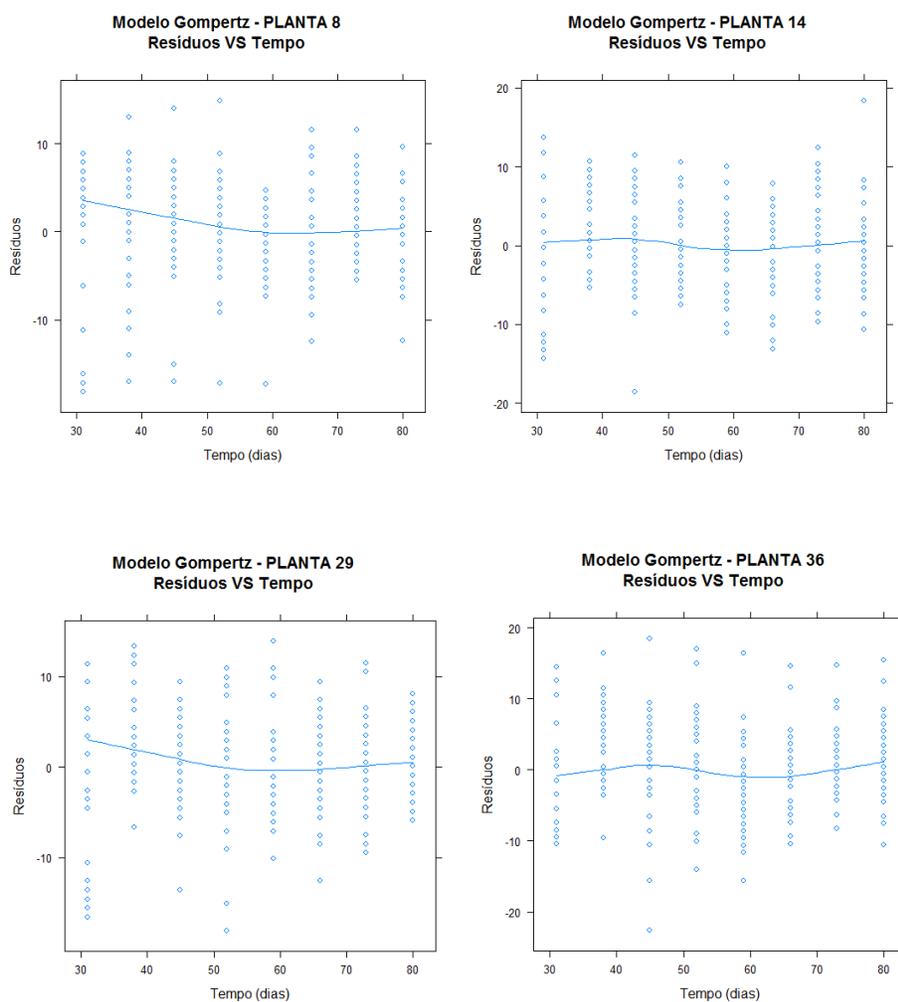


Gráfico 7 Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo Gompertz ao longo do tempo para as quatro plantas de ameixeira avaliadas

Tabela 5 Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Gompertz para as quatro plantas

Planta	Estatística do Teste Durbin Watson (Dw)
8	1,2675
14	1,3884
29	1,4905
36	1,2855

Como no modelo polinomial, observa-se que as estatísticas não estão próximas de 2, ou seja, há evidências de auto-correlação entre os resíduos.

Na Tabela 6 estão apresentadas as estimativas dos coeficientes de regressão do modelo Logístico com erros independentes. Todas as estimativas foram significativas pelo teste t (1%).

Tabela 6 Estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Plantas	Coeficientes de regressão		
	A	B	C
8	71,2478	-1,8579	0,0331
14	54,9801	-2,4243	0,0519
29	52,0404	-2,1711	0,0522
36	48,4812	-2,9419	0,0670

Todos os parâmetros foram significativos em 1%.

Observa-se que o modelo Logístico com erros independentes também parece ter superestimado o diâmetro máximo (parâmetro A) do fruto, para a planta

“8”, apesar de a estimativa ser menor que a do modelo de Gompertz. Percebe-se também da planta “8”, que o parâmetro relativo à taxa de crescimento do fruto (parâmetro C) está subestimado. Do mais, as estimativas dos parâmetros das outras plantas estão próximas uma das outras, dando indícios de não haver diferenças significativas entre elas, o que é confirmado pela representação gráfica dos intervalos de confiança dos parâmetros, os quais estão sobrepostos para as plantas “14”, “29” e “36” (Gráfico 8).

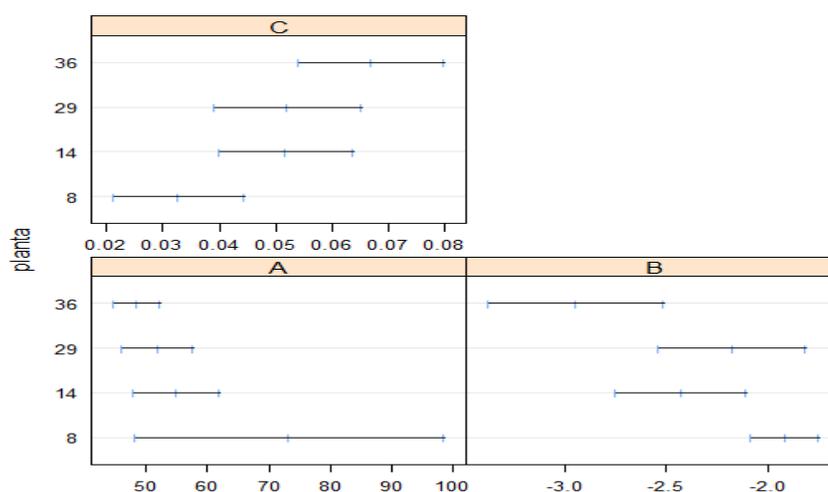


Gráfico 8 Representação gráfica dos intervalos de confiança das estimativas dos parâmetros para o ajuste do modelo Logístico, com erros independentes, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

A representação gráfica dos valores médios do diâmetro do fruto (Gráfico 9) e o ajuste pelo modelo Logístico, com erros independentes, sugerem que no tempo considerado no estudo o crescimento do fruto em diâmetro foi sempre crescente. Também foi verificado um bom ajuste pelo modelo Logístico para todas

as plantas, evidenciado pelo coeficiente de determinação ajustado (R_a^2) acima de 0,70.

Estudando o diâmetro externo longitudinal de frutos de coqueiro, variedade anã-verde, Benassi et al. (2007) verificaram que a função logística foi a que melhor se ajustou ao modelo de crescimento.

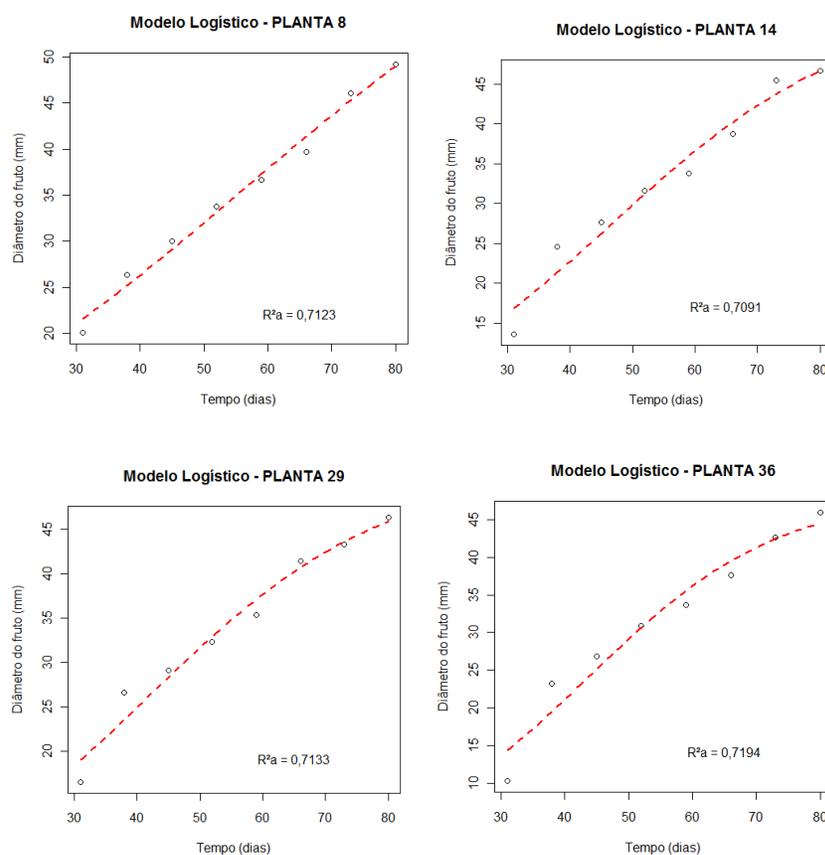


Gráfico 9 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Logístico com estrutura de erros independentes e estimativas do coeficiente de determinação ajustado (R_a^2) para as quatro plantas

Pelo Gráfico 10, verifica-se que após o ajuste da curva Logística com erros independentes, as plantas “8” e “29” também não apresentam resíduos que seguem uma distribuição normal.

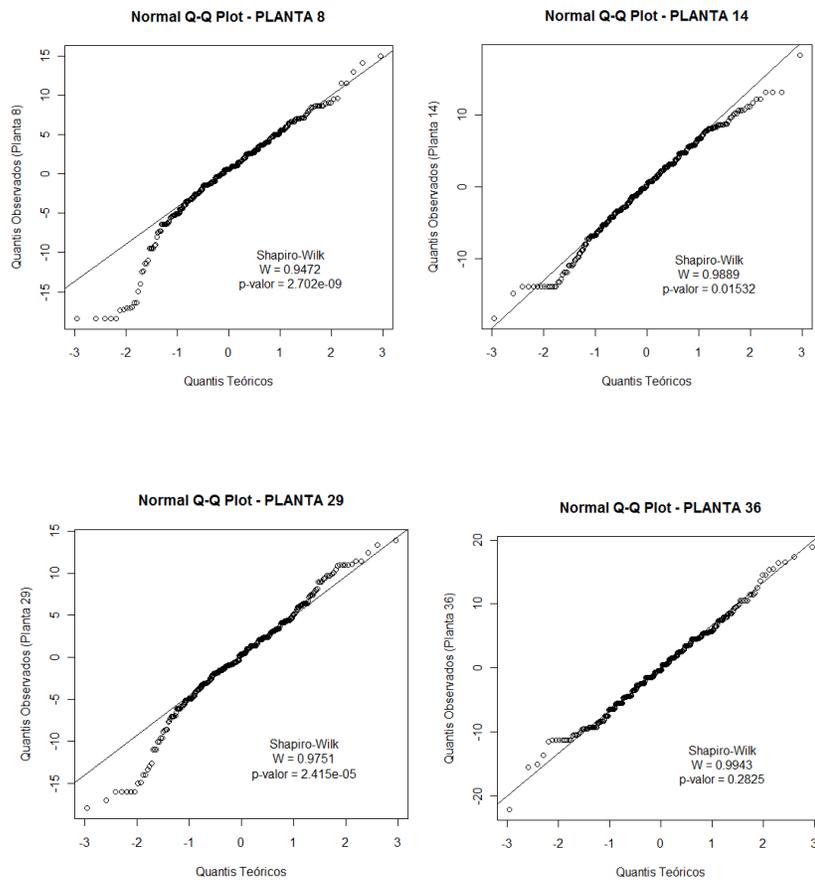


Gráfico 10 Representação gráfica da relação entre os quantis observados e os quantis teóricos esperados relativos aos resíduos do ajuste do modelo Logístico, com erros independentes, para as quatro plantas; estatística do teste de normalidade de Shapiro-Wilk

No Gráfico 11 tem-se a representação gráfica dos resíduos do ajuste do modelo Logístico com erros independentes ao longo do tempo, para todas as plantas. Pode-se notar a ocorrência de uma amplitude bem grande dos resíduos variando de -20 até 20 em um conjunto de dados que variam em média de 10 mm a 45 mm. Pelo teste de Durbin Watson verifica-se que os resíduos não são independentes para modelo Logístico, conforme Tabela 7.

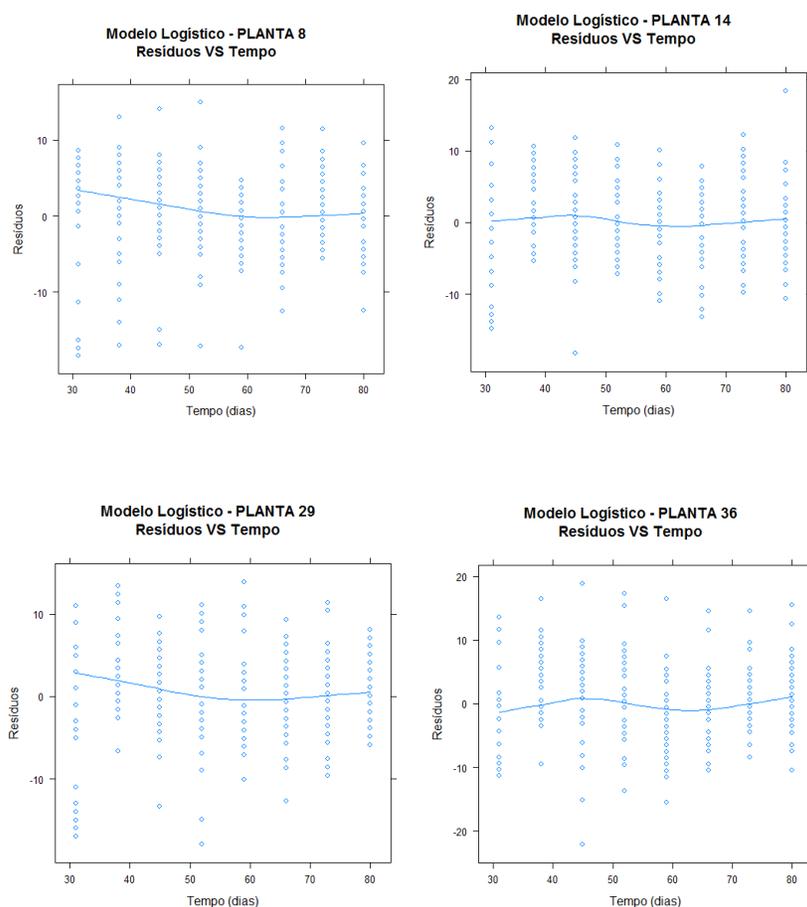


Gráfico 11 Representação gráfica e tendência de comportamento dos resíduos do ajuste do modelo Logístico ao longo do tempo para as quatro plantas de ameixeira avaliadas

Tabela 7 Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Logístico para as quatro plantas

Planta	Estatística do Teste Durbin Watson (Dw)
8	1,2628
14	1,3708
29	1,4795
36	1,2531

Como nos modelos anteriores, observa-se que as estatísticas não estão próximas de 2, ou seja, há evidências de auto-correlação entre os resíduos.

Na Tabela 8 tem-se as estimativas dos parâmetros para o modelo Gompertz com estrutura de erros auto-correlacionados AR(1). Tem-se que todos os parâmetros do modelo são significativos em 1% pelo teste t , com exceção dos parâmetros (A) e (C) da planta “8” que foram significativos aos 5% pelo teste t .

Tabela 8 Estimativas dos parâmetros para o ajuste do modelo Gompertz, com estrutura de erros autorregressiva de ordem 1, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Plantas	Coeficientes de regressão			
	A	B	C	Φ_1
8	93,3393	-0,9046	0,0168	0,3623
14	60,5561	-1,3016	0,0331	0,2992
29	55,1556	-1,2105	0,0363	0,2545
36	51,0493	-1,7321	0,0467	0,3560

Observa-se que os parâmetros relativos à planta “8” continuam superestimando (A) e subestimando (C), mesmo adotando-se a estrutura de erros

auto-correlacionados AR (1), sendo assim, as inferências nos parâmetros não serão muito precisas para essa planta.

Entretanto, observa-se que o ajuste do modelo Gompertz com estrutura de erros auto correlacionados AR (1) teve um melhor ajuste para todas as plantas que os modelos lineares e não lineares com erros independentes, segundo o valor do coeficiente de determinação ajustado (R_a^2) acima de 0,7325 (Gráfico 12) e mais, os resíduos não estão mais correlacionados, ou seja, estão independentes, conforme a Tabela 9.

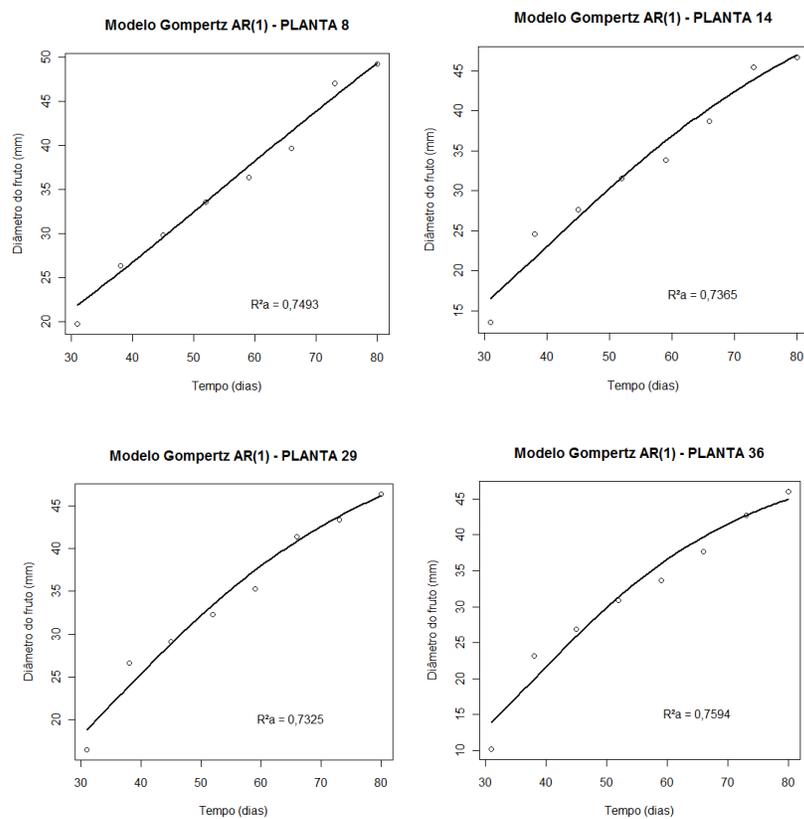


Gráfico 12 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz com estrutura de erros autorregressiva de ordem 1 e estimativa do coeficiente de determinação ajustado (R^2_a) para as quatro plantas

Tabela 9 Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Gompertz AR(1) para as quatro plantas

Planta	Estatística do Teste Durbin Watson (Dw)
8	2,0258
14	1,9713
29	2,0050
36	2,0396

Observa-se que as estatísticas estão muito próximas de 2, ou seja, não há evidências de que os resíduos estejam correlacionados.

Na Tabela 10, verifica-se que todos os parâmetros do modelo Logístico com erros auto-correlacionados AR (1) foram significativos em 1% pelo teste t .

Tabela 10 Estimativas dos parâmetros para o ajuste do modelo Logístico, com estrutura de erros auto-regressiva de ordem 1, ao diâmetro dos frutos da ameixeira “Gulfblase” para as quatro plantas

Plantas	Coeficientes Estimados			
	A	B	C	Φ_1
8	72,1614	-1,8516	0,0325	0,3651
14	55,1829	-2,4301	0,0517	0,3079
29	51,7590	-2,1859	0,0528	0,2601
36	48,5132	-2,9419	0,0670	0,3721

Observa-se que, mesmo considerando os erros auto-correlacionados AR (1), os parâmetros (A) e (C) da planta “8” do modelo Logístico também continuam superestimados e subestimados, respectivamente.

O ajuste do modelo Logístico com erro auto-correlacionado também foi melhor, para todas as plantas, do que os modelos com erros independentes, dado pelo (R_a^2) acima de 0,7310, conforme o Gráfico 13. Como no modelo não linear Gompertz AR(1), tem-se que os erros não estão mais correlacionados, conforme a Tabela 11.

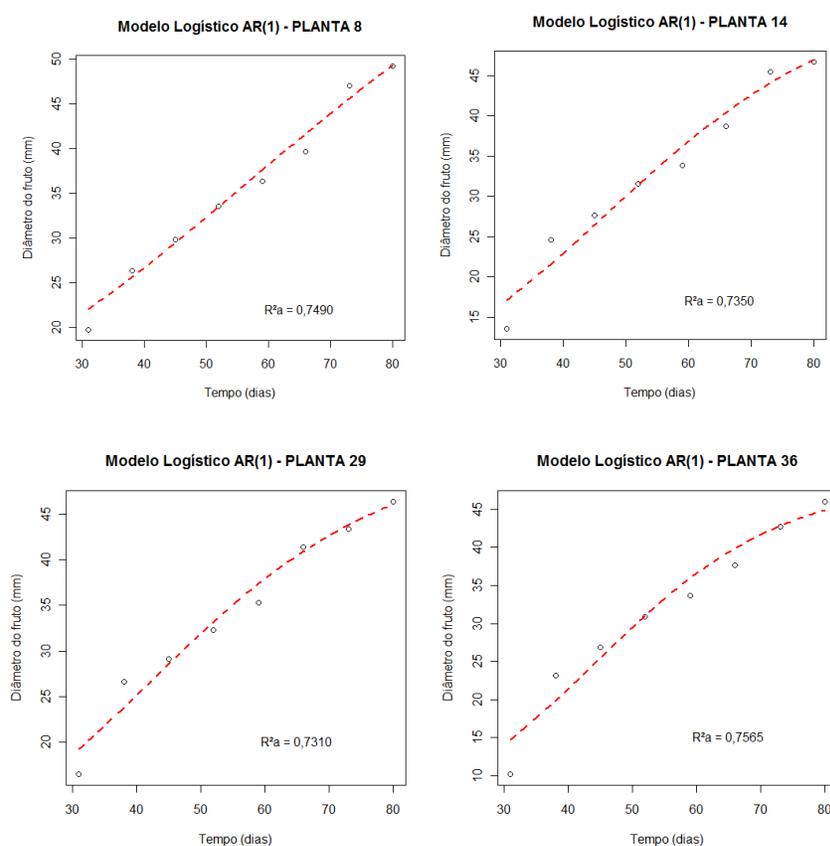


Gráfico 13 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Logístico com estrutura de erros autorregressiva de ordem 1 e estimativa do coeficiente de determinação ajustado (R_a^2) para as quatro plantas

Tabela 11 Estatísticas do teste de Durbin Watson dos resíduos do modelo não linear Logístico AR(1) para as quatro plantas

Planta	Estatística do Teste Durbin Watson (Dw)
8	2,0282
14	1,9739
29	2,0074
36	2,0481

Observa-se que as estatísticas estão muito próximas de 2, ou seja, não há evidências de que os resíduos estejam correlacionados.

Portanto, a modelagem dos erros melhorou a qualidade de ajuste dos modelos não lineares, o que já era de se esperar, pois à medida que se aumenta o número de parâmetros em um modelo, espera-se melhor ajuste; nos modelos estudados passou-se de três para quatro parâmetros. Além disso, com a modelagem dos erros, conseguiu-se retirar a dependência que havia entre os resíduos, sendo assim, os resíduos passaram a ser independentes.

4.4 Seleção dos modelos ajustados

Pelos avaliadores da qualidade de ajuste dos modelos para a planta “8”, conforme Tabela 12, verifica-se que tanto o modelo Gompertz quanto o Logístico, com estrutura de erros autorregressivos de ordem 1 são os que possuem maior R_a^2 ; menores desvios padrão, DPR e por fim, o critério de informação de Akaike (AIC) mostrou que, embora os modelos com estrutura de erros auto correlacionados têm o número de parâmetros maior, as estimativas do critério (AIC) são menores que as estimativas dos modelos com estrutura de

erros independentes sendo, portanto, os modelos de melhor qualidade de ajuste, com uma ligeira vantagem para o modelo Gompertz com erros autorregressivos AR(1).

Tabela 12 Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_a^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC), teste de Durbin-Watson (DW), para os modelos ajustados para a planta “8” da cultivar “Gulfblase” de ameixa

Planta	Modelo	Critérios de Seleção		
		R_a^2	DPR	AIC
8	Quadrático	0,7134	5,784	5,5101
	Gompertz	0,7132	5,794	5,5136
	Logístico	0,7123	5,804	5,5187
	Gompertz AR (1)	0,7493	5,409	5,3762
	Logístico AR (1)	0,7490	5,412	5,3772

No Gráfico 14 está a representação gráfica do ajuste dos dois modelos não lineares com estrutura de erros auto regressivos. Observa-se que as curvas quase que se sobrepõem uma à outra, o que já era esperado dado que todas as estimativas dos critérios de seleção estão muito próximas.

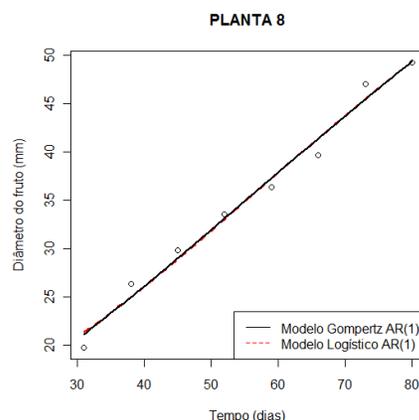


Gráfico 14 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz e Logístico com estrutura de erros autorregressiva de ordem 1 para a planta “8”

Porém, tem-se que o modelo Gompertz com estrutura de erros auto-correlacionados, os parâmetros (A) e (C) são os que apresentam estimativas superestimada e subestimada, de maior e menor valor, respectivamente, sendo assim, é preferível adotar-se o Logístico com erros auto-correlacionados AR (1) como o melhor modelo para o ajuste do diâmetro dos frutos da ameixeira para a planta “8”.

Para as outras plantas, pelo resultado da Tabela 13 verifica-se que, os modelos não lineares auto regressivos AR(1) também são os que possuem melhores estimativas dos critérios de seleção, novamente com uma ligeira vantagem ao modelo Gompertz.

Tabela 13 Estimativas dos critérios de seleção: coeficiente de determinação ajustado (R_a^2), desvio padrão residual (DPR), critério de informação de Akaike (AIC), teste de Durbin-Watson (DW), para os modelos ajustados para as plantas “14”, “29” e “36” da cultivar “Gulfblase” de ameixa

Planta	Modelo	Critérios de Seleção		
		R_a^2	DPR	AIC
14	Quadrático	0,7147	6,437	5,7233
	Gompertz	0,7125	6,472	5,7353
	Logístico	0,7091	6,509	5,7471
	Gompertz AR (1)	0,7365	6,185	5,6441
	Logístico AR (1)	0,7350	6,204	5,6504
29	Quadrático	0,7167	5,736	5,4935
	Gompertz	0,7158	5,755	5,5001
	Logístico	0,7133	5,780	5,5086
	Gompertz AR (1)	0,7325	5,574	5,4362
	Logístico AR (1)	0,7310	5,590	5,4420
36	Quadrático	0,7302	6,425	5,7204
	Gompertz	0,7264	6,480	5,7350
	Logístico	0,7194	6,562	5,7626
	Gompertz AR (1)	0,7594	6,068	5,6060
	Logístico AR (1)	0,7565	6,103	5,6177

No Gráfico 15, observa-se que os modelos quase se coincidem para todas as plantas, ficando difícil de julgar qual o melhor modelo em questão. Pelos critérios de seleção dos modelos não se observa diferenças marcantes entre os modelos sem auto-correlação e, entre os com auto-correlação; em todos os critérios as estimativas foram bastante semelhantes. Melhores ajustes foram verificados para os modelos com auto-correlação, como maior coeficiente de determinação, menor desvio e menor AIC.

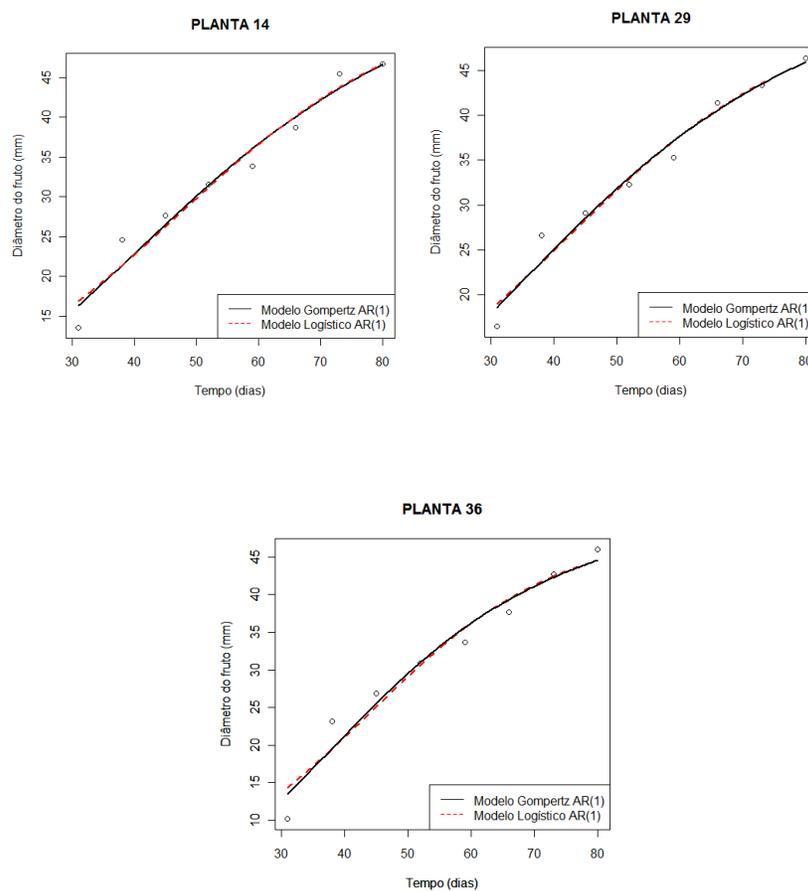


Gráfico 15 Representação gráfica dos valores médios do diâmetro dos frutos da ameixeira “Gulfblase” em função do tempo de avaliação, da curva de ajuste do modelo Gompertz e Logístico com estrutura de erros autorregressiva de ordem 1 para as plantas “14”, “29” e “36”

Porém, seguindo a rigor os critérios de seleção dos modelos, tem-se que o melhor modelo que descreve os diâmetros dos frutos ao longo do tempo para as plantas “14”, “29” e “36” é o modelo Gompertz com erros autocorrelacionados AR (1).

De modo geral, verificou-se que a planta “8” difere das demais, pois para ambas as funções não lineares estudadas, embora tenham mostrado um bom ajuste aos dados dessa planta, tendem a superestimar o diâmetro máximo (assíntota A) dos frutos da ameixeira. O mesmo ocorreu para as outras plantas, mas em menor escala. Isto pode ser devido ao período de avaliação e ao número de avaliações realizadas. Nesse período de avaliação o diâmetro dos frutos sempre foi crescente.

Pelos resultados obtidos não foi possível detectar o comportamento do tipo duplo-sigmóide sugerido na literatura para frutos que contenham caroço (BARBOSA et al., 1993; DEJONG; GOUDRIAAN, 1989).

Sugere-se que em trabalhos futuros com frutos da ameixeira “Gulfblase”, por ser um cultivar com frutos de ciclo curto, que as avaliações iniciam-se antes de 31 dias após a frutificação e que se use um maior número de intervalos de avaliação, como por exemplo, particionando os intervalos de coleta dos dados em períodos mais curtos, como de três em três dias e avançando mais ao final do ciclo para se atingir um diâmetro máximo de crescimento.

5 CONCLUSÕES

O modelo de regressão polinomial e os modelos não lineares descrevem de modo adequado o diâmetro do fruto durante sua fase de crescimento.

O modelo Logístico e o de Gompertz com estrutura de erros autorregressivos de ordem 1 descrevem melhor o crescimento dos frutos da ameixeira ao longo do tempo de avaliação.

REFERÊNCIAS

ÁLVAREZ, I. A.; CASTRO, P. R. C. Crescimento da parte aérea de cana crua e queimada. **Scientia Agrícola**, Piracicaba, v. 56, n. 4, p. 1069-1079, 1999.

BARBOSA, W. et al. Desenvolvimento dos frutos e das sementes de pêssegos subtropicais de diferentes ciclos de maturação. **Pesquisa Agropecuária Brasileira**, Brasília, v. 28, n. 6, p. 701-707, jun. 1993.

BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: J. Wiley, 1988. 365 p.

BENASSI, A. C. et al. Caracterização biométrica de frutos de coqueiro, *Vocos nucifera* L. variedade anã-verde, em diferentes estádios de desenvolvimento. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 29, n. 2, p. 302-307, ago. 2007.

BRUNA, E. D. Curva de crescimento de frutos de pêssego em regiões subtropicais. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 29, n. 3, p. 685-698, maio/jun. 2007.

CALEGARIO, N. et al. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal eucaliptus plantation case study. **Forest Ecology and Management**, Amsterdam, v. 204, n. 1, p. 11-20, Jan. 2005.

CASTRO, L. A. S. et al. **Implantação do pomar**. Pelotas: EMBRAPA Clima Temperado, 2003. 49 p.

_____. **Introdução: sistemas de produção**. 2. ed. Pelotas: EMBRAPA Clima Temperado, 2005. 5 p.

CASTRO NETO, M. T.; REINHARDT, D. H. Relações entre parâmetros de crescimento do fruto da manga cv. Haden. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 25, n. 1, p. 36-38, abr. 2003.

CHAGAS, P. C. **Cultivares de ameixas de baixa exigência em frio para regiões subtropicais do Estado de São Paulo**. 2008. 122 p. Dissertação (Mestrado em Fitotecnia) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2008.

CHARNET, R. et al. **Análise de modelos de regressão linear: com aplicações**. 2. ed. Campinas: UNICAMP, 2008. 356 p.

DEJONG, T. M.; GOUDRIAAN, J. Modeling peach fruit growth and carbohydrate requirements: reevaluation of the doublesigmoid growth pattern. **Journal of the American Society for Horticultural Science**, Alexandria, v. 114, n. 5, p. 800-804, Sept. 1989.

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

GALLANT, A. R. **Nonlinear statistical models**. New York: J. Wiley, 1987. 624 p.

GODOY, A. R.; CARDOSO, A. I. I. Curva de crescimento e qualidade de frutos de melão rendilhado sob cultivo protegido. **Revista Ceres**, Viçosa, MG, v. 50, n. 289, p. 220-303, 2003.

GRAYBILL, F. A.; IYER, H. K. **Regression analysis: concepts and applications**. Belmont: Duxbury, 2004. 699 p.

GRUMBERG, I. P. **Variedades de durazneros y ciruelos que se cultivan em el país**. Buenos Aires: Universidad de Buenos Aires, 1944. 453 p.

GUJARATI, D. **Econometria básica**. 4. ed. São Paulo: Campus, 2006. 812 p.

GURJÃO, K. C. O. et al. Desenvolvimento de frutos e sementes de tamarindo. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 28, n. 3, p. 351-354, dez. 2006.

HARTLEY, H. O. The modified Gauss-Newton method for the fitting of nonlinear regression function by least squares. **Technometrics**, Washington, v. 3, p. 269-280, 1964.

JAUER, A. et al. Análise de crescimento da cultivar de feijão pérola em quatro densidades de semeadura. **Revista da Faculdade de Zootecnia, Veterinária e Agronomia**, Uruguaiana, v. 10, n. 1, p. 1-12, 2003.

MARTINS, L. P. et al. Desenvolvimento de frutos de cirigueira (*Spondias purpurea* L.). **Revista Brasileira de Fruticultura**, Jaboticabal, v. 25, n. 1, p. 11-14, abr. 2003.

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

MOTULSKY, H.; CHRISTOPOULOS, A. **Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting**. San Diego: GraphPad, 2003. 351 p.

MOURA, M. L. et al. Crescimento e desenvolvimento de frutos do tomateiro “Santa Clara” e do seu mutante natural “Firme”. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 6, p. 1284-1290, nov./dez. 2004.

MUÑOZ, A. R. M. Pudriciones en fruta de exportación y mercado interno. **Revista Frutícola**, Curicó, v. 3, n. 3, p. 83-86, 1982.

OLIVEIRA, R. S. **Cultivo orgânico de ameixeiras em Delfim Moreira, MG**. 2009. 64 p. Dissertação (Mestrado em Fitotecnia) - Universidade Federal de Lavras, Lavras, 2009.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing.** Vienna: R Foundation for Statistical Computing, 2010. Disponível em: <<http://www.R-project.org>>. Acesso em: 10 nov. 2010.

RATKOWSKY, D. A. **Nonlinear regression modeling.** New York: M. Bekker, 1983. 276 p.

RENCHEER, A. C.; SCHAALJE, G. B. **Linear models in statistics.** New Jersey: J. Wiley, 2008. 672 p.

RESZCZYNSKY, A. Estudio de madurez de exportación en ciruelas El Dorado. In: SIMPÓSIO SOBRE MANEJO, CALIDAD Y FISILOGIA DE POSTCOSECHA DE FRUTAS, 2., 1977, San Felipe. **Anais...** San Felipe: Miscelaneas, 1977. p. 120-126.

SCHABENBERGER, O.; PIERCE, F. J. **Contemporary statistical models for the plant and soil sciences.** Boca Raton: CRC, 2002. 753 p.

SEARLE, S. R. **Linear models.** New York: J. Wiley, 1971. 532 p.

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.

STATISTICAL ANALYSIS SYSTEM INSTITUTE. **SAS procedures guide for computers.** Cary, 1999. v. 3, 373 p.

STEPHENS, M. A. EDF statistics for goodness of fit and some comparisons. **Journal of the American Statistical Association**, New York, v. 69, n. 347, p. 730-737, Sept. 1974.

TORRELLARDONA, S. D. **Frijo conservación de la fruta.** Barcelona: AEDOS, 1983. 369 p.

WEINBERGER, J. H. Plums. In: JANICK, J.; MOORE, J. N. **Advance in moderns fruit breeding**. West Lafayette: Purdue University, 1975. p. 336-347.

APÊNDICE A - Rotinas R

```
#####
# Rotina do programa R utilizado para a obtenção das estimativas dos #
# parâmetros do modelo polinomial. #
#####

dadosdis <- read.table("dadosfinal.txt", h=T)
attach(dadosdis)

# Pacotes utilizados

require(car)
require(lattice)
require(nlme)
require(stats)
require(base)

# Ajuste polinomial para o diam – Planta “8”

reg1 <- lm(diam ~ dias + I(dias^2),
data=dadosdis[dadosdis$planta=="8",])

summary(reg1)
anova(reg1)

# Representação Gráfica

rel <- summary(reg1)
diasmed <- c(31,38,45,52,59,66,73,80)
mdiam <- tapply(diam, list(planta,dias), mean)
diam8 <- mdiam[1,]
plot(diam8~diasmed, xlab="Tempo (dias)",
ylab="Diâmetro do fruto (mm)",
data=dadosdis[dadosdis$planta=="8",],main="Polinomial - PLANTA 8")
curve(rel$coef[1]+rel$coef[2]*x+rel$coef[3]*x^2, add=T
```

```

, col="1", lty=1, lwd=2)
text(65,22,
"R²a = 0,7134")
# Análise dos Resíduos

# Normalidade

shapiro.test(residuals(reg1, type="pearson"))

qqnorm(reg1$residuals,
main="Normal Q-Q Plot - PLANTA 8",
xlab="Quantis Teóricos",
ylab="Quantis Observados (Planta 8)")
qqline(reg1$residuals, col="1")
text(1.5, -10,
"Shapiro-Wilk
W = 0.9499
p-valor = 5.605e-09")
op <- par(fig=c(.1,.5,.4,0.9), new=TRUE)
hist(reg1$residuals, probability=T,
col="white", xlab="", ylab="", main="", axes=F)
lines(density(reg1$residuals), col="2", lwd=2)

# Homocedasticidade e Independência

xyplot(residuals(reg1)~dadosdis[dadosdis$planta=="8"],$dias,
xlab="Tempo (dias)", ylab="Resíduos",
type=c("p","smooth"), main="Modelo Polinomial - PLANTA 8
Resíduos VS Tempo ")

# As demais plantas seguem o mesmo esquema de estimação e construção dos
gráficos.

# Cálculo para todas as plantas usando o lmList

reg.diam <- lmList(diam ~ (dias + I(dias^2))|factor(posicao), data=dadosdis)
summary(reg.diam)
str(intervals(reg.diam))
intervals(reg.diam)[,1:3]

```

```
#####
# Rotina do programa R utilizado para a obtenção das estimativas dos #
# parâmetros do modelo Gompertz. #
#####

# Ajuste do modelo Gompertz para o diam – Planta “8”

require(nlme)
require(lattice)

getInitial(diam ~ SSGompertz(dias, A~planta,
B~planta, C~planta),
data=dadosdis[dadosdis$planta=="8",])

reg2 <- nls(diam ~ A*exp(-exp(-(B+C*dias))),
start=list(A=A0, B=B0, C=C0),
data=dadosdis[dadosdis$planta=="8",])

summary(reg2)
resulg <- summary(reg2)

# Cálculo do coeficiente de determinação ajustado (R2a)

SQE <- resulg$sigma^2*resulg$df[2]
SQT <- varr(dadosdis[dadosdis$planta=="8",]$diam)*
(length(dadosdis[dadosdis$planta=="8",]$diam)-1)
R2 <- 1 - SQE/SQT
R2A <- 1-((320-1)*(1-R2))/(320-2)
R2A

# Representação Gráfica

diasmed <- c(31,38,45,52,59,66,73,80)
mdiam <- tapply(diam, list(planta,dias), mean)
diam8 <- mdiam[1,]
plot(diam8~dias1, xlab="Tempo (dias)",
ylab="Diâmetro do fruto (mm)",
```

```

data=dadosdis[dadosdis$planta=="8",],main="Modelo Gompertz - PLANTA 8")
x<-seq(1,1100)
curve(resulg$coef[1] * exp(-exp(-(resulg$coef[2]+resulg$coef[3]*x))),add=T,
col="1", lty=1, lwd=2)
text(65,22,
"R²a = 0,7132")

```

```
# Análise dos resíduos
```

```
# Normalidade
```

```
shapiro.test(residuals(reg2, type="pearson"))
```

```

qqnorm(residuals(resulg),
main="Normal Q-Q Plot - PLANTA 8",
xlab="Quantis Teóricos", ylab="Quantis Observados (Planta 8)")
qqline(residuals(reg2))
text(1.2, -12,
"Shapiro-Wilk
W = 0.9484
p-valor = 3.723e-09")

```

```
# Homocedasticidade e independencia
```

```

xyplot(residuals(reg2)~dadosdis[dadosdis$planta=="8",]$dias,
xlab="Tempo (dias)", ylab="Resíduos",
type=c("p","smooth"), main="Modelo Gompertz - PLANTA 8
Resíduos VS Tempo ")

```

As demais plantas seguem o mesmo esquema de estimação e construção dos gráficos.

```
# Cálculo para todas as plantas.
```

```
getInitial(diam ~ SSGompertz(dias, A~posicao+planta, B~posicao+planta,
C~posicao+planta), data=dadosdis)
```

```

reggomp<- nlsList(diam ~ A * exp(-exp(-(B+C*dias)))|planta,
start = list(A=59, B = -1, C = 0.05),

```

```

data=dadosdis)
summary(reggomp)

# nlsList

plot(intervals(reggomp))
intervals(reggomp)

#####
# Rotina do programa R utilizado para a obtenção das estimativas dos #
# parâmetros do modelo Logístico. #
#####

# Ajuste do modelo Logístico para o diam – Planta “8”

getInitial(diam ~ SSlogis(dias,A~planta,B~planta,
C~planta),data=dadosdis[dadosdis$planta=="8",])

reg3 <- nls(diam ~ A/(1+exp(-(B+C*dias))), start=list(A=A0,
B=B0, C=C0), data=dadosdis[dadosdis$planta=="8",])

summary(reg3)
resull <- summary(reg3)

# Cálculo do coeficiente de regressão ajustado (R2a)

SQE <- resull$sigma^2*resull$df[2]
SQT <- varr(dadosdis[dadosdis$planta=="8",]$diam)*
(length(dadosdis[dadosdis$planta=="8",]$diam)-1)
R2 <- 1 - SQE/SQT
R2A <- 1-((320-1)*(1-R2))/(320-2)
R2A

# Representação Gráfica

diasmed <- c(31,38,45,52,59,66,73,80)
mdiam <- tapply(diam, list(planta,dias), mean)

```

```

diam8 <- mdiam[1,]

plot(diam8~diasmed, xlab="Tempo (dias)",
ylab="Diâmetro do fruto (mm)",
data=dadosdis[dadosdis$planta=="8",],main="Modelo Logístico - PLANTA 8")
x<-seq(1,1100)
curve(result$coef[1]/(1 + exp(-(result$coef[2]+result$coef[3]*x))),add=T,
col="1", lty=1, lwd=2)
text(65,22,
"R²a = 0,7123")
# Análise dos resíduos

# Normalidade

shapiro.test(residuals(result, type="pearson"))

qqnorm(residuals(result),
main="Normal Q-Q Plot - PLANTA 8",
xlab="Quantis Teóricos",
ylab="Quantis Observados (Planta 8)")
qqline(residuals(result))
text(1.2, -15,
"Shapiro-Wilk
W = 0.9472
p-valor = 2.703e-09")
# Homocedasticidade e independencia

xyplot(residuals(reg3)~dadosdis[dadosdis$planta=="8",]$dias,
xlab="Tempo (dias)", ylab="Resíduos",
type=c("p","smooth"), main="Modelo Logístico - PLANTA 8
Resíduos VS Tempo ")

# As demais plantas seguem o mesmo esquema de estimação e construção dos
gráficos.

# Cálculo para todas as plantas.

getInitial(diam ~ SSlogis(dias, A~posicao+planta, B~posicao+planta,
C~posicao+planta), data=dadosdis)

```

```
reglog <- nlsList(diam ~ A/(1+exp(-(B+C*dias)))|planta,  
start = list(A=54, B=1, C=0.01),  
data=dadosdis)
```

```
summary(reglog)
```

```
# nlsList
```

```
plot(intervals(reglog))  
intervals(reglog)
```

APÊNDICE B – Rotina SAS

```
#####
# Rotina do programa SAS utilizado para a obtenção das estimativas dos #
# parâmetros dos modelos não lineares Gompertz e Logístico com estrutura de #
# erros autorregressiva de ordem um. #
#####
```

```
# Ajuste do modelos não lineares Gompertz e Logístico para o diam – Planta “8”
```

```
data dadosdis ;
input diam dias;
datalines;
30    31
25    31
25    31
...   ...
49    80
51    80
48    80
;
/*sem erros autorregressivos*/
proc model data=dadosdis maxiter=1000;
diam= (escrever função);
parms A=A0 B=B0 C=C0;
fit diam/ outall out=saida dw prl=wald;
run;
/*com erros AR (1)*/
proc model data=dadosdis maxiter=1000;
diam= (escrever função);
parms A=A0 B=B0 C=C0;
fit diam/ outall out=saida dw prl=wald;
% ar (y,1);
run;
```

```
# As demais plantas seguem o mesmo esquema de estimação.
```