



Universidade Estadual da Paraíba  
Centro de Ciências e Tecnologia  
Departamento de Estatística

**José Severino da Silva Neto**

# **Comparação de modelos de crescimento aplicados a cultura do mamoeiro**

Campina Grande  
Dezembro de 2014

José Severino da Silva Neto

# Comparação de modelos de crescimento aplicados a cultura do mamoeiro

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientadora:

Ana Patricia Bastos Peixoto

Co-orientadora:

Gláucia Amorim Faria

Campina Grande

Dezembro de 2014

É expressamente proibida a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação.

S586c Silva Neto, José Severino da.

Comparação de modelos de crescimento aplicados a cultura do mamoeiro [manuscrito] / Jose Severino da Silva Neto. - 2014.  
34 p. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2014.

"Orientação: Profa. Dra. Ana Patricia Bastos Peixoto, Departamento de Estatística".

"Co-Orientação: Profa. Dra. Glauca Amorim Faria., Departamento de Matemática".

1. Modelos não lineares. 2. Curvas de crescimento. 3. Critério de seleção. 4. Mamoeiro. I. Título.

21. ed. CDD 519.5

José Severino da Silva Neto

## Comparação de modelos de crescimento aplicados a cultura do mamoeiro

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 09/12/14

### Banca Examinadora:



Prof<sup>a</sup> Dr<sup>a</sup> Ana Patricia Bastos Peixoto  
UEPB - DE/CCT  
Orientadora



Prof. Dr. Tiago Almeida de Oliveira  
UEPB - DE/CCT  
Examinador



Prof. Dr. Ricardo Alves de Olinda  
UEPB - DE/CCT  
Examinador

# Dedicatória

*A DEUS que me deu forças para enfrentar as dificuldades e conseguir chegar até aqui.  
Aos meus pais, Antonio José e Maria Lusinete,  
aos meus irmãos, Márcia Marcielle e Marcio Maciel,  
a minha namorada Bruna Brito, por todo amor, cuidado, apoio, confiança e paciência.*

# Agradecimentos

A DEUS, por ter me concedido a graça da vida e por me amparar em todos os momentos, por me dar forças para seguir enfrente e ter permitido em minha vida mais uma vitória.

A toda a minha família, em especial meus pais Antonio José da Silva e Maria Lusinete dos Santos Silva, que são os responsáveis pelo que sou hoje.

Aos meus irmãos, Márcia Marciele e Marcio Maciel pelo carinho, apoio e dedicação.

A minha namorada Bruna Brito, pela paciência, apoio, incentivo e por todo amor.

A minha orientadora Ana Patricia Bastos Peixoto, pela paciência, amizade e dedicação.

A UEPB por proporcionar condições para meu aprendizado, durante todo período do curso. Especialmente, aos professores do departamento que tanto contribuíram para minha formação.

A todos os meus amigos de curso, em especialmente Dark Maciel e Klaine Clemente, Arnete Campos por toda amizade, confiança, incentivo, por todos os momentos vividos juntos que jamais esquecerei.

Gostaria de agradecer também a todos os meus amigos Manoel Messias, Aclécio Pereira, Rivaildo Costa, Manoel Messias Santos (in memorian), Rubeilson Santos, que sempre me deram forças para chegar até aqui.

A todos que acreditaram em mim e contribuíram de forma direta e indireta, o meu muito obrigado.

*“Que os vossos esforços desafiem as impossibilidades,  
lembrai-vos de que as grandes coisas do homem foram  
conquistadas do que parecia impossível.”*

Charles Chaplin.

# Resumo

Por ser crescente a necessidade de mais tecnologia na cultura do mamoeiro, se tornam necessários experimentos mais precisos e interpretação do seu desenvolvimento mais detalhada, atribuído a isso, modelos de crescimento tem sido utilizados para investigar tais comportamentos em diversas culturas. Os dados utilizados para esse trabalho, foram oriundos de dois ensaios de uniformidade com mamoeiro, em casa de vegetação, na área experimental da Embrapa Mandioca e Fruticultura Tropical, em Cruz das Almas-BA. As variáveis avaliadas foram altura da planta em centímetros e o diâmetro do caule em centímetros. Nesse contexto, o objetivo deste trabalho foi ajustar curvas de crescimento associados a altura e diâmetro do mamoeiro ao longo dos dias. Os dados em estudo foram submetidos ao ajuste dos modelos Logístico, Gompertz e Von Bertalanffy. O método utilizado para o ajuste foi o iterativo Gauss-Newton. A interpretação biológica dos parâmetros, o coeficiente de determinação, o intervalo de confiança a 95% das estimativas, o critério de Informação de Akaike e o critério de Informação de Bayes, foram utilizados para selecionar o modelo adequado aos dados. Dentre os modelos ajustados o modelo Gompertz foi o mais indicado para representar a curva de crescimento referente a altura e diâmetro da cultura do mamoeiro.

**Palavras-chave:** Modelos não lineares; Curvas de crescimento; Critério de seleção

# Abstract

Because it is a growing need for more technology in the papaya tree, become necessary more accurate interpretation of his experiments and detailed development assigned to it, growth models have been used to investigate such behavior in various crops. The data used for this work were from two uniformity tests with papaya, in a greenhouse at Embrapa experimental area Cassava and Tropical Fruits, in Cruz das Almas, Bahia. The variables were plant height in centimeters and stem diameter in centimeters. In this context, the aim of this work was to adjust growth curves associated with height and diameter of papaya over the days. The data in the study were submitted to the adjustment of Logistic models, Gompertz and von Bertalanffy. The method used for the adjustment was the iterative Gauss-Newton. The biological interpretation of parameters, the coefficient of determination, confidence interval of 95% of the estimates, the criterion of Akaike information criterion and the Information Bayes, were used to select the appropriate model to the data. Among the models set the Gompertz model was the most suitable to represent the growth curve for height and diameter of papaya.

**Keywords:** Nonlinear models, Growth curves, Selection criteria

# Sumário

Lista de Figuras

Lista de Tabelas

<b>1</b>	<b>Introdução</b>	p. 12
<b>2</b>	<b>Fundamentação Teórica</b>	p. 14
2.1	Marco Histórico . . . . .	p. 14
2.2	Modelos de regressão linear . . . . .	p. 15
2.3	Modelos de regressão não linear . . . . .	p. 15
2.3.1	Métodos de estimação . . . . .	p. 16
2.4	Seleção e qualidade do ajuste . . . . .	p. 18
2.4.1	Coeficiente de determinação . . . . .	p. 18
2.4.2	Critério de Informação de Akaike (AIC) . . . . .	p. 19
2.4.3	Critério de informação (BIC) . . . . .	p. 20
2.5	Análise de resíduos . . . . .	p. 20
2.5.1	Distância de Cook . . . . .	p. 21
2.6	Curvas de crescimento . . . . .	p. 23
<b>3</b>	<b>Aplicação</b>	p. 25
3.1	Material e métodos . . . . .	p. 25
3.2	Resultados e discussão . . . . .	p. 26
<b>4</b>	<b>Conclusão</b>	p. 31



# Lista de Figuras

- 1 Ajuste dos modelos logístico, gompertz e Von Bertalanffy para a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro aos longo dos dias. p. 28
- 2 Gráfico dos quantis normais para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro ao longo dos dias para os modelos logístico, gompertz e von bertalanffy. . . . . p. 29
- 3 Gráfico dos resíduos estudentizados para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro ao londo dos dias para os modelos logístico, gompertz e von bertalanffy. . . . . p. 29
- 4 Gráfico das distâncias de Cook para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro para os modelos logístico, gompertz e von bertalanffy. . . . . p. 30

# Lista de Tabelas

1	Médias da altura e diâmetro das plantas de mamoeiro ao longo dos dias	p. 25
2	Estimativas dos parâmetros para o modelo Logístico, Gompertz e Von Bertalanffy, erro padrão da estimativa (E.P.E.), valor- $p$ para o teste $t$ e intervalos de confiança (IC) de 95% para altura em função dos dias . .	p. 27
3	Estimativas dos parâmetros para o modelo Logístico, Gompertz e Von Bertalanffy, erro padrão da estimativa (E.P.E.), valor- $p$ para o teste $t$ e intervalos de confiança (IC) de 95% para diâmetro em função dos dias .	p. 27
4	Coefficiente de determinação ( $R_a^2$ ) e valores de AIC e BIC para os modelos ajustados referente a altura e o diâmetro do mamoeiro. . . . .	p. 28

# 1 Introdução

O mamoeiro (*Carica papaya* L.), originário da América Central, é uma planta cultivada em regiões tropicais e subtropicais, estando disseminado praticamente em todo o território nacional, onde existem milhares de hectares propícios ao seu desenvolvimento. Por ser uma planta de crescimento rápido e contínuo, com floração e frutificação concomitantes e ininterruptas, necessita de adubações e suprimento de água constantes em todo o seu ciclo. Nas últimas décadas o Brasil se vem destacando como um dos principais produtores mundiais de mamão (FAO, 2012). Os principais entraves ao crescimento contínuo da cultura do mamão são o reduzido número de cultivares disponíveis, o que tem resultado em uma uniformidade nos plantios comerciais e uma conseqüente redução da variabilidade genética.

O melhoramento da cultura do mamoeiro pode estar, associado a boas práticas de manejo, contribuir substancialmente para o aumento da produtividade e para a melhoria das características qualitativas do fruto, objetivo que pode ser alcançado, em parte, por meio do acesso a informações relativas à herança das principais características agronômicas e da exploração da variabilidade genética.

A análise de regressão é uma das técnicas estatísticas mais utilizadas para pesquisar e modelar a relação existente entre duas ou mais variáveis, a aplicação da mesma tem-se destacado em diferentes áreas como por exemplo: industrial, econômica, biológica, química entre outras. Os modelos de regressão utilizados são classificados em duas categorias mais gerais: os modelos de regressão lineares e os não lineares. A diferença desses dois modelos está na forma como os parâmetros aparecem nos mesmos, se os parâmetros aparecem de forma linear na equação matemática que determina a relação entre as variáveis, ele será classificado na categoria dos lineares, caso contrário na dos não lineares.

Os estudos de curvas de crescimento são importantes uma vez que podem ser realizados em situações em que a informação contida numa sequência de pontos tamanho-tempo, do crescimento de uma espécie seja reduzida e sintetizada num conjunto de parâmetros que tenham interpretação prática e preditiva (PRADO, 2011). Alguns requisitos devem ser satisfeitos para que uma função de crescimento seja descritiva, dentre os quais a interpretação biológica dos parâmetros, um ajuste com pequenos desvios e o grau de

dificuldade do ajuste.

O objetivo deste trabalho foi obter estimativas para os parâmetros dos modelos não lineares de curva de crescimento Logístico, Gompertz e Von Bertalanffy, aplicado nos dados proveniente da cultura da mamoeiro, compará-los entre si, obtendo-se assim o modelo que melhor se ajusta ao padrão de crescimento.

## 2 Fundamentação Teórica

Esta seção tem como objetivo maior retratar de forma clara e pontual os aspectos históricos referentes aos modelos de regressão não linear, bem como estudar curvas de crescimento, utilizando-se de livros, artigos, teses e dissertações, com o objetivo de se obter informações precisas a cerca do assunto abordado.

### 2.1 Marco Histórico

O termo regressão foi introduzido por Francis Galton (1822-1911). Em seu estudo sobre altura de pais e filhos, observou que a altura dos filhos de pais com altura extrema tendia a regredir na direção da altura média (dos pais). Dando continuidade aos estudos de Galton, Karl Pearson coletou mais de mil registros das alturas de indivíduos pertencentes a grupos de famílias altas e de famílias baixas e verificou a Lei de *Regressão Universal*.

Atualmente a análise de regressão é uma das técnicas estatísticas mais utilizadas para pesquisar e modelar a relação existente entre duas ou mais variáveis. Por meio desta técnica é possível avaliar a existência e o grau de dependência entre as variáveis aleatórias analisadas. Esse grau de dependência é possível ser verificado por meio da análise de correlação, que consiste na medição do grau ou intensidade de associação entre duas variáveis. Quando se pode demonstrar que a variação de uma variável está de algum modo associada com a variação da outra, então podemos dizer que as duas variáveis estão correlacionadas.

A aplicação da análise de regressão tem-se verificado em diferentes áreas, tais como industrial, econômica, biológica, química entre outras. O modelo de regressão explica a existência da relação funcional de uma ou mais variáveis de interesse em função de outras variáveis explicativas, sendo a variável resposta a principal variável em estudo no modelo de regressão. Em geral, estima-se uma reta de regressão e, a partir desta reta, são feitas possíveis previsões sobre o comportamento da variável de interesse, para isto é necessário estimar os parâmetros do modelo de regressão. Este estudo é feito por meio de uma função matemática, onde se têm as variáveis explanatórias representadas pelos  $x_i$ , ou seja, variáveis que explicam o comportamento da variável resposta, representada por

$y$  (HOFFMANN, 2006).

Outra classe de modelos de regressão são os não lineares que é uma forma de análise observacional 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. Os dados são ajustados geralmente pelo método dos mínimos quadrados ou por algum método de aproximações sucessivas. As principais técnicas desenvolvidas para os modelos de regressão não linear no início da década de 70 se restringiam à suposição de normalidade para a variável resposta.

## 2.2 Modelos de regressão linear

Na regressão linear considera-se que a relação da resposta às variáveis é uma função linear de alguns parâmetros. Os modelos de regressão que não são uma função linear dos parâmetros se chamam modelos de regressão não linear. Sendo uma das primeiras formas de análise regressiva a ser estudada rigorosamente, e usada extensamente em aplicações práticas. Isso acontece porque modelos que dependem de forma linear dos seus parâmetros desconhecidos, são mais fáceis de ajustar que os modelos não lineares aos seus parâmetros, isso porque as propriedades estatísticas dos estimadores resultantes são fáceis de determinar.

Ajustar modelos de regressão linear é um dos métodos estatísticos utilizados quando se quer estudar o comportamento de variáveis numéricas e os efeitos lineares produzidos por elas, sendo representada da seguinte forma:

$$Y_i = \theta_0 + \theta_1 X_i + \varepsilon_i, \quad i = 1, \dots, n. \quad (2.1)$$

em que  $X_i$  representa cada observação da variável explicativa  $X$ ;  $\theta_0$  representa o coeficiente linear da reta, ou seja, representa o ponto inicial para a variável  $Y$ ;  $\theta_1$  representa o coeficiente angular da reta, ou seja, o grau que a reta faz com o eixo  $X$ , e define também o quanto aumenta, ou diminui, o valor de  $Y$  em relação a  $X$ ; por fim,  $\varepsilon_i$ , sendo o erro associado a cada observação em relação a reta de regressão linear.

## 2.3 Modelos de regressão não linear

Os modelos não lineares têm uma base teórica, os parâmetros dos modelos fornecem um maior conhecimento sobre o fenômeno em estudo do que os modelos lineares e geral-

mente fornecem um bom ajuste, com menos parâmetros do que os modelos lineares. A transformação de um modelo não linear em um modelo linear nos parâmetros, se por um lado facilita o processo de ajuste, por outro implica em fazer suposições não realísticas sobre o termo dos erros (distribuição normal com variância constante); além disso, perde-se informação sobre os erros padrão dos parâmetros originais. Além disso, existem modelos que são intrinsecamente não lineares, isto é, não podem ser linearizados por transformação.

Embora usa-se variáveis contínuas como variáveis independentes, não há razão para que as variáveis independentes nos modelos não lineares sejam contínuas. Ao contrário, podemos fazer uso de variáveis dummy para indicar a presença ou ausência de um grupo, ou codificar diferenças entre indivíduos (dados de medidas repetidas).

Alguns modelos podem parecer não lineares nos parâmetros, mas são intrinsecamente lineares, pois, as devidas transformações, podem se tornar modelos de regressão lineares em novos parâmetros com base em Gujarati (2006) . Os modelos não lineares são linearizados para facilitar a obtenção das estimativas dos parâmetros (BATES e WATTS, 1988). Diz-se que um modelo de regressão é intrinsecamente não linear nos parâmetros, quando ele não é linear e nem intrinsecamente linear.

O modelo de regressão não linear pode ser escrito como:

$$Y_i = f(X_i, \boldsymbol{\theta}) + \varepsilon_i, \quad i = 1, \dots, n \quad (2.2)$$

em que  $Y_i$  é a variável resposta,  $f(\cdot)$  é uma função não linear contínua, com forma conhecida, do vetor de variáveis explicativas  $X_i$ , e dos parâmetros desconhecidos  $\boldsymbol{\theta}$ . Neste trabalho os erros aleatórios,  $\varepsilon_i$  são os erros experimentais não observável diretamente, considerados independentes e identicamente distribuídos (*i.i.d*), normais com média zero e variância constante  $\sigma^2$ , isto é,  $\varepsilon_i \sim N(0, \sigma^2 I)$ .

A função esperança ou função de regressão é dada por  $E(Y_i|x_i) = f(X_i, \boldsymbol{\theta})$ . De acordo com Bates e Watts (1988) para os modelos de regressão não linear, pelo menos uma das derivadas parciais da função  $f$  em relação aos parâmetros depende de pelo menos um dos parâmetros do modelo, ou seja, pode-se dizer que um modelo é não linear quando ele é não linear nos parâmetros .

### 2.3.1 Métodos de estimação

Nem sempre os modelos não lineares são expressos numa forma paramétrica adequada, que facilite a convergência rápida dos processos iterativos utilizados na estimação dos

seus parâmetros, sendo necessário procurar, em muitos casos, uma reparametrização mais apropriada. Rawlings, Pantula e Dickey (1998) relataram que para os modelos de regressão não lineares, o sistema de equações normais não é resolvido facilmente, sendo necessários métodos iterativos na obtenção dessas estimativas.

O método de Gauss-Newton, também conhecido como método da linearização, usa uma expansão em série de Taylor para aproximar o modelo de regressão não linear com termos lineares. Essas iterações geralmente conduzem a uma solução para o problema de regressão não linear. Uma característica interessante do método de Gauss-Newton ocorre quando a função esperança é linear.

Sejam  $y_1, \dots, y_n$  variáveis aleatórias independentes, e suponha que os dados a serem analisados sejam representados pelo vetor  $\mathbf{y} = (y_1, \dots, y_n)^T$ . A estimativa do vetor de parâmetros  $\boldsymbol{\theta}$  é calculada pelo método dos mínimos quadrados que, nesse caso, coincide com o método de máxima verossimilhança, pois o modelo tem respostas normais independentes com a mesma variância. A estimativa de  $\boldsymbol{\theta}$  é, então, calculada, minimizando-se a função quadrática

$$S(\boldsymbol{\theta}) = \sum_{i=1}^n \{y_i - \eta_i(\boldsymbol{\theta})\}^2.$$

Derivando-se  $S(\boldsymbol{\theta})$  em relação a  $\boldsymbol{\theta}$ , obtém-se

$$\frac{\partial S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 2 \sum_{i=1}^n \{y_i - \eta_i(\boldsymbol{\theta})\} \frac{\partial \eta}{\partial \boldsymbol{\theta}}.$$

A estimativa  $\hat{\boldsymbol{\theta}}$  do vetor de parâmetros  $\boldsymbol{\theta}$  é obtida igualando-se  $\frac{\partial S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$  a zero para  $i = 1, \dots, p$ . Em geral, as equações  $\frac{\partial S(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_r} = 0$  não são lineares e têm que ser resolvidas, numericamente, por processos iterativos.

A abordagem sugerida por Gauss Newton, foi utilizar uma aproximação linear à função esperança, para reiterativamente melhorar o  $\boldsymbol{\theta}$  como valor inicial para  $\hat{\boldsymbol{\theta}}$  e continuar melhorando as estimativas, até não existir mais mudança. Esse método iterativo usa uma expansão em série de Taylor para aproximar o modelo de regressão não linear com termos lineares e, então, aplica mínimos quadrados para estimar os parâmetros. Iterações desses passos geralmente conduzem a uma solução para o problema de regressão não linear. O processo é repetido um determinado número de vezes, até que o vetor de estimativas não se altere mais dentro de precisão pré-estipulada.

Alguns desses métodos introduzem modificações para obter uma convergência mais rápida como o caso do método de Marquardt (1963), bastante usado em investigação pes-

queira. Outros métodos usam o desenvolvimento de Taylor de segunda ordem (método de Newton-Raphson), procurando-se assim uma melhor aproximação. Outros ainda, combinam as duas modificações.

Dentre os muitos existentes, um critério para verificar a convergência poderia ser

$$\sum_{r=1}^p \left( \frac{\theta_r^{(m+1)} - \theta_r^{(m)}}{\theta_r^{(m)}} \right)^2 < \xi,$$

tomando-se para  $\xi$  um valor suficientemente pequeno. A convergência em geral, depende dos valores iniciais para os parâmetros do vetor  $\beta$ . Isso pode evitar que problemas relacionados com a estrutura paramétrica do modelo, tais como, não linearidade acentuada, prejudiquem a convergência do processo iterativo.

## 2.4 Seleção e qualidade do ajuste

Para saber se o modelo selecionado explica razoavelmente bem o comportamento da variável resposta deve-se testar a qualidade do ajuste do modelo, que é dada por estatísticas que medem a proporção da variação na variável resposta que é explicada pelo modelo.

Sabe-se que a seleção de um modelo deve basear-se em um ajuste adequado aos dados amostrais e também na sua coerência com o evento biológico estudado. Selecionar o melhor modelo significa selecionar a melhor estrutura para as médias (parte fixa) e a melhor estrutura de covariâncias. A seleção e a avaliação de um modelo são tarefas simples, mas de extrema importância, pois determinarão o modelo que melhor prediz a variável dependente de forma parcimoniosa, isto é, com um menor número de parâmetros. Para isso são utilizadas ferramentas analíticas como testes de hipóteses e critérios de informação, como descritos a seguir.

### 2.4.1 Coeficiente de determinação

O coeficiente de determinação, conhecido por  $R^2$ , é o quadrado do coeficiente de correlação momento-produto de Pearson. Triola (1999) explica que existe correlação entre duas variáveis quando uma delas está relacionada com a outra de alguma forma. Relação, conforme Kerlinger (1990), “é um ir junto de duas variáveis”, é o que as duas variáveis têm em comum.

Segundo Kennedy (2008), supõe-se que o coeficiente de determinação representa a proporção da variação da variável dependente que é explicada pela variação da variável independente. Para o modelo não linear não é tão simples a aplicação desta definição, no entanto, um dos critérios para sua definição requer a presença do intercepto no modelo, na qual, nem sempre está presente no modelo de regressão. Deste modo, consideraremos o  $R_a^2$  como sendo uma medida próxima ao  $R^2$ .

Assim segue-se que:

$$R_a^2 = 1 - \frac{SQE(\hat{\theta})}{\|y - \hat{y}\|^2}$$

em que,  $SQE(\hat{\theta})$  é a soma dos quadrados dos resíduos avaliados em  $\hat{\theta}$ ,  $y$  é o valor observado e  $\hat{y}$  indica o valor predito. Entretanto, apenas o valor do  $R_a^2$  não serve como um critério adequado para verificação de ajuste dos modelos, em muitos casos o ajuste de modelos não lineares, é comum a obtenção de  $R_a^2$  assintóticos altos e similares (REZENDE et al., 2007).

### 2.4.2 Critério de Informação de Akaike (AIC)

A inferência estatística propôs em estimar o modelo verdadeiro para os dados. O modelo com menor valor de AIC é aquele que minimiza a distância de Kulbach-Leibler, por consequência, pode ser considerado como aquele mais próximo do modelo real (hipotético), com base em Akaike (1973) citado por Burnham e Anderson (2002). Sendo assim, a partir de uma série de modelos concorrentes escolhe-se aquele que apresentar o menor valor de AIC.

O conceito de entropia da informação é baseado no critério de informação de Akaike (AIC), esse critério oferece uma medida relativa da informação perdida quando um modelo é usado para descrever a realidade. As principais vantagens do AIC residem na possibilidade de comparar modelos com diferentes famílias de distribuições (modelos não aninhados) e por não precisar de mais inferências sobre o modelo para corroborar em seu resultado (BURNHAM, ANDERSON, 2004). Ademais, é de fácil avaliação, pois o melhor modelo é o que possuir o menor valor do AIC. De acordo com Burnham e Anderson, 2002, este critério tem sido amplamente usado em diferentes áreas do conhecimento, como biometria, econometria, pesquisa operacional, séries temporais, regressão espacial, e entre outras áreas.

O AIC apresenta algumas lacunas apesar de suas vantagens, como por exemplo: a im-

possibilidade de comparar modelos com variáveis respostas em transformações diferentes (logaritmo, exponencial, linear, etc) e com dados desbalanceados. O de maior importância é a impossibilidade de realizar um teste de hipótese, portanto o AIC é um critério de informação e não um teste estatístico onde seria possível observar se o AIC de um modelo é significativamente menor do que outro modelo.

Akaike (1974), mostrou que o viés é dado assintoticamente por  $p$ , em que  $p$  é o número de parâmetros a serem estimados no modelo, sendo assim definiu seu critério de informação como:

$$AIC = -2\log\ell(\hat{\theta}) + 2p$$

em que,  $p$  é o número de parâmetros e  $\log\ell(\hat{\theta})$  é o logaritmo natural da função de verossimilhança avaliada em  $\hat{\theta}$  e pode ser interpretado como a medida do ajuste do modelo.

### 2.4.3 Critério de informação (BIC)

Na metodologia Bayesiana têm-se a possibilidade de ser usada na identificação de modelos, onde foi mencionada por Akaike, sendo que alguns estatísticos o apresentam como o Critério de Informação Bayesiana. O valor mínimo do BIC estimado para um determinado modelo de um subsistema é utilizado para selecionar modelos ótimos em diversos campos da estatística.

O critério de informação bayesiano BIC (Bayesian Information Criterion) é um critério de seleção de modelo proposto por Schwarz em 1978 e modificado por Carlin e Louis em 2000, para ser aplicado considerando a densidade a posteriori dos parâmetros do modelo ajustado. O mesmo pondera entre a função log-verossimilhança maximizada e o número de parâmetros do modelo. O modelo será escrito da seguinte forma:

$$BIC = -2\log L + (p + 1)\log(n)$$

em que,  $p$  é a dimensão do vetor de parâmetros e  $n$  é o tamanho da amostra em que é o valor  $l$ .

## 2.5 Análise de resíduos

A análise dos resíduos tem uma importância fundamental na variação da qualidade dos ajustes de modelos. Basicamente, essa análise fornece evidências sobre possíveis violações nas suposições do modelo, tais como a de normalidade, homocedasticidade, e quando for

o caso ainda fornece indícios da falta de ajuste do modelo proposto. Além dos testes de significância e adequação já apresentados, a análise de resíduos vem completar o elenco de procedimentos que devem ser realizados após o ajuste de qualquer modelo.

Na análise de resíduos, verificam-se as pressuposições sob o modelo em estudo. Uma ou mais observações são ditas discrepantes (*outliers*) se seus resíduos são muito grandes (em valor absoluto) em relação aos demais (DRAPER e SMITH, 1998); (PRUDENTE, 2009). Já os pontos influentes são observações que, embora não apresentem resíduos grandes, podem alterar significativamente as estimativas dos parâmetros do modelo escolhido. Assim, para detectar a presença de pontos influentes ou de *outliers*, podem ser aplicadas técnicas com base nos resíduos ordinários e resíduos padronizados.

O resíduo para a  $i$ -ésima observação é obtido por meio da função  $r_i = y_i - \hat{\mu}_i$ , que mensura a diferença entre o valor observado e o valor ajustado, chamado de resíduo ordinário da variável resposta do modelo. Então pode-se afirmar que modelos bem ajustados deverão apresentar pequenos resíduos e caso contrário modelos mal ajustados apresentarão grandes resíduos. De acordo com Cordeiro e Neto (2006), os resíduos ordinários não são muito informativos, pois não apresentam variância constante  $Var(r_i) = \sigma^2(1 - h_{ii})$ , pois depende dos valores de  $h_{ii}$ .

Para garantir a independência do numerador e denominador, faz-se a padronização dos resíduos, obtendo o chamado o resíduo estudentizado, definido como

$$z_i^* = \frac{e_i}{\sqrt{\hat{\sigma}_i^2(1-h_{ii})}}, \quad i = 1, 2, \dots, n$$

Prudente (2009), afirmou que os resíduos projetados são melhores que os resíduos ordinários e padronizados, no que se refere a análise e diagnósticos dos modelos não lineares, pois os resíduos projetados superam os resíduos ordinários em diversos aspectos e muitas das técnicas de diagnóstico utilizadas na regressão linear são, também, aplicáveis aos mesmos.

### 2.5.1 Distância de Cook

A disposição dos pontos no espaço das variáveis de regressão é importante na determinação das propriedades do modelo. Em particular, as observações remotas podem, potencialmente, exercer o efeito de uma “alavanca” nas estimativas dos parâmetros, nos valores previstos e nas estatísticas utilizadas.

A matriz jacobiana da transformação  $\tilde{\mathbf{X}} = F(\boldsymbol{\theta}) = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$  é útil na identi-

ficação dessas observações influentes. Os valores dos elementos  $h_{ij}$  da matriz  $\tilde{\mathbf{X}}$  podem ser interpretados como a intensidade da alavancagem exercida pelos valores observados ( $y_i$ ) sobre os valores ajustados ( $\hat{y}_i$ ). Portanto, a inspeção dos elementos da matriz  $\tilde{\mathbf{X}}$  pode revelar pontos potencialmente influentes devido à sua localização no espaço das variáveis independentes.

Os elementos da diagonal da matriz  $\tilde{\mathbf{X}}$  são denotados por  $h_{ii}$ . Pode-se demonstrar que  $\sum_{i=1}^N h_{ii} = P$ , em que  $P$  é o número de parâmetros e  $N$  é o tamanho da amostra. Portanto, a magnitude média dos elementos  $h_{ii}$  da diagonal da matriz  $\tilde{\mathbf{X}}$  é  $P/N$ . Assim, uma forma aproximada, mas eficaz, de verificar se a  $i$ -ésima observação  $y_i$  é um ponto de grande alavancagem, é verificar se o elemento  $h_{ii}$  da diagonal da matriz  $\tilde{\mathbf{X}}$  é maior que  $2P/N$ , ou seja, se o valor de  $h_{ii}$  associado à  $i$ -ésima observação  $y_i$  é duas vezes maior que a média de todos os  $h_{ii}$  da diagonal da matriz  $\tilde{\mathbf{X}}$ . Caso seja constatado que um ou mais valores de  $h_{ii}$  são maiores que  $2P/N$ , então pode-se concluir que estas observações específicas são pontos de grande alavancagem (MYERS, et al. 2002).

Cook (1977) sugeriu a utilização de uma medida do quadrado da distância entre a estimativa dos mínimos quadrados  $\hat{\boldsymbol{\theta}}$  e uma estimativa obtida excluindo-se o  $i$ -ésimo ponto,  $\hat{\boldsymbol{\theta}}_{(i)}$ . Em geral, a medida desta distância, denominada distância de Cook, pode ser expressa como sendo:

$$D_i = \frac{\left(\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}\right)^\top (\mathbf{X}^\top \mathbf{X}) \left(\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}\right)}{P \times MS_E},$$

em que  $MS_E$  é a média dos quadrados do erro. Portanto, pode-se interpretar  $D_i$ , também, como a medida da mudança no vetor dos valores ajustados quando não usa-se a observação  $i$  para estimar  $\boldsymbol{\theta}$ .

A versão escalar da expressão para a estatística  $D_i$  proposta por Atkinson (1985) é,

$$D_i = \frac{r_i^2 h_{ii}}{P(1 - h_{ii})^2}$$

sendo  $r_i$  o  $i$ -ésimo resíduo ordinário estudentizado.

A distância de Cook prevê uma ordenação das observações em termos da sua influência sobre o vetor das estimativas dos coeficientes. A intenção não é aplicar um teste formal, e sim fornecer uma ajuda para detectar as observações influentes. Cook e Weisberg (1982) afirmam que é conveniente analisar casos em que  $D_i > 0,5$  e é sempre importante analisar casos em que  $D_i > 1$ . Esta análise consiste em verificar se a observação é realmente influente ou se é consequência de um modelo inadequado.

## 2.6 Curvas de crescimento

A análise dos dados sobre crescimento é importante em muitas áreas de pesquisa segundo Seber e Wild (2003). Na biologia, o objetivo está no estudo do crescimento de animais e plantas na tentativa de compreender seus mecanismos essenciais. Na química, o interesse está na formulação do produto de uma reação química ao longo do tempo. Na agricultura, o interesse está em conhecer como a produção cresce, quão rápido ela cresce, e como esses fatores se adequa a condições ou tratamentos ambientais, sendo assim, fica óbvia as vantagens econômicas e administrativas. Na medicina, o interesse está no estudo do crescimento infantil, bem como no crescimento de tumores e os efeitos de tratamentos sobre tais crescimentos. Os cientistas sociais têm grande interesse em estudar o crescimento da população, no fornecimento de alimentos e na demanda de energia.

O estudo do crescimento de vegetais possui uma larga aplicação na pesquisa agropecuária, pois possibilita ao pesquisador comparar o comportamento de uma planta ou de suas partes de interesse, diante das diversas situações experimentais empregadas. Observa-se que os modelos não-lineares proporcionam vantagens em relação aos modelos lineares, os quais podem auxiliar o pesquisador no processo de tomada de decisão por apresentarem parâmetros com interpretação biológica. Dentre essas funções não-lineares, pode-se citar os modelos: Logístico (NELDER, 1961), Gompertz (WINSOR, 1932), e Von Bertalanffy (Von BERTALANFFY, 1957).

É de total importância estudar curvas de crescimento, uma vez que podem ser realizadas em situações em que a informação contida numa sequência de pontos tamanho-tempo, do crescimento de uma espécie seja reduzida e sintetizada num conjunto de parâmetros que tenham interpretação prática e preditiva (PRADO, 2011). Para que uma função de crescimento seja descritiva é preciso que alguns requisitos sejam satisfeitos, em uma relação peso-idade, dentre os quais a interpretação biológica dos parâmetros (confiabilidade), um ajuste com pequenos desvios (precisão) e o grau de dificuldade do ajuste (operacionalidade).

Nas curvas de crescimento a escolha do modelo que melhor as descreve é baseada na qualidade de seu ajustamento aos dados, na interpretação biológica dos parâmetros e nas dificuldades computacionais para obtenção do ajustamento. Tendo sido identificado o modelo que melhor se ajustou ao padrão de crescimento, os parâmetros que o descrevem devem ser estudados, identificando-se os fatores ambientais que podem influenciá-los,

possibilitando ajustes no padrão de crescimento.

As formas mais gerais dos modelos são:

i) Logístico

$$y_t = \alpha(1 + \beta e^{-kt})^{-m} + \varepsilon \quad (2.3)$$

ii) Gompertz

$$y_t = \alpha e^{-\beta e^{-kt}} + \varepsilon \quad (2.4)$$

iii) Von Bertalanffy

$$y_t = \alpha(1 - \beta e^{-kt})^3 + \varepsilon \quad (2.5)$$

em que,  $y_t$  é a resposta média,  $\alpha$  é o parâmetro que representa o valor assintótico da curva, isto é, o valor médio livre das variações sazonais,  $\beta$  é uma constante que está relacionada aos valores iniciais, indicando a proporção do crescimento assintótico a ser ganho,  $k$  é a taxa de crescimento da variável de interesse, determina a eficiência do crescimento,  $t$  é a tempo em dias,  $m$  representa o parâmetro que define a forma da curva nos modelos e consequentemente determina o ponto na qual passa de uma fase de crescimento inibitório, indicando o ponto que ele passa a crescer com menor eficiência.

## 3 Aplicação

Nesta seção encontram-se as principais metodologias que serviram de base para este trabalho, tanto no que se refere ao ajuste dos modelos, quanto na utilização das técnicas de diagnósticos.

### 3.1 Material e métodos

Os dados utilizados para esse trabalho, foram oriundos de dois ensaios de uniformidade com mamoeiro, em casa de vegetação, na área experimental da Embrapa Mandioca e Fruticultura Tropical, em Cruz das Almas-Bahia. A cidade está situada a 12°40 19 de Latitude Sul e 39°06 22 de Longitude Oeste de Greenwich, tendo 220 m de altitude. O clima é tropical quente e úmido, Aw a Am, segundo a classificação de Köppen, com temperatura média anual de 24,5° C e umidade relativa de 80% (ALMEIDA, 1999). Os ensaios foram instalados no mês de dezembro de 2004, um com a cultivar Golden (material fixado), do grupo Solo, e o outro, com o híbrido Tainung N° 1, do grupo Formosa.

As avaliações ocorreram de janeiro a maio de 2005, período entre o desbaste das plantas e o aparecimento da primeira flor funcional. Os dados foram coletados semanalmente até, aproximadamente, três meses após o plantio, período esse em que a muda se encontra apta para ir ao campo (18, 26, 32, 39 e 54 dias). Após este período, as avaliações passaram a ser mensais devido ao crescimento se tornar mais lento (84, 114 e 144 dias). As seguintes variáveis foram avaliadas: altura de planta em centímetros e o diâmetro de caule em centímetros.

Tabela 1: Médias da altura e diâmetro das plantas de mamoeiro ao longo dos dias

Dias	Altura (cm)	Diâmetro (cm)
18	14,90	0,17
26	19,91	0,32
32	28,31	0,53
39	37,24	0,69
54	71,44	1,46
84	97,64	2,06
114	106,26	2,25
144	116,48	2,53

Utilizou-se os modelos de regressão não linear Logístico (2.3), Gompertz (2.4) e Von Bertalanffy (2.5) afim de inferir sobre a relação existente entre a altura e o diâmetro das plantas do mamoeiro aos longo dos dias e comparar qual modelo melhor representa essa relação.

Para a estimação dos parâmetros do modelo não linear foi utilizada a técnica dos mínimos quadrados ordinários, com o uso do método iterativo de Gauss-Newton (BATES; WATTS, 1988). Dentre as estatísticas fornecidas pelo procedimento de estimação, foram obtidos intervalos de confiança para os parâmetros, bem como o coeficiente de determinação assintótico ( $R_a^2$ ). Para a comparação dos modelos foi utilizado os critérios de informação AIC e BIC. Após obtenção das estatísticas citadas acima, procedeu-se a análise gráfica do modelo selecionado para verificar a qualidade do ajuste. Os ajustes inferenciais foram realizados com o uso do *software* R versão 3.1.1 (R DEVELOPMENT CORE TEAM, 2012)

## 3.2 Resultados e discussão

Tal como em regressão linear, é extremamente importante, em regressão não linear, ter à disposição de ferramentas que permitam uma avaliação do ajuste do modelo em uma dada aplicação.

Nas Tabelas 2 e 3 são apresentadas as estimativas encontradas para os parâmetros dos modelos não lineares Logístico (2.3), o Gompertz (2.4) e o Von Bertalanffy (2.5), conforme a metodologia estabelecida, bem como os valores do erro padrão da estimativa, o valor-p e os intervalos de confiança a 95%. Percebe-se que quase todos os parâmetros da função foram significativos, pois se encontram dentro do intervalo de confiança assintótico de 95% e estes, por sua vez, não incluem a constante zero.

Na Figura 1 são apresentados os ajustes de cada modelo em relação aos dados, verificando-se a relação existente entre para a altura das plantas (a) e o diâmetro do caule (b) da cultura do mamão ao longo dos dias.

Visto que realizadas as análises univariadas, empregou-se critérios para verificar a qualidade do ajuste, além de comparar os modelos. Os coeficientes de determinação obtidos foram altos para os modelos em que houve convergência, o que indica que todos os modelos apresentaram bom desempenho. Segundo Freitas (2005), em estudos com curvas de crescimento animal os modelos não-lineares entre os quais o de Gompertz produziu ajustes com  $R^2$  acima de 92%, sendo considerados como um bom ajuste. Foram observados

Tabela 2: Estimativas dos parâmetros para o modelo Logístico, Gompertz e Von Bertalanffy, erro padrão da estimativa (E.P.E.), valor- $p$  para o teste  $t$  e intervalos de confiança (IC) de 95% para altura em função dos dias

Modelos	Parâmetros	Estimativas	E.P.E.	Valor- $p$	IC (95%)
Logístico	$\theta_1$	111,0000	3,0750	<0,0001	[103,2261; 120,1490]
	$\theta_2$	3,1740	0,2927	0,0001	[2,4683; 4,1589]
	$\theta_3$	0,0649	0,0074	0,0003	[0,0479; 0,0918]
Gompertz	$\theta_1$	115,5000	3,8580	<0,0001	[106,3708; 127,3815]
	$\theta_2$	4,9520	0,7906	0,0015	[3,3982; 8,2624]
	$\theta_3$	0,0406	0,0047	0,0003	[0,0293; 0,0551]
Von Bertalanffy	$\theta_1$	118,4000	5,1350	<0,0001	[107,0313; 135,4703]
	$\theta_2$	1,0280	0,1457	0,0009	[0,7424; 1,6310]
	$\theta_3$	0,0326	0,0045	0,0008	[0,0219; 0,0463]

Tabela 3: Estimativas dos parâmetros para o modelo Logístico, Gompertz e Von Bertalanffy, erro padrão da estimativa (E.P.E.), valor- $p$  para o teste  $t$  e intervalos de confiança (IC) de 95% para diâmetro em função dos dias

Modelos	Parâmetros	Estimativas	E.P.E.	Valor- $p$	IC (95%)
Logístico	$\theta_1$	2,3613	0,0806	<0,0001	[2,15587; 2,6136]
	$\theta_2$	3,6559	0,4308	0,0004	[2,6606; 5,2304]
	$\theta_3$	0,0736	0,0104	0,0009	[0,0483; 0,1118]
Gompertz	$\theta_1$	2,4607	0,0815	<0,0001	[2,2638; 2,7159]
	$\theta_2$	6,2659	1,1465	0,0028	[4,0831; 11,5516]
	$\theta_3$	0,0435	0,0050	0,0003	[0,0313; 0,0600]
Von Bertalanffy	$\theta_1$	2,5250	0,0986	<,0001	[2,2975; 2,8471]
	$\theta_2$	1,2275	0,1807	0,0011	[0,8727; 2,0415]
	$\theta_3$	0,0347	0,0044	0,0005	[0,0243; 0,0490]

na Tabela 4 que o melhor ajuste foi o do modelo Gompertz, o qual apresentou os menores valores para o critério de informação AIC, BIC, e o maior valor para  $R_a^2$ , conforme os dados ajustados. O AIC não é uma prova a cerca do modelo, mas uma ferramenta para a seleção de modelos, não é um teste de hipóteses, e como tal não há significância e nem valor de probabilidade associado (valor- $p$ ) (BURNHAM; ANDERSON, 2002).

Com base em Floriano et al., (2006) O uso do critério de Akaike (AIC) mostrou-se adequado como critério de seleção de modelos para representar uma série temporal de dados de altura de árvores. Vários autores, em diversas áreas do conhecimento, baseiam-se nos critérios citados acima para selecionar o melhor modelo. De acordo com Peixoto (2013), quando se ajusta um modelo a um conjunto de dados é imprescindível que as estimativas obtidas a partir do modelo proposto sejam resistentes a pequenas perturbações nos dados ou no modelo.

Se o modelo ajustado não apresentar uma boa descrição dos dados observados, esse

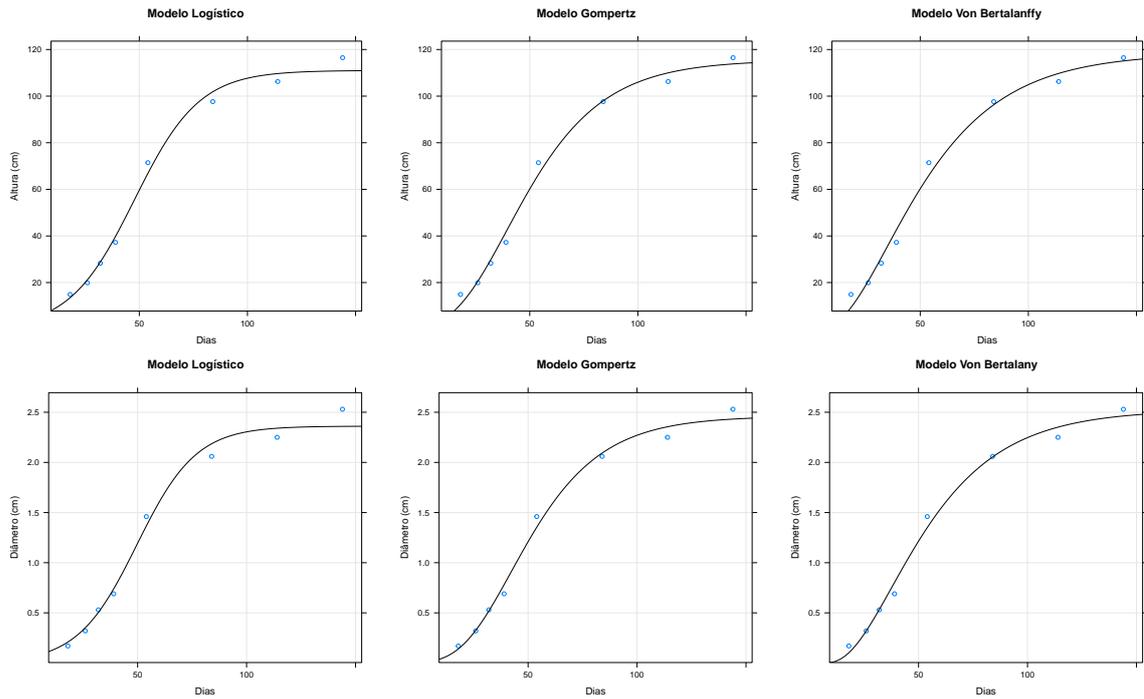


Figura 1: Ajuste dos modelos logístico, gompertz e Von Bertalanffy para a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro aos longo dos dias.

Tabela 4: Coeficiente de determinação ( $R_a^2$ ) e valores de AIC e BIC para os modelos ajustados referente a altura e o diâmetro do mamoeiro.

Variável	Modelo	$R^2$	AIC	BIC
Altura	Logístico	0,9891	50,2981	50,6159
	Gompertz	0,9910	49,9089	50,2267
	Von Bertalanffy	0,9874	51,8337	52,1515
Diâmetro	Logístico	0,9841	-7,4742	-7,1564
	Gompertz	0,9905	-11,2261	-10,9083
	Von Bertalanffy	0,9902	-10,7999	-10,4822

pode conduzir a inferências incorretas. Desda forma, é essencial analisar as suposições dos erros, a presença de possíveis pontos influentes e a validação do ajuste do modelo adotado. Para essa etapa, métodos de análise de sensibilidade e resíduos são ferramentas essenciais para detectar anomalias do modelo ajustado. Após a escolha do melhor modelo ajustado aos dados, deu-se a continuidade com a análise gráfica referentes a homocedasticidade e normalidade dos resíduos, apresentados. Ao analisarmos os gráficos da Figura 2 observa-se que as variáveis altura das plantas e diâmetro do caule aparentemente seguem uma certa normalidade. Na Figura 3 pode-se abservar a princípio que não há nenhuma suspeita da presença de pontos discrepantes nos dados observados, pois nenhum dos valores está fora do intervalo  $(-2, 2)$ , com base nos gráficos dos resíduos estudentizados.

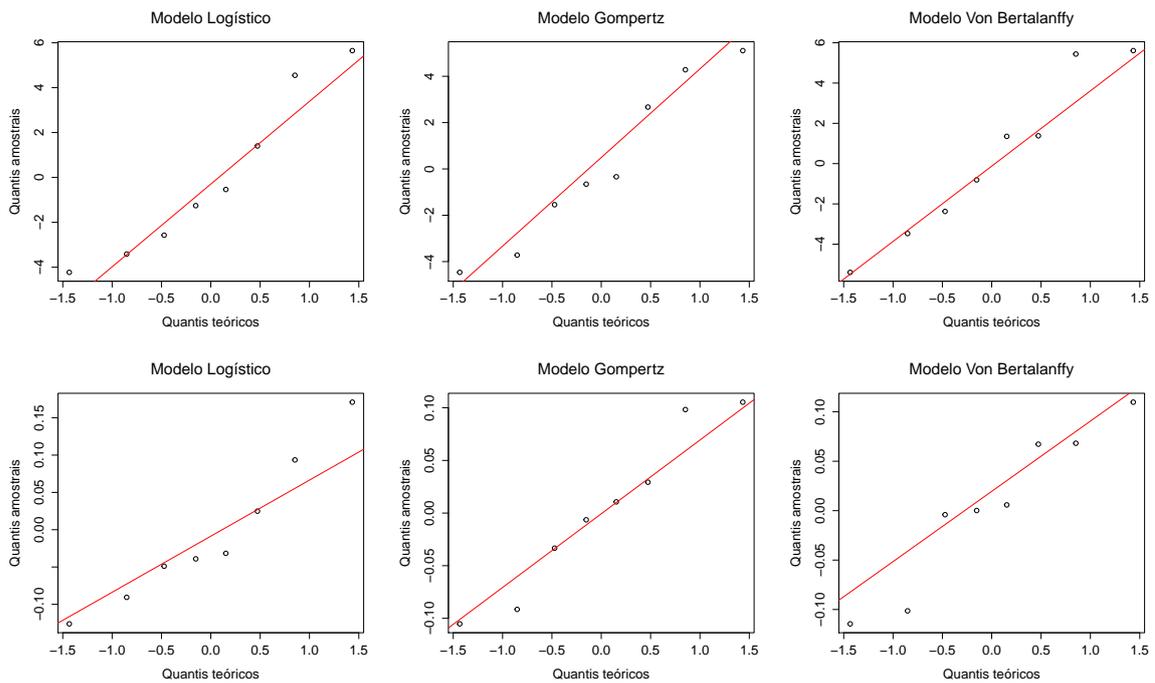


Figura 2: Gráfico dos quantis normais para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro ao longo dos dias para os modelos logístico, gompertz e von bertalanffy.

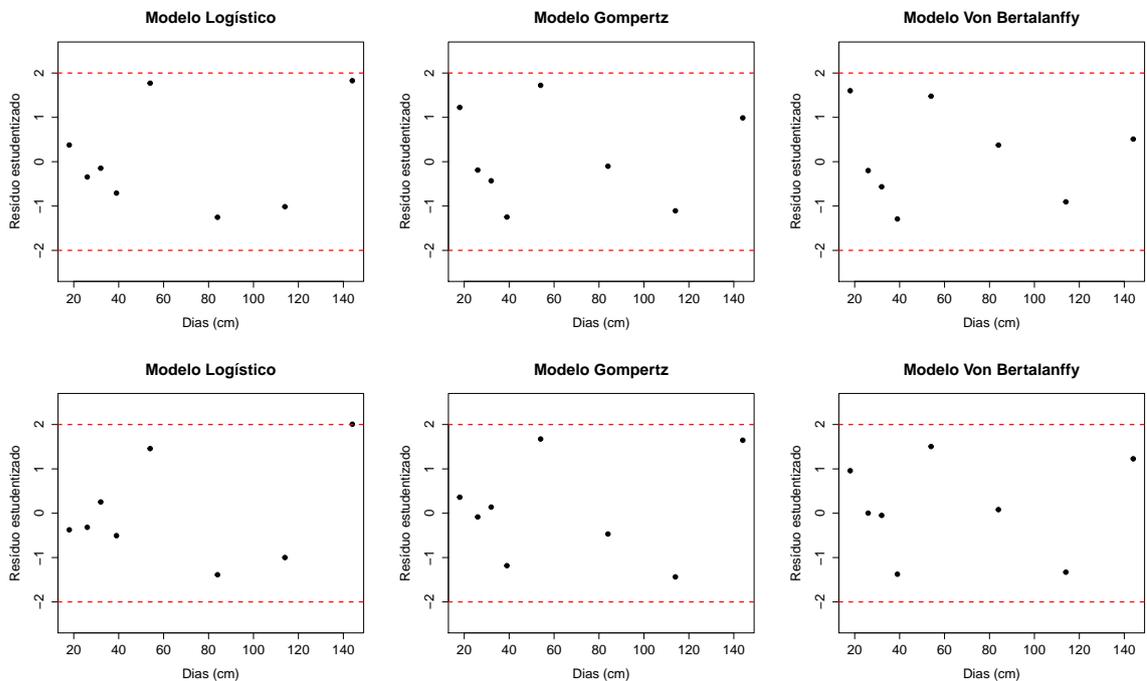


Figura 3: Gráfico dos resíduos estudentizados para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro ao longo dos dias para os modelos logístico, gompertz e von bertalanffy.

Na Figura 4 as observações em destaque, tanto para a altura da planta, como para o diâmetro do caule, podem ser consideradas observações influentes sobre todos os valores ajustados, pois os valores encontrados pela distância de Cook,  $D_i$ , foram maiores que a unidade absoluta ( $< 1$  em módulo) na maiorias das situações apresentadas. Deve-se realizar o ajuste sem estas observações, e verificar se ocorre alteração no ajuste do modelo e se a diferença entre os  $\theta$ 's, não foi maior do que 1 (em módulo), fato que não justifica a remoção das mesmas.

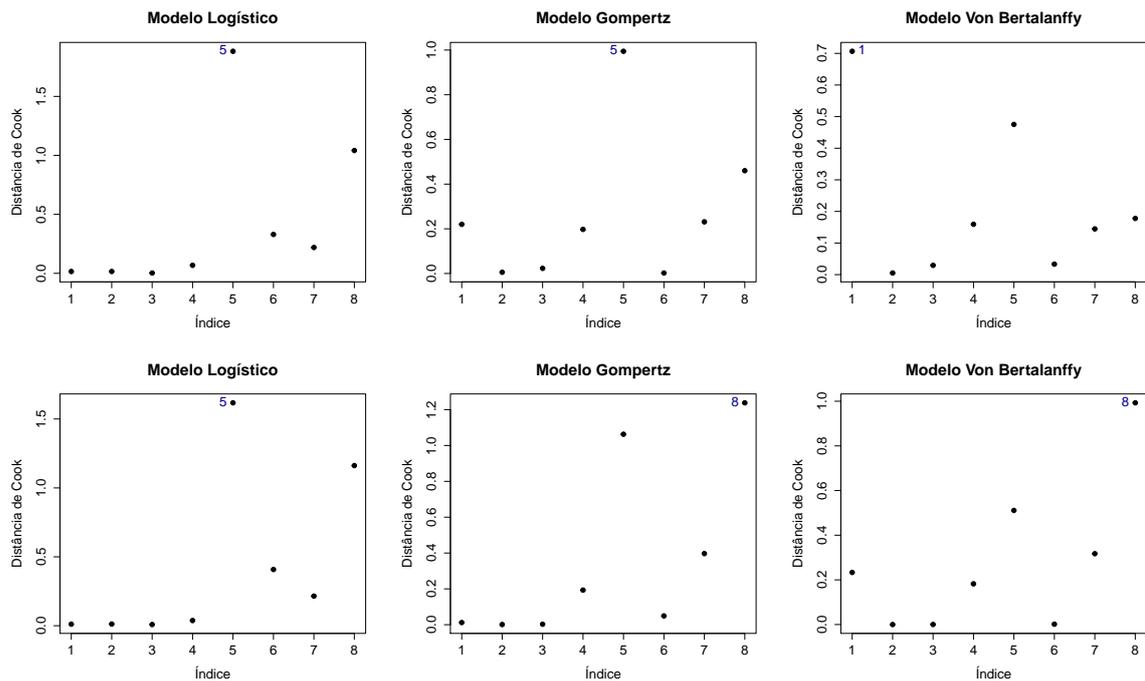


Figura 4: Gráfico das distâncias de Cook para a relação existente entre a altura (a) das plantas e o diâmetro (b) do caule do mamoeiro para os modelos logístico, gompertz e von bertalanffy.

Cordeiro, Prudente e Demétrio (2009), utilizaram a técnica de diagnóstico baseada no gráfico da distância de Cook a fim de diagnosticar o ajuste do modelo exponencial utilizado para explicar a fração de cloro disponível em um produto manufaturado.

## 4 Conclusão

O presente trabalho teve como principal objetivo verificar a importância dos ajustes de modelos não lineares apresentado através das curvas de crescimento, por serem simples e de fácil interpretação, têm-se mostrado adequado para descrever a espécie em estudo. A utilização de modelos não lineares foi adequada, pois estes apresentam parâmetros com interpretação biológica, sendo importantes na descrição e estudo do crescimento de frutos. Os modelos ajustados, Logístico, Gompertz e o Von Bertalanffy apresentaram-se apropriado para descrever curvas de crescimento para os dados apresentados. Sendo assim terminadas as análises dos dados, através dos avaliadores da qualidade de ajuste o modelo Gompertz apresentou melhor ajuste, sendo o mais indicado para representar curvas de crescimento médio entre as variáveis estudadas altura da planta em centímetros e o diâmetro do caule em centímetros aplicada na cultura do mamoeiro.

## 5 Referências Bibliográficas

- ALMEIDA, O. A. **Informações meteorológicas do CNP: mandioca e fruticultura tropical**. Cruz das Almas: Embrapa- CNPMF, 1999. 35 p. (Documentos, 34).
- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Boston, v.19, n.6, p.716-723, 1974.
- ATKINSON, A.C. **Plots, transformations, and regression**. New York: Oxford University Press, 1985. 280p.
- AMARAL, M. T. R. de. **Abordagem bayesiana para curva de crescimento com restrições nos parâmetros**. 2008, 111 p. Dissertação (Mestrado). RECIFE-PE.
- BARROS, L. C.; BASSANEZI, R. C. **Tópicos de Lógica Fuzzy e Biomatemática. Coleção IMECC**. Textos didáticos, v.5. Campinas-SP, 2006.
- BATES, D. M.; WATTS, D. G. Relative curvature measures of nonlinearity (With discussion). **Journal of the Royal Statistical Society**, Ser. B, v.42, n. 1, p. 1-25, 1980.
- BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: Wiley series in probability e mathematical statistics, 1988. 365p.
- BURNHAM, K. P.; ANDERSON, D. R. **Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach**. 2d ed. New York: Springer-Verlag. 2002.
- COOK, R.D. Detection of influential observations in linear regression. **Technometrics**, Alexandria, v.19, p.15-18, 1977.
- COOK, R.D.; WEISBERG, S. **Residuals and influence in regression**. New York: Chapman & Hall, 1982. 280p.
- COOK, R. D.; TSAI, C. L.; WEI, B. C. Bias in nonlinear regression. **Biometrika**, London, v. 73, n. 3, p. 615-623, Dec. 1986.
- CORDEIRO, G.M.; PRUDENTE, A.A.; DEMETRIO, C.G.B.; Uma revisão dos modelos normais não lineares. **Revista Brasileira de Biometria**. São Paulo, v.27, n.3, p.

360-393, 2009.

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

EMILIANO, P. C. de. Fundamentos e aplicações dos critérios de informações: Akaike e Bayseano. 2009, 92p . Dissertação (Mestrado). Lavras - Minas Gerais.

Food and Agriculture Organization of the United Nations - FAO. Agricultural data. <http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567ancor>. 21 Nov. 2014.

FERREIRA, M.C.C.S **Modelos de Regressão: uma aplicação em Medicina Dentária**. 2013 1p. Dissertação (Mestrado em Estatística, Matemática e Computação)-Lisboa, LIS.

FLORIANO, E. P. et al. Ajuste e seleção de modelos tradicionais para série temporal de dados de altura de árvores. **Ciência Florestal**, Santa Maria, v.16, n. 2, p.177-199, 2006.

FREITAS, A.R. Curvas de Crescimento na Produção Animal. **Revista Brasileira de Zootecnia**, v.34, n.3, p. 786-795, 2005

GUJARATI, D. Econometria básica. 4. ed. São Paulo: Campus, 2006. 806p.

HOFFMANN, R.; VIEIRA, S. **Análise de Regressão: Uma Introdução à Econometria**. São Paulo: Hucitec, 2006.

MYERS, R.H.; MONTGOMERY, D.C. VINING, G.H. **Generalized Linear Models with applications in Engineering and the Sciences**. John Wiley & Sons, New York, 2002, 544p.

KENNEDY, Peter. A Guide to Econometrics. Massachusetts: Blackwell Publishing, 2008.

KERLINGER, FRED. N. V. **Metodologia da Pesquisa em Ciências Sociais um tratamento conceitual**. São Paulo: Editora da Universidade de São Paulo, 1990.

MARQUARDT. D. W. **An algorithm for least squares estimation of nonlinear parameters**. Journal of the Society for Industrial and Applied Mathematics. v. 11, n. 2, p. 431-441, 1963.

NELDER, J.A. The fitting of a generalization of the logistic curve. **Biometrics**, v.17, p.89-94, 1961.

PEIXOTO, A. P. B. **Análise da dinâmica do potássio e nitrato em colunas de solo não saturado por meio de modelos não lineares e multiresposta**. Tese (Doutorado)-Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2013.

PRADO, T. K. L. **Modelos não-lineares no crescimento de frutos de coqueiro**. 2011, 93 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária), Universidade Federal de Lavras, Lavras, MG.

PRUDENTE, A. A. **Modelos não-lineares de regressão: Alguns aspectos de teoria assintótica**. 2009.

R Development Core Team (2012). R: A language and environment for statistical computing. Disponível em: <http://www.R-project.org/>. Acesso em: 10 set. 2014.

RAWLINGS, J. O.; PANTULA, S.G.; DICKEY, D. A. **Applied regression analysis**. 2nd ed. New York: Springer, 1998. 659 p.

REZENDE, D.M.L.C.; MUNIZ, J.A.; FERREIRA, D.F.; SILVA, F.F.; AQUINO, L.H.de. Ajuste de modelos de platô de resposta para a exigência de zinco em frangos de corte. **Ciência e Agrotecnologia**, Lavras, v.31, p.468-478, 2007.

SEBER, G.A.F.; WILD, C.J. **Nonlinear regression**. New York: J. Wiley, 1989. 768p

SILVA, N.A.M.; MUNIZ, J.A.; SILVA, F.F.; AQUINO, L.H.; GONCALVES, T. M. Aplicação do método Bayesiano na estimação de curvas de crescimento de animais da raça Nelore. *Revista Ceres*, Viçosa, v.54, p.192-199, abr./maio 2008.

SOUZA, E. M. **Modelagem não linear da extração de zinco em um solo tratado com lodo de esgoto**. 2008, 95p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária)-Universidade Federal de Lavras, Lavras, MG.

TRIOLA, M. F. **Introdução à Estatística**. Rio de Janeiro: LTC, 1999.

Von BERTALANFFY, L. Quantitative laws in metabolism and growth. **Quarterly Review of Biology**, v.32, p.218, 1957.

WINSOR, C.P. The Gompertz curve as a growth curve. **Proceedings of the National Academy of Science**, v.18, p.1-17, 1932.