



**WALLEFF DA SILVA E SILVA**

**CRITÉRIOS DE SELEÇÃO E QUALIDADE DE AJUSTE EM  
REGRESSÃO NÃO LINEAR:  
UMA ABORDAGEM DE MONTE CARLO**

**LAVRAS - MG**

**2021**

**WALLEFF DA SILVA E SILVA**

**CRITÉRIOS DE SELEÇÃO E QUALIDADE DE AJUSTE EM REGRESSÃO NÃO  
LINEAR:  
UMA ABORDAGEM DE MONTE CARLO**

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

Prof. DSc. Tales Jesus Fernandes  
Orientador

**LAVRAS - MG  
2021**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Silva, Walleff Silva e.

Crítérios de seleção e qualidade de ajuste em regressão não linear: uma abordagem de Monte Carlo / Walleff Silva e Silva. - 2021.

63 p. : il.

Orientador(a): Tales Jesus Fernandes.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2021.

Bibliografia.

1. Modelos de regressão não linear. 2. Curvas de Crescimento.  
3. Avaliadores de qualidade de ajuste. I. Fernandes, Tales Jesus. II. Título.

**WALLEFF DA SILVA E SILVA**

**CRITÉRIOS DE SELEÇÃO E QUALIDADE DE AJUSTE EM REGRESSÃO NÃO  
LINEAR: UMA ABORDAGEM DE MONTE CARLO**

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

APROVADA em 26 de Março de 2021.

Prof. DSc. Joel Augusto Muniz      UFLA  
Prof. DSc. Patrícia de Siqueira Ramos   Unifal

Prof. DSc. Tales Jesus Fernandes  
Orientador

**LAVRAS - MG  
2021**

*Dedico esse trabalho as minhas três mães: Antonia Regi (mãe), Francisca Roseno (Avó) e Odilia Maria (Tia).*

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, a Ele toda honra e glória desse momento, por todas as fraqueza e dificuldade ao qual em suplicas e orações Ele me levantou.

Agradeço, também, a minha família, em especial a minha mãe que em todos os momentos esteve ao meu lado, que sempre me apoiou em seguir meus sonhos, que me ensinou a buscar o que queria e nunca desistir por mais difícil o caminho. Mãe, você é e sempre será meu exemplo.

Agradeço ao meu orientador Tales, a quem aceitou orientar-me nesse projeto mesmo no meio do percurso. Com toda sua simplicidade, sabedoria e dedicação, dedicou sua atenção e acreditou no potencial do projeto guiando-me para este momento, obrigado professor.

A Universidade Federal de Lavras, pelo ensino e oportunidade de poder cursar esse mestrado e em especial ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, todos os professores, coordenadores, secretaria e ao grupo de estudo NLIN, meu eterno agradecimento.

Agradecer a CNPq pela patrocínio da bolsa que garantiu todo meu ensino mesmo longe de casa. Aos meus amigos de longa data, amigos de infância que sempre me apoiaram e aos amigos que conquistei ao longo da caminhada, mesmo aqueles que em algum momento seguiram outros caminhos. Gostaria de citar cada um que fez parte e me ajudou de alguma forma nesse mestrado, a lista é extensa e não quero ser negligente com nenhum, por esse motivo reservo a vocês todo meu agradecimento e todo meu carinho. Obrigado por serem minha família quando precisei e no momento que quis desistir, nunca esquecerei tal ato de amizade.

A cada pessoa que torceu por esse momento, cada abraço dedicado e acolhido, cada "Bom dia, Boa tarde ou Boa noite" que recebi, cada lágrima de tristeza e de felicidade, cada segundo ao longo desses anos, eu agradeço, obrigado a todos os envolvidos!

*"Os dias prósperos não vêm por acaso; nascem de muita fadiga e persistência."  
(Henry Ford)*

## RESUMO

Modelos de regressão não linear são difundidos na literatura e possuem grande utilização principalmente por suas vantagens em relação aos modelos de regressão linear. Pode-se citar a parcimônia, aplicabilidade e interpretação prática dos seus parâmetros como fatores determinantes no uso destes modelos, o que reflete diretamente em diferentes estudos na agricultura, biologia, economia, engenharias, entre outros. A utilização de curvas de crescimento é uma das principais aplicações dos modelos de regressão não linear para analisar o desenvolvimento ao longo da vida de determinado ser vivo. Existem na literatura muitos modelos para este propósito, com suas respectivas particularidades. Para saber qual dentre estes modelos fornece o melhor ajuste aos dados, utiliza-se como critério diferentes avaliadores de qualidade de ajuste, entretanto, não há como classificar qual melhor critério para a seleção de modelos deve ser utilizado em um estudo específico. Diferentes estudos apresentam o coeficiente de determinação, critério de informação de Akaike, critério de informação bayesiano como os mais úteis para avaliar um ajuste e dentre estes há certos questionamentos quanto a sua utilização para seleção de modelos de regressão não linear. Assim, o presente estudo tem como objetivo simular via Monte Carlo quatro cenários considerando as equações dos modelos de regressão não linear logístico, Gompertz, von Bertalanffy e Brody para avaliar a eficiência dos critérios de seleção em determinar o modelo que inicialmente gerou os dados. Foram utilizados quatro cenários de simulação Monte Carlo, sendo cada um considerado como padrão os modelos logístico, Gompertz, von Bertalanffy e Brody. Em seguida foram ajustados 4 modelos para cada um dos cenários e calculados os avaliadores de qualidade mais encontrados na literatura a fim de selecionar o modelo adequado. Os resultados demonstram que os avaliadores índice assintótico, desvio médio absoluto e coeficiente de determinação apresentam eficiência superior na escolha do ajuste adequado dentre os demais avaliadores estudados para os quatro cenários simulados.

**Palavras-chave:** Curvas de crescimento. Índice assintótico. Desvio médio absoluto. Medidas de não linearidade

## ABSTRACT

Nonlinear regression models are widespread in the literature and are widely used mainly for their advantages over linear regression models. Parsimony, applicability and practical interpretation of its parameters can be cited as determining factors in the use of these models, which directly reflect in different studies in agriculture, biology, economics, engineering, among others. The use of growth curves is one of the main applications of nonlinear regression models to analyze the development over the life of a certain living being. There are many models in the literature for this purpose, with their unique peculiarities. In order to know which one of these models offer the best fit to the data, different evaluated quality of fit criteria are used, however, there is no way to classify which best criterion for a selection of models should be used in a specific study. Different studies present the coefficient of determination, Akaike's information criterion, Bayesian information criterion as the most useful for assessing an adjustment and there are some questions regarding its use for the selection of non-linear regression models. Thus, the present study aims to simulate via Monte Carlo four scenarios considering as equations of the logistic nonlinear regression models, Gompertz, von Bertalanffy and Brody to assess the efficiency of the selection criteria in determining the model that created it generated the data. Four Monte Carlo simulation scenarios were used, each one being considered as standard the logistic models, Gompertz, von Bertalanffy and Brody. Then, 4 models were adjusted for each of the scenarios and the quality assessments most found in the literature were calculated in order to select the appropriate model. The results demonstrate that the asymptotic index, mean absolute deviation and determination coefficient raters show superior efficiency in choosing the appropriate adjustment among the other raters studied for the four simulated scenarios.

**Keywords:** Growth curves. Asymptotic index. Mean absolute deviation. Non-linearity measures

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>8</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>10</b>
<b>2.1</b>	<b>Modelos de regressão</b>	<b>10</b>
<b>2.2</b>	<b>Regressão não linear</b>	<b>11</b>
<b>2.3</b>	<b>Estimação dos parâmetros em modelos não lineares</b>	<b>15</b>
<b>2.4</b>	<b>Estimação de parâmetros por métodos iterativos</b>	<b>17</b>
<b>2.5</b>	<b>O estudo de curvas de crescimento</b>	<b>18</b>
<b>2.6</b>	<b>Crêterios de qualidade de ajuste</b>	<b>20</b>
<b>2.6.1</b>	<b>Medidas de não linearidade</b>	<b>25</b>
<b>2.7</b>	<b>Simulação de Monte Carlo (SMC)</b>	<b>28</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>30</b>
<b>3.1</b>	<b>Simulação</b>	<b>30</b>
<b>3.2</b>	<b>Métodos</b>	<b>33</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>35</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>48</b>
<b>6</b>	<b>APLICAÇÃO NA DESCRIÇÃO DO CRESCIMENTO DE EUCALIPTOS</b>	<b>49</b>
<b>6.1</b>	<b>Introdução</b>	<b>49</b>
<b>6.2</b>	<b>Material e Métodos</b>	<b>50</b>
<b>6.2.1</b>	<b>Material</b>	<b>50</b>
<b>6.2.2</b>	<b>Métodos</b>	<b>50</b>
<b>6.3</b>	<b>Resultados e Discussão</b>	<b>51</b>
<b>6.4</b>	<b>Conclusão</b>	<b>54</b>
<b>7</b>	<b>REFERÊNCIAS</b>	<b>56</b>

## 1 INTRODUÇÃO

Os modelos de regressão são muito utilizados com aplicação em diversas áreas. O principal objetivo da análise de regressão é identificar e quantificar algumas relações funcionais entre duas ou mais variáveis, onde uma variável é dita dependente da outra. Predição, estimação de parâmetros e controle são outros objetivos da análise de regressão e estes modelos podem ser classificados de três formas: linear, não linear e linearizáveis.

De acordo com Seber e Wild (1989), o termo linear é utilizado para indicar que o modelo é linear em relação aos seus parâmetros, ou seja, as derivadas parciais da função em relação aos parâmetros não dependem destes, enquanto que, os modelos não lineares, pelo menos uma das derivadas parciais em relação a qualquer parâmetro depende de algum parâmetro do modelo e não existe transformação capaz de torná-los lineares.

Os modelos não lineares apresentam algumas vantagens em relação aos lineares, tais como a origem na teoria e princípios biológicos, físicos, químicos e aplicados em situações que podem ser descritas por um menor número de parâmetros. Entretanto, há desvantagens, já que requer algoritmos iterativos e uso de valores iniciais para obtenção das estimativas dos parâmetros, permitem apenas aproximações em vez de inferências mais precisas e requerem sólida compreensão acerca do fenômeno em estudo (SCHABENBERGER & PIERCE, 2002).

Por suas vantagens, o modelo de regressão não linear é frequentemente utilizado por pesquisadores em diversas áreas do conhecimento, tais como, agricultura, biologia, econometria, engenharia, química etc. Na maioria das vezes as formulações de possíveis modelos são deduzidas a partir de suposições teóricas inerentes ao fenômeno que têm-se interesse modelar, e os parâmetros resultantes são interpretáveis. O modelo de regressão não linear com erros normais é um dos métodos mais utilizados quando a resposta é contínua, porque é de aplicação fácil e tem um grande número de teorias desenvolvidas para ele.

Dentre as aplicações dos modelos de regressão não lineares pode-se citar as curvas de crescimento. Kshirsagar e Smith (1995) afirmam que existe uma relação funcional de um efeito de tratamento e alguma variável indicadora de crescimento, esta variável pode ser ganho de peso e consegue-se modelar esta relação.

No estudo de curvas de crescimento, em geral, existe um grande número de modelos que são ajustados e geralmente o interesse do estudo é saber qual dos modelos possui o melhor ajuste e capacidade de predição da variável resposta. É preciso ter em mente que não existem modelos verdadeiros, há apenas modelos aproximados da realidade que buscam explicar o

fenômeno com perda de informações. São utilizados na literatura vários critérios para seleção do modelo mais adequado, de modo que não há um consenso sobre quais critérios de qualidade são mais eficientes nesta seleção.

De acordo com Donatelli e Konrath (2006), os resultados de medição afastam-se do valor verdadeiro da grandeza medida (valor do mensurando) por uma quantidade denominada erro de medição. A aplicação da simulação de Monte Carlo é uma forma de estudar situações em ambientes computacionalmente controlados e diminuir esses erros de medição.

Simulação de Monte Carlo refere-se ao emprego de modelos artificiais para representar processos reais de geração de dados, de modo a obter uma maior compreensão de tais processos (Barreto e Howland, 2005). Por meio da simulação, podem-se criar amostras grandes e estudar o comportamento em situações controladas, examinar propriedades de diferentes métodos de estimação e situações distantes do ideal.

Portanto, o presente trabalho tem como objetivo avaliar via simulação Monte Carlo os critérios de seleção mais eficientes para os modelos de regressão não linear Logístico, Gompertz, von Bertalanffy e Brody na descrição de curvas de crescimento.

## 2 REFERENCIAL TEÓRICO

### 2.1 Modelos de regressão

O modelo estatístico de regressão que procura explicar uma relação entre as variáveis  $\mathbf{X}$  e  $\mathbf{Y}$ , ou seja, uma função de  $\mathbf{X}$  que explique  $\mathbf{Y}$ , admitindo um erro aleatório aditivo é representado por:

$$\mathbf{Y}_i = f(\mathbf{X}_i; \theta) + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (2.1)$$

em que,  $\mathbf{Y}_i$  é o vetor com a variável resposta (variável dependente),  $\mathbf{X}_i$  é o vetor de uma ou mais variáveis independentes;  $\theta$  é o vetor de parâmetros do modelo;  $f(\cdot)$  é a função (ou relação) que acredita-se existir entre as variáveis e  $\varepsilon_i$  é o vetor de erros associados ao modelo, os quais por suposição seguem uma distribuição normal com média zero e variância constante,  $\varepsilon_i \sim N(0, \sigma^2)$ .

Segundo Draper e Smith (1998) os modelos de regressão podem ser classificados como:

- **modelos lineares:** são modelos cuja a derivada parcial em relação a qualquer parâmetro não depende de nenhum dos parâmetros do modelo, são lineares em seus parâmetros;
- **modelos não lineares:** nestes modelos pelo menos uma das derivadas parciais em relação a qualquer parâmetro depende de algum parâmetro do modelo e não existe transformação capaz de torná-los lineares, são não lineares em seus parâmetros;
- **modelos linearizáveis:** Em sua forma inicial são não lineares, mas podem se tornar lineares através de uma transformação.

A principal diferença entre os modelos lineares e não lineares está relacionado às suas formulações. O modelo linear objetiva explicar a melhor relação, se existir, entre as variáveis de dado fenômeno por meio de um conjunto de observações. No caso não linear, as formulações desse modelo são garantidas a partir de suposições teóricas e seus parâmetros são interpretáveis.

A escolha desses modelos a ser utilizado vai depender do fenômeno em estudo, em algumas situações um modelo linear pode ser adequado em outras o não linear pode ser mais apropriado. A forma que as variáveis se comportam ao explicar o fenômeno é que vai definir a utilização do modelo adequado.

A seguir, é detalhado o modelo de regressão não linear e a sua utilização na modelagem de curvas de crescimento.

## 2.2 Regressão não linear

Os modelos lineares são importantes e resolvem grande parte dos problemas, no entanto, os modelos não lineares destacam-se por terem interpretação prática dos parâmetros, aplicabilidade e melhores ajustes em fenômenos biológicos. Parcimônia, interpretabilidade e predição são algumas vantagens dos modelos não lineares.

A forma clássica de um modelo de regressão não linear é a mesma da expressão (2.1), a diferença está no resultado das derivadas da forma funcional do modelo  $f()$  em relação ao parâmetro  $\theta$ . Segundo Bates e Watts (1980) para os modelos de regressão não linear, pelo menos uma das derivadas parciais da função  $f()$  em relação aos parâmetros depende de algum dos parâmetros do modelo.

A utilização de modelos de regressão não linear é bem vasta, principalmente pelo formato sigmoideal destas curvas, este comportamento representado por medidas que representam o desenvolvimento ao longo do tempo. Os modelos logístico, Gompertz, von Bertalanffy e Brody são úteis e bem difundidos no ajuste de curvas de crescimento e, importante ressaltar, o comportamento desses ajustes pode variar de acordo com o fenômeno estudado, tendo cada um suas características próprias.

A seguir, será apresentado os modelos logístico, Gompertz, von Bertalanffy e Brody com suas características na descrição de curvas de crescimento.

### Modelo logístico

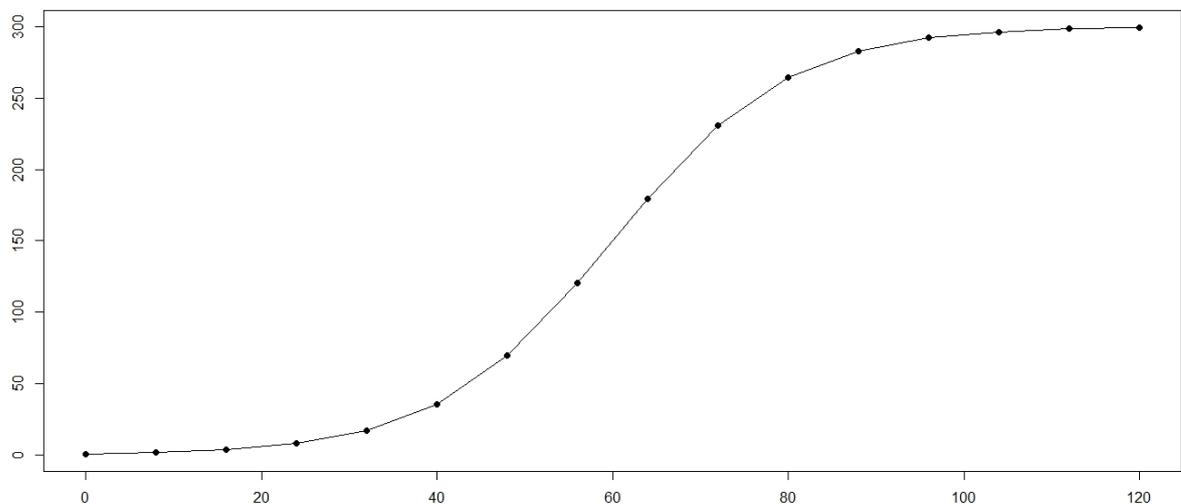
A função Logística foi proposta, inicialmente, por Verhust, em 1838, para o estudo de crescimento de populações humanas (Seber; Wild, 1989). Segundo Fernandes et al. (2015) muitos autores têm usado diferentes parametrizações para os modelos não lineares que descrevem curvas de crescimento, mas a que apresentou melhor desempenho para o modelo logístico devido, principalmente, à interpretação prática de todos os parâmetros foi:

$$Y_i = \frac{\alpha}{1 + e^{k(\beta - x_i)}} + \varepsilon_i \quad (2.2)$$

em que,  $i = 1, 2, \dots, n$ ;  $Y_i$  é a  $i$ -ésima observação da variável dependente,  $x_i$  é a  $i$ -ésima observação da variável independente;  $\alpha$  é o valor assintótico, ou seja, o valor esperado para o crescimento máximo do objeto em estudo;  $\beta$  é a abscissa do ponto de inflexão, ou seja, a partir dele o crescimento é desacelerado;  $k$  é um índice de maturidade ou precocidade, está associado ao

crescimento, quanto maior for seu valor, menos tempo será necessário para o objeto em estudo atingir o valor assintótico ( $\alpha$ );  $\varepsilon_i$  é o erro aleatório associado à  $i$ -ésima observação o qual pressupõe-se que seja independente e identicamente distribuído seguindo uma distribuição normal de média zero e variância constante, ou seja,  $\varepsilon_i \sim N(0, \sigma^2)$  (FERNANDES et al., 2014; FERNANDES et al., 2015; FREITAS, 2005; MUNIZ; NASCIMENTO; FERNANDES, 2017).

Esse modelo apresenta a reta  $y = \alpha$  como assíntota horizontal e curva sigmoideal simétrica no ponto de inflexão, no qual ocorre a mudança de concavidade da curva. A abscissa do ponto de inflexão é dada por  $x = \beta$ , que se comprova pelo valor da variável independente que anula a derivada de segunda ordem da função, o parâmetro  $k$  dará a forma da curva, para  $\alpha > 0$  e  $k > 0$ , quanto maior for o valor de  $k$ , maior será a inclinação da reta tangente no ponto de inflexão, sendo  $k = \frac{4}{\alpha} y'_{pi}$ , em que  $y'_{pi}$  é a primeira derivada calculada no ponto de inflexão (MISCHAN; PINHO, 2014).



Fonte: Do autor

**Figura 1:** Representação gráfica do modelo logístico

### Modelo Gompertz

O modelo foi proposto por Gompertz, em 1825, com objetivo de estudar as tábuas de vida. Wright, em 1926, utilizou o modelo para obtenção de curvas de crescimento e Medawar, em 1940, utilizou para descrever o crescimento do coração de galinhas.

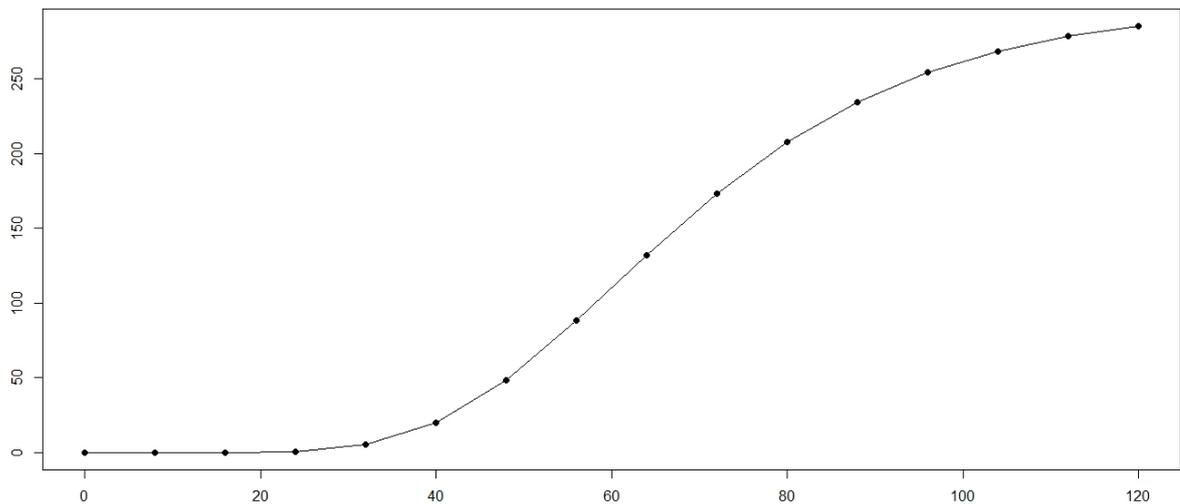
Esse modelo é muito semelhante ao modelo logístico, observando um crescimento exponencial inicialmente e, a seguir, assintótico. Segundo Fernandes et al. (2015), dentre as diferentes parametrizações do modelo Gompertz, a que apresenta melhor desempenho é:

$$Y_i = \alpha e^{-e^{k(\beta - x_i)}} + \varepsilon_i \quad (2.3)$$

A interpretação dos parâmetros é semelhante a anteriormente apresentada para o modelo logístico, inclusive a assíntota horizontal.

De acordo com Mischan e Pinho (2014) o parâmetro  $\beta$  é uma medida de posição do valor inicial de  $y$  em relação a assíntota, representa a ideia da razão entre a assíntota e quantidade que falta para atingi-la. Ainda segundo o autor, o parâmetro  $k$  representa o índice de crescimento, pode ser obtido por  $k = \frac{e}{\alpha} y'_{pi}$ ; em que  $y'_{pi}$  representa a primeira derivada calculada no ponto de inflexão.

O modelo Gompertz também apresenta curva com caráter sigmoidal, porém não é simétrica em relação ao ponto de inflexão, o que favorece muitos estudos de crescimento animal (SEBER; WILD, 1989).



Fonte: Do autor

**Figura 2:** Representação gráfica do modelo Gompertz

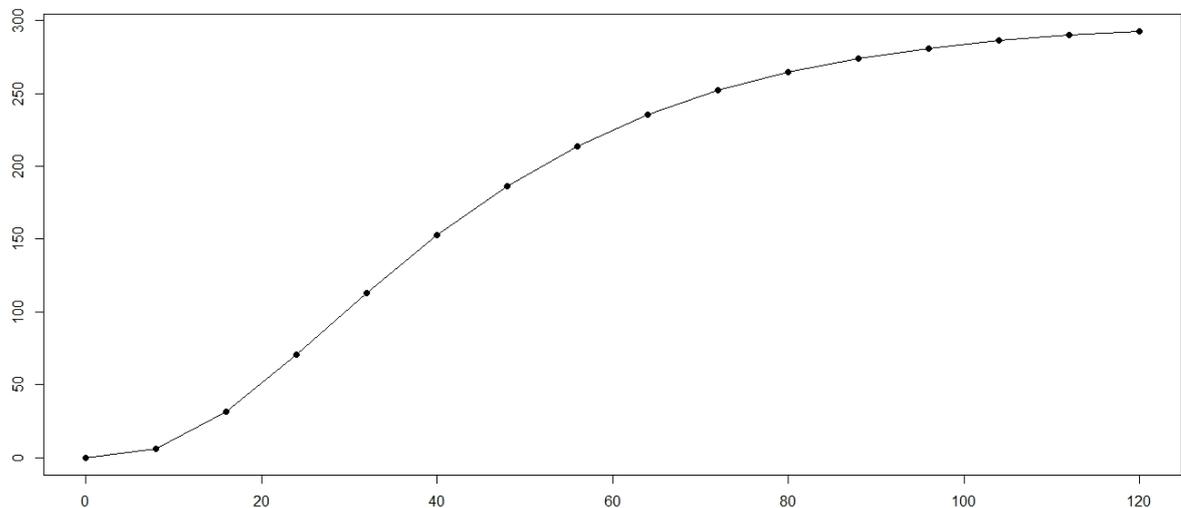
### Modelo von Bertalanffy

Ludwing von Bertalanffy em 1957 propôs esse modelo para utilização na área ecológica com fins a modelar o crescimento de peixes e crustáceos em função do tempo. Dentre as parametrizações deste modelo, pode-se encontrar na literatura esta:

$$Y_i = \alpha \left( 1 - \frac{e^{k(\beta - x_i)}}{3} \right)^3 + \varepsilon_i \quad (2.4)$$

A interpretação para  $Y_i$ ,  $x_i$ ,  $\alpha$ ,  $k$  e  $\varepsilon_i$  é a mesma apresentada anteriormente no modelo logístico. A equação (2.4) apresenta uma parametrização em que o parâmetro  $\beta$  pode ser interpretado como a abscissa do ponto de inflexão. Este modelo apresenta formato de curva sigmoidal e não é simétrico em relação ao ponto de inflexão.

Para um mesmo valor de  $\beta$ , o parâmetro  $k$  é inversamente proporcional à localização da abscissa do ponto de inflexão, assim, quanto maior o valor de  $k$ , maior a velocidade de crescimento (MISCHAN; PINHO, 2014).



Fonte: Do autor

**Figura 3:** Representação gráfica do modelo Von Bertalanffy

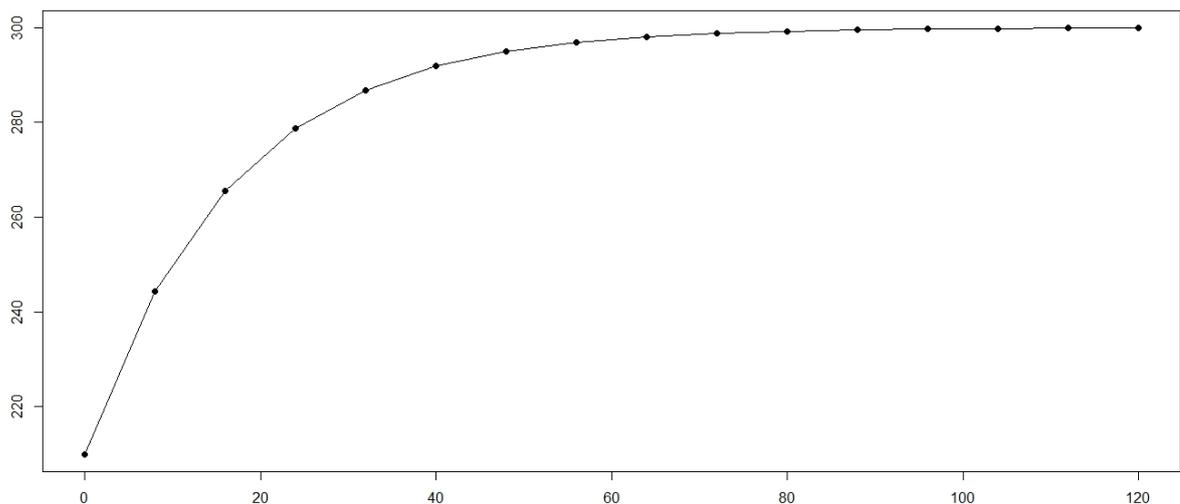
### Modelo Brody

Esse modelo foi proposto por Brody em 1945 e é usado para descrever o crescimento de organismos com taxas de crescimento decrescentes desde o início até a sua fase final, onde há uma estabilidade do crescimento (BRODY, 1945; MISCHAN; PINHO, 2014).

Esse modelo não apresenta comportamento sigmoidal, assim, fornecendo melhores ajustes na descrição de curvas de crescimento quando estas não apresentam este padrão, já começando com a taxa máxima de crescimento e diminuindo posteriormente. Sua parametrização pode ser descrita por:

$$Y_i = \alpha[1 - \beta e^{-kx_i}] + \varepsilon_i \quad (2.5)$$

A interpretação dos parâmetros é a mesma do modelo logístico, com exceção do parâmetro  $\beta$  que, neste modelo, não possui interpretação prática por não existir o ponto de inflexão da curva, portanto, o parâmetro  $\beta$  é uma medida de posição do valor inicial de  $y$  em relação à assíntota  $\alpha$ , esse valor representa a razão entre a assíntota e a quantidade que falta para atingi-la.



Fonte: Do autor

**Figura 4:** Representação gráfica do modelo Brody

### 2.3 Estimação dos parâmetros em modelos não lineares

Existem na literatura diversos métodos para a estimação de parâmetros, tanto para modelos lineares e não lineares, dentre os mais difundidos têm-se o método da máxima verossimilhança, dos momentos e o de mínimos quadrados, sendo o método de mínimos quadrados ordinários, o mais utilizado.

O método dos mínimos quadrados ordinários é aquele cuja estrutura dos erros não viola nenhuma das pressuposições, ou seja,  $\varepsilon \sim N(0, \sigma^2 I_n)$ . Em geral, assume-se que os erros são va-

riáveis aleatórias independentes e identicamente distribuídas, normais com média 0 e variância constante  $\sigma^2 I_n$ , em que  $I_n$  representa a matriz identidade de ordem  $n$ . O objetivo é encontrar o melhor ajuste para um conjunto de dados, que minimiza a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Matricialmente, a soma de quadrados dos resíduos,  $SQR(\theta)$ , do modelo (2.1) pode ser escrita como:

$$SQR(\theta) = \varepsilon' \varepsilon = [\mathbf{Y} - f(\mathbf{X}; \theta)]' [\mathbf{Y} - f(\mathbf{X}; \theta)], \quad (2.6)$$

Então,

$$\begin{aligned} SQR(\theta) &= [\mathbf{Y}' - f(\mathbf{X}; \theta)]' [\mathbf{Y} - f(\mathbf{X}; \theta)] \\ &= \mathbf{Y}' \mathbf{Y} - \mathbf{Y}' f(\mathbf{X}; \theta) - [f(\mathbf{X}; \theta)]' \mathbf{Y} + [f(\mathbf{X}; \theta)]' f(\mathbf{X}; \theta) \\ &= \mathbf{Y}' \mathbf{Y} - 2\mathbf{Y}' f(\mathbf{X}; \theta) + [f(\mathbf{X}; \theta)]' f(\mathbf{X}; \theta) \end{aligned}$$

Na expressão de  $SQR(\theta)$  para encontrar o estimador de mínimos quadrados é necessário diferenciar  $SQR(\theta)$  com relação ao vetor de parâmetros e o resultado da derivada igualar a zero, ou seja,

$$\frac{\partial SQR(\theta)}{\partial \theta} = -2\mathbf{Y}' \frac{\partial f(\mathbf{X}; \theta)}{\partial \theta} + \frac{\partial [f(\mathbf{X}; \theta)]'}{\partial \theta} \times f(\mathbf{X}; \theta) + [f(\mathbf{X}; \theta)]' \times \frac{\partial f(\mathbf{X}; \theta)}{\partial \theta}$$

em que  $\frac{\partial f(\mathbf{X}; \theta)}{\partial \theta}$  é conhecida como matriz jacobiana,  $n \times p$ , sendo  $n$  o número de observações e  $p$  a quantidade de parâmetros do modelo, contendo as derivadas parciais. Então:

$$-2\mathbf{Y}' \frac{\partial f(\mathbf{X}; \theta)}{\partial \theta} + 2[f(\mathbf{X}; \hat{\theta})]' \frac{\partial f(\mathbf{X}; \theta)}{\partial \theta} = 0$$

Chamando a matriz jacobiana de  $\mathbf{X}$  e aplicando a propriedade de matrizes  $(A'B)' = B'A$ , temos:

$$\mathbf{X}' f(\mathbf{X}; \hat{\theta}) = \mathbf{X}' \mathbf{Y} \quad (2.7)$$

Dessa forma, obtêm-se um sistema de equações normais não lineares, no qual a matriz jacobiana  $f(\mathbf{X}; \hat{\theta})$  depende de  $\theta$ , logo esse sistema não possui uma solução explícita para  $\hat{\theta}$ . Assim, as propriedades ótimas dos estimadores para os modelos de regressão não linear são válidas somente assintoticamente, isto é, somente quando a amostra é suficientemente grande.

Pode-se explicar isso, porque o sistema não possui solução explícita, necessitando de métodos iterativos para a estimação dos parâmetros.

## 2.4 Estimação de parâmetros por métodos iterativos

Não há forma fechada para a solução do sistema de equações normais de um modelo não linear, sendo necessário utilizar métodos numéricos iterativos (DRAPER e SMITH, 1998). A maioria dos métodos de estimação para modelos não lineares baseiam-se em métodos iterativos.

Há uma vasta literatura que aborda métodos iterativos para a obtenção dessas estimativas, dentre estas, o método Gauss-Newton consiste em aproximações do sistema não linear por um sistema linearizado. Para que possa ser aplicado o método dos mínimos quadrados e estimar os parâmetros, o método de Gauss-Newton aproxima o modelo de regressão não linear com termos lineares utilizando uma expansão em série de Taylor de primeira ordem.

Elias et al. (2016) discorrem que na utilização de uma abordagem de iteração numérica para encontrar raízes reais de uma equação, pode-se obter seus resultados dividindo o processo em três principais etapas: definição inicial de um intervalo, precisão e número de iterações.

Para convergência do resultado aproximado é necessária a obtenção de um intervalo no qual pode-se afirmar que, dentro deste, existe uma raiz da equação. Adota-se uma determinada tolerância para a aproximação do zero da equação localizado no intervalo definido. Por fim, para que o programa não entre em um *loop* infinito, escolhe-se um número máximo de iterações.

Essas iterações no método de Gauss-Newton são repetidas um determinado número de vezes até que o vetor de estimativas permaneça inalterado dentro de uma precisão pré-estipulada. Assim, é de suma importância a escolha dos valores iniciais, para que este consiga convergir com rapidez e seus resultados sejam precisos. Ratkowsky (1983) discute procedimentos para obtenção de bons valores iniciais para algumas classes de modelos (modelos de crescimento, modelos de regressão assintóticos etc.).

Ressaltam Mazucheli et al. (2002), que a rapidez na convergência depende da complexidade do modelo em estudo e, principalmente, da qualidade dos valores iniciais, necessários em qualquer método iterativo.

## 2.5 O estudo de curvas de crescimento

A curva de crescimento descreve uma sequência de medidas sobre determinada característica (peso, altura, volume ou diâmetro) de alguma espécie em função do tempo. O formato desse tipo de curva apresenta características exponencial ou sigmoideal (em forma de S). Segundo Mischan e Pinho (2014), o crescimento de seres vivos mostra um comportamento distinto, começando devagar, passando para uma fase exponencial e tendendo a estabilizar no final.

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 tenha interpretação prática e preditiva (PRADO et al. 2020).

Na literatura, vários trabalhos utilizam os modelos de regressão não lineares para ajuste de curvas de crescimento. Apresenta-se alguns desses estudos na tabela a seguir.

Tabela 2.1 – Principais estudos de curvas de crescimento utilizando modelagem de regressão não linear.

<b>Crescimento de frutos</b>	<b>Crescimento Vegetal</b>	<b>Crescimento animal</b>
Fernandes et al. (2014)	Pereira et al. (2016)	Taddei et al. (2010)
Muianga et al. (2016)	Bem et al. (2017)	Rocha-Silva et al. (2016)
Sari et al. (2019)	Kleinpaul et al. (2018)	Miranda et al. (2019)
Diel et al. (2019)	Rezende et al. (2020)	Botelho et al. (2019)
Silva et al. (2020)	Jane et al. (2020)	Fernandes et al. (2019)

Destaca-se os seguintes estudos:

Para o crescimento de frutos, Sari et al. (2019) descreveram o comportamento produtivo do tomate ao longo do tempo (colheitas) através de modelos não lineares Brody, Gompertz, logístico e von Bertalanffy, selecionando o melhor modelo com base no valor do coeficiente de determinação e na não linearidade paramétrica. Através do resultado baixo para a não linearidade paramétrica, o modelo logístico foi selecionado como melhor, para o comportamento produtivo dos genótipos.

Quanto ao crescimento vegetal, Jane et al. (2019) compararam os ajustes do modelo linear Polinomial com os modelos de regressão não linear logístico e Gompertz na descrição do crescimento em altura de pimenta da cultivar Doce, a comparação da qualidade de ajuste dos modelos foi feita por meio de análise gráfica e dos avaliadores: coeficiente de determinação ajustado ( $R_{ajs}^2$ ), desvio padrão residual (DPR) e o critério de informação de Akaike corrigido (AICc). Conforme os avaliadores de ajuste, o modelo logístico foi o mais adequado.

O estudo Miranda et al. (2019) exemplifica a utilização de curvas de crescimento animal. Os autores utilizaram os modelos não lineares na descrição do crescimento de tilápia do Nilo criada em tanques-rede. O modelo de Von Bertalanffy apresentou-se como mais adequado para a descrição da tilápia do Nilo em tanques-rede, pois tanto para o peso quanto para o crescimento apresentou o maior coeficiente de determinação ajustado e menor AIC.

Como apresentado nos parágrafos acima, a escolha do modelo que apresenta o melhor ajuste é um fator importante para a descrição e obtenção do objetivo de qualquer estudo. O "melhor" modelo é definido através da utilização dos critérios de seleção e qualidade de ajuste, no entanto, em regressão não linear não há um consenso na utilização destes critérios, a literatura apresenta diferentes avaliadores para selecionar curvas de crescimento.

Sari et al. (2018) utilizaram o critério de informação de Akaike (AIC), critério de informação bayesiano de Schwarz (BIC) e o coeficiente de determinação ( $R^2$ ) para a escolha de modelos não linear na análise de dados de várias colheitas. Lopes et al. (2016) utilizaram o quadrado médio do resíduo, coeficiente de determinação, erro médio de predição e desvio médio absoluto dos resíduos como critérios de seleção de modelos nas análises de dados longitudinais em bovinos Nelore Mocho por meio de modelos não lineares.

Não existe um critério específico para a seleção de um modelo mais adequado em regressão não linear, Archontoulis et al. (2015) no seu estudo de modelos e aplicações de regressão não linear em pesquisa agrícola sugerem uma diferenciação na utilização dos critérios de seleção. Segundo os autores, caso os modelos sejam aninhados utilizar os avaliadores AIC, BIC, teste F e teste de razão de verossimilhança é mais adequado para escolha do modelo, em caso contrário utiliza-se somente o AIC e BIC como critério de decisão.

Existe, ainda, autores que questionam a utilização de alguns avaliadores na modelagem de curvas de crescimento. Spiess e Neumeyer (2010) abordam a utilização do  $R^2$  e põem em prova a efetividade desse avaliador na seleção dos ajustes de modelos regressão não lineares na farmacologia e pesquisa bioquímica. De acordo com os autores, a confiabilidade na seleção de modelos utilizando o  $R^2$  é fraca, visto que seus resultados geralmente apresentam variação apenas na quarta casa decimal.

É necessário estudar a utilização desses avaliadores em modelos de regressão não linear e quais destes são eficientes na modelagem de curvas de crescimento.

## 2.6 Critérios de qualidade de ajuste

Um modelo representa, de forma simplificada, algum problema ou situação da vida real tendo como objetivo ilustrar aspectos do problema. No entanto, mais de um modelo pode se adequar a um mesmo fenômeno, pois cada pesquisador tem liberdade para modelar o fenômeno de acordo com a metodologia que desejar. Portanto, segundo Eloy et al. (2018), quando vários modelos são ajustados a um conjunto de dados é importante, e de interesse do estudo, saber qual dos modelos possui o melhor ajuste e capacidade de predição da variável resposta.

Entretanto, na maioria das vezes, a escolha dos melhores modelos não é tão simples; existe um grande número de avaliadores de qualidade de ajuste na literatura, muitas vezes há discordância entre eles para a definição do melhor modelo a ser utilizado para descrever o ajuste, assim, uma análise que considere simultaneamente todos os avaliadores de qualidade, em um ambiente controlado, é necessária para estabelecer qual modelo melhor se ajusta aos dados.

De acordo com Navarro e Myung (2004), ao avaliar um modelo, há vários fatores a considerar. Em termos gerais, podem ser usados métodos estatísticos para medir a suficiência descritiva de um modelo (ajustando-o aos dados e testando esses ajustes), como também sua generalização e simplicidade (usando ferramentas de seleção de modelos). Porém, a qualidade de um modelo também depende de sua interpretabilidade, de sua consistência com outros e de sua plausibilidade global. Isso implica em julgamentos inerentemente subjetivos, mas não menos importantes.

Pode-se encontrar na literatura uma gama de critérios de seleção e qualidade de ajuste para modelos, a maioria dos critérios de seleção são destinados para os modelos lineares e alguns destes são adequados para a utilização em modelos não lineares. No entanto, a literatura sobre os critérios de seleção para modelos não lineares ainda não é ampla, a escolha de avaliadores para seleção de modelos é livre para cada pesquisador e a confiabilidade na escolha do modelo por vezes é duvidosa pois não há estudos que comprovem qual melhor critério de seleção para estes modelos. Faz-se necessário um estudo para entender melhor sobre esses avaliadores e a eficiências dos mesmos.

A forma mais simples de escolha de modelos baseia-se no uso do coeficiente de determinação ( $R^2$ ) da regressão, que estima o ajustamento, e no quadrado médio dos resíduos (QME) que estima o afastamento, como recomendado por Sit (1994). Geralmente, a análise gráfica dos resíduos também é usada para decidir sobre o melhor modelo conforme descrito por Busab (1986). Outros critérios de seleção de modelos são usados em programas computacionais

como o  $R^2$  ajustado, a estatística Cp de Mallows, o critério de informação de Akaike (*Akaike's information criterion* – AIC), o critério bayesiano de Schwarz (*Schwarz's Bayesian criterion* – SBC) também chamado de critério de informação bayesiano de Schwarz (*Bayesian information criterion*– BIC) e o critério de predição de Amemiya (*Amemiya's prediction criterion* – PC), ainda, o critério bootstrap de Efron (*Efron bootstrap criterion* – EBC), o critério de validação cruzada (*cross-validation criterion* – CVC) e o mínimo comprimento de descrição (*minimum description length* – MDL).

A seguir são apresentados os critérios de avaliação de qualidade que serão estudados neste trabalho.

### **Coefficiente de determinação**

Charnet et al. (1999) esclarecem que o coeficiente de determinação  $R^2$  é a proporção da variabilidade dos Ys observados explicada pelo modelo considerado em que o valor de  $R^2$  pertence ao intervalo [0;1].

$$R^2 = 1 - \frac{SQE}{SQTotal},$$

em que SQE é a soma dos quadrados dos resíduos e SQTotal é a soma dos quadrados totais. Quanto mais próximo de 1 for o resultado de  $R^2$ , melhor o ajuste do modelo considerado.

### **Coefficiente de determinação ajustado**

O  $R^2$  ajustado ( $R^2_{ajs}$ ) pondera o coeficiente de determinação ( $R^2$ ) pelo número de variáveis explicativas e pelo número de observações da amostra. É particularmente útil quando desejamos comparar modelos de regressão múltipla que preveem a mesma variável dependente, pois penaliza aquele modelo com maior número de variáveis independentes (MAIA, 2017). No entanto, o  $R^2_{ajs}$  perde a interpretação prática do  $R^2$ , servindo apenas como critério de comparação entre modelos. É calculado da seguinte forma:

$$R^2_{ajs} = 1 - \frac{n-1}{n-p} \times (1 - R^2),$$

em que  $n$  é o número de observações,  $p$  o número de parâmetros e  $R^2$  o coeficiente de determinação. O resultado obtido varia de 0 a 1, o modelo que apresentar maior  $R^2$  ajustado será considerado como o ajuste mais adequado.

### Quadrado médio dos resíduos

O quadrado médio dos resíduos de um modelo de regressão é obtido por meio de:

$$QME = \frac{SQE}{(n - p - 1)},$$

em que SQE é a soma dos quadrados dos resíduos. O QME sempre decresce conforme  $p$  aumenta. Esse eventual aumento ocorre quando a redução do QME em adicionar um coeficiente para o modelo não é suficiente para compensar a perda nos graus de liberdade do denominador.

Assim, minimizar QME e maximizar  $R_{ajs}^2$  são equivalentes. Em outras palavras, para a seleção de modelos, deseja-se o modelo com menor QME.

### $C_p$ de Mallows

Segundo Hair Júnior et al. (1998), a estatística  $C_p$  de Mallows é delineada contra o número de parâmetros ( $p$ ); quanto mais próximo for  $C_p$  de  $p$ , menos tendenciosas são as estimativas dos parâmetros e melhor é o modelo. A estatística  $C_p$  de Mallows é dada por:

$$C_p = \frac{SQE}{S^2} - (n - 2p),$$

em que  $p$  é número de parâmetros do modelo;  $S^2$  é quadrado médio do erro para o modelo completo; SQE é soma de quadrados do erro para o modelo com  $p$  parâmetros, incluindo, se houver, o intercepto.

Um valor baixo de  $C_p$  indica que o modelo é relativamente preciso e possui variância pequena na estimativa verdadeira dos coeficientes da regressão e na previsão de respostas futuras. Assim como, modelos com baixa capacidade preditiva e tendência, têm valor de  $C_p$  maior do que  $p$  (MONTGOMERY; RUNGER, 2008).

### Critério de informação de Akaike (AIC)

Desenvolvido por Hirotugu Akaike sob o nome de "um critério de informação" em 1971 e proposto em Akaike (1974) é uma medida relativa da qualidade de ajuste de um modelo estatístico estimado. Fundamentada no conceito de entropia, oferecendo uma medida relativa das informações perdidas, quando um determinado modelo é usado para descrever a realidade. "Akaike encontrou uma relação entre a esperança relativa da informação de Kullback-Leibler e a função suporte maximizada, permitindo uma maior interação entre a prática e a teoria, em seleção de modelos e análises de conjuntos de dados complexos" (BURNHAM; ANDERSON,

2002).

O AIC é um critério que avalia a qualidade do ajuste do modelo paramétrico, estimado pelo método da máxima verossimilhança. Akaike (1974) definiu seu critério de informação como:

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

em que  $\log L(\hat{\theta})$  é a função suporte maximizada e  $k$  é o número de parâmetros.

Dado um conjunto de dados e vários modelos concorrentes, pode classificá-los de acordo com seu AIC, com aqueles tendo os menores valores de AIC sendo os melhores (Burnham; Anderson, 2002).

### **Critério de informação de Akaike corrigido (AICc)**

Bozdogan (1987) propôs a seguinte correção para o AIC, indicado para pequenas amostras e respostas com distribuição normal:

$$AICc = -2\log L(\hat{\theta}) + 2(k) + 2\frac{k(k+1)}{n-k-1},$$

A seleção do modelo mais adequado é semelhante ao AIC, os menores valores de AIC classificam os melhores modelos. O AICc não é calculado quando  $n - p - 1 \leq 0$ .

### **Critério de informação bayesiano (BIC)**

O critério de informação bayesiano (BIC), também chamado de Critério de Schwarz, foi proposto por Schwarz (1978), e é um critério de avaliação de modelos definidos em termos de probabilidade a posteriori, sendo assim chamado porque Schwarz deu um argumento Bayesiano para prová-lo. O Critério Bayesiano de Schwarz (BIC) tem como pressuposto a existência de um “modelo verdadeiro” que descreve a relação entre a variável dependente e as diversas variáveis explanatórias entre os diversos modelos sob seleção. Assim o critério é definido como a estatística que maximiza a probabilidade de se identificar o verdadeiro modelo dentre os avaliados. Então o Critério de informação Bayesiano (BIC) é dado por:

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

em que  $L(\hat{\theta})$  é a função suporte maximizada,  $k$  o número de parâmetros e  $n$  o número de observações.

Andrade Filho et al. (2010) afirmam que o critério BIC penaliza mais o número de parâmetros e tende a selecionar modelos com menor número de parâmetros, assim, dos modelos avaliados usando o BIC, aquele que minimizar o valor do BIC é o melhor modelo para os dados.

AIC e BIC apesar de conceitualmente diferentes acerca dos modelos em avaliação, utilizam o mesmo critério estatístico, o máximo da função de verossimilhança como medida do ajustamento, entretanto, definem valores críticos diferentes.

### **Raiz do erro quadrático médio**

A raiz do erro quadrático médio (RMSD), proposto por Fox(1981), consiste em avaliar a magnitude média dos erros de previsão, tem valor sempre positivo e quanto mais próximo de zero, maior a qualidade de predição em relação aos dados observados. Essa medida é capaz de apontar o quanto as predições se adequam bem aos dados observados, este índice não faz distinção quanto aos dados serem subestimados ou superestimados (Jacovides e Kontoyiannis, 1995).

A raiz do erro quadrático médio (RMSD) é calculada da seguinte forma:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2},$$

Em que  $y_i$  são os valores reais e  $\hat{y}_i$  os valores previstos no modelo. A desvantagem é que bastam alguns valores discrepantes para que ocorra aumento significativo em sua magnitude (Stone, 1993).

### **Desvio médio absoluto (DMA)**

Definida como a média das diferenças absolutas entre os valores reais ( $y_i$ ) e os valores previstos no modelo ( $\hat{y}_i$ ), dividido por n (número de observações). Pode ser obtida da seguinte forma:

$$DMA = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n},$$

Se  $y_i = \hat{y}_i$  obtemos DMA = 0, ou seja, é preferível o modelo que apresentar menor DMA.

### **Índice Assintótico**

O índice assintótico (IA), descrito por Ratkowsky et al. (1990), combina os critérios de desvio padrão absoluto (DPA) - calculado pela raiz quadrada do quadrado médio do resíduo, desvio médio absoluto (DMA) e coeficiente de determinação ajustado ( $R_{aj}^2$ ). Assim temos:

$$IA = (DPA - DMA) - R_{ajs}^2,$$

A modelagem que apresentar o IA mais próximo de zero é considerado como o modelo de melhor ajuste.

### Erro de predição médio ( $EP_M$ )

O erro de predição médio ( $EP_M$ ) é obtido calculando-se a média dos erros de predição com EP dado por:

$$EP = 100 \left( \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \right),$$

em que:  $y_i$  é o valor observado e  $\hat{y}_i$  o valor estimado. Este valor leva um sinal, designando se a função subestimou (+) ou superestimou (-) o peso observado, e sua magnitude mensura a aderência dos dados estimados aos dados observados. Para seleção de ajustes é preferível modelos com baixos  $EP_M$ .

#### 2.6.1 Medidas de não linearidade

As medidas de não linearidade são usadas para avaliar a adequabilidade de um modelo não linear quanto à aproximação de um modelo linear. Em outras palavras, objetiva-se verificar se o grau de não linearidade de um ajuste não linear é suficientemente pequeno para que as técnicas usuais de estimação, desenvolvidas para a regressão linear, sejam utilizadas como uma boa aproximação para o modelo não linear.

As medidas de vício de Box e as medidas de curvatura de Bates e Watts são as ferramentas mais utilizadas na avaliação da não linearidade de um modelo de regressão não linear encontradas na literatura. Box (1971) propôs uma fórmula para estimar os vícios dos estimadores de mínimos quadrados de um modelo de regressão univariado.

Bates e Watts (1980) apresentaram novas medidas de não linearidade baseadas no conceito geométrico de curvatura. Eles provaram que a não linearidade de um modelo pode ser decomposta em duas componentes: a não linearidade intrínseca e a não linearidade devida ao efeito de parâmetros.

Obtenção fácil de estimativas por mínimos quadrados, estimadores aproximadamente não-viciados, normalmente distribuídos com variância mínima, mesmo em pequenas amostras, valores de previsões mais precisos, convergência mais rápida de métodos iterativos, estimadores

com propriedades similares às propriedades ótimas de modelos lineares são algumas vantagens em se buscar modelos com comportamento próximo do comportamento linear.

### Vício de Box

Segundo Zeviani et al. (2012), o processo inferencial em modelos não lineares é baseado em argumentos assintóticos (grandes amostras) e é fortemente influenciado pela não linearidade do modelo, portanto, cuidados devem ser tomados ao se fazer inferências. Para uma avaliação da qualidade das inferências, vários pesquisadores propuseram medidas para quantificar a não linearidade e avaliar as suas consequências.

Box (1971) propôs uma estatística para avaliar o vício dos estimadores de mínimos quadrados dos parâmetros de um modelo de regressão não linear univariado, dada por:

$$B(\hat{\theta}) = -\frac{\sigma^2}{2} [\sum_{i=1}^n F(\theta)F^t(\theta)]^{-1} \sum_{i=1}^n F(\theta) \text{tr} [(\sum_{i=1}^n F(\theta)F^t(\theta))^{-1} H(\theta)]$$

em que  $F(\theta)$  é o vetor ( $p \times 1$ ) da primeira derivada do modelo e  $H(\theta)$  é a matriz ( $p \times p$ ) de segunda derivada, ambos com respeito aos elementos de  $\theta$ , avaliados em  $x_i, i = 1, \dots, n$ . Na prática, são usados  $\hat{\theta}$  e  $\sigma^2$  no lugar das quantidades desconhecidas e  $\text{tr}$  indica a operação traço.

A porcentagem de viés é uma medida mais interessante, pois é livre de escala e é dada por:

$$\%B(\hat{\theta}) = \frac{100 \times B(\hat{\theta})}{\hat{\theta}}$$

Segundo o autor, um vício aceitável seria aquele em que  $\%B(\hat{\theta}_i) < 1\%$ , valores de vício acima de 1%, em valor absoluto, indicam comportamento não linear. Em situações mais severas, deve-se fazer a correção das estimativas com translação dos intervalos de confiança na mesma quantidade.

A importância em se avaliar os vícios reside no fato deles indicarem qual ou quais parâmetros do modelo são os maiores responsáveis pelo distanciamento do comportamento linear (ZEVIANI et al., 2012). Uma vez conhecidos esses parâmetros, pode-se buscar por uma reparametrização que possa reduzir a não linearidade. O vício da nova parametrização pode ser avaliado por meio de outra estatística.

### Curvatura de Bates e Watts

Um tratamento explicitamente mais geométrico para a não linearidade foi apresentado por Bates e Watts (1980), que propuseram a quantificação da não linearidade em duas compo-

mentes. A não linearidade foi decomposta em duas componentes, como provaram os autores, a não linearidade intrínseca (IN) e aquela devida ao efeito de parâmetros (PE).

De acordo com os autores, a não linearidade intrínseca (IN) mede a curvatura do espaço de estimação no espaço amostral, em que o termo espaço de estimação se refere a todas as possíveis soluções do problema de mínimos quadrados. Enquanto a não linearidade devida ao efeito de parâmetros é consequência da falta de uniformidade do sistema de coordenadas no espaço de estimação.

Segundo Zeviani et al. (2012), essas medidas independem da escala da variável resposta e são úteis tanto para avaliar o comportamento de diferentes conjuntos de dados como de reparametrizações de um mesmo modelo.

A medida de não linearidade intrínseca em um modelo de regressão linear é zero, visto que o espaço de estimação é uma reta, um plano ou um hiperplano. Em um modelo não linear o espaço de estimação é curvilíneo, e (IN) mede a extensão dessa curvatura, dessa forma o valor da medida de não linearidade intrínseca não varia de acordo com o parâmetro escolhido.

Zeviani et al. (2012) explicam que a PE mede a falta de uniformidade do sistema de coordenadas da superfície esperada na vizinhança do local da solução e é dependente da parametrização do modelo. Corroboram Fernandes et al. (2015) que, diferente da medida de não linearidade intrínseca, a PE é sensível a alterações nos parâmetros dos modelos, portanto uma reparametrização do modelo pode significativamente alterar o valor da medida de não linearidade paramétrica, afetando a confiabilidade das estimativas.

A significância estatística de (IN) e (PE) é avaliada comparando seus valores com  $1/2\sqrt{F}$ , em que  $F = F_{p,n-p}^{\alpha}$  é o valor crítico obtido a partir de uma distribuição  $F$  com  $p$  e  $(n-p)$  graus de liberdade. Diel et al. (2019) explicam que o menor valores de não linearidade intrínseca (IN) e devida ao efeito de parâmetros (PE) indicam uma maior aproximação linear do modelo.

De acordo com Fernandes et al. (2015), grandes valores de curvatura intrínseca indicam intensa não linearidade da variável resposta, assim como um alto valor da curva paramétrica indica que a parametrização do modelo é responsável por uma maior distância da forma linear.

Portanto, em um modelo não linear com um dado valor de (IN), o valor de (PE) aumenta à medida que o seu comportamento se afasta do linear, uma vez que (PE) mede a extensão do comportamento não linear causado pela parametrização. Assim, se a não linearidade intrínseca (IN) for pequena, não significativa, pode ser que a não linearidade devido ao efeito de parâme-

tros (PE) seja alta, significativa, então, uma reparametrização muitas vezes pode reverter esse fato.

Uma boa aproximação linear garante que obtenhamos características como: estimadores não viesados, normalmente distribuídos, com variância mínima mesmo em pequenas amostras (MAZUCHELI; ACHCAR, 2002; SEBER; WILD, 2003; ZEVIANI et al., 2012; FERNANDES et al., 2015; DIEL et al., 2019).

## 2.7 Simulação de Monte Carlo (SMC)

O fundamento básico da simulação é a amostragem aleatória de valores para uma variável, a partir da distribuição de probabilidade da mesma. Para fazer amostragem em uma distribuição de probabilidade, precisa-se basicamente de um conjunto de números aleatórios e uma forma de transformar estes números em outro conjunto de números com a mesma distribuição.

A ideia de fazer várias simulações e analisar seus resultados baseia-se em um teorema bastante conhecido no campo da inferência estatística clássica: o teorema do limite central (GRINSTEAD e SNELL, 2012). De forma resumida, esse teorema propõe que, se gerarmos um número suficientemente grande de amostras e calcularmos alguma estatística a partir delas (média, mediana etc.), o valor médio destas estatísticas indicará o valor real desse parâmetro na população.

Meireles et al. (2017) discorrem, que em vez de coletar dados, formular uma hipótese e testá-la, as simulações permitem-nos gerar dados estipulando como eles devem se comportar. Por conta da possibilidade de controlar o número de simulações e as características das variáveis simuladas, uma das potencialidades do método é justamente emular um experimento em seu formato tradicional. Ao repetir o processo de geração de dados e utilizar aleatoriedade para investigar como pequenas variações alteram ou não um resultado, é possível examinar como pequenas mudanças na implementação de uma simulação afetam seus resultados (AXERLRD, 1997; MOONEY, 1997).

De acordo com Donatelli et al. (2005), a qualidade dos resultados obtidos na simulação de Monte Carlo irá depender dos seguintes fatores:

- representatividade do modelo matemático;
- qualidade da caracterização das variáveis de entrada;

- características do gerador de números pseudo-aleatórios utilizado;
- número de simulações realizadas (N);
- procedimento de definição do intervalo de abrangência

De acordo com Evans e Olson (1998), a correta identificação das distribuições de probabilidade dos dados de entrada é fundamental no processo de simulação. Vale ressaltar, também, que o número N de medições simuladas possui forte influência no erro amostral esperado para as estimativas obtidas por SMC.

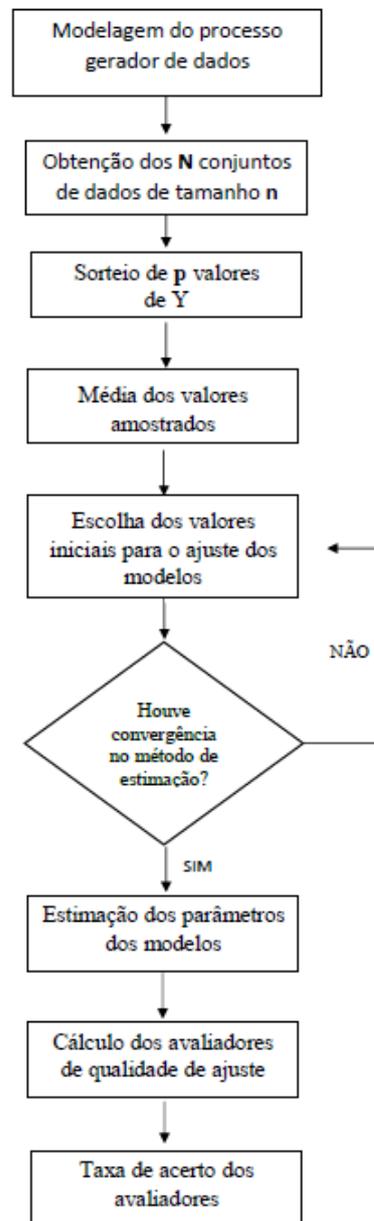
Para Moraes et al. (2016), a simulação de Monte Carlo é um método que consiste essencialmente em estabelecer uma amostragem artificial ou simulada, na medida em que um grande número de problemas econômicos, as observações para explicação de algum problema não encontram acesso junto à realidade por ser sua obtenção excessivamente custosa ou fisicamente impossível. Ela é umas das mais poderosas ferramentas, utilizadas na criação de estimativas (AMORIM et al., 2018). A SMC possibilita experimentar o efeito de intervenções, planos de contingência e proporciona maior suporte aos tomadores de decisão (YANG; TIAN, 2012).

Conforme Pagliarussi (2018), o uso de simulações permite a visualização dinâmica e a repetição de situações interessantes. De acordo com Kennedy (1998), uma razão importante para o uso de experimentos de Monte Carlo é propiciar o entendimento dos conceitos de amostragem repetida e propriedades da distribuição de amostragem de um estimador. No entanto, Martins et al. (2010) ressaltam que a simulação de Monte Carlo não fornece como resultado uma recomendação explícita para se tomar uma decisão, e sim um detalhamento para as possibilidades de resultado por meio de uma distribuição de frequência.

### 3 MATERIAL E MÉTODOS

#### 3.1 Simulação

Os dados analisados foram selecionados através de amostragens realizadas a partir de conjuntos de dados simulados pelo método de Monte Carlo utilizando equações de modelos de regressão não linear. A Figura 5 apresenta o fluxo de um processo para geração de modelos de regressão não linear por meio da simulação de Monte Carlo.



Fonte: Do autor

**Figura 5:** Estrutura do processo de simulação de Monte Carlo

em que:  $N$  representa o número de observações da variável dependente ( $Y_i$ ) que deseja-se gerar,  $n$  é o número de observações da variável independente ( $X_i$ ) e  $p$  é o tamanho da amostra que irá ser sorteada de  $N$  para representar  $Y_i$  na estimação.

Um estudo de Monte Carlo se inicia com a construção do modelo que permita ao computador imitar o processo gerador de dados, incluindo seu componente estocástico. A ideia da modelagem do processo gerador de dados consiste em utilizar a estrutura de algum modelo de regressão não linear (logístico, Gompertz, von Bertalanffy ou Brody) para a construção do modelo de simulação. Assim, após a escolha do modelo que irá gerar os dados do estudo, é necessário especificar alguns valores para que a estrutura do processo gerador de dados possa realizar esta simulação.

Todo modelo, seja de regressão linear ou não linear, é composto de parâmetros (que são estimados) e variáveis (dependentes e independentes). Dessa forma, o processo gerador dos dados, responsável pela obtenção dos valores que compõem as informações da variável dependente do modelo ( $Y_i$ ) necessita que as estimativas dos parâmetros e os valores da variável independente ( $X_i$ ) do modelo previamente escolhido sejam conhecidas e especificadas. Portanto, informações prévias dos modelos utilizados na geração dos dados deram suporte para que essas simulações alcançassem os objetivos almejados neste estudo.

Conhecendo os valores reais da variável independente ( $X_i$ ) que representa a função do tempo nas curvas de crescimento e os valores estimados dos parâmetros, a ideia foi utilizar o processo gerador de dados para simular 1.000 ( $N$ ) realizações da variável dependente do modelo selecionado inicialmente e, a estes valores, foram adicionados  $N$  valores de  $\varepsilon_i$ , o erro aleatório do modelo, gerados através de um processo de números pseudos aleatórios assumindo uma distribuição normal com média zero e variância  $\sigma^2$  (constante e conhecida), utilizando a função *rnorm* do *software* R. Vale ressaltar que é possível gerar erros a partir de uma distribuição de probabilidade diferente da normal.

Dessa forma, nesse estudo foram especificados 15 valores ( $n$ ) para a variável independente ( $X_i$ ) em termos de um erro aleatório ( $\varepsilon_i$ ), que produziram 1000 valores da variável dependente ( $Y_i$ ) para cada  $X_i$  por meio do processo de simulação de Monte Carlo. Seguindo esse racional, foram gerados dados em quatro cenários e os valores observados em cada cenário provém das simulações feitas utilizando as equações dos modelos logístico, Gompertz, von Bertalanffy e Brody. O esquema de cenários, o modelo simulado e os ajustes realizados em cada cenário estão definidos na Tabela 3.1.

Tabela 3.1 – Cenários simulados, modelo simulado e ajuste realizado em cada cenário.

Cenário	Modelo simulado	Ajuste realizado
1	logístico	logístico Gompertz von Bertalanffy Brody
2	Gompertz	logístico Gompertz von Bertalanffy Brody
3	von Bertalanffy	logístico Gompertz von Bertalanffy Brody
4	Brody	logístico Gompertz von Bertalanffy Brody

Os termos do erro aleatório geraram 1.000 valores seguindo distribuição normal com média zero e um desvio padrão proporcional à média, definido por  $\tau \times \bar{X}_i$ . O valor  $\tau$  foi definido após levantamento de referências bibliográficas sobre curvas de crescimento em diferentes cenários de modelagem de curvas de crescimento animal, vegetal e fruto, que disponibilizavam a informação do desvio padrão para a variável resposta. A partir desses valores observados de desvio em cada estudo foi calculada a proporção média dos desvios padrões em relação às médias em cada idade. Após o levantamento obteve-se como estimativa de  $\hat{\tau}$  o valor de 0,15.

Para o processo de simulação é necessário entender que o conjunto de dados depende crucialmente dos valores obtidos para o termo de erro. Um conjunto diferente de N valores de  $\varepsilon_i$  iria alterar significativamente os valores de  $Y_i$  obtidos para o mesmo problema.

Cada amostra obtida nos quatro cenários de simulação foi usada para estimar os valores dos parâmetros ( $\hat{\alpha}$ ,  $\hat{\beta}$  e  $\hat{k}$ ) dos modelos de regressão não linear que foram comparados. Portanto, para as 1.000 simulações dos 15 valores de  $X_i$  obtiveram-se 1.000 estimativas para gerar as curvas de crescimento. No entanto, para interesse prático, utilizar as 1.000 estimativas para comparação de curvas e os respectivos avaliadores não é um processo vantajoso e demanda muito tempo.

Portanto, a população do estudo representa os 1.000 valores simulados da variável dependente em função de cada  $X_i$ . Antes de ajustar os candidatos a modelo ideal e estimar seus parâmetros foi realizada uma amostragem aleatória de 20 valores (p) de  $Y_i$  na população para

cada ponto  $X_i$ , em seguida calculou-se a sua média. Esse resultado das médias formou o conjunto final de valores de  $Y_i$  (15 médias) para que a estimação dos parâmetros dos modelos fosse realizada e posteriormente a seleção do ajuste adequado fosse feita pelos avaliadores de qualidade de ajuste.

De forma ilustrativa, suponha que se deseja estudar o crescimento de bovinos, têm-se 1.000 bois no rebanho, selecionam-se aleatoriamente 20 deles e realiza-se a pesagem dos mesmos de 8 em 8 dias totalizando 15 pesagens. Calcula-se a média dos 20 pesos em cada idade e ajustam-se os 4 modelos estudados a este peso médio. Analisa-se qual modelo apresenta melhor ajuste com base nos avaliadores de qualidade. Como o "modelo original" é conhecido, repete-se a situação anterior 100 vezes para saber a porcentagem de acerto para cada avaliador na seleção do modelo correto.

### 3.2 Métodos

Com base na amostra de médias obtidas para cada um dos 4 cenários de simulação, foram ajustadas as curvas de crescimento e estimados os parâmetros dos modelos logístico, Gompertz, von Bertalanffy e Brody, havendo convergência, utilizando a função *nls* do *software* R. Cada cenário de simulação foi repetido 100 vezes, calculadas as estimativas dos parâmetro e selecionado o ajuste adequado por meio dos 13 avaliadores anteriormente citados.

A estimação dos parâmetros dos modelos foi realizada utilizando o algoritmo de convergência de Gauss-Newton ou método da linearização, que consiste na utilização de uma expansão da série de Taylor para aproximar o modelo de regressão não linear em termos lineares para, então, aplicar o método dos mínimos quadrados, encontrando assim a estimação dos parâmetros (MAZUCHELI; ACHCAR, 2002).

Utilizando todos os critérios para seleção do modelo adequado, descritos no referencial teórico, foi considerado como preferível o ajuste, dentre os quatro modelos ajustados (logístico, Gompertz, von Bertalanffy e Brody), aquele que estiver em conformidade com a conclusão da maioria dos avaliadores supracitados o ideal é que este ajuste selecionado como adequado seja o mesmo cuja a equação foi utilizada como base para a simulação dos dados amostrados no cenário em estudo.

Foi calculada a taxa de acerto para cada um dos avaliadores de qualidade de ajuste que selecionaram corretamente o ajuste ao qual gerou os dados amostrados, dentre as 100 repe-

tições do processo de simulação. Os critérios de seleção mais eficientes foram aqueles que apresentaram as maiores porcentagens de distinção do modelo correto.

Todas as análises e a parte computacional foram realizadas pelo *software* estatístico R (R DEVELOPMENT CORE TEAM 2021), utilizando os pacotes que estão disponíveis para as estimação dos parâmetros, testes estatísticos e gráficos.

#### 4 RESULTADOS E DISCUSSÃO

Os resultados aqui apresentados foram obtidos seguindo o raciocínio de simulação apresentado na seção anterior. Inicialmente os cenários simulados foram resumidos em medidas descritivas. A Tabela 4.1 apresenta estes resultados para os quatros modelos simulados (logístico, Gompertz, von Bertalanffy e Brody) e os valores dos parâmetros escolhidos para cada cenário de simulação.

Tabela 4.1 – Análise descritiva das simulações e respectivos parâmetros escolhidos para a equação de cada modelo simulado.

Modelo ( $\alpha;\beta;k$ )	Medidas	Resultado
Logístico (300;60;0,1)	Mínimo	0,7418
	Máximo	299,2582
	1° Quartil	14,8927
	3° Quartil	285,1073
	Média	150,0000
	Mediana	150,0000
	Desvio Padrão	128,5466
Gompertz (300;60;0,05)	Mínimo	0,0000
	Máximo	239,4010
	1° Quartil	4,0770
	3° Quartil	239,4010
	Média	124,8220
	Mediana	110,3710
	Desvio Padrão	116,2590
von Bertalanffy (300;1;0,04)	Mínimo	0,0000
	Máximo	292,6540
	1° Quartil	102,3000
	3° Quartil	275,9000
	Média	184,4312
	Mediana	224,8455
	Desvio Padrão	107,5118
Brody (300;0,3;0,06)	Mínimo	210,0000
	Máximo	299,9328
	1° Quartil	284,8000
	3° Quartil	299,6000
	Média	285,2514
	Mediana	297,4697
	Desvio Padrão	25,3580

Cada modelo simulado apresenta um comportamento de curva diferente, como foi exemplificado pelas figuras apresentadas na seção 2. O comportamento, em números, de cada curva de crescimento nas simulações feitas estão detalhadas na Tabela 4.1 através da sua caracterização e escolha de cada parâmetro.

O parâmetro  $\alpha$ , que representa a assíntota (ponto mais alto que a curva atinge), foi fixado no valor 300 para melhor explicação dos resultados em todas as simulações dos modelos realizadas, enquanto os demais parâmetros ( $\beta$ ,  $k$ ) foram alterados de acordo com as características de curvas e do próprio modelo.

Pôde-se observar a simetria do modelo logístico através dos resultados média e mediana iguais (150, ambos), como também o crescimento mais acelerado no modelo Brody, seu valor mínimo já é alto e apresenta baixo desvio, já que seu valores permanecem com baixa variação a partir de determinado ponto.

Realizada todas as 4 simulações dos modelos de regressão não linear, obtidas as médias para cada cenário (logístico, Gompertz, von Bertalanffy e Brody) e replicadas as mesmas condições 100 vezes, foram calculados os avaliadores de qualidade de ajuste para todos os modelos dentro destes cenários. Assim, a Tabela 4.2 apresenta as estimativas médias para os parâmetros dos modelos logístico, Gompertz, von Bertalanffy e Brody.

Tabela 4.2 – Estimativa média para os parâmetros dos modelos logístico, Gompertz e von Bertalanffy, considerando os quatros cenários de simulação.

Modelo Simulado ( $\alpha, \beta, k$ )	Modelos Ajustados ( $\alpha_0; \beta_0; k_0$ )	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$
Logístico (300;60;0,1)	Logístico (250;40;0,02)	300,8145	60,11411	0,09940
	Gompertz (250;40;0,02)	310,8931	53,97854	0,06366
	von Bertalanffy (250;40;0,02)	379,0493	50,80720	0,02792
	Brody (250;40;0,02)	N/A	N/A	N/A
Gompertz (300;60;0,05)	Logístico (200;40;0,03)	282,4438	66,79975	0,08329
	Gompertz (200;40;0,03)	301,4185	60,17520	0,04980
	von Bertalanffy (200;40;0,03)	395,5582	60,63712	0,02307
	Brody (200;40;0,03)	N/A	N/A	N/A
von Bertalanffy (300;1;0,04)	Logístico (250;0,3;0,02)	284,8101	39,81492	0,07508
	Gompertz (250;0,3;0,02)	294,5482	31,79921	0,04839
	von Bertalanffy (250;0,3;0,02)	300,6005	27,48322	0,04002
	Brody (250;0,3;0,02)	373,9251	5,28607	0,01535
Brody (300;0,3;0,06)	Logístico (250;0,3;0,04)	300,4052	-12,42273	0,07188
	Gompertz (250;0,3;0,04)	300,6736	-16,33309	0,06539
	von Bertalanffy (250;0,3;0,04)	300,7735	-17,77549	0,06330
	Brody (250;0,3;0,04)	300,9916	-20,89928	0,05918

A Tabela 4.2 apresenta o modelo simulado juntamente com os parâmetros escolhidos ( $\alpha$ ,  $\beta$ ,  $k$ ) para que fosse possível gerar os dados na equação de cada modelo de simulação. Ela apresenta, também, os modelos ajustados em cada cenário simulado e os valores iniciais ( $\alpha_0$ ,  $\beta_0$ ,  $k_0$ ) em cada ajuste. Por fim, a estimação média de cada parâmetro ( $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{k}$ ), em todos os

cenários e para cada ajuste que obteve convergência na estimação, para os modelos no qual não houve convergência os resultados dos parâmetros estão identificados como N/A (Não aplicado).

Nota-se através dos resultados listados na Tabela 4.2 que os cenários de simulação dos modelos logístico e Gompertz não obtiveram convergência para o ajuste do modelo Brody em suas repetições. Uma possível explicação seria o formato sigmoidal desses modelos, enquanto o modelo Brody apresenta formato de curva diferente dos demais modelo de regressão não linear.

Para todas as simulações os demais resultados dos ajustes de seus respectivos modelos se adequaram bem e apresentaram estimativas médias dos parâmetros próximas aos valores iniciais simulados. É importante perceber que os valores iniciais dos ajustes foram escolhidos bem abaixo dos valores iniciais simulados, em todos os cenários, isso foi feito para verificar o quão preciso seria o modelo para estimar seus parâmetros.

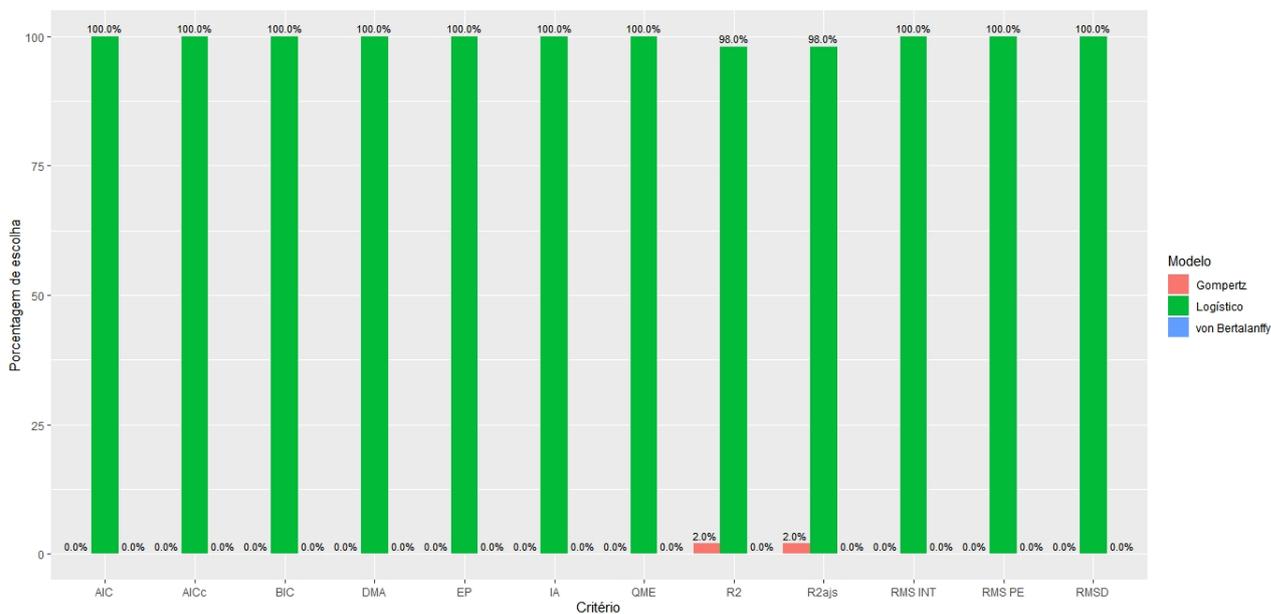
Observando as estimativas médias para os ajustes de modelos utilizando os valores simulados, percebe-se que para o cenário de simulação do modelo von Bertalanffy, a estimativa do parâmetro  $\hat{\beta}$  no ajuste do respectivo modelo foi superestimada, sendo 1 o valor real enquanto a estimativa média desse parâmetro no ajuste do modelo von Bertalanffy utilizando os dados simulados deste mesmo modelo foi de 27,48, enquanto os demais parâmetros obtiveram estimativas médias próximas as originais.

Outro ponto importante a ser frisado é, apesar da estimativa do parâmetro  $\hat{\beta}$  do modelo Brody apresentar valores médios negativos em suas estimativas, isto não impacta nos resultados de seleção do modelo adequado e escolha do melhor avaliador, visto que o parâmetro  $\beta$  para este modelo não possui interpretação prática como os demais modelos aqui estudados.

Após os ajustes de todos os modelos para cada cenário em todas as suas 100 repetições, a avaliação do modelo mais adequado foi realizada pelos critérios de qualidade de ajuste, citados na seção 2. A Figura 6 apresenta o percentual de escolha do modelo adequado para a simulação considerando a equação do modelo logístico em suas 100 replicações.

Cada coluna discriminada pela cor, na Figura 6, representa o modelo ajustado aos dados simulados e suas respectivas porcentagens de escolha por cada avaliador de qualidade de ajuste ao modelo adequado aquela simulação.

Percebe-se que nesse cenário não houve discrepâncias em relação à escolha do modelo mais adequado entre os avaliadores apresentada na Figura 6, considerando a simulação através da equação do modelo logístico.



Fonte: Do autor

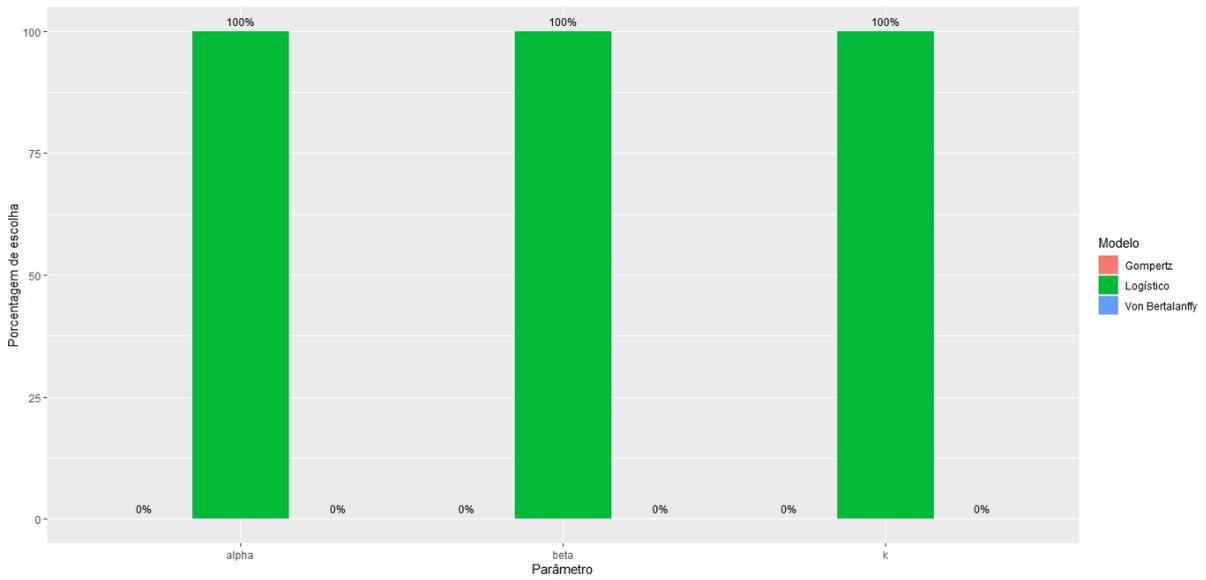
**Figura 6:** Gráfico do percentual de escolha dos avaliadores para os ajustes de modelos considerando para o modelo simulado logístico

Em sua totalidade, os avaliadores escolheram o ajuste do modelo logístico como mais adequado para representar o cenário 1 de simulação, o qual utilizou a própria equação do modelo logístico para gerar os dados, representando 100% das escolhas em suma dos critérios listados, exceto para os avaliadores coeficiente de determinação ( $R^2$ ) e coeficiente de determinação ajustado ( $R^2_{ajs}$ ), em 98% das repetições estes avaliadores consideraram o modelo logístico como preferível, a porcentagem complementar foi destinada ao modelo Gompertz, o qual apresenta comportamento mais próximo ao modelo logístico.

Quanto a utilização desses avaliadores na escolha de um modelo de regressão não linear, Deprá et al. (2016), constatou que por meio dos valores de  $R^2$  e DMA, no ajuste do modelo logístico, indicaram que este modelo explicou a maior parte da variabilidade nas variáveis altura de planta e número de folhas por planta para os genótipos de cultivares crioulas de milho e de progênies de meios-irmãos maternos em função da soma térmica.

Os resultados do vício de Box são apresentados em um gráfico a parte, a Figura 7 traz as informações de percentual de escolha para menores vícios em cada parâmetro dos modelos ajustados.

Novamente, foi perceptível a concordância do ajuste mais adequado em relação aos parâmetros e o vício de Box, todos selecionaram o ajuste do modelo logístico como mais adequado, obtendo 100% da porcentagem de escolha a esse ajuste nas repetições. Assim o ajuste

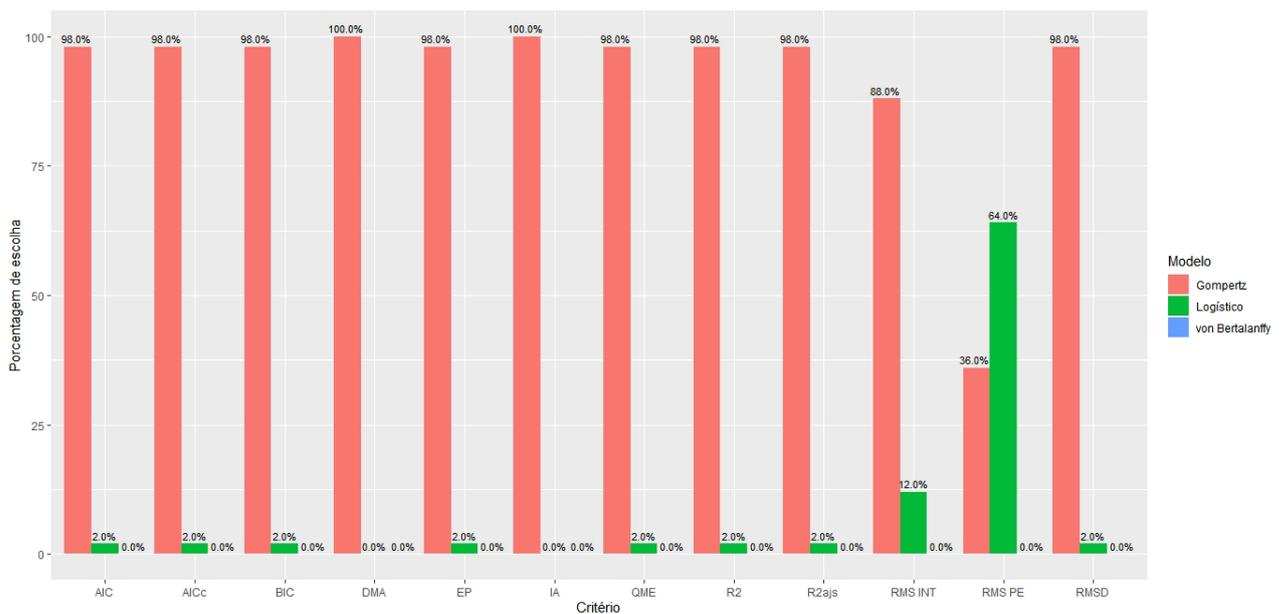


Fonte: Do autor

**Figura 7:** Gráfico do percentual do vício de Box para escolha do ajuste e modelos considerando o modelo simulado logístico

do modelo logístico no cenário de simulação para este mesmo modelo apresentou para todos os parâmetros menores vícios de Box.

Já para o cenário 2 de simulação, através da equação do modelo Gompertz, podem-se perceber algumas mudanças em relação aos resultados ao apresentado no cenário anterior, como mostra a Figura 8.



Fonte: Do autor

**Figura 8:** Gráfico do percentual de escolha dos avaliadores para os ajustes de modelos considerando o modelo simulado Gompertz

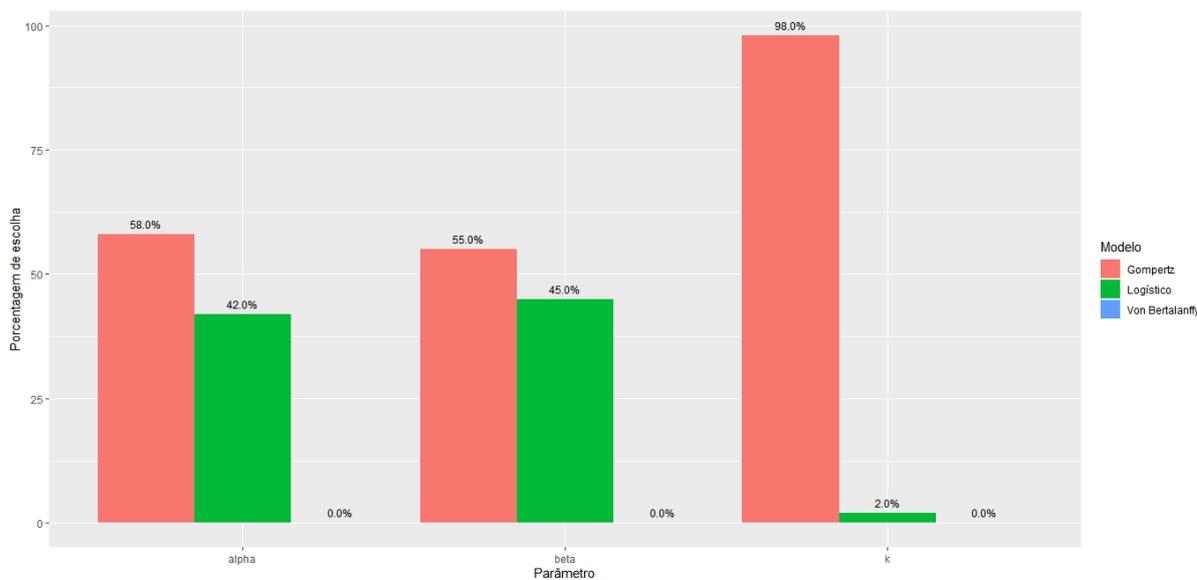
A Figura 8 deixa evidente que o ajuste do modelo Gompertz foi o mais adequado em representar a simulação do cenário 2, reflexo das altas porcentagens de escolha desse modelo pela maioria dos avaliadores aqui apresentados. Dos 12 avaliadores contidos nessa figura, 11 deles apresentaram maiores porcentagens de escolha para este modelo, todos estes com porcentagens superior a 85% de escolha.

Pode-se destacar, mais evidente que a simulação anterior, que os avaliadores Índice Assintótico (IA) e Desvio Médio Absoluto (DMA) foram levemente mais precisos que os demais para esse cenário. Em 100% das repetições estes selecionaram o ajuste do modelo Gompertz como a "equação original" de simulação dos dados, enquanto os avaliadores Critério de Informação de Akaike (AIC), Critério de Informação de Akaike Corrigido (AICc), Critério de informação Bayesiano (BIC), Erro de predição médio ( $EP_M$ ), Quadrado Médio dos Resíduos (QME),  $R^2$ ,  $R^2_{ajs}$  e Raiz do erro quadrático médio (RMSD) selecionaram o mesmo modelo somente em 98% das repetições.

Drumond et al. (2013) no estudo de crescimento para codornas de corte explicam que considerando os critérios para qualidade de ajuste do modelo, observou que o  $R^2$  não seria um bom indicador para escolha dos modelos, pois todos apresentaram valores altos e semelhantes. Sendo assim, selecionaram os modelos que mais se ajustaram aos dados através do Índice Assintótico. Assim, foram observado menores valores de IA para o modelo de Gompertz em codornas machos e para o logístico em fêmeas.

As Curvaturas de Battes e Watts (Intrínseca e Paramétrica) não chegaram à mesma conclusão na escolha do ajuste adequado, enquanto a curvatura intrínseca (RMS INT) selecionou corretamente o modelo ideal em 88% das repetições, a da curvatura paramétrica (RMS PE) não chegou à mesma conclusão em aceitar o modelo Gompertz como a equação original, as avaliações da curvatura paramétrica concluem que o ajuste logístico seria o modelo ideal para 64% das repetições, somente em 36% das replicações o ajuste do modelo Gompertz seria satisfatório para esse avaliador.

Considerando a escolha do modelo adequado através do vício de Box, a Figura 9 demonstra que em todos os parâmetros os menores vícios obtidos, em maiores percentuais, foi o ajuste do modelo Gompertz. Considerando os menores vícios o parâmetro  $k$ , que representa a primeira derivada calculada no ponto de inflexão, apresentou o maior percentual de escolha para o modelo Gompertz, 98% dos casos. Por sua vez parâmetro  $\beta$  considerou o ajuste do modelo Gompertz como menor vício em 55% dos ajustes, enquanto para a assintótica, parâmetro



Fonte: Do autor

**Figura 9:** Gráfico do percentual do Vício de Box para escolha do ajuste e modelos considerando o modelo simulado Gompertz

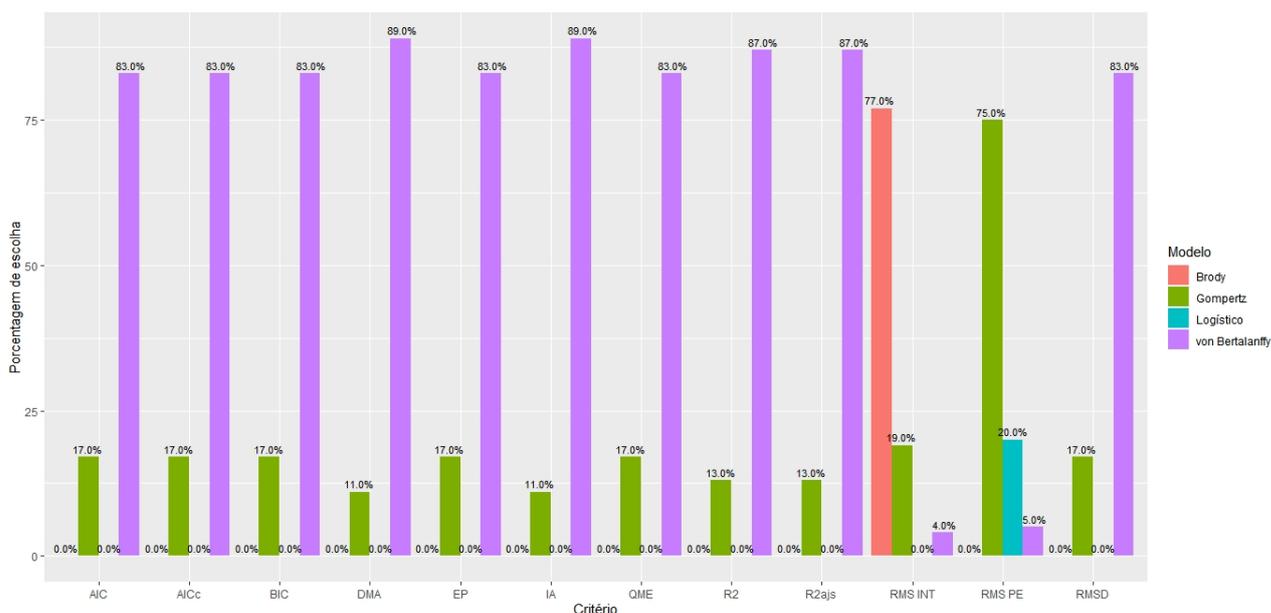
$\alpha$ , 58% dos ajustes apresentaram menores vícios para o ajuste da equação original dos dados, o modelo Gompertz.

Através da simulação do cenário 3 considerando a equação do modelo von Bertalanffy e em seguida ajustando as 100 repetições de simulação, pode-se encontrar os resultados para os avaliadores de qualidade de ajuste que são apresentados na Figura 10.

Os avaliadores de qualidade de ajuste para esse cenário, em sua maioria, concordaram que o modelo adequado e o qual simulou os dados iniciais foi o von Bertalanffy, destaca-se o comportamento de escolha para cada avaliador e suas porcentagens de escolha do modelo adequado.

Apesar da alta taxa de acerto da maioria dos avaliadores, nesse cenário, em selecionar o ajuste ideal o qual provem da equação que simulou os dados, é nítida uma maior precisão de acerto para os avaliadores DMA e IA pela alta porcentagem de seleção do "modelo correto", ambos concordam em 89% das simulações que o modelo o qual gerou os dados é o von Bertalanffy, porcentagem superior a todos os demais avaliadores aqui abordados.

A utilização do IA em estudos de curvas de crescimento e sua efetividade na escolha do modelo adequado é ressaltado no estudo de Veloso et al. (2015), os mesmos utilizaram o índice assintótico para seleção do modelo adequado por se tratar de um critério de avaliação mais completo, assim, pôde-se observar que o melhor ajuste foi apresentado pela curva proposta por



Fonte: Do autor

**Figura 10:** Gráfico do percentual de escolha dos avaliadores para os ajustes de modelos considerando o modelo simulado von Bertalanffy

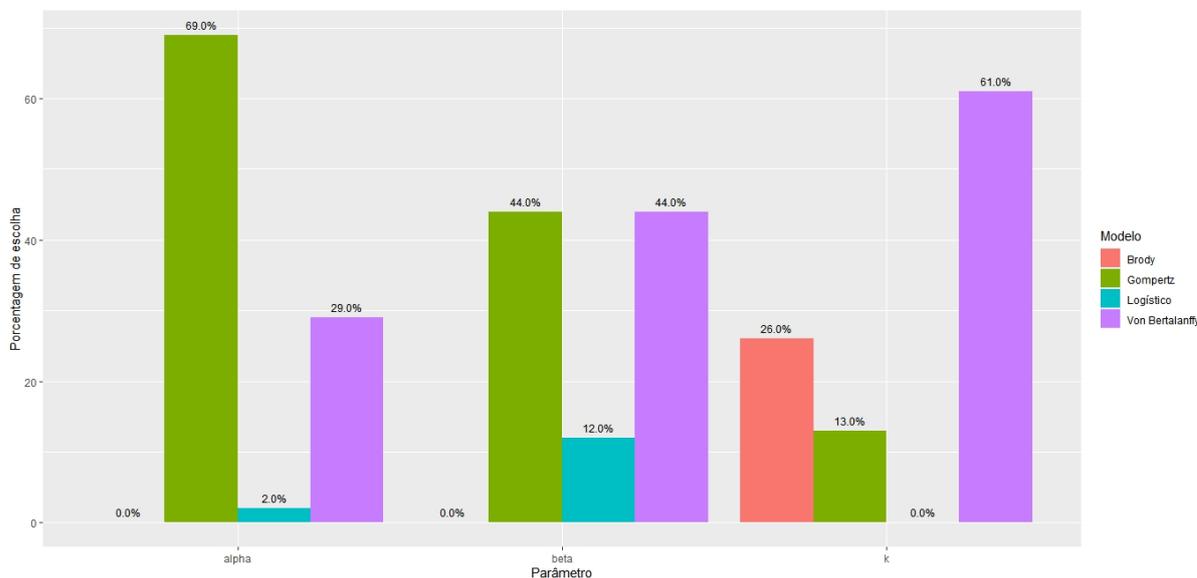
von Bertalanffy, sendo este o recomendado para descrever o padrão de crescimento das aves dos genótipos de frangos tipo caipira avaliados no estudo.

O coeficiente de determinação e sua forma ajustada ( $R^2$  e  $R^2_{ajs}$ ) demonstram uma precisão de 87% de acerto na escolha do ajuste do modelo von Bertalanffy. Enquanto os avaliadores AIC, AICc, BIC, EP, QME e RMSD formaram um grupo de comportamento semelhante quanto a seleção do ajuste adequado, em 83% das repetições estes consideraram a equação do modelo von Bertalanffy como a matriz de geração de dados simulados.

Araújo et al. (2020) utilizaram os modelos Gompertz e logístico na determinação de Curva de crescimento e desempenho de cabras leiteiras alimentadas com torta de mamona destoxicada por diferentes soluções alcalinas. Observou que ambos os modelos tiveram bons ajustes, contudo, o modelo logístico apresentou maior  $R^2$  e, ao mesmo tempo, menor índice assintótico (IA).

Destaca-se o quão imprecisas foram as curvaturas de Battes e Watts em selecionar o modelo correto que simulou os dados deste cenário. Segundo as medidas de curvatura intrínseca e a paramétrica, os dados provêm de uma equação do modelo von Bertalanffy em apenas 4 e 5%, respectivamente, das 100 repetições ajustadas. O acerto desses avaliadores na escolha da equação correta que gerou os dados neste cenário foi a menor entre todos os simulados. Segundo esses avaliadores, a curvatura intrínseca aponta que em 77% das repetições os dados apresen-

tam um comportamento característico do modelo Brody e 75% das repetições pela curvatura paramétrica tem comportamento semelhante ao modelo Gompertz.



Fonte: Do autor

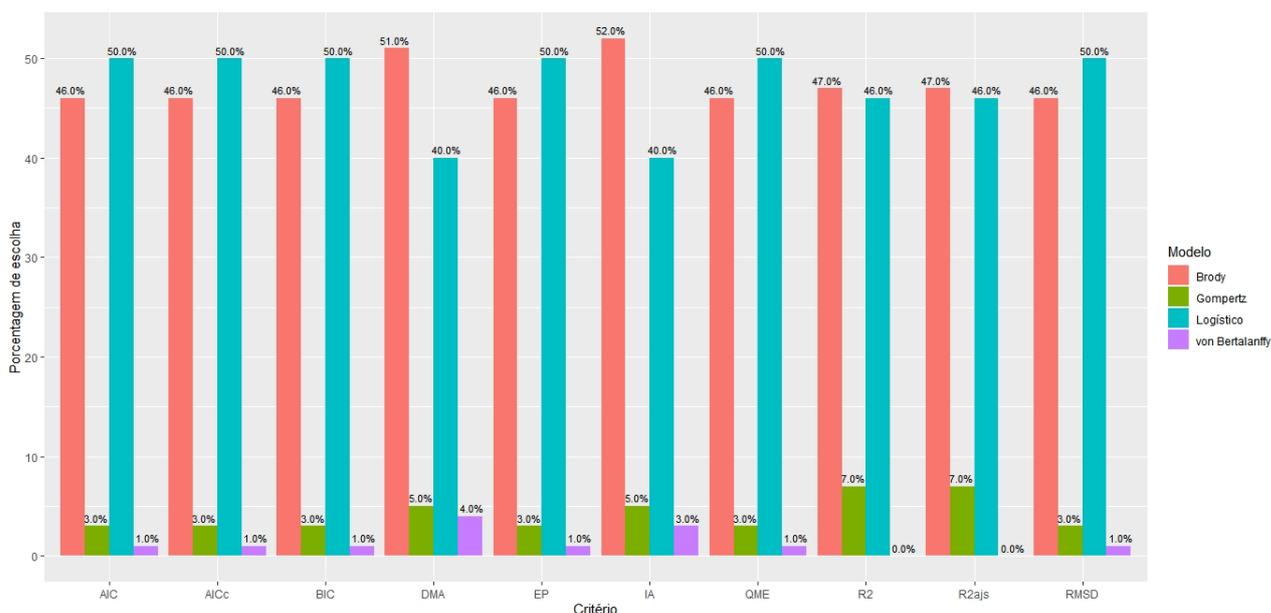
**Figura 11:** Gráfico do percentual do Vício de Box para escolha do ajuste e modelos considerando o modelo simulado von Bertalanffy

A Figura 11 apresenta os percentuais para os ajustes com menores vícios por parâmetros. O parâmetro  $k$  neste cenário apresentou a maior taxa de acerto considerando o menor vício, em 61% dos ajustes o modelo von Bertalanffy foi considerado ideal para representar os dados simulados. Para o vício de  $\alpha$  somente em 29% das simulações, para o modelo inicialmente simulado, apresentou menores resultados de Vício de Box, enquanto para esse mesmo parâmetro 69% dos ajustes apresentam menores vícios para o modelo Gompertz.

Por fim, para o parâmetro  $\beta$ , não houve preferência de ajuste considerando o menor vício sendo a seleção do "modelo real", von Bertalanffy, foi responsável por 44% das repetições ajustadas, empatado com os ajustes do modelo Gompertz. Assim, os resultados demonstram que o Vício de Box não foi eficiente na seleção do ajuste adequado neste cenário de simulação.

O cenário 4 de simulação, considerando a simulação através da equação do modelo Brody, inclusive por apresentar um padrão de crescimento diferente dos demais, trouxe resultados chave para avaliação dos critérios de qualidade de ajuste e definição do objetivo desse estudo, os resultados desse cenário são apresentado na Figura 12.

De acordo com os resultados obtidos pelos avaliadores de qualidade de ajuste neste cenário, pode-se verificar que a maioria destes foram equivocados em relação à seleção do modelo inicial de simulação. Segundo os avaliadores e os maiores percentuais de escolha o ajuste do



Fonte: Do autor

**Figura 12:** Gráfico do percentual de escolha dos avaliadores para os ajustes de modelos considerando o modelo simulado Brody

modelo logístico seria o ideal, enquanto o modelo Brody que gerou os dados simulados só foi considerado corretamente apenas pelos avaliadores DMA, IA,  $R^2$  e  $R^2_{ajs}$ .

O grupo composto por AIC, AICc, BIC, EP, QME e RMSD apresentaram maiores escolha para o modelo logístico em termos percentuais, em 50% das repetições este avaliadores acreditaram que os dados originários são provenientes de uma equação do modelo logístico, enquanto somente em 46% do total de simulações esses avaliadores acertaram em relação ao modelo adequado, selecionar o modelo Brody como o correto. O  $R^2$  e  $R^2_{ajs}$  foram assertivos em 47% das repetições ao escolher como adequado o modelo Brody neste cenário.

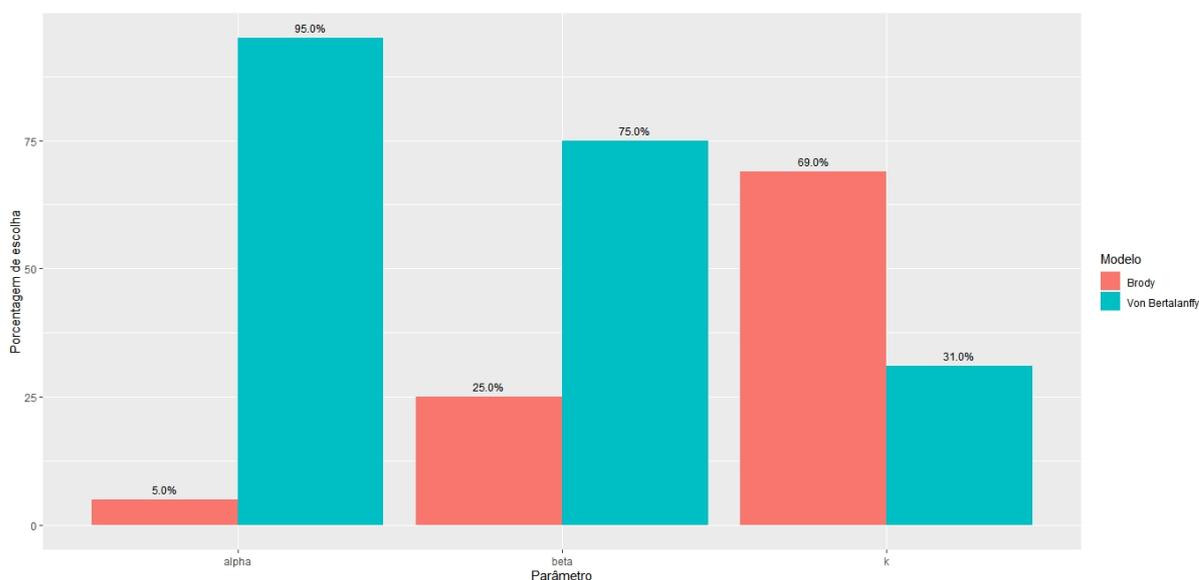
Veloso et al. (2016) recomendam utilizar outros avaliadores de qualidade de ajuste ao invés de utilizar o  $R^2_{ajs}$ , pois as diferenças entre os  $R^2_{ajs}$  dos diferentes modelos usados no estudo de curvas de crescimento foram irrisórias na seleção. Percebeu-se que os melhores grupos de ajuste ao crescimento de frango de corte foram os que apresentaram, de forma geral, menores valores de QME, AIC, BIC e IA considerados simultaneamente.

Em relação a acurácia da seleção, o Desvio Médio Absoluto (DMA) selecionou o modelo Brody como adequado em 51% das simulações, sendo a diferença para o segundo modelo mais escolhido (logístico) como preferível de 11%, pode-se dizer que este avaliador foi mais assertivo que os anteriores.

A utilização do DMA como avaliador de ajuste adequado para curvas de crescimento foi aplicado por Lopes et al. (2016) em seu estudo sobre as análises de dados longitudinais em bovinos Nelore Mocho por meio de modelos não lineares. Os autores afirmam que dentre todos os modelos analisados, o modelo Brody apresentou menor variância residual, indicada pela pequena estimativa do desvio médio absoluto (DMA). Como também, o resultado do coeficiente de determinação selecionou bem o ajuste mais adequado.

Entretanto, o Índice Assintótico foi ainda mais preciso na escolha do modelo Brody para este cenário, visto que este inicialmente gerou os dados simulados. Este avaliador obteve porcentagem de acerto em 52% das repetições ajustadas, 12% de diferença para escolha do modelo logístico, considerado segundo ajuste mais adequado pelo mesmo avaliador. Portanto, este avaliador obteve a maior precisão frente aos demais estudados.

Para esse cenário foi observado que a curvatura de Battes e Watts não apresentou convergência em seus resultados, assim não possível avaliar o comportamento desse critério para a seleção do modelo adequado. Já os resultados para o Vício de Box só obtiveram convergência para os ajustes dos modelos von Bertalanffy e Brody e seus percentuais são apresentados na Figura 13.



Fonte: Do autor

**Figura 13:** Gráfico do percentual do Vício de Box para escolha do ajuste e modelos considerando o modelo simulado Brody

Os ajustes do modelo von Bertalanffy apresentou menores Vícios para os parâmetros  $\alpha$  e  $\beta$  em 95% e 75% neste cenário de simulações, respectivamente. Enquanto a escolha para o

modelo Brody foi satisfatória somente para o parâmetro  $k$ , o qual apresentou menores vícios em 69% das repetições.

Algumas literaturas discorrem sobre a utilização desses avaliadores e corroboram com os resultados aqui encontrados. Autores concordam que a seleção de modelos ajustados é bem mais eficiente através dos avaliadores  $R^2$ , DMA e IA, sendo o  $R^2$  o avaliador mais utilizado na literatura para seleção de modelos não lineares.

A efetividade do coeficiente de determinação na escolha do modelo adequado é questionada por alguns autores em estudo de curvas de crescimento. Spiess et al. (2010) em uma simulação de Monte Carlo para avaliar os resultados de  $R^2_{ajs}$  em ajustes de modelos não lineares na Farmácia e Biomedicina verificou que os valores  $R^2_{ajs}$  diminuíam e os valores AICc aumentavam com o erro mais alto, no entanto, a maior crítica dos autores quanto ao uso do  $R^2$  é devido a sua ineficiência na seleção dos modelos no cenário simulado, de modo que as diferenças entre o  $R^2$  dos diversos modelos ocorreram apenas na quarta casa decimal.

De modo geral, pode-se concluir que a maioria dos avaliadores estudados no presente trabalho escolheram corretamente em alguns dos cenários, através dos gráficos apresentados (Figs 6 a 13). No entanto, alguns avaliadores foram mais eficientes que outros ao selecionar a equação de ajuste mais adequada e que inicialmente simulou os dados ajustados. Para os avaliadores  $R^2$ ,  $R^2_{ajs}$ , AIC, AICc, BIC, EP, QME, RSMD, DMA e IA, pode-se comprovar que em baixas variações no formato da curva, como no caso do Logístico e Gompertz, a precisão ou eficiência desses avaliadores chegam a ser muito próximas.

Considerando todos os cenários estudados, destacam-se os avaliadores DMA e IA sua alta eficiência na escolha do modelo adequado. Estes avaliadores selecionaram sempre o modelo correto, em maiores percentuais, em todos os cenários. Foi mais expressiva esta superioridade nas simulações e ajustes das equações dos modelos von Bertalanffy e Brody em que os demais avaliadores não apresentaram bons resultados. Ressalta-se que o percentual de acerto destes avaliadores foi sempre acompanhado pelos avaliadores  $R^2$  e  $R^2_{ajs}$  que também mostraram-se eficientes e úteis como avaliadores para seleção de modelos de regressão não linear, principalmente por fornecerem uma ideia simples e clara do quanto a variação dos dados é explicada pelo modelo em estudo.

Os critérios AIC, AICc e BIC apresentaram sempre resultados semelhantes em sua escolha de modelo ideal, variando apenas a escala de comparação, mas sempre concordando na indicação do modelo mais adequado (mesmo quando erram). Assim, utilizar apenas um destes

critérios seria suficiente para selecionar o modelo de ajuste mais adequado. Como também, as conclusões de EP, QME e RMSD seguiram o mesmo padrão destes avaliadores anteriormente citados.

O Cp de Mallows não apresentou resultados significativos nesse estudo. Para todos os modelos e diferentes cenários de simulação, o seu valor permaneceu constante, este avaliador não mostrou-se eficiente na classificação destes modelos não lineares.

As medidas de não linearidade embora necessárias para a avaliar a adequabilidade das aproximações lineares no momento de estimação dos parâmetros, percebeu-se que as curvaturas de Battes e Watts bem como também o Vício de Box não foram eficazes na seleção do modelo mais adequado. A eficiência em selecionar modelos por meio destas medidas não foi satisfatória na maioria dos cenários de simulações, de modo que, não concluíram corretamente a equação original que gerou o dados.

Portanto, sugere-se que aos pesquisadores utilizarem estas medidas de não linearidade apenas para verificar a adequabilidade do modelo, verificando se as curvaturas de Battes e Watts são significativas e se os vícios são pequenos, por exemplo. Mas não devem ser utilizadas como critérios de seleção (quanto menor melhor), conforme comumente encontrados em trabalhos da literatura como Sari et al. (2018), Diel et al. (2018), Diel et al. (2020) e Silva et al. (2019).

## 5 CONCLUSÃO

Considerando os cenários de simulação de curvas de crescimento com os diferentes formatos e avaliando o ajuste dos modelos: logístico, Gompertz, von Bertalanffy e Brody, conclui-se que:

- Para a seleção do melhor modelo devem ser utilizados os avaliadores índice assintótico (IA), desvio médio absoluto (DMA) e coeficiente de determinação ( $R^2$ ) (ou  $R_{ajs}^2$  se tiver modelos com diferentes números de parâmetros) nesta ordem para decidir qual o modelo mais adequado na descrição dos dados. Caso julgue necessário mais algum avaliador, pode-se optar por apenas um dentre critério de informação de Akaike (AIC), critério de informação de Akaike corrigido (AICc) ou critério de informação de Bayesiano (BIC) e também por apenas um entre erro de predição médio (EP), quadrado médio dos resíduos (QME) e a raiz do erro quadrático médio (RMSD).
- As medidas de Curvatura de Battes e Watts e Vício de Box, embora importantes para a avaliação da adequabilidade do ajuste dos modelos, não são indicados para seleção do melhor modelo.

## 6 APLICAÇÃO NA DESCRIÇÃO DO CRESCIMENTO DE EUCALIPTOS

### 6.1 Introdução

A madeira plantada da espécie do gênero *Eucalyptus* tem grande potencial devido à sua disponibilidade em curto espaço de tempo e utilização na produção de matéria prima para celulose, carvão vegetal, moirões e postes. Segundo dados do IBGE, em 2018 o Brasil possuía em torno de 9,9 milhões de hectares de florestas plantadas, dos quais 7,5 milhões eram de eucalipto, sendo a região sudeste responsável por 42,3% da área plantada total no Brasil, principalmente nos estados de Minas Gerais (26,1%) e São Paulo (12,2%). Para a máxima utilização desses hectares, pesquisadores e madeireiros têm buscado formas mais eficientes principalmente no melhoramento e crescimento das árvores.

Dentre as características das árvores, a altura, assim como o diâmetro, constitui-se em uma importante variável sendo, esta, determinada ou estimada, servindo, para o cálculo do volume e para o cálculo de incrementos em altura e em volume. Retslaff et al. (2015) ressaltam que, em inventários florestais, determinar a altura de todas as árvores pode acarretar mais erros que estimá-la, em razão das dificuldades de visualização, além de ser uma atividade que demanda tempo e gera custos maiores. Mendonça et al. (2017) afirmam que o volume de madeira, tanto presente quanto futuro, são os itens mais importantes no planejamento florestal.

Especificadamente na área florestal, alguns trabalhos na literatura utilizam os modelos não lineares para a descrição da curva de crescimento. Venduscrolo et al. (2016) avaliou a modelagem por regressão não linear e por redes neurais artificiais para a estimativa da altura de árvores de eucaliptos. Coutinho et al. (2017) ajustaram os modelos de regressão não linear Weibull, Logístico, Gompertz, Schumacher e Chapman-Richards, para avaliar o padrão de crescimento e descrever a distribuição probabilística do incremento corrente anual em diâmetro de *Cryptomeria japonica*. Santos et al. (2017) através dos modelos não lineares estimou características como altura dominante, diâmetro médio, área basal e volume com casca, em função da idade.

Desta forma, o objetivo desta aplicação é selecionar dentre os modelos de regressão não linear logístico, Gompertz, von Bertalanffy e Brody o mais adequado para descrever o crescimento em volume de madeira dos híbridos *Eucalyptus urophylla x Eucalyptus grandis* de acordo com os avaliadores  $R^2$ ,  $R^2_{ajs}$ , DMA e IA.

## 6.2 Material e Métodos

### 6.2.1 Material

Foram utilizados dados retirados de Gonçalves et al. (2017) referentes a plantios de híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla* históricos e ativos que se encontravam a partir do segundo ciclo, com referência em dezembro de 2012. Para a estimativa das curvas de crescimento dos plantios de eucalipto foram utilizadas as informações sobre a idade (1 a 15 anos, totalizando 15 medidas no tempo) e a produção volumétrica ( $m^3/ha$ ) de madeira dos plantios monoclonais.

Os dados do presente estudo são provenientes do Inventário Florestal junto à empresa Fibria Celulose S.A., regional Aracruz, com sede no município de Aracruz, Espírito Santo, Brasil. A temperatura média da região é de 28°C e precipitação anual média de 1.200mm. A predominância de solo na região são o Latossolo Vermelho Amarelo Distrófico e Podzólico Vermelho Amarelo (Prefeitura Municipal de Aracruz, 2013).

### 6.2.2 Métodos

A análise de crescimento das plantas de *Eucalyptus grandis* x *Eucalyptus urophylla* foi feita por meio de modelos não lineares. Foram ajustados os modelos Logístico (6.1), Gompertz (6.2), von Bertalanffy (6.3) e Brody (6.4).

$$Y_i = \frac{\alpha}{1 + e^{k(\beta - x_i)}} + \varepsilon_i \quad (6.1)$$

$$Y_i = \alpha e^{-e^{k(\beta - x_i)}} + \varepsilon_i \quad (6.2)$$

$$Y_i = \alpha \left[ 1 - \frac{e^{k(\beta - x_i)}}{3} \right]^3 + \varepsilon_i \quad (6.3)$$

$$Y_i = \alpha [1 - \beta e^{(-kx_i)}] + \varepsilon_i \quad (6.4)$$

Para comparação e avaliação dos ajustes dos modelos utilizou-se os avaliadores que tiveram melhor eficiência na escolha dos modelos no estudo de simulação. O Índice Assintótico (IA), sendo escolhido o modelo mais adequado aquele que apresentar menor estimativa de IA.

O coeficiente de determinação ( $R^2$ ) e coeficiente de determinação ajustado ( $R_{ajs}^2$ ), sendo considerado o melhor modelo aquele que apresenta maior valor. Também, utilizou-se o Desvio Médio Absoluto (DMA) para comparação dos ajustes, sendo considerado como melhor modelo aqueles que apresentam os menores valores de DMA. Para ilustrar a estimativas dos modelos ajustados utilizou-se a análise gráfica.

As estimativas dos parâmetros e as figuras aqui apresentadas foram obtidas pelo método iterativo de Gauss-Newton implementado na função *nls* e *gnls* do programa R (R CORE TEAM, 2021). A significância dos parâmetros foi verificada por meio do teste t ao nível de 5%.

### 6.3 Resultados e Discussão

Inicialmente foram ajustados os modelos supracitados e feita a análise de resíduos. Os resultados encontrados na análise de resíduos estão contidos na Tabela 6.1. Os pressupostos que devem ser assumidos pelo erro (normalidade, independência e homogeneidade das variâncias) foram avaliados, respectivamente, pelos testes Shapiro-Wilk (SW), Durbin-Watson (DW) e Breusch-Pagan (BP), considerando o nível de significância de 5%.

Tabela 6.1 – Valores p dos teste Shapiro-Wilk (SW), Durbin-Watson (DW) e Breusch-Pagan (BP) utilizados para validação dos pressupostos da análise de resíduos para os modelos Logístico, Gompertz, von Bertalanffy e Brody, para o volume de madeira.

Teste	Logístico	Gompertz	von Bertalanffy	Brody
SW	0,7398	0,5211	0,8312	0,5980
DW	< <b>0,001</b>	<b>0,0356</b>	< <b>0,001</b>	< <b>0,001</b>
BP	0,2846	0,1863	0,5252	0,0795

Observa-se que os pressupostos de normalidade e homogeneidades de variância para o erro aleatório foram validados, no entanto, para todos os modelos e para todas as classificações dos sítios o teste Durbin-Watson foi significativo, violando o pressuposto de independência.

Jane et al. (2019) no estudo sobre o crescimento da planta de pimenta do cultivar Doce, perceberam que não houve necessidade da inclusão do parâmetro autorregressivo, no entanto, Muniz et al. (2017) adicionaram ao seu modelo esse parâmetro e verificaram uma melhoria significativas no ajuste de crescimento dos frutos de cacau, dentre elas a redução significativa do desvio padrão residual, tornando as estimativas mais confiáveis.

Nesta aplicação não será adicionada o parâmetro autorregressivo ao modelo. Os resultados de ajuste para ambos os casos são bem próximos e chegam à mesma conclusão, sendo assim, a estimação sem esse parâmetro é válida como exemplo de aplicação.

A Tabela 6.2 apresenta as estimativas e erros padrão dos parâmetros obtidos pelo método iterativo de Gauss-Newton, para o volume de madeira de *Eucalyptus grandis* x *Eucalyptus urophylla* para os modelos Logístico, Gompertz, von Bertalanffy e Brody.

Tabela 6.2 – Estimativas e erro padrão (EP) dos parâmetros e avaliadores de qualidade de ajuste ( $R^2$ ,  $R^2_{ajs}$ , IA, DMA) para os parâmetros dos modelos Logístico, Gompertz, von Bertalanffy e Brody, para a descrição do volume de madeira de *Eucalyptus grandis* x *Eucalyptus urophylla*.

Modelo	Parâmetro	Estimativa	Erro Padrão	$R^2$	$R^2_{ajs}$	IA	DMA
Logístico	$\hat{\alpha}$	245,9575	3,2838	0,9972	0,9968	0,0884	3,3719
	$\hat{\beta}$	6,1357	0,1110				
	$\hat{k}$	0,4307	0,0173				
Gompertz	$\hat{\alpha}$	263,7795	0,4835	<b>0,9999</b>	<b>0,9999</b>	<b>-0,8823</b>	<b>0,3023</b>
	$\hat{\beta}$	4,9471	0,0108				
	$\hat{k}$	0,2638	0,0020				
von Bertalanffy	$\hat{\alpha}$	277,3292	3,2450	0,9851	0,9992	-0,5032	1,6591
	$\hat{\beta}$	4,2707	0,0544				
	$\hat{k}$	0,2065	0,0054				
Brody	$\hat{\alpha}$	374,1954	37,2934	0,9923	0,9910	0,7861	5,8403
	$\hat{\beta}$	0,9144	0,1923				
	$\hat{k}$	0,0812	0,0137				

Pelos resultados expostos na Tabela 6.2 pode-se notar que o modelo Logístico apresentou estimativas de crescimento máximo mais baixas que os demais ajustes, enquanto a assíntota para o modelo Brody foram superestimadas. As estimativas de volume máximo de madeira varia entre  $245,9575 \text{ m}^3/\text{ha}$  e  $374,1954 \text{ m}^3/\text{ha}$ . Analisando as estimativas para o ponto de inflexão (parâmetro  $\hat{\beta}$ ), que indica o ponto de crescimento máximo em volume de madeira, o mesmo está entre o quarto e sexto ano.

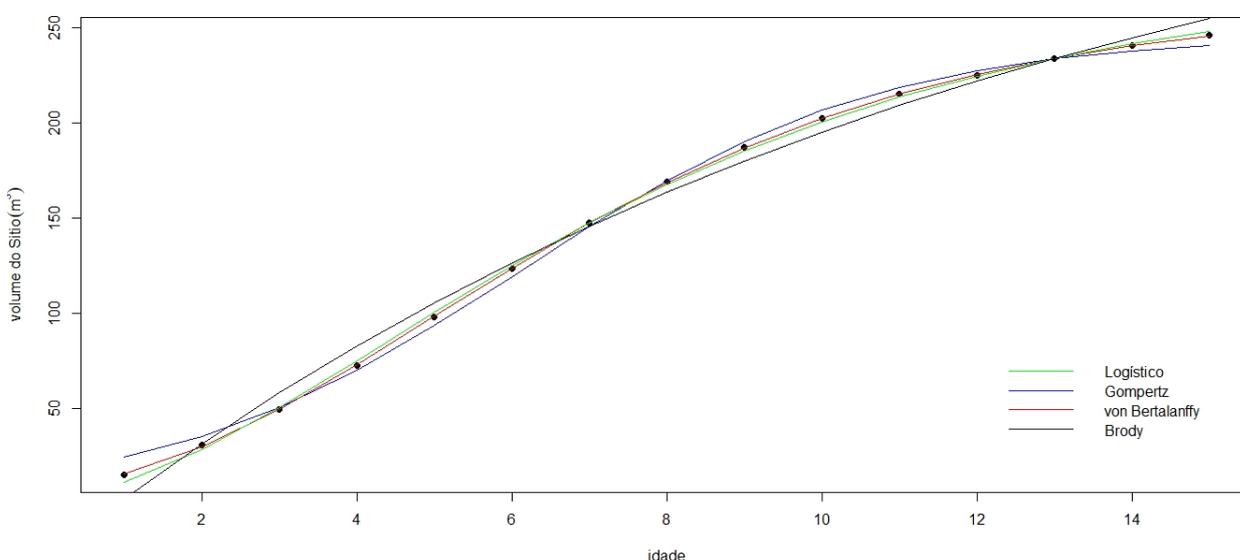
De acordo com os avaliadores de ajuste, a escolha do modelo Gompertz como mais adequado foi unânime para caracterizar o crescimento para o volume de madeira de *Eucalyptus grandis* x *Eucalyptus urophylla*. Este modelo apresentou altos valores de  $R^2$  e  $R^2_{ajs}$  (0,9999 em ambos os avaliadores) e menores DMA e IA (0,3023 e -0,8823, respectivamente), sendo este modelo o escolhido.

Lundgren et al. (2017) avaliando a influência do tipo de amostragem na estimativa de volume de madeira de Eucalipto, verificou o volume total de madeira que foi medido, aos 7,5 anos, forneceu o valor de 166,14 metros cúbicos para uma área de 2,4 h, resultado este, próximo as estimativas para o Sítio III. Corroborando Marangon et al. (2017), na dinâmica da distribuição diamétrica e produção de eucalipto em diferentes idades e espaçamentos, que aos 7 anos, o Sítio

atinge seu máximo de produtividade volumétrica acumulada, e aos 13 anos o seu mínimo para as três idades analisadas, possivelmente devido as árvores desta idade apresentarem maiores volumes individuais e menor distribuição diamétrica.

Coutinho et al. (2017) explicam que o crescimento das árvores não é regular ao longo da vida da árvore, para os indivíduos de criptoméria estudados notou-se diminuição do incremento em diâmetro à altura do peito ao longo dos anos, complementa que os maiores incrementos concentram-se na fase juvenil das árvores, onde naturalmente há maior vigor, quando comparado aos estágios de maturidade e senescência. Segundo Lima et al. (2017), em seu estudo, observando as estimativas da taxa de maturidade ou crescimento ( $\hat{k}$ ) verificou-se que ocorreram valores pequenos, principalmente quando se ajustou o modelo von Bertalanffy, esses valores maiores indicam que as plantas demoram mais tempo para atingir a maturidade.

A Figura 14 apresenta os ajuste de modelos para descrição do volume de madeira dos híbridos.



Fonte: Do autor

**Figura 14:** Gráfico dos ajustes dos modelos descrição do volume de madeira de *Eucalyptus grandis* x *Eucalyptus urophylla*

Pode-se observar pela Figura 14 que os modelos se ajustaram bem aos dados, havendo em alguns ajustes pontos de subestimação e em outros foram superestimados. Os modelos Brody e von Bertalanffy apresentaram esse comportamento ao longo de suas curvas, inicialmente subestimando as estimativas em alguns pontos e em outros passando a superestimar.

Os melhores ajuste de curva, visualmente, foram para o os modelos Gompertz e Logístico, sendo o Gompertz o modelo que melhor se ajustou aos dados e confirmado pelos avaliadores anteriormente apresentados seus resultados.

A subestimação e superestimação em estimativas da assíntota foram observadas em outros trabalhos, Santos et al. (2017) verificaram pontos de subestimativa e superestimativa na estimativa do volume de Tauari na Floresta Nacional do Tapajós, houve uma tendência de superestimação para os volumes de até  $5 m^3$  e subestimação a partir de  $10 m^3$ . Gomes et al. (2018) observaram que houve tendência de superestimação na altura comercial no estrato inferior e de subestimação no estrato superior para estimativas volumétricas de madeira na Floresta Nacional do Tapajós.

Nunes et al. (2017) utilizando os modelos de Hohenadl-Kren, Husch, Stoate e Schumacher-Hall para estimação volumétrica de um povoamento adensado de *Eucalyptus* sp, encontraram uma notória tendência de subestimativa para as árvores de menor diâmetro, enquanto que o modelo de Brenac apresentou ligeira tendência em superestimativa para as árvores com diâmetro à altura do peito (DAP) entre 6 e 8 cm.

Andrade et al. (2017) utilizaram um modelo volumétrico, para árvores de *Eucalyptus urophylla* e *Eucalyptus grandis* com idade entre 5 a 7 anos, desenvolvido a partir da equação de fator de forma adaptado ao modelo biomatemático de Gompertz, apresentando uma menor variação dos erros em torno do eixo zero, com tendência à subestimação de forma menos pronunciada.

A modelagem volumétrica resultou em estimativas semelhantes a outros trabalhos Santos et al. (2017) encontraram resultados satisfatórios utilizando o modelo Gompertz para o volume, apresentando média de crescimento semelhante aos encontrados no presente trabalho. Vendruscolo et al. (2017) verificaram que as equações obtidas com o modelo de Gompertz apresentaram estatísticas de ajustes iguais ligeiramente superiores a demais modelos para estimativa da altura de árvores de *Tectona grandis* L.f.

#### **6.4 Conclusão**

Pode-se concluir que o modelo Gompertz foi o mais adequado para descrever o crescimento em volume de madeira dos híbridos *Eucalyptus grandis* x *Eucalyptus urophylla*, obtendo estimativas satisfatórias para o volume de madeira.

Todos os avaliadores de qualidade de ajuste concordaram em selecionar este modelo como mais adequado. A estimativa para o volume máximo de madeira no modelo selecionado foi de  $263.78m^3/ha$  e este atingiu o ponto de máximo crescimento em volume entre o quarto e o sexto ano.

## 7 REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Boston, v.19, n.6, p.716-723, Dec. 1974.
- AMORIM, F. R., ABREU, P. H. C., PATINO, M. T. O., & TERRA, L. A. A. Análise dos Riscos em Projetos: Uma Aplicação do Método de Monte Carlo em uma Empresa do Setor Moveleiro. **Future Studies Research Journal: Trends and Strategies**, 10(2), 332-357, 2018.
- ANDRADE, V. C. L. de. Modelos Volumétricos de Dupla Entrada para Aplicar em Povoamentos Florestais Brasileiros. **Floresta Ambient**. [online]. 2017, vol.24, e00135415. Epub 10-Abr-2017.
- ANDRADE FILHO, M. G. de; AMARAL, M. T. R. do; STOSIC, B. D.; SANTOS, E. S. Uma abordagem bayesiana para curvas de crescimento com diferentes densidade a priori. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 2, p. 161-181, 2010.
- ARAÚJO, R. A. de; POMPEU, R. C. F. F.; ROGÉRIO, M. C. P.; SALLES, H. O; COSTA, C. dos S.; FONTINELE, R. G.; SILVA, L. N. C. da; NEIVA, J. N. M. Growth and performance curve of dairy goats fed with detoxified castor bean cake by different alkaline solutions. **Semina: Ciências Agrárias**, Londrina, v. 41, n. 6, suplemento 2, p. 3377-3390, 2020.
- ARCHONTOULIS, S.V.; MIGUEZ, F. Nonlinear Regression Models and Applications in Agricultural Research. **Agronomy Journal**. 105. 1. 2015.
- AXEROLD, R. Advancing the art of simulation in the social sciences. **Simulating Social Phenomena** pp 21-40, 1997.
- BARRETO, H.; HOWLAND, F. **Introductory econometrics. Using Monte Carlo simulation with Microsoft Excel**. Cambridge University PressCambridge, UK, 2005.
- BATES, D. M.; WATTS, D. G. Relative curvature measures of nonlinearity. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1–25, 1980.
- BEM, C. M. DE; CARGNELUTTI FILHO, A.; FACCO, G.; SCHABARUM, D. E.; SILVEIRA, D. L.; SIMÕES, F. M.; ULIANA, D. B. Growth models for morphological traits of sunn hemp. **Semina-Ciencias Agrarias**, v. 38, p. 2933, 2017.
- BOTELHO, H. A.; LEIRA; M. H.; MELO, C. C. V.; CAFÉ, M. B.; DE FREITAS, R. T. F. Determinação da curva de crescimento e alometria de pacu (*Piaractus mesopotamicus*), tambaqui

(*Colossoma macropomum*) e seus híbridos. **Agrarian**, Dourados, v. 12, n. 43, p. 97-103, set. 2019.

BOX, M. Bias in nonlinear estimation. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 171–201, 1971.

Bozdogan, H. Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions. **Psychometrika**. 52. 1987.

BRODY, S. **Bioenergetics and growth**. New York: Reinhold Publishing Corporation, 1945.

BURNHAM, K.P.; ANDERSON, D.R. **Model selection and multimodel inference: a practical information-theoretic approach**. New York: Springer, 488p, 2002.

BUSSAB, W. O. **Análise de variância e de regressão**. São Paulo: Atual, 147p. 1986.

CHARNET, R.; BONVINO, H; FREIRE C. A. L; CHARNET, E. M. R. **Análise de modelos de regressão linear com aplicações**. Campinas, SP: Unicamp, 1999.

COUTINHO, V. M.; CORTE, A. P. D.; SANQUETTA, C. R.; RODRIGUES, A. L; SANQUETTA, M. N. I. Modelagem do crescimento de *Cryptomeria japonica* por análise de tronco parcial. **Pesquisa Florestal Brasileira**, online, v.37, n.90, p. 93-98, 2017.

DEPRÁ, M. S.; LOPES, S. J.; NOAL, G.; REINIGER, L. R. S.; COCCO, D. T. Modelo logístico de crescimento de cultivares crioulas de milho e de progênes de meios-irmãos maternos em função da soma térmica. **Ciência Rural**, 46(1), 36-43. Epub 20 de outubro de 2016.

DIEL, M. I.; SARI, B. G.; KRYSCZUN, D. K.; OLIVOTO T.; PINHEIRO, M. V. M.; MEIRA, M.; SCHMIDT, D.; LÚCIO, A. D.; Nonlinear regression for description of strawberry (*Fragaria x ananassa*) production. **The Journal of Horticultural Science and Biotechnology**, 94:2, 259-273, 2019.

DIEL, M. I.; LÚCIO, A. D.; VALERA, O. V. S.; SARI, B. G.; OLIVOTO, T., PINHEIRO, M. V. M.; MELO, P. J. de; TARTAGLIA, F. L.; SHMIDT, D. Production of biquinho pepper in different growing seasons characterized by the logistic model and its critical points. **Ciência Rural**, 50(4), e20190477. Epub April 22, 2020.

DONATELLI, G. D., KONRATH, A.C. Considerações sobre a Aplicação da Simulação de Monte Carlo na Avaliação de Incertezas de Medição. **Revista Ciência & Tecnologia**. Editora: UNIMEP, Volume 13 , Nº 25/26, outubro de 2006.

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

DRUMOND, E. S. C.; GONÇALVES, F. M.; VELOSO, R. de C.; AMARAL, J. M.; BALOTIN, L. V.; PIRES, A. V.; MOREIRA, J. Curvas de crescimento para codornas de corte. **Ciência Rural**, 43(10), 1872-1877. 2013.

ELIAS, R. W. D.; SILVA, R. M. ; ALEMEIDA, J. K. B. ; VALADARES, D. C. G. . Comparação de Métodos Iterativos de Resolução de Equações Não Lineares Implementados no Octave. In: I SMNE - 1º Simpósio de Métodos Numéricos em Engenharia, 2016, Curitiba. **Anais do 1º Simpósio de Métodos Numéricos em Engenharia**. Curitiba: Editora Tecnologia, p. 101-105, 2016.

ELOY, R. A. O. **Avaliação da qualidade de ajuste e predição de modelos não lineares: uma aplicação em dados de crescimento de frutos de cacaueteiro**. Dissertação (Mestrado em Estatística e Experimentação Agrônômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2018.

EVANS, James R.; OLSON, David L. **Introduction to Simulation and Risk Analysis**. Upper Saddle River, New Jersey: Prentice Hall, 1998.

FERNANDES, T. J.; MUNIZ, J. A.; PEREIRA, A. A.; MUNIZ, F. R.; MUIANGA, C. A. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum Technology**, Maringá, v. 37, n. 4, p. 397–402, 2015.

FERNANDES, T. J.; PEREIRA, A. A.; MUNIZ, J. A.; SAVIAN, T. V. Seleção de modelos não lineares para a descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, p. 207–215, 2014.

FERNANDES, F. A.; FERNANDES, T. J.; PEREIRA, A. A.; MEIRELLES, S. L. C.; COSTA, A. C. Growth curves of meat-producing mammals by von Bertalanffy's model. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 54, 2019.

FOX, D. G. Judging Air Quality Model Performance. **Bulletin of the American Meteorological Society**, v. 62, p. 599-609, 1981.

FREITAS, A. R. Curvas de crescimento na produção animal. **Revista Brasileira Zootecnia**, v. 34, n. 3, p. 786–795, 2005.

GONCALVES, J. C.; OLIVEIRA, A. D. de; CARVALHO, S. P. C.; GOMIDE, L. R. Análise econômica da rotação florestal de povoamentos de eucalipto utilizando a simulação de Monte Carlo. **Ciência Florestal**, Santa Maria, v. 27, n. 4, p. 1339-1347, 2017.

GOMES, K. M. A. G.; RIBEIRO, R.; GAMA, J. R. V. Eficiência na estimativa volumétrica de madeira na Floresta Nacional do Tapajós. **Nativa**, Sinop, v.6, n.2, p. 170-176, mar./abr. 2018.

GRINSTEAD, C.; SNELL, J. **Introduction to probability**. Rhode Island: American Mathematical Society, 2012.

HAIR JUNIOR, F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. **Multivariate data analysis**. Upper Saddle River: Prentice Hall, 1998.

IBGE, INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção da Extração vegetal e Silvicultura**, Rio de Janeiro, v. 33, p. 1-8, 2018.

JACOVIDES, C. P; KONTOYIANNIS, H. Statistical procedures for the evaluation of evapotranspiration computing models. **Agricultural Water Management**. V.27, n. 3-4, July, pg. 365-371, 1995.

JANE, S. A.; FERNANDES, F. A.; SILVA, E. M; MUNIZ, J. A.; FERNANDES, T. J. Comparação dos modelos polinomial e não lineares na descrição do crescimento de pimenta. **Revista Brasileira de Ciências Agrárias**. Recife, v.14, n.4, 2019.

JANE, S. A.; FERNANDES, F. A.; SILVA, E. M.; MUNIZ, J. A.; FERNANDES, T. J.; PIMENTEL, G. V. Adjusting the growth curve of sugarcane varieties using nonlinear models. **Ciência Rural**, v. 50, n. 3, e20190408, 2020.

KENNEDY, P. G. E. Teaching Undergraduate Econometrics: A Suggestion for Fundamental Change. **American Economic Review**, vol. 88, issue 2, 487-91, 1998.

KLEINPAUL, J. A.; CARGNELUTTI FILHO, A.; SILVEIRA, D. L.; NEU, I. M. M.; BANDEIRA, C. T.; PROCEDI, A. Nonlinear Models for Plant Height of Rye Cultivars at Sowing Times. **Journal of Agricultural Science**, v. 10, p. 157, 2018.

KSHIRSAGAR, A. M.; SMITH, W. B. **Growth curves**. New York: Marcel Dekker, p.359, 1995.

LIMA, K. P.; MORAIS, A. R.; VIEIRA, N. M. B.; VILLA, F.; ANDRADE, M. J. B. Nonlinear models for use in description of Boron accumulation in different parts of Jalo beans. **Revista Brasileira de Biometria**, Lavras, v.35, n.4, p.834-861, 2017.

LOPES, F.; MAGNABOSCO, C.; Souza, F.; ASSIS, A.; BRUNES, L. Análises de dados longitudinais em bovinos Nelore Mocho por meio de modelos não lineares. **Archivos de Zootecnia**. 65. 123-129. 10.21071/az.v65i250.478. 2016.

LUNDGREN, W. J. C.; SILVA, J. A. A. da; FERREIRA, R. L. C. Estimação do Volume de Eucaliptos por Krigagem e Cokrigagem no Semiárido Pernambucano. **Floresta Ambient.** [online]. 2017, vol.24, e00140415. Epub 12-Set-2017.

MAIA, A. G. **Econometria: Conceitos e Aplicações**. Editora Saint Paul. 2017.

MARANGON, G. P.; COSTA, E. A.; ZIMMERMANN, A. P. L.; SCHNEIDER, P. R.; SILVA, E. A. Dinâmica da distribuição diamétrica e produção de eucalipto em diferentes idades e espaçamentos. **Revista de Ciências Agrárias**, v. 60, n. 1, p. 33-37, jan./mar. 2017.

MARTINS, V. L. M., WERNER, L., PINTO, F. T. Uso da Simulação de Monte Carlo para Avaliação da Confiabilidade de um Produto. **Simpósio de Administração da Produção, Logística e Operações Internacionais**. SIMPOI, São Paulo, SP, Brasil, 13. 2010.

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

MEIRELES, F; SILVA, D; CORREA, F. Simulações de Monte Carlo no ensino de Ciência Política. **Rev. Bras. Ciênc. Polít.**, Brasília, n. 24, p. 223-254, Dec. 2017.

MENDONÇA, A. R.; CALEGARIO, N.; SILVA, G. F.; CARVALHO, S. P. C. Growth and yield models for eucalyptus stands obtained by differential equations. **Scientia Agricola**, v. 74, n. 5, p. 264-370, 2017.

MIRANDA, L. F. M.; LIMA, K. P. de; SILVA, E. M.; FERNANDES, F. A.; MUNIZ, J. A.; MORAIS, A. R. de. Modelos não lineares na descrição do crescimento de tilápia do Nilo criada em tanques-rede. **Sigmae**. Alfenas. v.8, n.2, p 606-611, 2019.

MISCHAN, M. M.; PINHO, S. Z. de. **Modelos não lineares [recurso eletrônico]: funções assintóticas de crescimento**. São Paulo: Cultura Acadêmica, 184 p, 2014.

MONTGOMERY, C. D.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 2.ed. São Paulo: LTC, 2003.

MOONEY, C. Z. **Monte Carlo simulation**. Thousand Oaks: Sage Publications, 1997.

MORAES, L. P., MAIA, P. R. B., PINTO, A. C. F., KLOTZLE, M. C., & GOMES, L. L. Aplicação de Técnica de Redução de Variância no Prêmio de Opções Asiáticas de Eletricidade por Simulação de Monte Carlo. **Revista Economia & Gestão**, 16(43), 33-50, 2016.

MOTA, L. F. M.; ALCÂNTARA, D. C.; ABREU, L. R. A.; COSTA, L. S.; PIRES, A. V.; BONAFÉ, C. M.; SILVA, M. A.; PINHEIRO, S. R. F. . Crescimento de codornas de diferentes

grupos genéticos por meio de modelos não lineares. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, 67(5), 1372-1380. 2015.

MUIANGA, C. A.; MUNIZ, J. A.; NASCIMENTO, M. da S., FERNANDES, T. J.; SAVIAN, T. V. Descrição da curva de crescimento de frutos do cajueiro por modelos não lineares. **Revista Brasileira Fruticultura** [online]. vol.38, n.1, pp.22-32, 2016.

MUNIZ, J. A.; NASCIMENTO, M. d. S.; FERNANDES, T. J. Nonlinear models for description of cacao fruit growth with assumption violations. **Revista Caatinga**, Mossoró, v. 30, n. 1, p. 250–257, 2017.

NAVARRO, D. J.; MYUNG, In Jae. **Model evaluation and selection**. Columbus, USA: Ohio State University, Department of Psychology, 6p, 2004.

NUNES, J. S.; SOARES, T. S. Estimativas volumétricas para um povoamento adensado de *Eucalyptus* sp. em regime de curta rotação. **Revista de Agricultura Neotropical**, Cassilândia-MS, v. 4, n. 4, p. 77-86, out./dez. 2017.

PAGLIARUSSI, M. S. O ensino do modelo clássico de regressão linear por meio de simulação de Monte Carlo. **Revista de Contabilidade e Organizações**, v. 12, p. e152100, 28 dez. 2018.

PEREIRA, A. A.; MORAIS, A. R.; SCALCO, M. S.; FERNANDES, T. J. Modelagem não linear do crescimento em altura do cafeeiro irrigado e não irrigado em diferentes densidades. **IRRIGA**, v.1, n.1, p.140-149, 2016.

PRADO, T. K. L. do; SAVIAN, T. V.; MUNIZ, J. A. Ajuste dos modelos gompertz e logístico aos dados de crescimento de frutos de coqueiro anão verde. **Ciência Rural**, v. 43, n. 5, p.803–809, 2013.

PRADO, T.K.L. do; SAVIAN, T.V.; FERNANDES, T.J.; MUNIZ, J.A. Study of the growth curve of the internal cavity of fruit dwarf green coconut. **Revista Ciência Agronômica**, v. 51, n.3, p. 1-7, 2020

RATKOWSKY, D.A. **Nonlinear Regression Modeling: a Unified Practical Approach**. New York: Marcel Dekker, 1983.

RATKOWSKY, D.A. **Handbook of nonlinear regression models**. New York and Basel, Marcel Dekker, 241p, 1990.

- REZENDE, B. R.; LIMA, M. S. S.; SANTANA, H. A.; SANTOS, W. V. dos; SILVA, A. R. da. Modeling the Growth of Forage Legumes. **Journal of Agricultural Science**, vol. 12, n. 5, p. 139-144, 2020.
- RETSLAFF, F. A. S.; FIGUEIREDO FILHO, A.; DIAS, A. N.; BERNETT, L. G.; FIGURA, M. A. . Curvas de sítio e relações hipsométricas para *Eucalyptus grandis* na região dos Campos Gerais, Paraná. **CERNE**, Lavras , v. 21, n. 2, p. 219-225, 2015.
- ROCHA-SILVA, M.; ARAUJO, C. V.; PIRES, A. V.; DE PAULA, E. J. H.; FERREIRA, E. B.; DA SILVA, F. G. Curva de crescimento de codornas de corte por meio de modelos de regressão não-lineares. **Archives of Veterinary Science**, [S.l.], v. 21, n. 4, dec. 2016.
- SANTOS, A.; SILVA, S.; LEITE, H.; CRUZ, J. Influência da variabilidade edafoclimática no crescimento de clones de eucalipto no Nordeste baiano. **Pesquisa Florestal Brasileira**, v.37, n. 91, p. 259-268, 2017.
- SARI, B.; OLIVOTO, T.; DIEHL, M.; KRYSCZUN, D.; LÚCIO, A.; SAVIAN, T. Nonlinear Modeling for Analyzing Data from Multiple Harvest Crops. **Agronomy journal**. 110.10.2134/agronj2018.05.0307. 2018.
- SARI, B. G.; LÚCIO, A. D. ; SANTANA, C. S. ; SAVIAN, T. V. Describing tomato plant production using growth models. **SCIENTIA HORTICULTURAE**, v. 246, p. 146-154, 2019.
- SCHABENBERGER, O.; PIERCE, F.J. **Contemporary statistical models**. Boca Raton: CRC, LLC, 738p, 2002.
- SCHWARZ, G. Estimating the dimensional of a model. **Annals of Statistics**, Hayward, v.6, n.2, p.461-464, Mar. 1978.
- SEBER, G. A. F.; WILD, C. J. **Nonlinear regression**. New Jersey: John Wiley & Sons, Inc., Hoboken, 1989.
- SILVA, É. M. da; TADEU, M. H.; SILVA, V. P. da; PIO, R.; FERNANDES, T. J.; MUNIZ, J. A. Description of blackberry fruit growth by nonlinear regression models. **Revista Brasileira de Fruticultura**, v. 42, n. 2, e-177, 2020.
- SILVA, P. V., SAVIAN, T. V. Chanter model: nonlinear modeling of the fruit growth of cocoa/- Modelo Chanter: modelagem não linear do crescimento de frutos do cacauzeiro. **Ciencia Rural**, vol. 49, no. 11, 2019.

SIT, V. **Catalog of curves for curve fitting - Handbook 4**. Victoria: Ministry of Forests, Forest Science Research Branch, Biometrics information, 110p, 1994.

SPIESS, A.; NEUMEYER, N. **An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach**. *BMC pharmacology*. 10. 6. 2010.

STONE, R. J. Improved statistical procedure for the evaluation of solar radiation estimation models. *Solar Energy*, v. 51, n. 4, p. 289-291, 1993.

TADDEI, F. G.; HERREREA, D. R. Crescimento do caranguejo *Dilocarcinus pagei* Stimpson. *Boletim do Instituto de Pesca*. 36. 99-110. 2010.

VELOSO, R.C. ; PIRES, A.V. ; TORRES FILHO, R.A.; DRUMOND, E.C.S. ; COSTA, L.S.; AMARAL, J.M.; PEREIRA, I.G. Crescimento de genótipos de frangos tipo caipira. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v.67, n.5, p.1361-1371, 2015.

VELOSO, R.C.; WINKELSTROTER, L.K.; SILVA, M.T.P.; PIRES, A.V.; TORRES FILHO, R.A.; PINHEIRO, S.R.F.; COSTA, L.S.; AMARAL, J.M. Seleção e classificação multivariada de modelos não lineares para frangos de corte. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 68(1), 191-200. 2016.

VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S.; MOURA, J. P. V. M.; MAMORÉ, F. M. D.; SIQUEIRA, T. A. da S. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. *Revista Brasileira de Biometria*, São Paulo, v. 33, n. 4, p. 556-569, 2017.

YANG, W.; TIAN, C. Monte-Carlo simulation of information system project performance. *Systems Engineering Procedia*, 3(1), 340-345, 2012.

ZEVIANI, W. M.; SILVA, C. A.; CARNEIRO, W. J. de O.; MUNIZ, J. A. Modelos não lineares para a liberação de potássio de esterco animal em latossolos. *Ciência Rural* [online]. vol.42, n.10, pp.1789-1796, 2012.