



ÉDIPO MENEZES DA SILVA

**USO DE MODELOS MISTOS E PONTOS CRÍTICOS EM
DADOS DE CRESCIMENTO DE FRUTOS**

**LAVRAS – MG
2024**

ÉDIPO MENEZES DA SILVA

**USO DE MODELOS MISTOS E PONTOS CRÍTICOS EM DADOS DE
CRESCIMENTO DE FRUTOS**

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

Prof. Dr. Joel Augusto Muniz
Orientador

**LAVRAS – MG
2024**

**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, Édipo Menezes da.

Uso de modelos mistos e pontos críticos em dados de
crescimento de frutos / Édipo Menezes da Silva. – 2023.

94 p. : il.

Orientador: Joel Augusto Muniz.

Tese (Doutorado) - Universidade Federal de Lavras, 2023.

Bibliografia.

1. Variabilidade. 2. Estimação. 3. Verossimilhança restrita.
I. Muniz, Joel Augusto. II. Título.

ÉDIPO MENEZES DA SILVA

**USO DE MODELOS MISTOS E PONTOS CRÍTICOS EM DADOS DE
CRESCIMENTO DE FRUTOS**

**DEMONSTRATION OF CRITICAL POINTS IN NONLINEAR GROWTH MODELS
LOGISTIC AND GOMPERTZ, AND PROVISION OF R STATISTICAL SOFTWARE
SCRIPTS**

Tese 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 Doutor.

APROVADO em 23 de fevereiro de 2023.

Prof. Dr. Tales Jesus Fernandes	UFLA
Prof. Dr. Marcelo Angelo Cirillo	UFLA
Prof. Dr. Silvio de Castro Silveira	FEOL
Profa. Dra. Adriele Aparecida Pereira	UNIFAL

Prof. Dr. Joel Augusto Muniz
Orientador

**LAVRAS – MG
2024**

Dedico essa tese a mulher mais incrível e guerreira que eu já conheci, minha mãe!

AGRADECIMENTOS

Mais um ciclo se encerra e chegou a hora de olhar para trás e agradecer a todos aqueles que contribuíram para que esse dia chegasse.

Primeiro, agradecer a Deus por tudo que me ajudou a construir até aqui. A minha mãe, por ser a minha grande inspiração a sempre ser melhor.

Aos meus irmãos, por todo apoio ao cuidar da nossa mãe.

Aos amigos de São Vicente de Minas, em especial ao Beto e a Vivi, por estarem comigo, desde o início dessa louca jornada de estudos.

Ao Henrique e ao Victor, por me ajudarem em diversos momentos e pela amizade. Ao Felipe, pelas conversas sobre modelos não lineares e outros assuntos.

Ao Grupo de estudos NLIN, pelos ensinamentos e por me ajudar a entender melhor o mundo dos modelos não lineares de crescimento.

Ao meu grande mentor, professor Joel, pela paciência, dedicação e por me ensinar a ser um pesquisador. Ao professor Tales, por sempre contribuir com as minhas pesquisas.

A ex- secretária Nádia, do programa de pós-graduação em Estatística e Experimentação Agropecuária, pelos serviços prestados frente à secretária.

À Universidade Federal de Lavras (UFLA), pela oportunidade de realização do doutorado, e aos professores do Departamento de Estatística (DES).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo financiamento da bolsa de estudos.

Enfim, a todos aqueles que direta e indiretamente contribuíram para mais essa etapa da minha vida.

*Faça o teu melhor, nas condições que você tem,
enquanto não tem condições melhores para
fazer melhor ainda (Mario Sergio Cortella).*

RESUMO

Desde o surgimento do novo Coronavírus, no final de 2019, os modelos não lineares de crescimento ganharam destaque, tanto na mídia nacional quanto internacional. Isso se deve à tendência sigmoide observada na média móvel do número de casos, o que levou vários pesquisadores a utilizar esses modelos para compreender esse comportamento. Essa mesma forma de crescimento em "S" ou sigmoidal foi observada em diversos seres vivos. Enquanto alguns autores optaram por empregar modelos lineares para descrever esse crescimento, acabaram deixando de aproveitar as vantagens inerentes aos modelos não lineares. Embora os modelos lineares fossem mais fáceis de estimar, uma vez que não necessitavam de métodos iterativos, os modelos não lineares de crescimento ofereceram uma série de vantagens, incluindo simplicidade, interpretação biológica dos parâmetros e a possibilidade de extrapolar as inferências para além do intervalo das variáveis. Além disso, outros aspectos interessantes foram explorados, como os chamados "pontos críticos". Esses pontos foram determinados a partir das derivadas das funções de crescimento, da primeira até a quarta ordem, em relação à variável independente (geralmente o tempo). Os pontos críticos incluíram o ponto de aceleração máxima, ponto de inflexão, ponto de desaceleração máxima e ponto de desaceleração assintótica. Além dos pontos críticos, foi possível explorar a variação entre os indivíduos, analisando como cada um cresceu ao longo do tempo, considerando a inclusão de efeitos aleatórios nos parâmetros do modelo, caracterizando-o como um modelo misto. Com essa abordagem, objetivou-se, nesse estudo, estimar os pontos críticos dos modelos não lineares de crescimento e investigar a variabilidade entre os indivíduos, utilizando modelos mistos. Isso foi feito para um único estágio de desenvolvimento e para dois estágios de desenvolvimento, aplicando os modelos aos dados de crescimento de Amora Preta e coco Anão Verde. Os pontos críticos foram obtidos, por meio das derivadas e foram representados em um gráfico que mostrou a taxa de crescimento, permitindo visualizar a localização de cada ponto crítico. A modelagem mista para a cultivar Choctaw foi realizada, utilizando o modelo logístico simples com efeito aleatório nos parâmetros β_1 e β_2 para o comprimento, bem como o modelo duplo logístico misto, com efeito aleatório nos parâmetros β_1 e β_4 em relação ao diâmetro.

Palavras-chave: variabilidade; estimação; verossimilhança restrita; derivadas; comparação de modelos.

ABSTRACT

Since the emergence of the new Coronavirus in late 2019, nonlinear growth models have gained prominence both in national and international media. This is due to the sigmoidal trend observed in the moving average of the number of cases, which has led many researchers to employ these models to understand this behavior. This same sigmoidal growth pattern is observed in various living organisms. While some authors chose to use linear models to describe this growth, they ended up missing out on the inherent advantages of nonlinear models. Although linear models were easier to estimate since they did not require iterative methods, nonlinear growth models offered several advantages, including simplicity, biological interpretation of parameters, and the ability to extrapolate inferences beyond the range of variables. Additionally, other interesting aspects have been explored, such as the so-called "critical points". These points were determined by taking derivatives of the growth functions, from the first to the fourth order, with respect to the independent variable (usually time). The critical points included the point of maximum acceleration, inflection point, point of maximum deceleration, and asymptotic deceleration point. Beyond the critical points, it was possible to explore variability among individuals by analyzing how each one grew over time, considering the inclusion of random effects in the model parameters, characterizing it as a mixed-effects model. With this approach, the aim of this study was to estimate the critical points of nonlinear growth models and investigate variability among individuals using mixed-effects models. This was done for a single stage of development and for two stages of development, applying the models to the growth data of Blackberry and Dwarf Green Coconut. The critical points were obtained through derivatives and were represented on a graph that showed the growth rate, allowing the location of each critical point to be visualized. The mixed-effects modeling for the Choctaw cultivar was performed using the simple logistic model with random effects on parameters β_1 and β_2 for length, as well as the mixed-effects double logistic model with random effects on parameters β_1 and β_4 in relation to diameter.

Keywords: variability; estimation; restricted likelihood; models comparison.

IMPACTOS SOCIAIS, TECNOLÓGICOS, ECONÔMICOS E CULTURAIS

A pesquisa sobre modelagem do crescimento de plantas utilizando modelos não lineares e técnicas de modelagem mista pode ter um impacto significativo tanto na economia quanto na tecnologia agrícola. Em termos econômicos, compreender melhor os padrões de crescimento das culturas agrícolas, como a Amora Preta e o coco Anão Verde, pode levar a ganhos substanciais de eficiência na produção agrícola. Ao identificar os pontos críticos de crescimento e entender a variabilidade entre os indivíduos de uma cultura, os agricultores podem otimizar suas práticas de manejo, aumentando a produtividade e reduzindo os custos de produção. Isso não apenas beneficia os agricultores, aumentando sua renda e lucratividade, mas também pode contribuir para a estabilidade econômica das regiões agrícolas, gerando empregos e impulsionando o crescimento econômico local. Em relação à tecnologia agrícola, a aplicação de modelos não lineares e técnicas avançadas de análise de dados representa um avanço significativo. Ao integrar esses métodos computacionais sofisticados com o conhecimento tradicional sobre o cultivo de plantas, abre-se espaço para o desenvolvimento de novas tecnologias e práticas agrícolas mais eficientes e sustentáveis. Isso pode incluir o desenvolvimento de sistemas de monitoramento automatizado do crescimento das plantas, a implementação de algoritmos de otimização para o manejo agrícola e a criação de dispositivos e sensores inteligentes para coleta de dados em tempo real. Essas tecnologias têm o potencial de revolucionar a agricultura, tornando-a mais produtiva, sustentável e adaptada às necessidades específicas de cada cultura e região. Em última análise, isso pode impulsionar a inovação e o crescimento no setor agrícola, contribuindo para o desenvolvimento econômico e o bem-estar social.

SOCIAL, TECHNOLOGICAL, ECONOMIC AND CULTURAL IMPACTS

The research on modeling plant growth using nonlinear models and mixed modeling techniques can have a significant impact on both the economy and agricultural technology. Economically, gaining a better understanding of the growth patterns of agricultural crops, such as Blackberry and Dwarf Green Coconut, can lead to substantial gains in agricultural production efficiency. By identifying critical growth points and understanding variability among individuals within a crop, farmers can optimize their management practices, increasing productivity and reducing production costs. This not only benefits farmers by increasing their income and profitability but can also contribute to the economic stability of agricultural regions by generating jobs and driving local economic growth. In terms of agricultural technology, the application of nonlinear models and advanced data analysis techniques represents a significant advancement. By integrating these sophisticated computational methods with traditional knowledge of plant cultivation, it opens up space for the development of new, more efficient, and sustainable agricultural technologies and practices. This may include the development of automated plant growth monitoring systems, the implementation of optimization algorithms for agricultural management, and the creation of intelligent devices and sensors for real-time data collection. These technologies have the potential to revolutionize agriculture, making it more productive, sustainable, and tailored to the specific needs of each crop and region. Ultimately, this can drive innovation and growth in the agricultural sector, contributing to economic development and social well-being.

SUMÁRIO

	PRIMEIRA PARTE	10
1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	14
2.1	Amora preta	14
2.1.1	Choctaw	15
2.2	Coco Anão Verde	16
2.3	Crescimento de Frutos	16
2.4	Modelos de regressão	17
2.5	Modelo de regressão linear fixo e misto (MLM)	18
2.6	Modelo de regressão não linear de efeitos fixos	20
2.7	Modelo de regressão não linear de efeitos mistos	22
2.8	Métodos de estimação de parâmetros	24
2.8.1	Estimação de β e b	24
2.8.2	Estimação de β	25
2.9	Máxima Verossimilhança (MV)	26
2.10	Máxima verossimilhança restrita (MVR)	27
2.10.1	Mínimos quadrado	29
2.11	Métodos iterativos	30
2.12	Método de Seleção de modelo	31
2.12.1	Crítério de Informação de Akaike (AIC)	31
2.12.2	Crítério Informação de Bayesiano (BIC)	31
2.12.3	Coefficiente de determinação ajustado (Raj2)	31
2.12.4	Teste da razão de verossimilhança (TRV)	32
	REFERÊNCIAS	35
	SEGUNDA PARTE – ARTIGOS	39
	ARTIGO 1 - Evaluation of the critical points of the most adequate nonlinear model in adjusting growth data of ‘green dwarf’ coconut fruits	40
	ARTIGO 2 - Adjustment of mixed nonlinear models on Blackberry fruit growth	52
	ARTIGO 3 - Demonstration of critical points in nonlinear growth models Logistic and Gompertz, and provision of R statistical software scripts	71
	TERCEIRA PARTE	91
1	CONSIDERAÇÕES FINAIS	92

PRIMEIRA PARTE

1 INTRODUÇÃO

Nos últimos tempos, houve um aumento significativo na aplicação de modelos de crescimento. Esse aumento em sua utilização foi particularmente evidente, desde o surgimento da pandemia de COVID-19 em 2019. Isso se deve em parte ao fato de que o número acumulado médio de casos, ao longo do tempo, exibe uma tendência sigmoide. Essa característica de crescimento é bem capturada por modelos de regressão não lineares, em especial os pertencentes à família Richards, como o modelo Logístico, Gompertz, Von Bertalanffy e Brody, conforme discutido por Fernandes et al. em 2020.

No campo das ciências agrárias, uma diversidade de estudos tem se dedicado à utilização desses tipos de modelos (LIMA et al., 2019; SILVA et al., 2020; FRUHAUF et al., 2020) para a análise do crescimento de plantas e animais. Entretanto, os autores, dessas pesquisas, frequentemente não exploraram todas as potencialidades que esses modelos podem proporcionar, como, por exemplo, os pontos críticos do crescimento e as variações existentes entre os indivíduos.

O padrão de crescimento em organismos vivos frequentemente segue uma série de características reconhecíveis, sendo a mais comum a forma de "S" ou sigmoide. Isso significa que o crescimento começa de forma lenta, avança para uma fase de crescimento acelerado ou exponencial, atinge um ponto de inflexão, e depois se estabiliza até alcançar a maturidade. Em alguns cenários, o processo de crescimento pode, até mesmo, ocorrer em dois estágios, o que é denominado "duplo S" ou duplo sigmoide. Nesses casos, o estágio de desenvolvimento é repetido duas vezes (FERNANDES et al., 2017; FERNANDES et al., 2022; SILVA et al., 2020).

Os modelos de crescimento não lineares apresentam distinções que os separam dos modelos lineares, uma delas sendo a exigência do emprego de métodos iterativos para se aproximar das estimativas dos parâmetros. Muitos pesquisadores evitam a utilização desses modelos, em razão da necessidade desses métodos iterativos, que demandam valores iniciais para iniciar a aproximação. Em algumas situações, obter esses valores iniciais não é uma tarefa trivial. Em busca de soluções, a literatura descreve vários métodos utilizados para obter esses valores iniciais, como ilustrado nos estudos de Silva et al. (2020) e Silva et al. (2021), onde gráficos são empregados para identificar esses valores, e na pesquisa de Michan e Pinho (2014), que recorre a ferramentas matemáticas para essa determinação.

Ao estabelecer o conjunto inicial de valores para a estimativa dos parâmetros, os modelos não lineares de crescimento oferecem várias vantagens em comparação aos modelos

lineares. Algumas dessas vantagens incluem a menor utilização de parâmetros, a interpretação direta e prática dos parâmetros, bem como a capacidade de extrapolar inferências para além do âmbito das variáveis em questão (MICHAN; PINHO, 2014).

Além das vantagens mencionadas no parágrafo precedente, os modelos de crescimento não lineares possuem outras características de extrema relevância, como a identificação de pontos críticos, ao longo do processo de crescimento (MICHAN; PINHO, 2014). Esses pontos críticos delineiam fases distintas do ciclo de crescimento, enriquecendo as informações que podem ser obtidas ao longo desse ciclo. Os pontos críticos em questão englobam o ponto de máxima aceleração (PAM), o ponto de inflexão (PI), o ponto de máxima desaceleração (PDM) e o ponto de desaceleração assintótica (PDA) (MICHAN; PINHO, 2014; TEIXEIRA et al., 2021).

Os estudos conduzidos por Sari et al. (2018), Sari et al. (2019), Diel et al. (2019) e Teixeira et al. (2021) abordam, de maneira gráfica, o comportamento dos pontos críticos, sem realizar a estimativa precisa da localização de cada um desses pontos. Contudo, obter essas estimativas para cada ponto crítico (tanto em termos de abscissa quanto das ordenadas) pode ser altamente benéfico. Isso permitiria identificar com maior precisão as diferentes fases do crescimento, o que, por sua vez, pode ser instrumental para tomar decisões como a antecipação da colheita.

Neste trabalho, foram aplicadas as parametrizações dos modelos Gompertz e Logístico, como apresentadas por Fernandes et al. (2015) e o modelo Brody. A integração dos pontos críticos e a modelagem mista proporcionou *insights* valiosos, possibilitando uma análise mais abrangente das características do crescimento, que pode ser aplicada em situações práticas como a otimização do manejo agrícola.

Na maioria dos estudos, a consideração da repetição dos dados é frequentemente negligenciada, levando à modelagem apenas da média desses valores, como evidenciado nos trabalhos de Silva et al. (2020) e Silva et al. (2021). No entanto, outra informação potencialmente relevante para compreender o processo de crescimento é a avaliação da possível variação entre os indivíduos, como mencionado em Sari et al. (2019). Isso pode ser abordado por meio da caracterização de modelos não lineares de efeitos mistos.

De acordo com Searle, Casella e McCulloch (2009), para que um modelo seja considerado de efeitos mistos, é necessário que, pelo menos um termo do modelo de regressão seja aleatório, excluindo-se o erro, que sempre é tratado como aleatório. Portanto, ao adotar abordagens que consideram efeitos mistos, é possível capturar a variação entre os frutos, o que

enriquece a compreensão da heterogeneidade subjacente ao crescimento e permite uma modelagem mais completa e precisa.

Portanto, o propósito deste estudo foi empregar os modelos não lineares Logístico, Gompertz e Brody, simples e duplo. O objetivo consistia em estimar os pontos críticos, ao mesmo tempo em que se investigava a potencial variabilidade entre os indivíduos. Essas metodologias foram aplicadas a conjuntos de dados relacionados à Amora Preta e ao Coco Anão Verde, como intuito de compreender melhor os padrões de crescimento e suas nuances.

A estrutura, desta dissertação, é delineada da seguinte forma: A primeira seção abrange o referencial teórico e a revisão bibliográfica, fornecendo o embasamento necessário para a compreensão do conteúdo apresentado. Na segunda seção, é apresentado o artigo que se concentra na estimação dos pontos críticos, já publicados. A terceira seção explora a aplicação de modelos não lineares mistos em conjuntos de dados relativos ao crescimento da amora-preta. A quarta e última seção aborda a construção analítica dos pontos críticos, além de disponibilizar os comandos em R para a estimação e criação gráfica desses pontos.

2 REFERENCIAL TEÓRICO

A seguir foi apresentado toda a parte teórica que deu base aos artigos apresentados nesta tese.

2.1 Amora preta

A amoreira, pertencente à família Rosaceae - mesma família das rosas - faz parte do gênero *Rubus*. Essa planta dá origem à amora preta, um pequeno fruto com um peso variando entre 4 e 7 gramas, de coloração negra e sabor que pode variar entre ácido e doce-ácido. Um traço distintivo dessa espécie é a forma como seus frutos se agrupam, sendo classificados como frutos agregados. Esses agrupamentos consistem em dezenas de frutos verdadeiros, também conhecidos como minidrupas ou drupetes, cada um contendo uma semente no interior (ANTUNES; TREVISAN; PEREIRA, 2007).

Conforme Antunes (2002) relata, o fruto *in natura* é constituído por aproximadamente 85% de água, juntamente com 10% de carboidratos, minerais, vitaminas A e B, e cálcio. Isso proporciona diversos benefícios para a saúde humana, incluindo a regulação de hemorragias, o controle da pressão arterial, efeitos sedativos, propriedades antioxidantes, entre outros.

O cultivo da amora preta ("blackberry"), no Brasil, teve início por meio de um programa de melhoramento genético conduzido pela Embrapa Clima Temperado. Esse programa resultou no desenvolvimento de três cultivares: Tupy, Guarani e Cainganue. A primeira coleção dessas variedades foi implantada em 1974, no município de Canguçu, no Rio Grande do Sul (ANTUNES; TREVISAN; PEREIRA, 2007).

De forma oficial, em um ambiente de campo experimental, no estado de Minas Gerais, as primeiras coleções de amostras foram introduzidas em 1996. Isso ocorreu na Universidade Federal de Lavras (UFLA) e posteriormente na fazenda da EPAMIG, localizada em Caldas, ambos acontecendo no mesmo ano (ANTUNES, 2002).

Curi et al. (2015) conduziram um estudo abrangendo a amora preta e a framboesa em Lavras, Minas Gerais. Conduziu-se, este estudo, com o objetivo de avaliar várias características dessas espécies. A análise realizada por esses pesquisadores resultou na conclusão de que a maioria das cultivares apresenta um ciclo de crescimento superior a 100 dias. Notavelmente, a cultivar Brazos se destacou pela maior produtividade, tornando-se uma opção viável para os produtores da região.

Para o desenvolvimento desses frutos, condições climáticas específicas são necessárias. Isso inclui temperaturas moderadas, durante o verão, níveis adequados de luminosidade, precipitação pluviométrica apropriada, baixas temperaturas no inverno e, de preferência, altitudes superiores a 1000 metros ou, no mínimo, 200 horas de frio anual. Essas especificações climáticas elucidam por que as regiões sul e sudeste do Brasil são as mais propícias para o cultivo desse tipo de fruto (ANTUNES; TREVISAN; PEREIRA, 2007).

A produção da amora preta está em franca expansão no mercado brasileiro, impulsionada por suas notáveis características, que incluem a riqueza em antioxidantes e pigmentos. Esse aumento na adoção da cultura da amora preta também é atribuído ao baixo custo de implementação em comparação com outras variedades de frutos. Isso faz da amora preta uma fonte adicional de renda para a agricultura familiar (CAMPAGNOLO; PIO, 2012).

Embora seja recomendado para consumo *in natura*, a amora preta é suscetível a deterioração rápida, em decorrência de sua estrutura frágil, suscetibilidade a doenças e alto metabolismo. Isso resulta na possibilidade de perda considerável da produção (PALHARINI et al., 2015). Por essarazão, grande parte da colheita é direcionada para o processamento, sendo transformada em geleias, sucos, polpa para sorvetes e corantes naturais (FACHINELLO et al., 2011).

A seguir, será apresentada uma síntese das principais características da cultivar a ser analisada neste estudo.

2.1.1 Choctaw

A cultivar 'Choctaw' teve sua origem no Programa de Melhoramento da University of Arkansas, sendo desenvolvida em 1975, a partir do cruzamento entre as variedades 'Brazos' e 'Darrow', compartilhando a mesma linhagem com as cultivares Comanche e Caingangue. Notabiliza-se por sua alta prolificidade, postura ereta, notável produtividade e facilidade no desenvolvimento de hastes. Além disso, exibe uma notável resistência ao frio durante o inverno. Seus frutos, caracterizados por sementes pequenas e textura firme, possuem um peso médio de 5g e um perfil gustativo mais ácido do que doce. A floração se dá em outubro, enquanto a maturação ocorre em novembro (CURI et al., 2015).

Essa cultivar surgiu a partir da colaboração entre a Embrapa e a University of Arkansas. Sua floração tem início no final de agosto, com a colheita agendada para novembro. A planta se caracteriza por apresentar espinhos e hastes eretas. Apesar de seu sabor tender mais para o ácido do que para o doce, os frutos são recomendados para consumo *in natura*. Em um estudo

realizado em Lavras - MG, a cultivar 'Choctaw' demonstrou uma produtividade média de 6.650,00 kg/ha (CURI et al., 2015).

2.2 Coco Anão Verde

O coco anão verde, cientificamente conhecido como *Cocos nucifera* var. *nana*, é uma variedade de coqueiro que produz cocos menores em comparação com a variedade tradicional. Essa planta é nativa de regiões tropicais e é apreciada por suas características compactas e frutos de tamanho reduzido. O coco anão verde é frequentemente cultivado em jardins, quintais e até mesmo em vasos, tornando-o uma escolha popular para áreas com espaço limitado.

Uma das características mais marcantes do coco anão verde é o tamanho de suas palmeiras e frutos. As árvores geralmente atingem uma altura máxima de 6 a 9 metros, tornando-as mais gerenciáveis em comparação com as variedades de coqueiros tradicionais, que podem crescer até 30 metros ou mais. Os cocos anões verdes também têm frutos menores, mas igualmente saborosos e nutritivos. Eles têm uma casca fina e são conhecidos por seu conteúdo de água doce e carne macia.

2.3 Crescimento de Frutos

Nos últimos anos, tem havido um notável aumento, tanto na quantidade quanto na relevância de estudos voltados à modelagem do crescimento de frutos. Esse incremento pode ser, em parte, atribuído aos avanços tecnológicos em computação, que tornaram a modelagem mais acessível, por meio de modelos não lineares de crescimento. Anteriormente, esse tipo de modelagem era mais complexo, em razão da necessidade de se obter estimativas aproximadas, por meio de métodos iterativos (MICHAN; PINHO, 2014).

Em relação ao estudo conduzido por Benassi et al. (2006), seu objetivo era definir a curva de crescimento dos frutos de coco. Isso envolveu avaliar parâmetros como diâmetro longitudinal e transversal, bem como massas do fruto, albúmen sólido, albúmen líquido, casca, fibras, endocarpo e brácteas. Além disso, o estudo também buscou estabelecer a curva de acúmulo de água-de-coco no interior da noz e avaliar os estágios de formação, por meio de avaliações sensoriais. O modelo Logístico foi considerado apropriado para o diâmetro externo longitudinal e a massa sólida, enquanto para as demais variáveis, foi aplicada uma combinação do modelo Logístico com o exponencial.

No que se refere à análise de dados relacionados aos frutos de tamarindo, Gurjão et al. (2006) investigaram as mudanças fisiológicas ocorrentes, durante o desenvolvimento desses frutos. O propósito era obter informações para aprimorar, tanto o manejo prévio quanto o pós-colheita. Os resultados evidenciaram que o crescimento dos frutos seguiam um padrão sigmoide simples. O comprimento, a largura e espessura manifestaram duas fases distintas de crescimento, enquanto a massa da matéria seca revelou três fases distintas de crescimento.

Explorando as características físicas, químicas e fisiológicas dos frutos de gabioba, Silva et al. (2009) empregaram modelos lineares para examinar o desenvolvimento desses frutos visando a otimizar o manejo pré e pós-colheita. Embora tenham identificado um comportamento de crescimento sigmoide, os autores optaram por utilizar modelos de terceiro grau para caracterizar tal processo.

2.4 Modelos de regressão

Os modelos de regressão tiveram sua origem no Século XIX, com Francis Galton, que realizou um estudo para examinar a relação entre a altura dos pais e a altura dos filhos. Seu objetivo era investigar se a estatura dos pais influenciava a altura dos filhos. Contudo, ele constatou que quando o pai apresentava uma altura muito acima ou muito abaixo da média, a altura dos filhos tenderia a regredir para a média, o que deu origem ao termo "regressão" (DEMÉTRIO; ZOCCHI, 2006).

De maneira geral, os modelos de regressão consistem em equações matemáticas que estabelecem uma relação entre uma variável dependente (ou resposta) e uma ou mais variáveis independentes (ou regressoras). Essas variáveis são relacionadas, por meio de uma função, que pode ser linear ou não linear (RENCHEER; SHAALJE, 2008). A forma matemática do modelo é a seguinte:

$$Y_i = f(X_i) \quad (2.1)$$

Nessa expressão, Y_i denota o vetor que compreende as respostas observadas, X_i representa o vetor contendo as variáveis independentes e $f(.)$ é uma função, que pode ser tanto linear como não linear.

No entanto, em contextos de estudos experimentais ou observacionais, há um fator que escapa ao controle direto do pesquisador. Isso torna a equação 2.1 inadequada para descrever de maneira precisa o fenômeno sob investigação. A razão para isso é que a equação 2.1 é exata,

enquanto nos estudos experimentais e observacionais existe uma margem de erro inerente. Portanto, para acomodar essa variabilidade e incerteza, se faz necessário empregar um modelo estatístico, cuja representação é dada pela seguinte expressão.

$$Y_i = f(\mathbf{X}_i; \boldsymbol{\theta}) + \epsilon_i \quad (2.2)$$

Nesse contexto, Y_i , \mathbf{X}_i e $f(\cdot)$ possuem as mesmas definições conforme a equação 2.1. O vetor $\boldsymbol{\theta}$ é composto pelos parâmetros do modelo, enquanto ϵ_i representa o erro aleatório associado ao modelo. Como premissa inicial, assume-se que ϵ_i seja independente e provenha de uma distribuição normal com média zero e variância constante (HOFFMANN, 2015).

2.5 Modelo de regressão linear fixo e misto (MLM)

De acordo com Draper e Smith (1998), um modelo é considerado linear quando as derivadas parciais não estão sujeitas a nenhuma dependência em relação a parâmetros, como exemplificado a seguir:

1. $f(X; \theta) = \theta \rightarrow \frac{\partial f}{\partial \theta}$
2. $f(X; \theta; \alpha) = \theta + \alpha X \rightarrow \frac{\partial f}{\partial \theta} = 1$ e $\frac{\partial f}{\partial \alpha} = X$

Em contextos de dados longitudinais, ou seja, observações obtidas repetidamente do mesmo indivíduo, ao longo do tempo, é esperado que ocorram erros dependentes. Ignorar essa característica pode resultar em ajustes inadequados dos modelos (PINHEIRO; BATES, 2004). Laird e Ware (1982) propuseram uma abordagem alternativa para lidar com a situação mencionada no parágrafo anterior, conhecida como metodologia de modelos mistos, também chamados de modelos hierárquicos ou multiníveis. Essa abordagem foi inicialmente introduzida por Henderson (1975) e ganhou maior destaque, na década de 80, com o trabalho de Laird e Ware (1982).

Por definição, os modelos lineares mistos (MLM) consideram simultaneamente a presença de efeitos fixos e aleatórios, excluindo apenas o erro experimental, que é sempre considerado aleatório (PINHEIRO; BATES, 2004).

Dependendo dos objetivos do estudo, os MLM podem ser formulados em dois estágios, conforme discutido por Harvielle (1977), Laird e Ware (1982), Ware (1985) e apresentado por Medeiros et al. (2020).

- a) Primeiro Estágio: Considera o indivíduo i ($i=1,\dots,n$) e a medida no tempo t_{ij} , com $j(j=1,\dots,n_i)$. No primeiro estágio, é ajustado um modelo de regressão para cada indivíduo e se baseia no comportamento dos perfis individuais.
- b) Segundo Estágio: O objetivo desse estágio é verificação a variabilidade entre os indivíduos, caso deste trabalho, representado por:

$$\boldsymbol{\beta}_i = \mathbf{K}_i \boldsymbol{\beta} + \mathbf{b}_i \quad (2.3)$$

Nesse cenário, \mathbf{K}_i representa-se uma matriz de covariâncias conhecida ($q \times p$), $\boldsymbol{\beta}$ é o vetor de parâmetros desconhecidos de ordem p , que é o foco da estimativa. Além disso, \mathbf{b}_i denota o vetor de efeitos aleatórios, seguindo uma distribuição normal $N(0, \mathbf{D})$, em que \mathbf{D} representa a matriz de variâncias e covariâncias associadas à dependência e heterocedasticidade dos resíduos.

A equação geral desse modelo, segundo Searle; Cassela; McCulloch (2009), Lindstrom e Bates (1990) é a seguinte:

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i \quad (2.4)$$

Nesse contexto, \mathbf{Y}_i constitui um vetor das variáveis observadas, com dimensões $n \times 1$; \mathbf{X}_i denota a matriz de delineamento dos efeitos fixos, de dimensões $n \times p$; \mathbf{Z}_i é a matriz de incidência dos efeitos aleatórios (covariáveis), com dimensões $n \times q$; $\boldsymbol{\beta}$ é o vetor dos parâmetros dos efeitos fixos, possuindo dimensões $p \times 1$; \mathbf{b}_i representa o vetor dos parâmetros dos efeitos aleatórios, com dimensões $q \times 1$; e $\boldsymbol{\epsilon}_i$ indica o vetor das variáveis de erro aleatório, tendo dimensões $n \times 1$.

No que se refere aos parâmetros dos efeitos aleatórios (\mathbf{b}) e aos erros aleatórios ($\boldsymbol{\epsilon}_i$), como premissa inicial, assume-se que sigam uma distribuição normal com média zero e matriz de variância e covariância:

$$\mathit{var}(\boldsymbol{\epsilon}) = \mathbf{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \mathbf{R}$$

$$\mathit{var}(\mathbf{b}) = \mathbf{E}(\mathbf{b}\mathbf{b}') = \mathbf{D}$$

$$\mathit{var}(\mathbf{Y}) = \mathit{var}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon})$$

$$\mathit{var}(\mathbf{Y}) = \mathit{var}(\mathbf{X}\boldsymbol{\beta}) + \mathit{var}(\mathbf{Z}\mathbf{b}) + \mathit{var}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')$$

$$\text{var}(\mathbf{Y}) = \mathbf{0} + \mathbf{Z}\text{var}(\mathbf{b})\mathbf{Z}' + \mathbf{R}$$

$$\text{var}(\mathbf{Y}) = \mathbf{0} + \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R}$$

E a esperança de \mathbf{Y} é:

$$\mathbf{E}(\mathbf{Y}) = \mathbf{E}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon})$$

$$\mathbf{E}(\mathbf{Y}) = \mathbf{E}(\mathbf{X}\boldsymbol{\beta}) + \mathbf{E}(\mathbf{Z}\mathbf{b}) + \mathbf{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')$$

$$\mathbf{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$$

Com isso,

$$\mathbf{Y} \sim \mathbf{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R})$$

Reescrevendo esse modelo de forma matricial fica,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 & z_1 & 0 & \cdots & 0 \\ x_2 & 0 & z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & 0 & 0 & \cdots & z_n \end{bmatrix} \times \begin{bmatrix} \beta \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Como modelo hierárquico, o modelo linear de efeitos mistos pode ser reestruturado conforme a equação a seguir:

$$y|b \sim \mathbf{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R})$$

e

$$\mathbf{b} \sim \mathbf{N}(\mathbf{0}, \mathbf{D}) \quad (2.5)$$

$$y_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Z}_1\mathbf{u}_1 + \cdots + \mathbf{Z}_r\mathbf{u}_r + \boldsymbol{\epsilon}_i \quad (2.6)$$

2.6 Modelo de regressão não linear de efeitos fixos

No modelo delineado por Lindstrom e Bates (1990), os efeitos aleatórios constituem elementos dos componentes não lineares da estrutura do modelo, que pode ser expresso da seguinte forma:

$$y_i = f(\mathbf{X}_i; \boldsymbol{\beta}; \mathbf{Z}_i; \mathbf{b}_i) + \boldsymbol{\epsilon}_i \quad (2.7)$$

Na estrutura apresentada:

$y_i = [y_{i1}, \dots, y_{in_i}]'$ representa o vetor ($n_i \times 1$) das respostas do i -ésimo indivíduo, em que n_i indica o número de momentos temporais observados no i -ésimo indivíduo, os quais podem variar.

$x_i = [x_{i1}, \dots, x_{iw}]'$ corresponde a uma matriz conhecida ($n_i \times w$) de valores dentro do indivíduo, podendo incluir apenas os instantes de avaliação.

Os primeiros registros de estudos que abordam análise de regressão não linear remontam à década de 1920, com contribuições de Ronald Aylmer Fisher e W. A. Mackenzie. Contudo, devido à complexidade apresentada por esses modelos, apenas na década de 1970 eles começaram a ser explorados de maneira mais aprofundada. Esse fenômeno pode ser atribuído, em parte, às limitações das calculadoras da época, as quais não eram suficientemente avançadas para lidar com os cálculos complexos associados a tais modelos (DODGE, 2008; WYZYKOWSKI et al., 2015).

A análise de regressão não linear segue o mesmo princípio dos modelos lineares, essencialmente buscando examinar a relação funcional entre uma variável dependente e uma ou mais variáveis independentes. A distinção primordial reside na natureza da função que estabelece essa relação. Conforme destacado por Draper e Smith (1998), a diferenciação entre os modelos lineares e não lineares pode ser identificada, por meio das derivadas parciais em relação aos parâmetros. No caso da regressão linear, após derivar essas relações, as derivadas não dependem de nenhum parâmetro. Em contrapartida, nos modelos não lineares, as derivadas parciais em relação aos parâmetros sempre serão dependentes de pelo menos um parâmetro, como ilustrado no exemplo a seguir:

$$f(X; \beta; \alpha) = \theta + \beta^{\alpha X} \rightarrow \frac{\partial f}{\partial \theta} = 1, \frac{\partial f}{\partial \beta} = e^{\alpha X}, \frac{\partial f}{\partial \alpha} = \alpha X e^{\alpha X}$$

Ao lidar com o estudo do crescimento de organismos vivos, é comum esperar um padrão em forma de "S" ou sigmoide. Esse padrão sugere um início de desenvolvimento acelerado, seguido por uma mudança na velocidade de crescimento (ponto de inflexão) e, por fim, uma estabilização que resulta no término do crescimento (DRAPER, SHIMIT, 2011). Embora os modelos lineares se ajustem razoavelmente a esse tipo de curva, especialmente os modelos de terceiro grau, eles carecem de interpretação prática e direta de seus parâmetros. Além disso, seu alcance é limitado, permitindo apenas a estimativa de valores para variáveis dependentes ou independentes. Diante dessa lacuna, os modelos não lineares de crescimento surgem como uma

classe vantajosa, com destaque para aqueles pertencentes à família Richards (como Brody, Gompertz, Logístico e Von Bertalanffy).

Os modelos de crescimento da família Richards trazem consigo uma série de vantagens em relação aos modelos lineares, incluindo parcimônia, interpretabilidade de seus parâmetros e a capacidade de extrapolar inferências para além do domínio das variáveis explanatórias e de resposta. No entanto, até o virar do século 21, a utilização desses modelos era dificultada pela necessidade de empregar métodos iterativos para a obtenção das estimativas. Com o avanço tecnológico e o desenvolvimento de computadores mais poderosos, esses modelos tornaram-se mais acessíveis e, conseqüentemente, mais empregados.

Apesar desse progresso, ainda nos dias atuais, alguns pesquisadores optam por não adotar esses modelos não lineares. Eles preferem utilizar modelos lineares para descrever o crescimento de organismos vivos, o que pode parecer uma escolha razoável à primeira vista. No entanto, ao aprofundar a análise das possibilidades, torna-se evidente que os modelos não lineares de crescimento são mais adequados. Além de estimar as variáveis de resposta, esses modelos não lineares oferecem informações valiosas que podem ser exploradas em maior profundidade.

2.7 Modelo de regressão não linear de efeitos mistos

Os modelos não lineares de efeitos mistos constituem uma extensão dos modelos lineares de efeitos mistos (BATES; WATS, 1998). Essa classe de modelos possibilita a análise da variação entre os indivíduos, incorporando efeitos aleatórios aos modelos.

Conforme indicado por Bates e Watts (1990), os modelos não lineares mistos (MNLM) se mostram particularmente apropriados para lidar com dados de medidas repetidas, uma vez que conseguem capturar a dependência entre os indivíduos nesse tipo de estudo. Dessa forma, o pressuposto de independência, que é fundamental nos modelos clássicos, torna-se mais flexível nesse contexto.

A formulação do MNLM é expressa pela seguinte equação:

$$y_i = f(\mathbf{X}_i; \boldsymbol{\beta}_i; \mathbf{Z}_i; \mathbf{b}_i) + \epsilon_{ij} \quad (2.8)$$

O vetor $y_i [y_{i1}, y_{i2}, \dots, y_{in_i}]'$ representa as respostas observadas, onde n_i é o número de tempos observados para o i -ésimo indivíduo. A matriz $\mathbf{X}_i [x_{i1}, x_{i2}, \dots, x_{in_i}]'$ é de dimensão

$(n_i \times 1)$ e contém os valores intra-indivíduos. O vetor $\boldsymbol{\beta}$ é um conjunto de parâmetros de dimensão $(p \times 1)$, \mathbf{Z}_i são constantes conhecidas que relacionam \mathbf{y}_i com \mathbf{b}_i , onde \mathbf{b}_i é um vetor de coeficientes aleatórios de dimensão $(q \times 1)$. A função $f()$ é uma função não linear em relação aos parâmetros e $\boldsymbol{\varepsilon}_i$ é o erro aleatório do modelo, de dimensão $(n_i \times 1)$.

Os pressupostos para as análises dos modelos consideram que os $\boldsymbol{\varepsilon}_i$ são independentes e identicamente distribuídos com média zero e matriz de variâncias e covariâncias $\boldsymbol{\sigma}^2 \mathbf{R}_i (n_i \times n_i)$. A matriz \mathbf{R}_i pode assumir diversas estruturas de variância e covariância, sendo a combinação de $\boldsymbol{\sigma}^2 \mathbf{R}_i$ equivalente à matriz identidade $\boldsymbol{\sigma}^2 \mathbf{I}_n$ em alguns casos.

O vetor de parâmetros \mathbf{b}_i é considerado independente e identicamente distribuído, de acordo com uma distribuição normal com média zero e matriz de covariância $\boldsymbol{\sigma}^2 \mathbf{D}$. Vale ressaltar que \mathbf{b}_i representa a parte aleatória do modelo e é independente do vetor $\boldsymbol{\varepsilon}_i$ (PINHEIRO; BATES, 2004).

Reestruturando o modelo 2.8 de maneira matricial, podemos representar as observações y_{ij} com t_{ij} medidas repetidas da seguinte forma:

$$y_{ij} = f(\phi_i, x_{ij}) + \epsilon_{ij} \quad (2.9)$$

onde $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, n_i$.

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{bmatrix}, \boldsymbol{\epsilon}_i = \begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{in_i} \end{bmatrix}, \boldsymbol{\eta}_i(\boldsymbol{\phi}_i) = \begin{bmatrix} f(\boldsymbol{\phi}_i, \mathbf{x}_{i1}) \\ f(\boldsymbol{\phi}_i, \mathbf{x}_{i2}) \\ \vdots \\ f(\boldsymbol{\phi}_i, \mathbf{x}_{in_i}) \end{bmatrix}$$

Nesse contexto, o erro aleatório ϵ_i segue uma distribuição normal com média zero e matriz de variâncias e covariâncias $\boldsymbol{\sigma}^2 \mathbf{R}$, onde \mathbf{R}_i é uma matriz de dimensão $(n_i \times n_i)$ que representa as variâncias e covariâncias. Embora em muitas situações a matriz \mathbf{R}_i seja igual à matriz identidade \mathbf{I} , a inclusão de \mathbf{R}_i no modelo possibilita especificar a estrutura de variâncias e covariâncias de maneira mais geral. Isso permite adaptar o modelo para diferentes padrões de dependência e heterocedasticidade nos dados. Vale ressaltar que a inclusão de \mathbf{R}_i torna desnecessária a utilização da estrutura autoregressiva de ordem (p) (WYZYKOWSKI et al., 2015; LINDSTROM; BATES, 1990).

Em diversas situações, a matriz \mathbf{R}_i é escolhida de acordo com o conhecimento prévio sobre a estrutura dos erros, permitindo assim uma modelagem mais precisa da dependência e heterocedasticidade presentes nos dados.

$$\mathbf{y}_i = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}, \boldsymbol{\phi}_i = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_M \end{bmatrix}, \boldsymbol{\eta}_i(\boldsymbol{\phi}_i) = \begin{bmatrix} \eta_1(\phi_1) \\ \eta_1(\phi_2) \\ \vdots \\ \eta_M(\phi_M) \end{bmatrix}$$

$D = \text{diag}(D, D, \dots, D)$ e $R = \text{diag}(R_1, R_2, \dots, R_M)$. Agora o modelo geral é:

$$y|b \sim N(n(\boldsymbol{\phi}, \boldsymbol{\sigma}^2 \mathbf{R})), \boldsymbol{\phi} = A\boldsymbol{\beta} + Bb \quad (2.11)$$

$$b \sim N(0, \sigma^2 \tilde{D})$$

Em que,

$$B = \text{diag}(\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_M)$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}, A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{bmatrix}$$

2.8 Métodos de estimação de parâmetros

Na estimação dos parâmetros dos modelos de regressão, tanto lineares quanto não lineares, foram empregados dois dos principais métodos: mínimos quadrados e máxima verossimilhança.

2.8.1 Estimação de $\boldsymbol{\beta}$ e \mathbf{b}

Lindstrom e Bates (1990) destacam que quando os componentes de variância \mathbf{R} e \mathbf{D} são conhecidos, e $\boldsymbol{\eta}$ é uma função linear de $\boldsymbol{\beta}$ e \mathbf{b} ($\boldsymbol{\eta}(\mathbf{A}_i + \mathbf{B}\mathbf{b}) = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}$), onde $\mathbf{X} = [\mathbf{X}_1^T, \mathbf{X}_1^T, \dots, \mathbf{X}_M^T]$ e $\mathbf{Z} = \text{diag}(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_M)$, a estimação de $\boldsymbol{\beta}$ e \mathbf{b} é frequentemente realizada,

utilizando mínimos quadrados generalizados. Isso proporciona um método eficaz para estimar esses parâmetros em modelos não lineares de efeitos mistos.

$$\widehat{\beta}_{lin} = \widehat{\beta}_{lin}(\theta)(\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y} \quad (2.12)$$

e o preditor linear de $\widehat{\beta}_{lin} = \widehat{\beta}_{lin}(\theta) = \widetilde{\mathbf{D}} \mathbf{Z}^{-1} \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}) \widehat{\beta}_{lin}(\theta)$ onde $\mathbf{V} = \mathbf{A} + \mathbf{Z} \widetilde{\mathbf{D}} \mathbf{Z}^T$ e θ é um vetor contém os elementos de \mathbf{D} e os parâmetros em \mathbf{R}_i . O vetor de parâmetros θ inclui todos os parâmetros ligados a variância, exceto o σ , sem especificar como \mathbf{R}_i é parametrizado. As estimativas de $\widehat{\beta}_{lin}$ e $\widehat{\mathbf{b}}_{lin}$ conjuntamente maximizam a função:

$$g_{lin} = (\beta, b | y) = \frac{1}{2} \sigma^2 (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}b)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}b) - \frac{1}{2} \sigma^{-2} \mathbf{b}^T \widetilde{\mathbf{D}}^{-1} \mathbf{b} \quad (2.13)$$

Considerando β como um valor fixo e g como uma constante acrescida do logaritmo da função de densidade a posteriori de b , torna-se evidente que b maximiza g para um dado valor de β . Portanto, $\tilde{\beta}$ é o estimador de máxima verossimilhança associado a uma distribuição aproximada de y . Similar ao caso linear, essas estimativas podem ser obtidas mediante o acréscimo de "pseudo dados" ao vetor de observações, conforme detalhado por Wyzykowski et al. (2015) e Lindstrom e Bates (1990):

$$\tilde{y} = \tilde{\eta}(\mathbf{A}\beta + \mathbf{B}b) + \tilde{\epsilon} \quad (2.14)$$

$$\text{Em que, } A = \begin{bmatrix} R^{-1/2} y \\ 0 \end{bmatrix}, \tilde{\epsilon} \sim (0, \sigma^2 I) \text{ e } \tilde{\eta}(\mathbf{A}\beta + \mathbf{B}b) = \begin{bmatrix} R^{-1/2} (\mathbf{A}\beta + \mathbf{B}b) \\ \widetilde{\mathbf{D}}^{-1/2} b \end{bmatrix}.$$

2.8.2 Estimação de β

Na literatura, há uma variedade de métodos para estimar parâmetros, incluindo o método de momentos, mínimos quadrados, métodos bayesianos, entre outros. Entretanto, ao lidar com modelos não lineares mistos, os métodos de máxima verossimilhança restrita (MVR) e máxima verossimilhança (MV) são amplamente empregados.

2.9 Máxima Verossimilhança (MV)

Na literatura, existem vários métodos para estimar parâmetros em modelos estatísticos, sendo um dos mais fundamentais e amplamente utilizados o método da máxima verossimilhança (MV). O método da máxima verossimilhança é uma abordagem estatística que busca encontrar os valores dos parâmetros de um modelo probabilístico que maximizam a verossimilhança dos dados observados.

Autores renomados como Casella e Berger (2002), Cox e Hinkley (1974), e Pawitan (2001) destacaram a importância da verossimilhança em seus trabalhos. A ideia central por trás do método da máxima verossimilhança é encontrar os parâmetros que tornam os dados observados mais prováveis de terem sido gerados pelo modelo proposto. Isso é feito calculando a função de verossimilhança, que é uma medida da probabilidade dos dados observados sob as condições do modelo. A estimativa de máxima verossimilhança é obtida encontrando os valores dos parâmetros que maximizam essa função.

A maximização da função de verossimilhança pode ser realizada usando técnicas computacionais avançadas, como algoritmos de otimização. Uma vez que os parâmetros que maximizam a verossimilhança são encontrados, eles são considerados as estimativas de máxima verossimilhança dos parâmetros do modelo.

O método da máxima verossimilhança tem propriedades estatísticas bem estabelecidas, como eficiência assintótica e boas propriedades de estimativas em grandes amostras. Além disso, o método da máxima verossimilhança oferece uma abordagem sistemática e coerente para a estimativa de parâmetros em modelos estatísticos, o que o torna uma ferramenta poderosa e amplamente aplicável em diversas áreas da análise estatística.

De forma geral, a estrutura da função de verossimilhança é a seguinte:

$$\int p(y|b)p(d)db \quad (2.15)$$

mas a função esperança de η é não linear em b , com isso não é possível encontrar uma forma fechada para encontrar as estimativas (WYZYKOWSKI et al., 2015; LINDSTROM; BATES, 1990). Com essa dificuldade, é necessário aproximar os valores das estimativas por uma normal multivariada. Para isso aproxima-se o erro $y - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}b)$ por série de Taylor na vizinhança de b .

$$y - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}b) \approx y - [\eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\hat{b}) + \hat{\mathbf{Z}}b - \hat{\mathbf{Z}}\hat{b}]$$

onde,

$$\widehat{\mathbf{z}}_i = \widehat{\mathbf{z}}_i(\theta) = \frac{\partial \eta_i}{\partial \mathbf{b}_i^T} \Big|_{\widehat{\beta}, \widehat{b}} = \frac{\partial \eta_i}{\partial \boldsymbol{\phi}^T} \Big|_{\widehat{\beta}, \widehat{b}} \mathbf{B}_i,$$

$$\widehat{\mathbf{Z}} = \widehat{\mathbf{Z}}_1(\theta) = \text{diag}(\widehat{\mathbf{z}}_1, \widehat{\mathbf{z}}_2, \dots, \widehat{\mathbf{z}}_n) = \frac{\partial \boldsymbol{\eta}}{\partial \mathbf{b}^T} \Big|_{\widehat{\beta}, \widehat{b}} \quad (2.16)$$

Com $\widehat{\mathbf{Z}}$ dependendo de θ , pois \mathbf{B} e \mathbf{b} são. Então, a distribuição aproximada de y é

$$y - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widehat{\mathbf{b}}) + \widehat{\mathbf{Z}}\widehat{\mathbf{b}} - \widehat{\mathbf{Z}}\mathbf{b} | b \sim N(0, \sigma^2 \mathbf{R})$$

Sendo a distribuição aproximada de y ,

$$y | b - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widehat{\mathbf{b}}) + \widehat{\mathbf{Z}}\widehat{\mathbf{b}} - \widehat{\mathbf{Z}}\mathbf{b} | b \sim N(0, \sigma^2 \mathbf{R}) \quad (2.17)$$

Essa expressão, em conjunto com a distribuição de b (2.11), permite que se aproxime a distribuição marginal de y como:

$$y \sim N(\eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widehat{\mathbf{b}}) - \widehat{\mathbf{Z}}\widehat{\mathbf{b}}, \sigma^2 \widehat{\mathbf{V}}) \quad (2.18)$$

Com, $\widehat{\mathbf{V}} = \widehat{\mathbf{V}}(\theta) = \mathbf{R} + \widehat{\mathbf{Z}}\widehat{\mathbf{D}}\widehat{\mathbf{Z}}^T$ Aplicando o logarítmico na função de verossimilhança na equação 2.17, fica:

$$l_F(\beta, \sigma, \theta | y) = -\frac{1}{2} \log |\sigma^2 \widehat{\mathbf{V}}| - \frac{1}{2} \sigma^2 [y - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widehat{\mathbf{b}})]^T \widehat{\mathbf{V}}^{-1} [y - \eta(\mathbf{A}\boldsymbol{\beta} + \mathbf{B}\widehat{\mathbf{b}})] \quad (2.19)$$

Nesse caso, $\widehat{\beta}$ e $\widehat{\mathbf{Z}}$ dependem de θ e β^{MV} , σ^{MV} e θ^{MV} são os estimadores de máxima verossimilhança de β , σ e θ .

2.10 Máxima verossimilhança restrita (MVR)

A Máxima Verossimilhança Restrita (MVR) é uma extensão do método de Máxima Verossimilhança (MV) que incorpora restrições aos parâmetros do modelo, levando em consideração informações adicionais sobre os parâmetros estimados. Autores proeminentes, como Casella e Berger (2002), Cox e Hinkley (1974), e Pawitan (2001), exploraram a importância dessa abordagem em suas obras.

Na MVR, as restrições podem ser impostas de diferentes maneiras, como por meio de limites superiores ou inferiores nos parâmetros, relações lineares entre os parâmetros ou impondo que certas combinações lineares de parâmetros sejam iguais a valores pré-especificados. Essas restrições refletem conhecimento prévio sobre os parâmetros ou restrições teóricas que se deseja incorporar ao modelo.

A principal vantagem da MVR é a incorporação de informações adicionais ao processo de estimação dos parâmetros. Isso pode melhorar a precisão das estimativas, especialmente quando os dados observados não são suficientemente informativos por si só. Ao impor restrições, a MVR utiliza a verossimilhança não apenas para ajustar o modelo aos dados, mas também para atender às restrições impostas.

A Máxima Verossimilhança Restrita é uma ferramenta poderosa que combina a capacidade do método de Máxima Verossimilhança de encontrar estimativas pontuais eficientes com a incorporação de informações adicionais. Isso torna a MVR uma abordagem valiosa em situações em que se deseja fazer inferências que levem em consideração restrições teóricas ou conhecimento prévio sobre os parâmetros do modelo.

O método de Máxima Verossimilhança Restrita é particularmente útil em contextos nos quais os parâmetros do modelo têm interpretações específicas ou são limitados por restrições práticas. Ele permite incorporar essas informações no processo de estimação, tornando as estimativas mais realistas e coerentes com a natureza do problema em análise. Esse método é o mesmo usado para os estimadores de MV, exceto que o logaritmo da função de verossimilhança l_F é:

$$l_R = (\beta, \sigma, \theta | y) = -\frac{1}{2} \log |\sigma^2 \widehat{\mathbf{V}}^T \widehat{\mathbf{V}} \widehat{\mathbf{X}}| + l_F \quad (2.20)$$

Em que

$$\mathbf{X}_i = \widehat{\mathbf{X}}_i(\boldsymbol{\theta}) = \left(\frac{\partial \eta_i}{\partial \boldsymbol{\beta}_i^T} \Big|_{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}} \right) = \left(\frac{\partial \eta_i}{\partial \boldsymbol{\phi}_i^T} \Big|_{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}} \right) \mathbf{A}_i \quad (2.21)$$

e

$$\widehat{\mathbf{X}} = \widehat{\mathbf{X}}(\boldsymbol{\theta}) = \begin{pmatrix} \widehat{X}_1 \\ \widehat{X}_2 \\ \vdots \\ \widehat{X}_M \end{pmatrix} = \frac{\partial \eta_i}{\partial \boldsymbol{\beta}_i^T} \Big|_{\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{b}}}$$

Os estimadores β^{MV} , σ^{MV} e θ^{MV} são aqueles que maximizam l_R (LINDSTROM; BATES, 1990).

2.10.1 Mínimos quadrado

O método dos mínimos quadrados (MMQ) envolve a minimização da soma dos quadrados das diferenças entre os valores estimados e os dados observados. Sob determinadas condições, os estimadores obtidos, por meio desse método, são não enviesados e possuem variância mínima (DRAPER; SMITH, 1998). Para realizar ajustes em modelos utilizando esse método, diversos procedimentos estão disponíveis na literatura, sendo o método de mínimos quadrados um dos mais amplamente empregados. Gallant (1987), por meio da análise de resíduos, categorizou o método de mínimos quadrados em três abordagens distintas:

- a) Ordinários: Nessa categoria, os erros não violam nenhuma das pressuposições do modelo, ou seja, são distribuídos de acordo com $\epsilon_{ij} \sim N(0, \sigma^2)$.
- b) Ponderados: Nessa abordagem, abrangem-se situações em que os erros apresentam heterocedasticidade, ou seja, as variâncias dos erros não são constantes. Os erros são modelados como $\epsilon_{ij} \sim N(0, \mathbf{D}\sigma^2)$, em que \mathbf{D} é uma matriz diagonal ponderada pela variância σ^2 .
- c) Generalizados: Nessa categoria, os erros violam a suposição de homocedasticidade e/ou independência. Os erros são representados por $\epsilon_{ij} \sim N(0, \mathbf{M}\sigma^2)$, em que \mathbf{M} é a matriz de variâncias e covariâncias dos resíduos. A diagonal principal da matriz \mathbf{M} contém as variâncias, enquanto os demais elementos são as covariâncias entre ϵ_{ij} para $i \neq j$.

Essas categorias auxiliam na compreensão das diferentes formas pelas quais os erros podem se comportar em relação às suposições do modelo. A escolha da abordagem apropriada depende da análise cuidadosa dos dados e da seleção daquela que melhor se ajusta à realidade dos dados observados.

2.11 Métodos iterativos

Ao minimizar a soma de quadrados, a estimativa dos parâmetros $\hat{\beta}$ obtida como $(X'X)^{-1}X'Y$. Dado que sistemas de equações não lineares não possuem soluções explícitas para estimar os parâmetros, métodos iterativos são empregados para alcançar essa estimativa.

Esses métodos envolvem aproximações lineares da função $f(x, \beta)$ em cada etapa, continuando até que um critério de parada predefinido seja atingido.

Diversos métodos estão disponíveis na literatura, como Gradiente, Gauss-Newton, Newton e Levenberg-Marquardt. De acordo com Bates e Watts (1988), esses métodos se distinguem pela forma como $\beta - \beta^0$ é calculado, em que β^0 é o valor inicial para o processo iterativo e β é o valor aproximado. De maneira geral, os critérios de parada incluem:

Convergência: A iteração cessa quando a diferença entre os valores sucessivos de β cai abaixo de um limiar específico. **Número máximo de iterações:** Um número pré-determinado de iterações é definido e o processo para após atingir esse limite. **Alteração pequena nas funções objetivo:** O processo iterativo cessa quando a variação na função objetivo se torna negligível. Esses critérios direcionam os métodos iterativos e garantem que a estimativa dos parâmetros seja obtida de forma eficiente e precisa, mesmo em sistemas complexos de equações não lineares:

- a) Gradiente: $\beta - \beta^0 = X'\varepsilon$
- b) Gauss-Newton: $\beta - \beta^0 = (X'X)X'\varepsilon$
- c) Newton: $\beta - \beta^0 = G^{-1}X'\varepsilon$
- d) Maquardt: $\beta - \beta^0 = [X'X + \delta (X'X)^{-1}]^{-1}X'\varepsilon$, em que $(X'X)^{-1}$ é inversa generalizada.

O método de Gauss-Newton é o mais utilizado, este usa a expansão da série de Taylor de primeira ordem. Vale ressaltar que mesmo com o critério de parada pré-estabelecido não é garantida a convergência. Para que esse método convirja e essa convergência não leve a mínimos globais, é necessário bons valores iniciais, caso contrário a convergência não é garantida.

Na estimação dos parâmetros, o método de Gauss-Newton é o mais utilizado, este usa a expansão da série de Taylor de primeira ordem. Vale ressaltar mesmo que tendo um critério de parada pré estabelecido, não é garantida a convergência. Outra situação que se deve ter muita atenção, é a possibilidade dessa convergência ir para mínimos locais e é importante ter bons

valores iniciais para contornar tal situação. Estudos anteriores e experiência do pesquisador podem ajudar a ter bons valores para a primeira iteração.

2.12 Método de Seleção de modelo

Para selecionar o modelo que melhor descreve a relação funcional estabelecida, os métodos de Critério de Informação de Akaike (AIC), Critério de Informação Bayesiano (BIC), coeficiente de determinação ajustado (R_{aj}^2) e o teste da razão de verossimilhança (TRV)

2.12.1 Critério de Informação de Akaike (AIC)

O critério de informação de Akaike (AIC) (AKAIKE, 1974), é uma estatística de comparação entre modelos e foi estimado por meio da seguinte expressão:

$$AIC = -2\log L(\hat{\theta}) + 2p \quad (2.22)$$

em que $\log L(\hat{\theta})$ é o logaritmo da função de máxima verossimilhança e p é o número de parâmetros do modelo. Para menores valores de AIC, melhor esse modelo se ajusta aos dados.

2.12.2 Critério Informação de Bayesiano (BIC)

O critério de informação Bayesiano (BIC) (SCHWARZ, 1978), um utilizado para comparar modelos, da mesma forma que o AIC, sendo usado a seguinte fórmula:

$$BIC = -2\log L(\hat{\theta}) + p\log(n) \quad (2.23)$$

sendo $L(\hat{\theta})$ e p tem as mesmas definições do AIC e n é o número de observação da amostra. Menores valores de BIC, indica um melhor ajuste do modelo.

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

De acordo com Draper e Smith (1998), o coeficiente de determinação ajustado R_{aj}^2 é uma métrica frequentemente utilizada para a seleção do melhor modelo, sendo particularmente útil quando comparando modelos com diferentes números de parâmetros. Esse método penaliza os modelos que possuem um maior número de parâmetros, introduzindo uma correção em

relação ao coeficiente de determinação (R^2). Essencialmente, o R_{aj}^2 pondera o coeficiente de determinação pela quantidade de parâmetros presentes no modelo.

A expressão do coeficiente de determinação ajustado R_{aj}^2 é a seguinte:

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

Nesse contexto, R_{aj}^2 representa o coeficiente de determinação, n denota o número de observações, i está associado ao termo de interceptação da curva e p representa o número de parâmetros. Valores maiores de R^2 indicam um ajuste mais adequado do modelo, tornando-o preferível em comparação a outros modelos em análise.

2.12.4 Teste da razão de verossimilhança (TRV)

O teste de razão de verossimilhança é uma técnica estatística fundamental utilizada para avaliar a melhoria no ajuste de dois modelos, um dos quais é um caso especial do outro. Esse teste compara o valor da função de verossimilhança dos dois modelos, permitindo a determinação estatística se a inclusão de parâmetros adicionais no modelo mais complexo resulta em um ajuste significativamente melhor aos dados.

A ideia central do teste é comparar a diferença na deviance (ou, em casos específicos, na log-verossimilhança) entre os dois modelos. A estatística de teste é distribuída como uma qui-quadrado, com graus de liberdade igual à diferença no número de parâmetros entre os dois modelos, conforme descrito por Lehmann (1986) e abordado por Pinheiro e Bates (2004).

Autores como Neyman e Pearson foram pioneiros na formulação desse teste, contribuindo para o desenvolvimento dos fundamentos estatísticos por trás dele. A técnica de teste de razão de verossimilhança é amplamente utilizada em diferentes áreas da estatística, especialmente na modelagem estatística, inferência estatística e seleção de modelos.

Nesse contexto, consideramos que L_2 é o modelo mais abrangente e L_1 é o modelo mais restrito, com menos parâmetros. Vale ressaltar que L_2 possui uma verossimilhança maior do que L_1 , e o logaritmo natural da verossimilhança de L_2 é maior do que o logaritmo natural da verossimilhança de L_1 .

As hipóteses subjacentes ao teste de razão de verossimilhança são as seguintes:

H_0 : Os parâmetros adicionais introduzidos em L_2 não são estatisticamente significativos e podem ser eliminados, tornando-o igual a L_1 .

H_1 : Os parâmetros adicionais em L_2 são estatisticamente significantes e, portanto, há uma melhoria significativa no ajuste em comparação com L_1 .

O teste de razão de verossimilhança compara a diferença entre os valores logarítmicos das verossimilhanças dos dois modelos com a distribuição de uma estatística de teste, geralmente seguindo uma distribuição qui-quadrado com graus de liberdade igual à diferença no número de parâmetros entre os dois modelos. Esse teste permite tomar decisões estatísticas objetivas sobre a adequação e a complexidade dos modelos mistos em estudo

Dessa forma, a estatística do teste da razão de verossimilhança é formulada da seguinte maneira, conforme apresentado por Wyzykowski et al. (2015) e Pinheiro e Bates (2004):

$$L = -2 \log \frac{L_1}{L_2} = -2[\log(L_2) - \log(L_1)] \quad (2.24)$$

L_1 e L_2 representam as verossimilhanças dos modelos mais restrito e mais geral, respectivamente. A estatística de teste deve ser positiva e segue uma distribuição qui-quadrado, com graus de liberdade igual à diferença no número de parâmetros entre os dois modelos (*geral - restrito*). As condições de regularidade, conforme descritas por Pinheiro e Bates (2004), para n variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com função densidade de probabilidade (fdp), são as seguintes:

- a) O espaço paramétrico Θ é um intervalo aberto;
- b) A distribuição de probabilidade acumulada, $F(x)$, tem suporte comum;
- c) $f(x, \beta)$ admite a 3ª derivada contínua em β ;
- d) $\int_0^\infty f(x; \beta) dx$ pode ser derivado três vezes dentro do sinal da integral;
- e) A informação de Fisher, $I_{F(\beta)}$, satisfaz: $0 < I_{F(\beta)} < \infty$, em que:

$$I_{F(\beta)} = E \left[\frac{\partial \log[f(x; \beta)]}{\partial \beta} \right]^2 = -E \left[\frac{\partial^2 \log[f(x; \beta)]}{\partial \beta^2} \right]$$

- f) Existe a terceira derivada satisfazendo:

$$\left[\frac{\partial^3 \log(f(x; \beta))}{\partial \beta^3} \right] \leq M(x)$$

para $\forall x \in A$ em $E[M(x)] < \infty$.

Considerando as condições de regularidade, a estatística L segue assintoticamente uma distribuição qui-quadrado (χ_r^2), em que r é o número de graus de liberdade, calculado como a diferença entre a quantidade de parâmetros do modelo mais restrito e o modelo mais geral.

REFERÊNCIAS

- AKAIKE, H. New look at statistical-model identification. **IEEE Transactions on Automatic Control**, New York, v. Ac19, n. 6, p. 716-723, 1974.
- ANTUNES, L. E. C. Amora-preta: nova opção de cultivo no brasil. **Ciência Rural**, Santa Maria, v. 32, n. 1, p. 151–158, 2002.
- ANTUNES, L. E. C.; RASSEIRA, M. Aspectos técnicos da cultura da amora-preta. **Embrapa Clima Temperado**, 2004.
- ANTUNES, L. E. C.; TREVISAN, R.; PEREIRA, I. d. S. Produção de amora-preta. **Seminário brasileiro sobre pequenas frutas**, v. 4, p. 65–71, 2007.
- ARAÚJO, S.D. A região semiárida do nordeste do Brasil: questões ambientais e possibilidades de uso sustentável dos recursos. **Rios Eletrônica-Revista Científica da FASETE**, Paulo Afonso, v.5, n.5, p.88-98, 2011.
- AROUCHA, E. M. M. et al. Análise físico-química e sensorial de água-de-coco em função de estágio de maturação das cultivares de coco anão verde e vermelho. **Agropecuária científica no semiárido**, Campina Grande, v. 10, n. 1, p. 33-38, 2014.
- CAMPAGNOLO, M. A.; PIO, R. Enraizamento de estacas caulinares e radiculares de cultivares de amoreira-preta coletadas em diferentes épocas, armazenadas a frio e tratadas com aib. **Ciência Rural**, Santa Maria, v. 42, n. 2, p. 232–237, 2012.
- CURI, P. N., PIO, R., MOURA, P. H. A., TADEU, M. H., NOGUEIRA, P. V., PASQUAL, M. Produção de amora-preta e amora-vermelha em Lavras - MG. **Ciência Rural**, [s.l.], v. 45, n. 8, p.1368–1374, 28 abr. 2015.
- BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: J. Wiley, 1988. 365 p.
- BENASSI, A.C. et al. Caracterização biométrica de frutos de coqueiro, *Vocos nucifera* L. variedade anã-verde, em diferentes estádios de desenvolvimento. **Revista Brasileira de Fruticultura**, [s.l.], v. 29, p. 302-307, 2007.
- BROWN, H.; PRESCOTT, R. **Applied mixed models in medicine**. 2nd. ed. Chichester: J. Wiley, 2006. 455 p.
- CÂMARA, F.M.M. et al. Fenologia do coqueiro anão verde em região semiárida. **Agropecuária Científica no Semiárido**, Campina Grande, v.14, n.4, p.291-296, 2019.
- CURI, P. N. et al. Produção de amora-preta e amora-vermelha em lavras-mg. **Ciência Rural**, Santa Maria, v. 45, n. 8, p. 1368–1374, 2015.
- DEMÉTRIO, C.G.B.; ZOCCHI, S.S. **Modelos de regressão**. Piracicaba: ESALQ, 2006.

- DIEL, M. I. et al. Nonlinear regression for description of strawberry (*Fragaria x ananassa*) production. **The Journal of Horticultural Science and Biotechnology**, v. 94, n. 2, p. 259-273, 2019.
- DODGE, Y. **The concise encyclopedia of statistics**. New York: Springer-Verlag, 2008. 616 p.
- DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3rd ed. New York: J. Wiley, 1998. 706 p.
- FACHINELLO, J. C. et al. Situação e perspectivas da fruticultura de clima temperado no Brasil. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 33, p. 109–120, 2011.
- FERNANDES, T. J. et al. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum**. Technology, v. 37, n. 4, p. 397-402, 2015.
- FERNANDES, T. J.; PEREIRA, A. A.; MUNIZ, J. A. Double sigmoidal models describing the growth of coffee berries. **Ciência Rural**, Santa Maria, v. 47, n. 8, 2017.
- FERNANDES, F.A. et al. Panorama da fase inicial do crescimento dos números de casos e óbitos causados pela COVID-19 no Brasil. **Research, Society and Development**, [s.l.], v. 9, n. 10, p. e1539108560-e1539108560, 2020.
- FERNANDES, J.G. et al. Description of the peach fruit growth curve by diphasic sigmoidal nonlinear models. **Revista Brasileira de Fruticultura**, [s.l.], v. 44, 2022.
- FRÜHAUF, A. C. et al. Nonlinear models in the study of the cedar diametric growth in a seasonally dry tropical forest. **Revista Brasileira de Ciências Agrárias**, [s.l.], v. 15, n. 4, p. 1-8, 2020.
- GALLANT, A. R. **Nonlinear statistical models**. New York: John Wiley & Sons, 2009.
- GURJÃO, K.C.O. et al. Desenvolvimento de frutos e sementes de tamara. **Revista Brasileira de Fruticultura**, [s.l.], v. 28, p. 351-354, 2006.
- HARVILLE, D. A. Maximum likelihood approaches to variance component estimation and to related problems. **Journal of the American Statistical Association**, New York, v. 72, n. 358, p. 320-338, 1977.
- HENDERSON, C. H. Best linear estimation and prediction under a selection model. **Biometrics**, Washington, v. 31, n. 2, p. 423-447, 1975.
- HOFFMANN, R. **Análise de regressão: uma introdução à econometria**. [S.l.]: edusp, 2016. 393 p.
- Lehmann, E. L. **Testing Statistical Hypotheses**. New York: Wiley, 1986.
- LIMA, K. P. et al. Uso de modelos não lineares na descrição do acúmulo de boro em diferentes partes do feijoeiro cultivar jalo. **Revista Brasileira de Biometria**, Lavras, v. 35, n. 4, p.834–861, 2017.

- LIMA, K. P. et al. Ajuste de modelos não lineares para descrição do fruto cajá-manga. **Sig-mae**, v. 8, n. 2, p. 221-226, 2019.
- LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. **Biometrics**, Arlington, v. 38, n. 4, p. 963-974, 1982.
- LINDSTROM, M. J.; BATES, D. M. Nonlinear mixed effects models for repeated measures data. **Biometrics**, p. 673-687, 1990.
- MISCHAN, M. M.; PINHO, S. Z. **Modelos não lineares: funções assintóticas de crescimento**. São Paulo: Cultura Acadêmica, 2014.
- PALHARINI, M. C. A. et al. Efeito da temperatura de armazenamento na conservação pós-colheita de amora-preta. **Pesquisa Agropecuária Tropical**, Goiânia, v. 45, n. 4, p. 413-419, 2015.
- PINHEIRO, C. J.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer-Verlag, 2004. 528 p.
- RENCHER, A. C.; SCHAALJE G. B. **Linear models in statistics**. 2nd ed. New Jersey: J. Wiley, 2008. 672 p.
- SARI, B. G. et al. Nonlinear growth models: An alternative to ANOVA in tomato trials evaluation. **European Journal of Agronomy**, [s.l.], v. 104, p. 21-36, 2019.
- SARI, B. G. et al. Nonlinear modeling for analyzing data from multiple harvest crops. **Agro-nomy Journal**, [s.l.], v. 110, n. 6, p. 2331-2342, 2018.
- SEARLE, S. R.; CASELLA, G.; MCCULLOCH, Charles E. **Variance components**. New York: John Wiley & Sons, 2009.
- SILVA, E. P. et al. Caracterização física, química e fisiológica de gabirola (*Campomanesia pubescens*) durante o desenvolvimento. **Food Science and Technology**, [s.l.], v. 29, n. 4, p. 803-809, 2009.
- SILVA, E.S. et al. Desenvolvimento da parte aérea do coqueiro anão verde sob aplicação de fertilizantes. **Revista Brasileira de Agricultura Irrigada**, Fortaleza, v.11, n.4, p.1571-1577, 2017.
- SILVA, E. M. et al. Description of blackberry fruit growth by nonlinear regression models. **Revista Brasileira de Fruticultura**, v. 42, 2020.
- SILVA, E M. et al. Evaluation of the critical points of the most adequate nonlinear model in adjusting growth data of 'green dwarf' coconut fruits. **Revista Brasileira de Fruticultura**, v. 43, 2021.
- SCHWARZ, G. Estimating dimension of a model. **Annals of Statistics**, Hayward, v. 6, n. 2, p. 461-464, 1978.
- TEIXEIRA, Glaciane Lopes et al. Growth curves of campolina horses using nonlinear models. **Livestock Science**, [s.l.], v. 251, p. 104631, 2021.

WYZYKOWSKI, J. et al. Análise do diâmetro de copa do cafeeiro recepado utilizando um modelo não linear misto. **Revista Brasileira de Biometria**, São Paulo, v. 33, n. 3, p. 243-256, 2015.

SEGUNDA PARTE – ARTIGOS

**ARTIGO 1 - EVALUATION OF THE CRITICAL POINTS OF THE MOST
ADEQUATE NONLINEAR MODEL IN ADJUSTING GROWTH DATA OF 'GREEN
DWARF' COCONUT FRUITS**

Publicado na Revista Brasileira de Fruticultura

DOI: 10.1590/0100-29452021726

Normas do periódico (versão publicada)

Evaluation of the critical points of the most adequate nonlinear model in adjusting growth data of 'green dwarf' coconut fruits

Édipo Menezes da Silva¹, Ariana Campos Fruhauf², Edilson Marcelino Silva³,
Joel Augusto Muniz⁴, Tales Jesus Fernandes⁵, Victor Ferreira da Silva⁶

Abstract- 'Green Dwarf' coconut is a fruit of great economic interest, since all its components are used, in addition to water, its main component. It is a culture of humid tropics, widely produced in northeastern Brazil, being an important income source for the region. The phenology study of this type of fruit is extremely important, but there are few studies in literature. Regression models, especially nonlinear growth models, can be of great value to understand how fruit growth behaves. The scarcity of works of this nature may be linked to some difficulties in estimating parameters of nonlinear models, such as assigning initial values to the iterative process. Overcoming this difficulty, for regression analysis, linear or not, several steps need to be respected to ensure the validity of information. Much information can be extracted from nonlinear growth models, such as the asymptotic value, growth rate and critical points (maximum acceleration point, inflection point, maximum deceleration point and asymptotic deceleration point). The aim of this work was to describe the stages of nonlinear regression analysis and to estimate the critical points of 'Green Dwarf' coconut growth curves. After initial adjustments, the only unmet assumption was independence, adding a first order autoregressive term. Again, models were adjusted and all parameters were significant, with both models, Gompertz and Logistic, adjusting well to data, with slight advantage for the Logistic model with better adjustment quality criteria values, with maximum expected LED and LEDKP values of 21.4037 cm and 21.5478 cm, respectively. The x and y axis of critical points were estimated, with values that can help producers to make more objective decisions about the appropriate time to harvest coconut fruits, considering the most diverse uses of this type of fruit.

Index terms: Regression analysis, selection of models, goodness fit, sigmoid; fruit development.

Avaliação dos pontos críticos do modelo não linear mais adequado no ajuste aos dados de crescimento de frutos de coqueiro-anão-verde

Resumo- O coco-anão-verde é um fruto com grande interesse econômico, pois se aproveitam de todos os seus componentes, com o principal interesse na água. É uma cultura de trópicos úmidos, largamente produzido no Nordeste brasileiro, sendo uma importante fonte de renda para a região. O estudo da fenologia desse tipo de fruto pode ser de extrema importância, porém são poucos os trabalhos encontrados na literatura. Os modelos de regressão, com destaque para os não lineares de crescimento, podem ser de grande valia para entender como o crescimento do fruto ocorre. A escassez de trabalhos dessa natureza pode estar ligada a algumas dificuldades na estimação dos parâmetros dos modelos não lineares, como atribuir valores iniciais para o processo iterativo. Contornando essa dificuldade, para a análise de regressão, linear ou não, diversas etapas precisam ser respeitadas para garantir a validade de suas informações. É possível retirar-se diversas informações dos modelos não lineares de crescimento, como o valor assintótico, a taxa de crescimento e os pontos críticos (ponto de aceleração máxima, ponto de inflexão, ponto de desaceleração máxima e ponto de desaceleração assintótico). Com isso, o objetivo deste trabalho foi descrever as etapas da análise de regressão não linear e estimar os pontos críticos das curvas de crescimento de coco-anão-verde. Após os ajustes iniciais, o único pressuposto não atendido foi o de independência, sendo adicionado o termo autorregressivo de primeira ordem. Novamente foram ajustados os modelos e todos os parâmetros foram significativos, com ambos os modelos, Gompertz e Logístico, se ajustando bem aos dados, tendo uma ligeira vantagem para o modelo Logístico com melhores valores dos critérios de qualidade de ajuste. Com o valor máximo esperado para DEL e DELMP sendo 21,4037 cm e 21,5478 cm, respectivamente. As abscissas e ordenadas dos pontos críticos foram estimados, com valores que podem ajudar ao produtor a tomar decisões mais objetivas a respeito do momento adequado para colheita dos frutos de coco-anão-verde, considerando os mais diversos usos desse tipo de cultura.

Termos para indexação: Análise de regressão; Desenvolvimento de frutos; Modelo Gompertz; Modelo Logístico; Qualidade de ajuste.

Corresponding author:
ediposvm01@gmail.com

Received: October 14, 2020
Accepted: December 21, 2020

Copyright: All the contents of this journal, except where otherwise noted, is licensed under a Creative Commons Attribution License.



¹PhD student in Agricultural Statistics and Experimentation, Federal University of Lavras, Lavras-MG, Brazil. Email: ediposvm01@gmail.com (ORCID 0000-0002-1613-6522)

²PhD student in Agricultural Statistics and Experimentation, Federal University of Lavras, Lavras-MG, Brazil. Email: arianafruhauf@gmail.com (ORCID 0000-0001-8759-7758)

³PhD in Statistics and Agronomic Experimentation, Professor at the Federal University of Lavras, Lavras-MG, Brazil. Email: edilsonmg3@hotmail.com (ORCID: 0000-0002-2800-3495)

⁴PhD in Agronomy / Statistics and Agronomic Experimentation, Professor at the Federal University of Lavras, Lavras-MG, Brazil. Email: joamuniz@des.ufla.br (ORCID 0000-0002-1069-4136)

⁵PhD in Statistics and Agronomic Experimentation, Professor at the Federal University of Lavras, Lavras-MG, Brazil. Email: tales.jfernandes@ufla.br (ORCID: 0000-0002-1457-9653)

⁶PhD student in Agricultural Statistics and Experimentation, Federal University of Lavras, Lavras-MG, Brazil. Email: victor.est.ufop@gmail.com (ORCID: ORCID 0000-0003-1452-6831)

Introduction

Coconut tree (*Cocos nucifera* L.) is a plant that belongs to the Arecaceae family, reaching up to ten meters in height. It is a tree of great economic interest, since all its components are used: fruit, leaves, pulp, water, bark and fibers. The main economic activity involving this species is the production of green coconut, mainly for water production (SILVA et al., 2017). It is a culture of humid tropics; however, production is largely concentrated in semi-arid regions of northeastern Brazil, with high temperatures and little rain for most of the year (CÂMARA et al., 2019). Phenology in coconut farming is important in fruit development and production, allowing testing the viability of fruits over the years, in different seasons, types of soil and climate. However, despite the relevance of results, studies on the phenology of coconut crops are scarce and generic in northeastern Brazil (ARAÚJO et al., 2011; CÂMARA et al., 2019). Thus, growth models can be of great value to understand the phenology of this culture, since studying growth has become of great importance to help producers and show the best way to deal with growth phases under study. However, in practice, several authors, from the most varied areas, who work with these types of data and adjustment of regression models, do not take into account the stages of statistical modeling, which are considered essential to ensure validity of inferences. Regression analysis originated in the 19th century, with Francis Galton, which is a technique that allows us inferring the functional relationship between a dependent variable, generally designated as “Y” and one or more independent random variables, or “X” variables. When animal or plant growth data are analyzed, the functional relationship is usually described by a curve in the form of a sigmoid curve and is more adjusted by non-linear models, more specifically, nonlinear growth models (FERNANDES et al. 2019, RIBEIRO et al., 2018b), which have advantages over linear models, as they are more parsimonious, that is, they satisfactorily describe the relationship between variables with smaller number of parameters and with biological interpretation (MISCHAN; PINHO, 2014).

For linear or nonlinear regression models to be correctly used, one must consider the modeling steps starting with the choice of models to fit data, residual analysis, significance of parameters and adjustment adequacy. The first step of the procedure is the choice of candidate models, that is, through descriptive analysis, the researcher will know which models may be more suitable to describe the relationship under study (SARI et al., 2019a). After the initial choice, it is necessary to adjust the models by estimating parameters based on experimental data; however, in nonlinear models, there is an additional difficulty, as it is not possible to obtain such estimates in an algebraic way, requiring approximation

through the use of numerical methods, which need the attribution of good initial values to initiate the iterative process (MISCHAN; PINHO, 2014; ARCHONTOULIS; MIGUEZ, 2015; FRÜHAUF, A.C. et al., 2020).

Since models are adjusted, it is necessary to verify the residual vector to guarantee the results of hypothesis tests applied in studies. There are several hypothesis tests in literature to verify the validation of residual vector assumptions which can also be verified by the graphic analysis (RIBEIRO et al., 2018b; SILVA et al., 2020). For models to be considered fit to describe the functional relationship in study, in addition to the analysis of residues, it is essential to verify the significance of the parameters of models. If any parameter is considered non-significant, another model must be adjusted (RENCHER; SCHAALJE, 2008).

With residual vector assumptions and significant parameters, one more inspection is important, the verification of the adequacy of models, which can be performed through the adjusted determination coefficient (R^2_{aj}) and the parametric non-linearity measure (BATES; WATS, 1980). Models will be considered adequate when R^2_{aj} values are close to 1 and parametric non-linearity measures are below 1 (SARI et al., 2019b).

With all these steps completed, the various tested models can be reduced to a single model, said to be the “most adequate” (ARCHONTOULIS; MIGUEZ, 2015). For the selection of the “most adequate” model, some criteria are used to assess the quality of fit, which are used in several growth studies (RIBEIRO et al., 2018a, FERNANDES et al., 2019; SILVA et al., 2020), such as the Akaike information criterion (AIC) (AKAIKE, 1974), the Bayesian information criterion (BIC) (SCHWARZ, 1978), adjusted determination coefficient (R^2_{aj}), residual standard deviation (RSD) and the intrinsic (C) and parametric (C^p) nonlinearity measures.

Several authors have used these steps to study fruit growth (MUIANGA et al., 2016; FERNANDES et al., 2017; MUNIZ, J. A. et al., 2017; DIEI et al., 2019; JANE et al. 2019, RIBEIRO et al., 2018a, RIBEIRO et al., 2018b, SARI et al., 2019b, SILVA et al. 2020; JANE et al., 2020a).

Curves that represent growth models do not have extreme, maximum or minimum points, but some points are considered important from the physiological point of view, each having a specific meaning. The study of these points is performed through the derivatives of equations in relation to time (MISCHAN; PINHO, 2014).

Critical points have been satisfactorily studied by the following authors in the growth of different types of fruits (SARI et al., 2019a; SARI et al., 2019b; DIEI et al., 2020; SILVA et al., 2020). In this context, the aim of this work was to describe steps in adjusting Gompertz and Logistics models to growth data of the longitudinal diameter of ‘Green Dwarf’ coconut fruits, estimating the critical points of models.

Material and methods

Data were obtained from Benassi (2007) and the fieldwork was carried out in an 'Green Dwarf' coconut orchard (*Coco nucifera* L.) installed in the city of Bebedouro, state of São Paulo. The work was carried out in a seven-year-old coconut orchard with spacing of seven meters between rows and six meters between plants in a triangle shape. Eight different plants were evaluated, measuring three fruits, totaling 24 fruits, and the average was obtained to adjust models. From these fruits, measurements in relation to the longitudinal external diameter of harvested fruits (LED) and fruits kept on plants (LEDKP) were performed with the aid of caliper until 120 days of age, with measurement unit in centimeter (cm). The first measurement was performed 1 day after inflorescence opening (DAIO) and the last at 375 DAIO every 15 days, totaling 26 measurements over time.

Logistic (1) and Gompertz (2) nonlinear regression models were adjusted and the parameterization chosen was based on the work of Fernandes et al. (2015).

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

$$Y_i = \alpha e^{-e^{k(\beta - t_i)}} + \varepsilon_i \quad (2)$$

Where Y_i is the response variable or the diameter value taken at time t_i , with $i = 1, 2, \dots, 26$, α being the asymptotic value or maximum expected fruit diameter value, β is the abscissa of the inflection point, with k value being related to the fruit growth rate, and the higher the k value, the faster the fruit is expected to reach its adult stage (MISCHAN; PINHO, 2014) and ε_i is the random error, in which, as initial assumption, has normal distribution, with zero mean, independent and with constant variance, that is $\varepsilon_i \sim N(0, \sigma^2)$.

In nonlinear models, it is necessary to use an iterative method to find approximation of estimates. For this work, the iterative method used was that of Gauss-Newton because it was implemented in the "nls" function and the initial values were graphically obtained by the "manipulate" function (ALLAIRE, 2014), both packages of the R statistical software.

After initial adjustments, residual analysis was performed to verify whether assumptions were not violated. For this, the Shapiro-Wilk (SHAPIRO; WILK, 1965) tests for normality, Breusch-Pagan (BREUSCH; PAGAN, 1979) for homoscedasticity and Durbin-Watson (DURBIN; WATSON, 1951) for the independence of residues were used, and complementary to aforementioned tests, graphical analysis was used.

If data normality is not met, some transformation may be performed to data or a change in the choice of the model to be adjusted. If there is violation in the

homoscedasticity assumption, a viable solution would be to model the inverse of the variance. Finally, if residues are not independent, it is necessary to model dependence through the inclusion of an order- p autoregressive term (AR (p)).

The significance of parameters of models can be verified through the Student's t test, in which values lower than pre-established nominal level imply significant parameter; otherwise, this parameter is not significant. The confidence interval (CI) of parameters can also be used, which can be calculated with the following expression:

$$IC(\theta_i) = \hat{\theta}_i \pm t_{(v; \frac{\alpha}{2})} ep(\hat{\theta}_i)$$

Where θ_i is the i -th estimated parameter, $t_{(v; \frac{\alpha}{2})}$ the quantile of Student's t distribution, α the significance level adopted and v the degrees of freedom, $v = n - p$, n the size of the response vector, p the number of parameters of the model and ep is the standard error $ep(\theta_i)$ of the estimate, being $ep(\theta_i) = \sqrt{V(\theta)}$ and $V(\theta)$ an estimate of the variance of parameter θ_i obtained from the diagonal of the variances and covariance matrix. If this interval does not contain zero, the parameter can be considered significant.

When it comes to regression analysis, the ideal is to have several candidate models to test which is the most suitable for each situation, in the case of this work, models (1) and (2) were those indicated to be tested due to the data dispersion graph and study objectives. After verifying which models are candidates, the one that best fits these data is selected to then draw conclusions about its results. To select the model that best fits data, the following criteria were used:

- Corrected Akaike information criterion,

$$AIC_c = AIC + \frac{2p(p+1)}{n-p-1} \quad \text{with } n \text{ being the number of observation and the number of parameters;}$$

- Bayesian information criterion

$$BIC = -2 \ln(L(\hat{\theta})) + p \ln(n), \quad L(\hat{\theta}) \quad \text{being the likelihood function and } \ln \text{ is a natural logarithm operator.}$$

- Residual standard deviation $RSD = \sqrt{QME}$, QME is the residual mean square;

- Non-linearity measures (intrinsic and parametric);

- Adjusted determination coefficient

$$R_{adj}^2 = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - p} \right] \quad \text{and } R^2 \text{ the determination coefficient.}$$

Lower AIC_c , BIC, RSD values, non-linearity, intrinsic and parametric measures and higher R_{adj}^2 values characterize the most appropriate model to describe data under study. Together with the interpretations of parameters, the first four partial derivatives in relation to variable time of Logistic and Gompertz models were used, the first derivative being the growth rate. By

equalizing the second derivative to zero, the inflection point (ip) of the curve is obtained. The third derivative equal to zero allows finding the maximum acceleration point (map) and the maximum deceleration point (mdp)

and the fourth derivative brings the information of the asymptotic deceleration point (adp) (MISCHAN; PINHO, 2014). To find the coordinates of critical points using the parameterization of this work, expressions shown in Table 1 can be used:

Table 1. Expressions to find abscissa and ordinates of maximum acceleration points (map), inflection point (ip), maximum deceleration point (mdp) and asymptotic deceleration point (adp) of Logistic and Gompertz models using parameterization of this work.

Modelo	Ponto	pam	pi	pdm	pda
Logístico	Abscissa	$\frac{k\beta - 1,319}{k}$	$\frac{\alpha}{2}$	$\frac{k\beta + 1,319}{k}$	$\frac{k\beta + 2,2924}{k}$
	Ordenada	$\frac{\alpha}{3 + \sqrt{3}}$	β	$\frac{\alpha}{3 - \sqrt{3}}$	$\frac{\alpha}{1,101}$
Gompertz	Abscissa	$\frac{k\beta - 0,962}{k}$	$\frac{\alpha}{e}$	$\frac{k\beta + 0,962}{k}$	$\frac{k\beta + 1,7975}{k}$
	Ordenada	$\alpha e^{-\frac{2+\sqrt{5}}{2}}$	β	$\alpha e^{-\frac{2-\sqrt{5}}{2}}$	$\frac{\alpha}{e^{0,1657}}$

Analyses were performed using the R statistical software (R CORE TEAM, 2020), with the aid of the “nlme” (PINHEIRO et al., 2020), “car” (JOHN; SANFORD, 2019), “lmtest” (ACHIM; HOTHORN, 2002) and “manipulate” packages (ALLAIRE, 2014). For comparison purposes, 1% nominal level was adopted.

Results and discussion

As initial part of any regression analysis, models are adjusted considering all assumptions met, after these first adjustments, the residual analysis is performed. With results presented in Table 2, it was observed that the only assumption violated was that of independence, which was expected for the LED variable, since data were taken from the same fruit and as mentioned in the works by Ribeiro et al. (2018a) and Cassiano and Safadi (2015), observations performed on the same individual are generally autocorrelated. In a complementary way, graphic analysis corroborated the same results of hypothesis tests in Table 2, as can be seen in Figures 1, 2, 3 and 4. With this observation, models were again adjusted incorporating the first-order autoregressive term (AR (1)).

Table 2. P-value of Shapiro-Wilk (SW), Breusch-Pagan (BP) and Durbin-Watson (DW) tests for Logistic and Gompertz models adjusted to ‘Green Dwarf’ coconut growth data in relation to the longitudinal external diameter of harvested fruits (LED) and longitudinal external diameter of fruits kept on plants (LEDKP).

Feature	Model	SW	BP	DW
LED	Logistic	0,6102	0,0245	0,0041
	Gompertz	0,2598	0,0120	0,0000
LEDKP	Logistic	0,5124	0,0150	0,0000
	Gompertz	0,1496	0,0594	0,0000

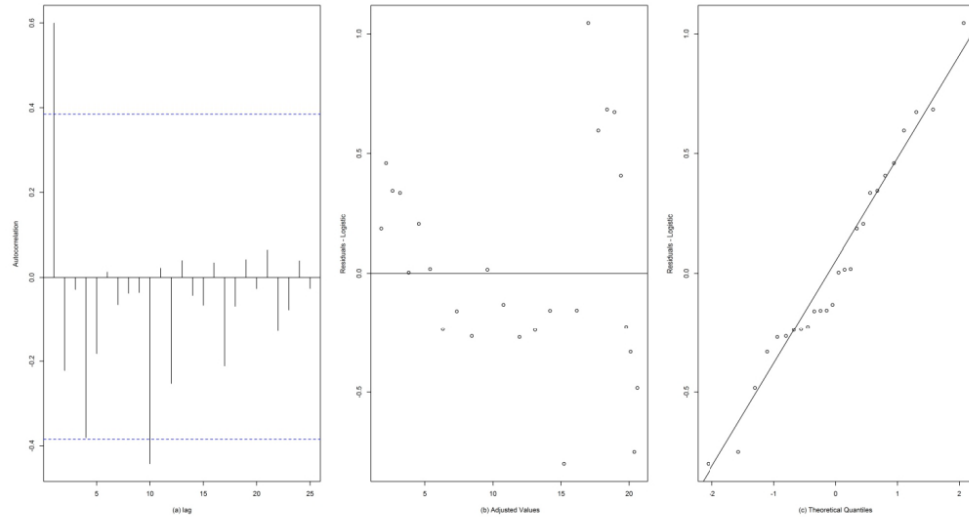


Figure 1. Residual analysis for LED (mm), where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residues and (c) residual values in relation to the theoretical quantiles for the Logistic model.

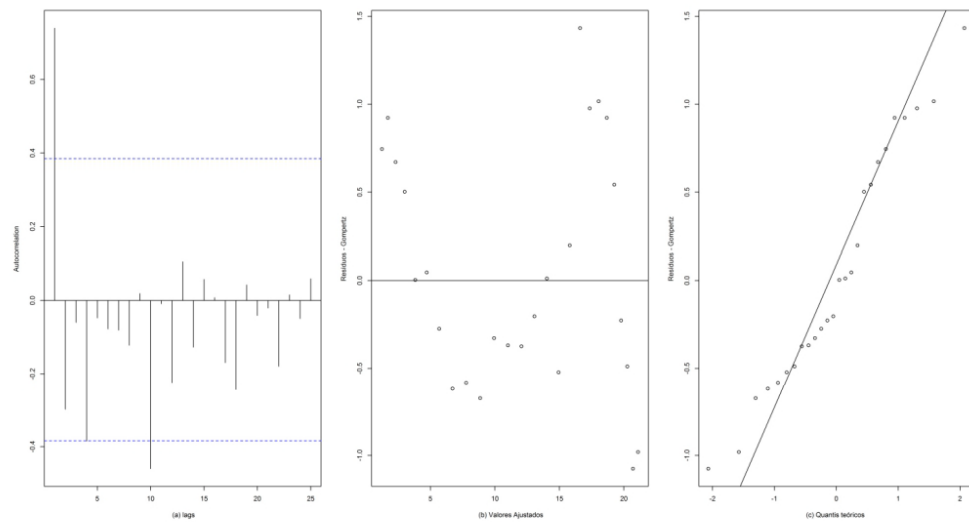


Figure 2. Residual analysis for LED (mm), where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residues and (c) residual values in relation to the theoretical quantiles for the Gompertz model.

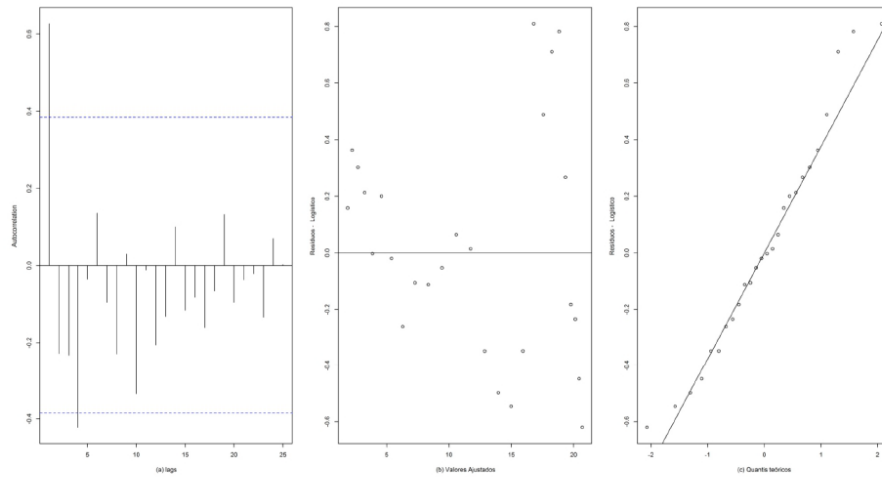


Figure 3. Residual analysis for LEDKP (mm), where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residues and (c) residual values in relation to the theoretical quantiles for the Logistic model.

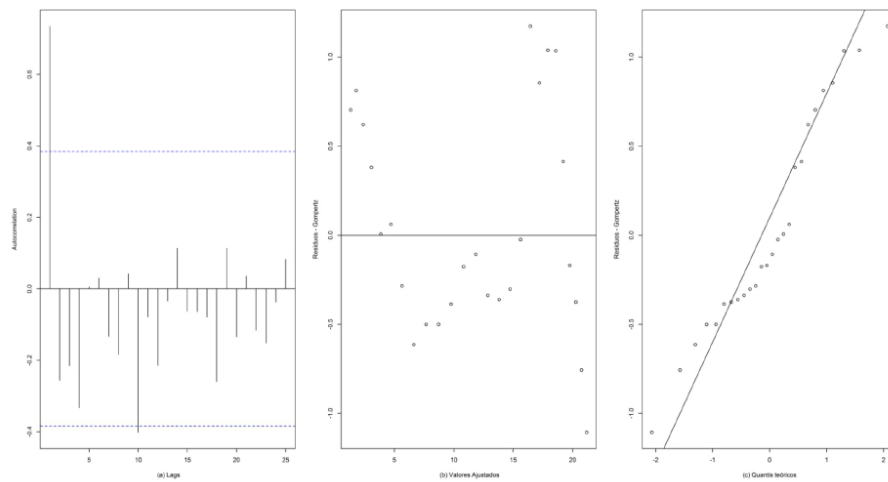


Figure 4. Residual analysis for LEDKP (mm), where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residues and (c) residual values in relation to the theoretical quantiles for the Gompertz model.

Before concluding that models are adequate and selecting the one that best fits data, the significance of parameters was tested via Student's t test and 95% confidence intervals. For both adjusted models, all parameters were significant and their respective confidence intervals did not contain zero, that is, after verifying adjustments, significance and CI of parameters, all models were adequate to describe the relationship between LED and LEDKP in relation to time (Table 4).

The results of criteria used to assess the adjustment quality of models, with R^2_{aj} values greater than 0.99 and parametric non-linearity less than 1, indicate good adjustments of models in relation to data (Table 3) for study with tomato and study with meat producing mammals, respectively. Sari et al. (2019a) and Fernandes et al. (2019) obtained similar values.

Table 3. Criteria for assessing adjustment quality: Akaike information criterion (AIC_c), Bayesian information criterion (BIC), residual standard deviation (RSD), intrinsic non-linearity (C'), parametric non-linearity (C^{θ}) and adjusted determination coefficient (R^2_{aj}), for comparing the Logistical and Gompertz models to 'Green Dwarf' coconut growth data in relation to the longitudinal external diameter of harvested fruits (LED) and longitudinal external diameter of fruits kept on plants (LEDKP).

Feature	Model	AIC_c	BIC	RSD	(C^{θ})	(C')	(R^2_{aj})
LED	Logistic	29,7492	34,9487	0,4630	0,2487	0,0414	0,9954
	Gompertz	38,7762	43,9758	0,7411	0,7625	0,0584	0,9804
LEDKP	Logistic	21,1359	26,3355	0,4199	0,2375	0,0363	0,9962
	Gompertz	31,2845	36,4841	1,4282	0,7604	0,051	0,9520

Table 4. Estimates of parameters of the Logistic model and their respective confidence intervals, lower limit (IC inf) and upper limit (IC sup), abscissa and ordinates of maximum acceleration points (map), inflection point (ip), maximum deceleration point (mdp) and asymptotic deceleration point (adp) of the Logistic model applied to 'Green Dwarf' coconut growth data in relation to the longitudinal external diameter of harvested fruits (LED) and longitudinal external diameter of fruits kept on plants (LEDKP).

Parameters	LED			LEDKP		
	IC inf	Estimates	IC sup	IC inf	Estimates	IC sup
Parameter α	20,2881	21,4037**	22,5175	20,4373	21,5478**	22,6601
Parameter β	152,2472	163,9767**	175,7062	155,3203	168,1374**	178,9545
Parameter k	0,0128	0,01464**	0,0166	0,0125	0,0143**	0,0160
phi	-----	0,6154	-----	-----	0,6787	-----
Critical points	LED			LEDKP		
map (x)		73,8811			75,8996	
map (y)		4,5231			4,5535	
ip (x)		163,9767			168,1374	
ip (y)		10,7018			10,7739	
mdp (x)		254,0726			260,3752	
mdp (y)		16,8805			16,9942	
adp (x)		320,5614			328,4451	
adp (y)		19,4402			19,5711	

** Significant at 1% probability.

Although both adjusted models were shown to be adequate to describe the relationship among variables, to make inferences about their estimates, the Logistic model was chosen as the one that best adhered to data, as it can be observed that this model presented higher R^2_{aj} values and lower AIC_c , BIC and RSD values (Table 3). Bates and Wats (1988) used intrinsic and parametric nonlinearity measures to evaluate the nonlinearity of the nonlinear model, and the lower the C^1 and C^0 values, the closer to linear is the model, and as C^1 and C^0 values were lower for the Logistic model in both LED and LEDKP, it seems to be the most appropriate model, corroborating AIC_c , BIC, RSD and R^2_{aj} values.

According to Diel et al. (2019), several studies use only adjustment quality evaluators as the only way to choose the best model, but another aspect must be taken into account, the estimated asymptote value (α), which cannot present sub or overestimated values, as this parameter has the most important interpretation. In this work, the estimated asymptote values of the Logistic model, for both LED and LEDKP, had small deviation from observed values, with LED values of 21.4037 cm and LEDKP values of 21.5478 cm (Figures 1 and 2, respectively). It should also be pointed out that it is necessary to use nonlinear regression models with some caution, because if the necessary steps of a good study are not respected, mistaken results can be obtained and, checking only criteria to assess the adjustment quality can be dangerous in the inferences of models, and the ideal situation is to check the bias of the estimates of parameters, mainly the horizontal asymptote, that is, parameter α (Table 4).

With another extremely important aspect and corroborating adjustment quality evaluators, the Logistic model can be considered the most suitable to describe the relationship proposed in this work, and this model was also chosen as the most adequate in the works of Prado, Savian and Muniz (2013), Prado et al. (2020), Muianga et al. (2016), Jane et al. (2020a) and Jane et al. (2020b).

Table 4 shows the estimates and confidence intervals for parameters of the external longitudinal diameter of harvested fruits (LED) and external longitudinal diameter of fruits kept on plants (LEDKP) and the critical points of the four derivatives of the Logistic model. Asymptotic values were 21.4037 cm and 21.5478 cm, respectively, for LED and LEDKP, which values were found at 375 DAIO, higher than those found by Silva et al. (2009), who found for 'Green Dwarf' coconut grown in the conventional system of 17.23 cm and in the organic system of 16.75 cm, which difference can be explained by the fact that in the work of Silva et al. (2009), the authors harvested fruits at 240 DAIO, a phase occurring between the inflection point and the maximum deceleration point found in this work.

According to Silva et al. (2009), Ferreira Neto et al. (2007) and Ribeiro, Costa and Aragão (2017), coconut fruits are harvested at 240 DAIO, as the fresh coconut water market makes this requirement due to the fact that the water taste is better in this period. In the work by Silva et al. (2009), the diameter value was 16.57 cm in the organic system and 17.23 cm in the conventional system and with estimates of this work at 240 DAIO, phase close to the maximum deceleration point (mdp), LED value was 16.11 cm and LEDKP value was 15.86 cm.

For the agroindustry, Ribeiro, Costa and Aragão (2017) point out that the fruit must be harvested between 8 and 9 months or between 240 DAIO and 270 DAIO, phase comprised between the maximum deceleration point and the asymptotic deceleration point found in this study. For exactly 270 DAIO, estimated LED and LEDKP values were 17.66 cm and 17.47 cm, respectively. The final phase of fruit harvesting takes place when they are dry, and according to Ribeiro, Costa and Aragão (2017), this phase occurs at 11 months or 330 DAIO. Mischan and Pinho (2014) point out that the asymptotic deceleration value (ADP) represented by 90% of the horizontal asymptote value (α), which corresponds to LED of approximately 329 DAIO and LEDKP of approximately 327 DAIO, being close to 330 DAIO needed for coconut fruits to be dry or 19.43 cm for LED and 19.57 cm for LEDKP.

The observed data, the adjusted curves and the curve that represents the growth rate over time (Figures 5 and 6), the highest point of the growth rate curve being the inflection point and as mentioned by Mischan and Pinho (2014), this point is exactly half of the upper horizontal asymptote of the Logistic model, which for LED occurred at approximately 163 DAIO or 10.70 cm (Figure 5) and for LEDKP, 167 DAIO or 10.77 cm (Figure 6). In addition to information about the curve that represents the growth rate, the adjusted curve adhered well to observed data, corroborating the initial hypothesis that data have sigmoidal shape. Other important points for the production of 'Green Dwarf' coconut in the growth rate curve are map (in yellow color), mdp (in blue color) and adp (in pink color) (Figure 5).

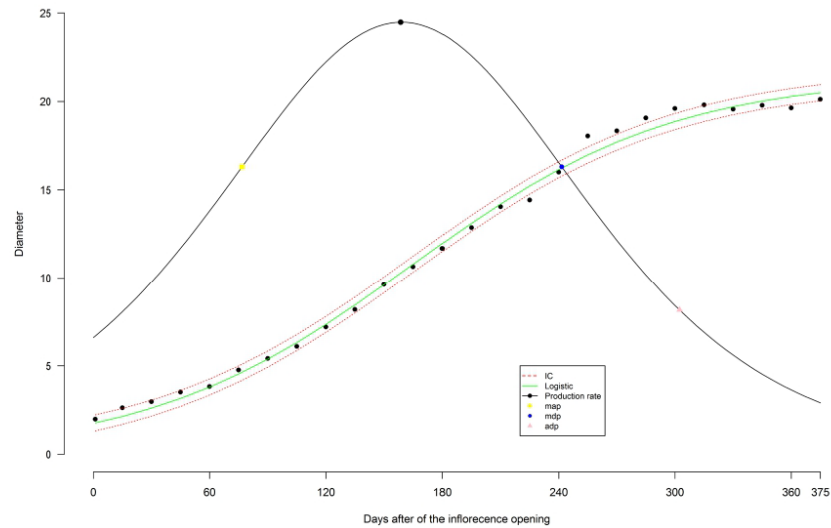


Figure 5. Growth rate with the inflection point (ip), maximum acceleration point (map), maximum deceleration point (mdp), asymptotic deceleration point (adp), and confidence interval (CI) of the estimated Logistic model curve to LED data of 'Green Dwarf' coconut fruits.

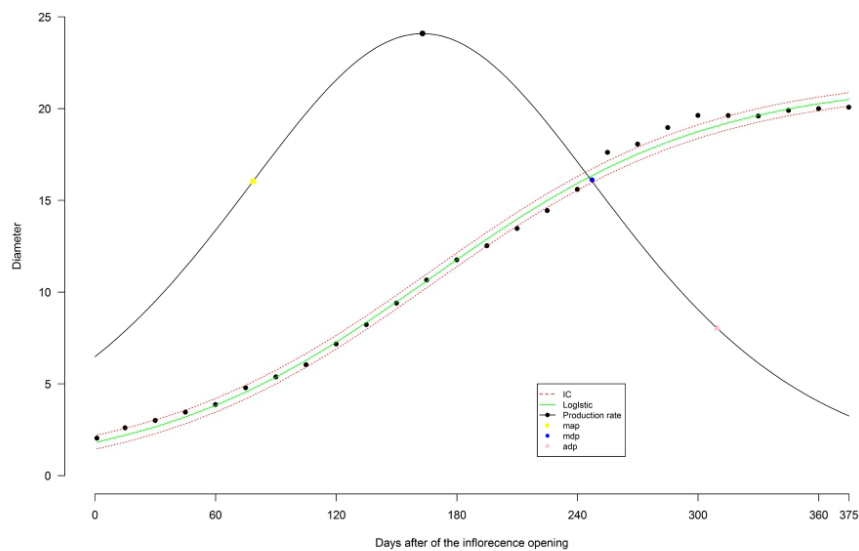


Figure 6. Growth rate with the inflection point (ip), maximum acceleration point (map), maximum deceleration point (mdp), asymptotic deceleration point (adp), and confidence interval (CI) of the estimated Logistic model curve to LEDKP data of 'Green Dwarf' coconut fruits.

Conclusion

The Logistic and Gompertz models were adequate to describe the relationship between LED and LEDKP; however, due to criteria used to assess the adjustment quality and the analysis of the horizontal asymptote, the Logistic model was more adequate, with asymptotic values for both variables of approximately 21 cm.

For studies with 'Green Dwarf' coconut, the use of critical points of models is of great help for the researcher to find the harvest point of fruits with the best water flavor, for the use of agro-industry and the use of dry coconut in the food industry. The asymptotic values were 21.4037 cm for LED and 21.5478 cm for LEDKP.

References

- ACHIM, Z.; TORSTEN, H. Diagnostic checking in regression relationships. **R News**, v.2, n.3, p.7-10, 2002. Available from: <https://CRAN.R-project.org/doc/Rnews/>.
- ALLAIRE, J.J. **Manipulate**: interactive plots for R studio. R package version 1.0.1. 2014. Available from: <https://CRAN.R-project.org/package=manipulate>.
- ARAÚJO, S.D.A região semiárida do nordeste do Brasil: questões ambientais e possibilidades de uso sustentável dos recursos. **Rios Eletrônica-Revista Científica da FASETE**, Paulo Afonso, v.5, n.5, p.88-98, 2011.
- ARCHONTOULIS, S.V.; MIGUEZ, F.E. Nonlinear regression models and applications in agricultural research. **Agronomy Journal**, Madison, v.107, n.2, p.786-798, 2015.
- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Boston, v.19, n.6, p.716-723, 1974
- BATES, D.M.; WATTS, D.G. Relative curvature measures of nonlinearity. **Journal of the Royal Statistical Society: Series B. Methodological**, Oxford, v.42, n.1, p.1-16, 1980.
- BENASSI, A.C.; RUGGIERO, C.; MARTINS, A.B.G.; SILVA, J.A.A. Caracterização biométrica de frutos de coqueiro, *Vocos nucifera* L. variedade anã-verde, em diferentes estádios de desenvolvimento. **Revista Brasileira de Fruticultura**, Jaboticabal, v.29, n.2, p.302-307, 2007.
- BREUSCH, T.S.; PAGAN, A.R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica**, New York, v.47, n.5, p.1287-1294, 1979.
- CÂMARA, F.M.M.; ALVES, A.A.; CELEDÔNIO, W.F.; OLIVEIRA, L.M.; PEREIRA, G.A.; MENDONÇA, V. Fenologia do coqueiro anão verde em região semiárida. **Agropecuária Científica no Semiárido**, Campina Grande, v.14, n.4, p.291-296, 2019.
- CASSIANO, F.R.; SÁFADI, T. Modelos de crescimento animal para tempos irregulares. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v.50, p.1114-1119, 2015.
- DIEL, M.I.; SARI, B.G.; KRYSCZUN, D.K.; OLIVORO, T.; PINEHIRO, M.V.M.; MEIRA, D.; SCHMIDT, D.; LUCIO, A.D. Nonlinear regression for description of strawberry (*Fragaria x ananassa*) production. **The Journal of Horticultural Science and Biotechnology**, Warnick, v.94, n.2, p.259-273, 2019.
- DURBIN, J.; WATSON, G.S. Testing for serial correlation in last squares regression. **Biometrika**, London, v.38, n.1/2, p.159-177, 1951.
- 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**, Maringá, v.37, n.4, p.397-402, 2015.
- FERNANDES, T.J.; PEREIRA, A.A.; MUNIZ, J.A. Double sigmoidal models describing the growth of coffee berries. **Ciência Rural**, Santa Maria, v.47, n.8, 2017.
- 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, p.1-8, 2019.
- FERREIRA NETO, M.; GHEYI, H.R.; FERNANDES, P.D.; HOLANDA, J.S. de; BLANCO, F.F. Emissão foliar, relações iônicas e produção do coqueiro irrigado com água salina. **Ciência Rural**, Santa Maria, v.37, p.1675-1681, 2007.
- JANE, S.A.; FERNANDES, F.A.; SILVA, E.M.; MUNIZ, J.A.; FERNANDES, T.J. Comparison of the polynomial and logistic models on description of pepper growth. **Revista Brasileira de Ciências Agrárias**, Fortaleza, v.14, e.7180, 2019.
- JANE, S.A.; FERNANDES, F.A.; MUNIZ, J.A.; FERNANDES, T.J. Nonlinear models on description of sugarcane height and diameter variety RB92579. **Revista Ciência Agronômica**, Fortaleza, v.51, n.4, e.20196660, 2020a.

- JANE, S. A. et al. Adjusting the growth curve of sugarcane varieties using nonlinear models. *Ciência Rural*, Santa Maria, v. 50, n. 3, 2020b.
- JOHN, F.; SANDFOR W. **An {R} companion to applied regression**. 3rd ed. Thousand Oaks: Sage, 2019 Available from: <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- FRÜHAUF, Ariana Campos et al. Nonlinear models in the study of the cedar diametric growth in a seasonally dry tropical forest. *Revista Brasileira de Ciências Agrárias*, Fortaleza, v. 15, n. 4, p. 1-8, 2020
- MISCHAN, M.M.; PINHO, S.D. **Modelos não lineares: funções assintóticas de crescimento**. São Paulo: Cultura Acadêmica, 2014.
- MUIANGA, C.A.; MUNIZ, J.A.; NASCIMENTO, M.S.; FERNANDES, T.J.; SAVIAN, T.V. Por descrição da curva de crescimento de frutos do cajueiro modelos não lineares. *Revista Brasileira de Fruticultura*, Jaboticabal, v.38, n.1, p.22-32, 2016.
- MUNIZ, J. A., et al. Nonlinear models for description of cacao fruit growth with assumption violations. *Revista Caatinga*, v. 30, n. 1, p. 250-257, 2017.
- PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D. **Linear and nonlinear mixed effects models**. 2020. Available from: <https://CRAN.R-project.org/package=nlme>.
- PRADO, T.K.L.; SAVIAN, T.V.; MUNIZ, J.A. Ajuste dos modelos Gompertz e Logístico aos dados de crescimento de frutos de coqueiro anão verde. *Ciência Rural*, Santa Maria, v.43, n.5, p.803-809, 2013.
- PRADO, T.K.L.; SAVIAN, T.V.; FERNANDES, T.J.; MUNIZ, J.A. Study on the growth curve of the internal cavity of 'Dwarf green' coconut fruits. *Revista Ciência Agronômica*, Fortaleza, v.51, n.3, p.e20154591, 2020.
- R CORE TEAM. **R: A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, Vienna, 2020. Disponível em: <https://www.R-project.org/>; RENCHER, A.C.; SCHAALJE, G.B. **Linear models in statistics**. New York: John Wiley & Sons, 2008.
- RENCHE, A. C.; SHAALJE, G. B. *Linear models in statistics*. John Wiley Sons, 2008. RIBEIRO, T.D.; SAVIAN, T.V.; FERNANDES, T.J.; MUNIZ, J.A. The use of the nonlinear models in the growth of pears of 'Shinseiki' cultivar. *Ciência Rural*, Santa Maria, v.456, 48, n.1, 2018a.
- RIBEIRO, T.D.; MATTOS, R.V.P.; MORAIS, A.R.; MUNIZ, J.A. Description of the growth of pequi fruits by nonlinear models. *Revista Brasileira de Fruticultura*, Jaboticabal, v.40, n.4, 2018b.
- RIBEIRO, F.E.; COSTA, E.F.N.; ARAGÃO, W. **Árvore do conhecimento**. Disponível em: <http://www.agencia.cnptia.embrapa.br/gestor/coco/arvore/CONT000gl5m652b02wx5ok0xkgyq5dmanf8w.html>. Acesso em: 18 jun. 2020.
- SARI, B.G.; LÚCIO, A.D.C.; SAMANTA, C.S.; SAVIAN, T.V. Describing tomato plant production using growth models. *Scientia Horticulturae*, Amsterdam, v.246, p.146-154, 2019a.
- SARI, B.G.; LÚCIO, A.D.C.; SAMANTA, C.S. OLIVOTO, T., DIEI, M.I.; KRYSZCZUN, D.K. Nonlinear growth models: An alternative to ANOVA in tomato trials evaluation. *European Journal of Agronomy*, Amsterdam, v.104, p.21-36, 2019b.
- SCHWARZ, G. Estimating the dimension of a model. *The Annals of Statistics*, Durham, v.6, n.2, p.461-464, 1978.
- SHAPIRO, S.S.; WILK, M.B. An analysis of variance test for normality (complete samples). *Biometrika*, Cambridge, v.52, n.3-4, p.591-611, 1965.
- SILVA, E.S.; LIMA, F.N.R.; PINHEIRO NETO, L.G.; MENEZES, A.S.; MOREIRA, F.J.C.; LOPES, F.G.N. Desenvolvimento da parte aérea do coqueiro anão verde sob aplicação de fertilizantes. *Revista Brasileira de Agricultura Irrigada*, Fortaleza, v.11, n.4, p.1571-1577, 2017.
- SILVA, E.M.; TADEU, M.H.; SILVA, V.F.; PIO, R.; FERNANDES, T.J.; MUNIZ, J.A. Description of blackberry fruit growth by nonlinear regression models. *Revista Brasileira de Fruticultura*, Jaboticabal, v.42, n.2, p.1-8, 2020.
- SILVA, D.L.V.; ALVES, R.E.; FIGUEIREDO, R.W.; MACIEL, V.T.; FARIAS, J.M.; AQUINO, R.L. Características físicas, físico-químicas e sensoriais da água de frutos de coqueiro anão verde oriundo de produção convencional e orgânica. *Ciência e Agrotecnologia*, Lavras, v.33, n.4, p.1079-1084, 2009.

**ARTIGO 2 - ADJUSTMENT OF MIXED NONLINEAR MODELS ON
BLACKBERRY FRUIT GROWTH**

Publicado na Revista Brasileira de Fruticultura

DOI: 10.1590/0100-29452023665

Normas do periódico (versão publicada)



Economy - Original Article - Edited by: Alessandro Dal Col'Lucio

Adjustment of mixed nonlinear models on Blackberry fruit growth

Édipo Menezes da Silva¹, Maraísa Hellen Tadeu¹, Edilson Marcelino da Silva¹, Rafael Pio¹, Tales Jesus Fernandes¹, Joel Augusto Muniz¹

¹ Federal University of Lavras (UFLA) Lavras, MG, Brazil.

*Corresponding author: ediposvm01@gmail.com

Abstract: Blackberry fruits belong to the *genus Rubus*, are fruits more cultivated in temperate climate in the summer, with low luminosity and low temperature in the winter. These fruits have as characteristic the quickperishingafter harvest and regression models, more specifically, nonlinear models, single or double sigmoid growth curve, are more recommended to model the growth of living beings. Several authors have used these models, considering only the average data of individuals under study; however, they do not consider the variability between them. One way to better capture the variability between individuals is by using mixed-effects nonlinear models that, by definition, combine the fixed and random part in the same model. Data used in this work were diameter and length of 'Choctaw' blackberry fruits, . The random effects of models were tested on parameters, with some steps, in order to reach the most appropriate model. For fixed-effects models, the least squares method was used, and for mixed models, the restricted likelihood was used. To reach the model that best fits data, the fit quality criteria (R^2 , AIC_c and TRV) were used. For fruit diameter, the simple sigmoid nonlinear model was the logistic with random effect in β_1 and β_2 , and for fruit length, the model was the Logistic + Logistic, with random effect in β_1 and β_4 .

Index terms: 'Choctaw' cultivar; Restricted likelihood; Variability; Estimation.

Ajuste de modelos não lineares mistos no crescimento dos frutos de amora-preta

Resumo: Os frutos de amora-preta pertencem ao gênero *Rubus*, são frutos mais cultivados em clima moderado, no verão, com pouca luminosidade e baixa temperatura no inverno. Esses frutos têm por característica perecerem de forma rápida após a colheita, e os modelos de regressão, mais especificamente, os não lineares, de crescimento sigmoide simples ou duplo, são mais recomendados para se modelar o crescimento de seres vivos. Diversos autores utilizam esses modelos, levando em consideração apenas os dados médios dos indivíduos em estudo; contudo, não levam em consideração a variabilidade existente entre eles. Uma forma de captar melhor a variabilidade existente entre os indivíduos é utilizando os modelos não

Rev. Bras. Frutic., v.45, e-665 DOI: <https://dx.doi.org/10.1590/0100-29452023665>

Received 05 Nov, 2022 • Accepted 16 Mar, 2023 • Published Nov/Dec, 2023. Jaboticabal - SP - Brazil.



All the contents of this journal, except where otherwise noted, is licensed under a *Creative Commons Attribution 4.0 International (CC BY 4.0)*.

lineares de efeitos mistos que, por definição, combinam a parte fixa e aleatória no mesmo modelo. Os dados utilizados, neste trabalho, foram o diâmetro e o comprimento dos frutos de amora-preta, cultivar Choctaw. Os efeitos aleatórios dos modelos foram testados nos parâmetros, com algumas etapas, a fim de chegar ao modelo mais adequado. Para os modelos de efeitos fixos, o método de mínimos quadrados foram utilizados, e para os modelos mistos a verossimilhança restrita. Para se chegar ao modelo que melhor se ajuste aos dados, os critérios qualidade de ajuste (R^2_{aj} , AIC_c and TRV) foram utilizados. Para o diâmetro dos frutos, o modelo não linear sigmoide simples foi o logístico, com efeito aleatório no β_1 and β_2 ; já, para o comprimento, foi o Logístico + Logístico, tendo o efeito aleatório no β_1 and β_4 .

Termos de indexação: Cultivar Choctaw; Verossimilhança restrita; Variabilidade; Estimação.

Introduction

Small fruits, such as blackberries, strawberries, raspberries, gooseberries, blueberries, among others, have been growing in importance, quality and volume. The cultivation of these species is characterized by the low cost of implantation, production and good economic return. The fresh consumption of fruits maintains their physicochemical, nutritional and biological properties. However, most of the production is destined to the production of processed products such as jellies, juices, pulp for ice cream, among others. The consumption in the processed form is justified by the fact that these fruits have quick perishing, that is, after harvesting, their shelf life is short (TADEU et al, 2015).

Blackberry belongs to the genus *Rubus*, is a small fruit, with good adaptation to temperate climate. Its cultivation for marketing began in Europe in the 16th century, later taken to the USA and arriving in Brazil in the 1970 (RASEIRA; FRANZON, 2012). This is a crop with quick financial return for the producer, but little explored in Brazil, with plantations occurring in the states of Rio Grande do Sul, Espírito Santo, Rio de Janeiro and southern Minas Gerais (FACHINELLO et al., 2011).

These fruits are subjectively harvested by color and size (CAVALINI et al., 2006). The conservation time of fresh fruits is short, so identifying the optimal harvest time is very important, increasing the consumption time of fresh fruits. With the rapid loss

of post-harvest quality, there is a very large limitation for consumption in the fresh form; therefore, it is very important to use techniques to mitigate losses.

Thus, choosing the ideal harvest time is very important and knowing the growth and development curve of this fruit can be useful in this process, helping to reduce the rapid loss of quality after the harvest period, which is the main problem for fruits intended for fresh consumption.

Linear or non-linear regression models can be of great value in obtaining statistical models to objectively describe how this growth pattern occurs, highlighting non-linear regression models that, in the case of simple sigmoid, have less parameters and with biological interpretation (MAZZINI et al., 2003).

However, in several studies with longitudinal data, such as those carried out by Muianga et al., (2016), Ribeiro et al., (2018a), Silva et al., (2020), Miranda et al. (2021) and Silva et al., (2021b), researchers do not take into account the variability among researched fruits. Among the various ways to verify the variability among fruits, modeling via mixed models can be highlighted (WYZYKOWSKI et al., 2015; SARI et al., 2018; SARI et al., 2019b). Mixed models are the result of a technique where the fixed part and the random part are combined (PINHEIRO; BATES, 2004), and the advantage of modeling in this way is being able to capture the variability among fruits under study, with the inclusion

of random effect on one or more parameters. SARI et al. (2018) studied the random effect of growth, verifying the inclusion of random effects to the parameters of nonlinear growth models and concluded that the inclusion of random effects is efficient.

As in fixed-effects models, for the modeling of random effects, it is necessary to assign initial values for the method to interactively calculate the estimates up to a stopping criterion and, as in fixed models, it is not an easy task. Estimation can be performed by some methods, the restricted likelihood being the most used (LINDSTROM; BATES, 1990). Restricted likelihood is a generalization of likelihood, but this method takes into account the fixed part and the random part in the model (PINHEIRO; BATES, 2004).

There are advantages in mixed models in relation to fixed-effect ones, that in addition to verifying the variation among individuals, these models are more flexible in relation to assumptions, incorporating dependence and heteroscedasticity to the model via the variance-covariance matrix (PINHEIRO; BATES, 2004).

In literature, most studies on fruit growth are carried out with simple sigmoid models; however, in some cases, this growth occurs in two development stages. In the first stage, growth is slow at the beginning, changing to fast and stabilizing at the end, and in a second stage, the same characteristics are found, becoming a double sigmoid model (FERNANDES et al, 2017; SILVA et al., 2020, FERNANDES et al., 2022).

Thus, the aim of the present study was to verify the growth pattern through non-linear sigmoid single and double regression models, adding, if necessary, one or more parameters of random effects, applying this methodology to data obtained from 'Choctaw' blackberry in relation to fruit diameter and length.

Material and methods

Data were extracted from (TADEU et al., 2015) and represent the results of the experiment carried out at the Federal University of Lavras, Lavras, MG, from January 2012 to January 2014. The climate in the region is Cwb type (mesothermal or high-altitude tropical climate), with dry winter and rainy summer, according to the Koppen classification. The variables under study were diameter and length of 'Choctaw' blackberries. Seedlings were planted in 2009 with "T" vertical cordon with 60 cm of distance and 80 cm of height and 3.0 m x 0.5 m spacing.

Tadeu et al. (2015) carried out the initial study aiming to compare the effect of conventional (control) and drastic summer pruning. From this study, diameter and length measurements of the 'Choctaw' cultivar were taken, in mm. Such measurements were collected with the aid of digital caliper (model King Tools 150 mm, Cia, São Paulo, SP). The choice of fruits took into consideration the development stage, that is, fruits whose petals and anthers already fell, at the same time, were chosen considering longitudinal data, that is, the same fruits had measurements verified until harvest. Twelve measurements were collected over time, the first being three days after anthesis (DAA) and the last after thirty-five DAA.

The parameterization used in Table 1 for model adjustments was taken from Fernandes et al. (2015), as follows. In this parameterization, β_1 and β_4 represent the maximum expected growth in each growth curve, β_2 and β_5 the inflection point in the respective curves, except for the Brody model, which does not have an inflection point, β_3 and β_6 the parameters linked to the growth rate in each curve, representing the fixed part and $b_1, b_2, b_3, b_4, b_5, b_6$ are the random effects parameters associated with each of the fixed-effects parameters, with b_1 and b_2 associated with β_1 and β_4 , b_3 and b_5 associated with β_2 and β_5 and b_4

and b_6 associated with β_3 and β_6 , respectively. $\sigma^2 \mathbf{D}$, $\sigma^2 \mathbf{D}$, the matrix of variances and co-
 With b_i the random effects of the model independent and identically distributed, $b_i \sim N(0,$ variances and ϵ_{ij} the random error associated with the model, with ϵ_{ij} .

Table 1 - Expressions of Logistic, Gompertz, Logistic + Logistic, Gompertz + Gompertz, Logistic + Brody and Gompertz + Brody adjusted models of fixed and random effects to describe the diameter and height growth of 'Choctaw' blackberry.

Model	Random effect	Expression
Logistic	-----	$\frac{\beta_1}{1 + e^{\beta_3(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	β_1	$\frac{(\beta_1 + b_1)}{1 + e^{\beta_3(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	β_3	$\frac{\beta_1}{1 + e^{(\beta_3 + b_3)(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	$\beta_1\beta_3$	$\frac{(\beta_1 + b_1)}{1 + e^{(\beta_3 + b_3)(\beta_1 - x_{ij})}} + \epsilon_{ij}$
	$\beta_1\beta_2$	$\frac{(\beta_1 + b_1)}{1 + e^{\beta_3((\beta_2 + b_2) - x_{ij})}} + \epsilon_{ij}$
Gompertz	-----	$\beta_1 e^{-e^{\beta_3(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	β_1	$(\beta_1 + b_1) e^{-e^{\beta_3(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	β_3	$\beta_1 e^{-e^{(\beta_3 + b_3)(\beta_2 - x_{ij})}} + \epsilon_{ij}$
	$\beta_1\beta_2$	$(\beta_1 + b_1) e^{-e^{\beta_3((\beta_2 + b_2) - x_{ij})}}$
Brody +Logistic	-----	$\beta_1(1 - e^{\beta_3(\beta_2 - x)}) + \frac{(\beta_4 - \beta_1)}{1 + e^{\beta_5(\beta_6 - x_{ij})}} + \epsilon_{ij}$
	$\beta_1\beta_4$	$(\beta_1 + b_1)(1 - e^{\beta_3(\beta_2 - x_{ij})}) + \frac{((\beta_1 + b_1) - (\beta_4 + b_4))}{1 + e^{\beta_5(\beta_6 - x_{ij})}} + \epsilon_{ij}$
Logistic+ Logistic	-----	$\frac{\beta_1}{1 + e^{\beta_3(\beta_2 - x_{ij})}} + \frac{(\beta_4 - \beta_1)}{1 + e^{\beta_5(\beta_6 - x_{ij})}} + \epsilon_{ij}$
	$\beta_1\beta_4$	$\frac{(\beta_1 + b_1)}{1 + e^{\beta_3(\beta_2 - x_{ij})}} + \frac{((\beta_4 + b_4) - (\beta_1 + b_1))}{1 + e^{\beta_5(\beta_6 - x_{ij})}} + \epsilon_{ij}$

At first, in order to have prior knowledge of how these measures behave over time, scatter plots and boxplots were built with the ggplot2 package of the R software (WICHAM, 2016). Subsequently, fixed-effects nonlinear regression models were adjusted, without

verifying the assumptions of residuals. Since these fixed-effects models were adjusted, the following steps were used to verify in which parameters the random effect will be included. Step 1 is a method taken from Pinheiro and Bates (2004), which consists of

making individual confidence intervals (95%) in parameters illustrated in Figure 1. If all intervals overlap each other, it is an indication that there is no need to include the random effect in the parameter; however, if at least one does not overlap with the others, there is a need to include this effect in the parameter.

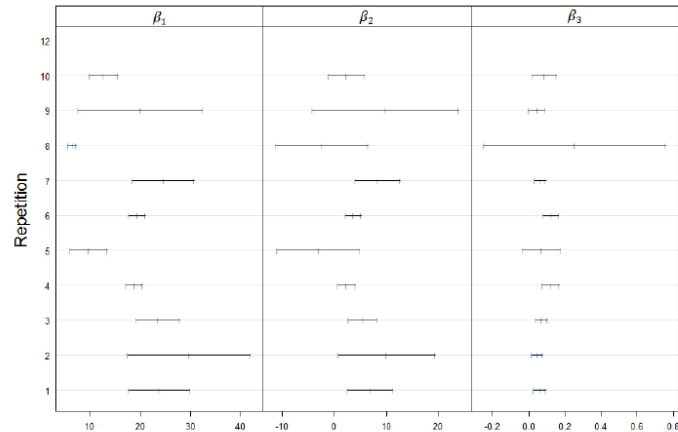


Figure 1- Graphic representation of the 95% confidence intervals for estimates of the logistic non-linear model parameters.

Step 2, readjust the models combining the inclusion of random effects to parameters using the comparison criteria between models (R_{aj}^2 , AIC_c , and TRV), thus reaching a decision on which parameters need to be included (WYZYKOWSKI, et al. , 2014).

After choosing the two models to describe diameter and length, these models were adjusted once again, adding the structure of the variance-covariance matrix of the random effects to parameters. The matrices chosen were varPower and varExp, both implemented in the nlme package (PINHEIRO et al. (2021)) and the first-order autoregressive term (AR1).

To adjust the models, the R software (R DEVELOPMENT CORE TEAM, 2022) was used, estimating the parameters of fixed-effects models via ordinary least squares and for mixed models restricted maximum likelihood, using the nlme packages (PINHEIRO et al., 2021). The initial values for the iterative method to start the iteration was a graphic analysis with the nlsList and manipulate

functions (Allaire, 2014; Diel et al., 2019).

To verify the homogeneity of variances and normality of residues, the Breusch-Pagan (BREUSCH; PAGAN, 1979) and Shapiro-Wilk (SHAPIRO; WILK, 1965) tests were used, respectively. To assess the existence of residual autocorrelation, the Durbin-Watson test was used (DURBIN; WATSON, 1951). In addition to hypothesis testing, graphical analysis was used to verify these assumptions. To verify the fit quality of models, the parametric non-linearity measure was used, and as mentioned by Sari et al. (2019a), values of parametric nonlinearity measures less than 1 indicate that the parameterization used is acceptable. Models were compared using the following criteria:

i) Adjusted determination coefficient (R_{aj}^2), obtained by: $R_{aj}^2 = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - p} \right]$,

where R^2 is the determination coefficient, n is the number of times in which measurements were taken and p is the number of model parameters. ii) Corrected Akaike

Information Criterion (AKAIKE, 1974), given by: $AIC_c = AIC + \frac{2p(p+1)}{n-p-1}$, where

$AIC = n \ln\left(\frac{SQR}{n}\right) + 2p$, where is the sum of

squares of residuals, p is the number of model parameters, n is the sample size and \ln is the natural logarithmic operator, iii) Likelihood ratio test given by: $L = -2\log(L_1/L_2)$. After choosing the models with the best fit quality values, the critical points were estimated for fruit diameter and length, according to Silva et al. (2021b). For all analyses, the significance level adopted was 5%.

Results and discussion

Initially, the exploratory analysis of data was carried out in order to verify the behavior of data over time. As observed in Table 2, the averages of the two characteristics studied at 35 DAA were 17.60 and 15.80 mm, respectively for diameter and length. These two averages can be used as the initial value of the asymptotic parameter of models (MISCHAN; PINHO, 2014; FRÜKAUF, 2022).

Table 2- Descriptive statistics (average and variance) for diameter and length of ‘Choctaw’ blackberry fruits, at each time, in mm.

Time	Average		Variance	
	Length	Diameter	Length	Diameter
1	5.52	4.34	0.24	0.31
5	7.38	6.87	1.22	1.82
8	7.71	8.62	4.57	4.68
10	9.73	9.71	4.52	11.50
13	10.40	11.10	6.43	13.00
16	11.50	11.70	8.14	14.90
20	11.60	11.70	12.10	16.40
22	12.70	12.50	12.50	18.70
24	13.10	13.30	14.80	18.90
28	14.30	14.00	24.50	28.60
31	16.00	15.20	29.60	33.70
35	17.60	15.80	34.60	46.60

Figures 2 and 3 show the scatter plots and boxplots of growth data (diameter and length) of individual fruits over time. It is possible to observe that, in both cases, at the beginning of the productive cycle, fruits have more homogeneous growth and, over time, there is greater growth variation (heterogeneous growth).

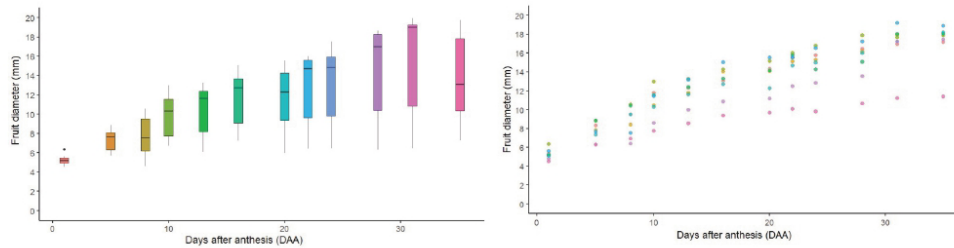


Figure 2 - Scatter plot and boxplot of ‘Choctaw’ blackberry diameter data over time.

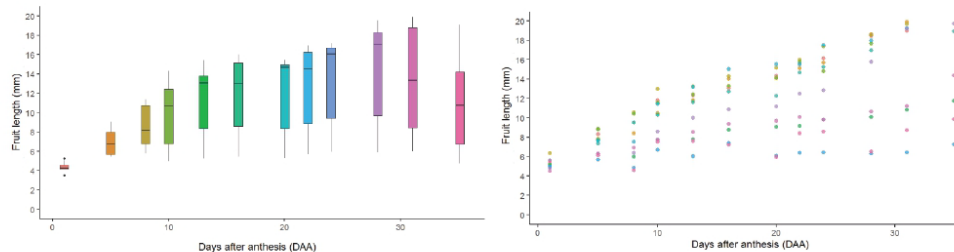


Figure 3 - Scatter plot and boxplot of ‘Choctaw’ blackberry length data over time.

The increase in data variability can be considered an indication of the need to model data heterogeneity and add a random effect to maximum growth parameters of models under study (PINHEIRO; BATES, 2004) or also adjust data quantiles with non-linear growth models via quantile regression (OLIVEIRA A. C. R. et al. (2021)).

Figure 4 shows the confidence intervals (CI) individually for each fruit. For CI that correspond to the adjustment of the Brody + Logistic and Double Logistic fixed models,

several fruits did not have individual adjustments with convergence. For fruits in which it was not possible to obtain estimates, it was not possible to draw conclusions regarding parameters with need to include a random effect. In the other adjustments (Gompertz and Logistic), there was convergence of all individual adjustments and the assumption that in parameters β_1 and β_2 it is possible to add a random effect due to the non-overlapping of some of CIs (MEDEIROS S. D. S. et al. (2020), WYZYKIOWSKI et al (2015) and PINHEIRO; BATES (2004)).

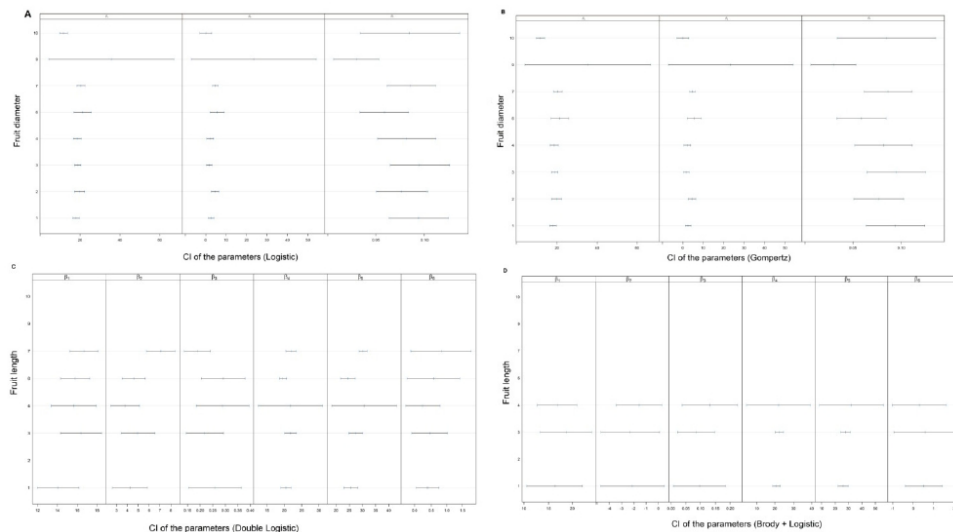


Figure 4 - Graphic representation of 95% confidence intervals of parameters of Logistic (A) and Gompertz (B) models for diameter growth and Logistic + Logistic (C) and Brody+Logistic (D) models for 'Choctaw' blackberry fruit length.

Non-linear fixed models were adjusted for fruit diameter and length via ordinary least squares, since, as mentioned by Wzykiowski et al., (2015) and Medeiros et al., (2020), the first adjustment is performed considering that the residual vectors are independent, identically distributed by a zero mean normal and constant variance. After the adjustment, considering that all assumptions were met, the analysis of residuals is necessary to prove or not the hypotheses of the vector of residues.

The results of normality (SW), homoscedasticity (BP) and independence (DW) tests are shown in Table 3, and for fixed-effect models, it was possible to observe that all models had breach of normality assumptions, homogeneous and independent variances (p -value < 0.05), with graphic analyses (Figures 5, 6, 7 and 8), corroborating the tests adopted. With the breach of assumptions, modeling via mixed models, in addition to adding random effects to parameters, the incorporation of the first-order autoregressive term

and heterogeneity modeling was necessary, according to Fernandes et al. (2014), Frühauf et al. (2022), Medeiros et al. (2020), Ribeiro et al. (2018a), Ribeiro et al. (2018b), Silva et al. (2020), Silva et al. (2021a), Silva et al. (2021b) and Wyzykiowski et al. (2015).

Table 3 – P-value of Shapiro-Wilk (SW), Breusch-Pagan (BP) and Durbin-Watson (DW) tests for

the Logistic, Gompertz, Brody + Logistic and Double Logistic models of fixed effects adjusted to growth data of 'Choctaw' blackberry in relation to diameter and length.

Characteristic	Adjusted model	SW	BP	DW
Diameter	Gompertz	<0.001	<0.001	0.031
	Logistic	<0.001	<0.001	0.023
Length	Brody + Logistic	<0.001	<0.001	0.012
	Logistic + Logistic	<0.001	<0.001	0.026

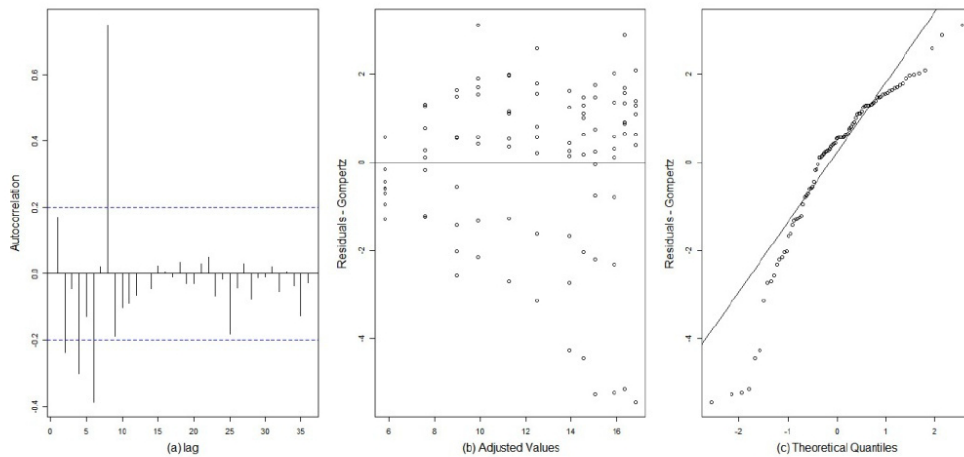


Figure 5 - Residual analysis for diameter growth, where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for the simple Gompertz model.

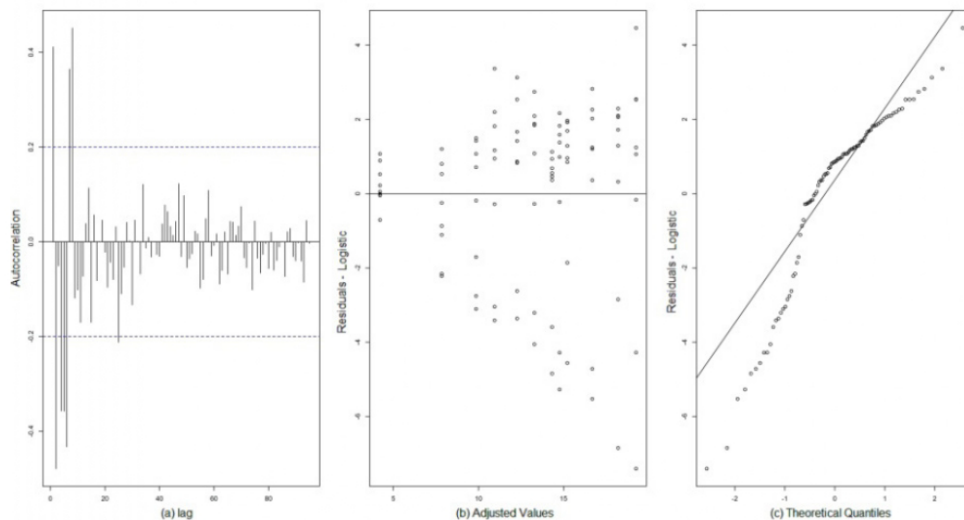


Figure 6 - Residual analysis for diameter growth, where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for the simple logistic model.

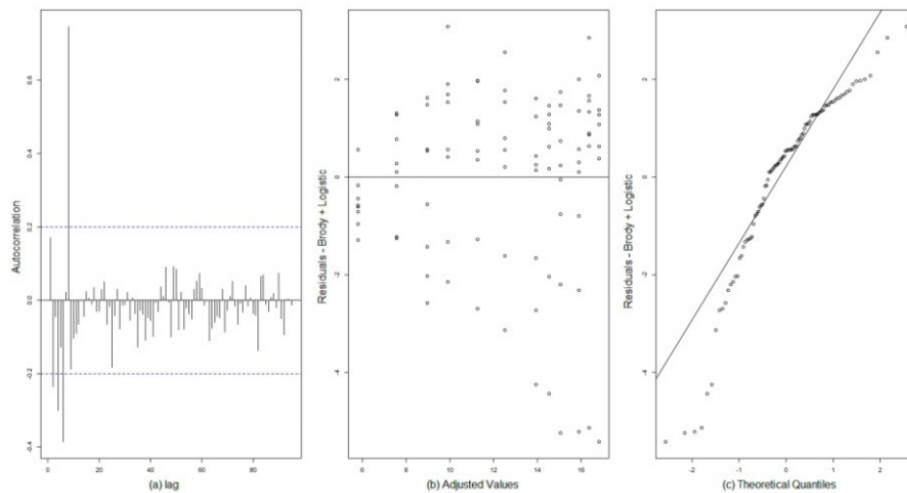


Figure 7 - Residual analysis for length growth, where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for the Brody + Logistic model.

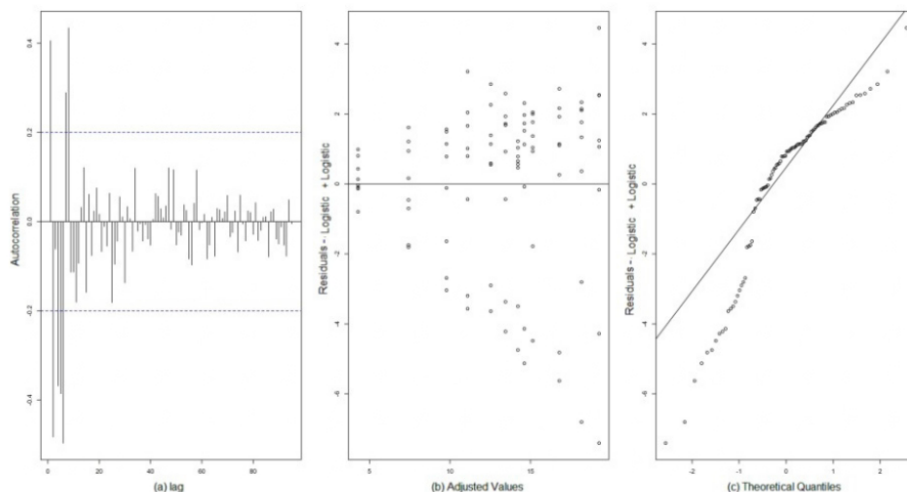


Figure 8 - Residual analysis for length growth, where (a) represents the autocorrelation graph by lags, (b) represents the adjusted values in relation to residuals and (c) the residual values in relation to the theoretical quantiles for the Logistic + Logistic model.

When evaluating the parametric non-linearity (c^{θ}) of Gompertz, Logistic, Double Logistic and Brody + Logistic models, all non-linearity values were less than 1, being considered that the models are adequate to describe the functional relationship between time and diameter and time and length (SARI et al. (2019a)).

Table 4 shows the fit quality evaluators of models for models that had convergence with the inclusion of random effects to parameters, where all had the inclusion of the first-order autoregressive term and some combinations with the varPower and varExp structure.

Table 4 - Criteria to assess the fit quality: adjusted determination coefficient (R_{aj}^2), Akaike information criterion (AICc), and the likelihood ratio test (TRV), to compare Logistic, Gompertz, Double Logistic and Brody + Logistic models with growth data from 'Choctaw' blackberry fruits in relation to diameter and length.

Characteristic	Adjusted model	Parameters	R_{aj}^2	AICc	TRV
Length	Brody + Logistic	a and a1	0.98	437.15	-208.21
		a and a1*	0.98	435.54	-206.43
		a and a1**	0.98	437.22	-207.27
	Logistic + Logistic	a and a1	0.98	426.65	-202.99
		a and a1*	0.98	425.41	-201.37
		a and a1**	0.98	427.37	-202.34
Diameter	Gompertz	a	1.00	275.18	-132.46
		k	1.00	277.37	-133.55
		a and b	1.00	269.97	-127.85
	Logistic	a and b*	0.97	263.87	-124.80
		a and b**	0.97	265.92	-125.83
		a	1.00	289	-139.36
Diameter	Logistic	k	0.8	405.52	-197.63
		a and b	1.00	282.4	-134.06
	a and k	1.00	284.25	-142.32	
	a and b*	1.00	269.68	-127.71	
a and b**	1.00	271.17	-128.45		

* varPower; ** varExp;

When comparing the results for fruit length with Brody + Logistic and Logistic + Logistic models with the varPower weight matrix, the highest adjusted determination coefficient and likelihood ratio test values, as well as the lowest AICc values were found. In the study by Silva et al. (2020), the Logistic + Logistic fixed-effects model was found to be the best model for 'Choctaw' blackberry data.

For diameter results, similarly to length, models adjusted with the inclusion of the varPower weight matrix were those with the best quality indicators for Gompertz and Logistic models. For length, the Logistic + Logistic model showed the best adjustment, as found in the study by Silva et al. (2020) and for diameter, the logistic model, the same model found by Muianga et al. (2016) for cashew data.

When defining the two best models for fruit diameter and length, the analysis of their residues was carried out and results are shown in Figures 9 and 10.

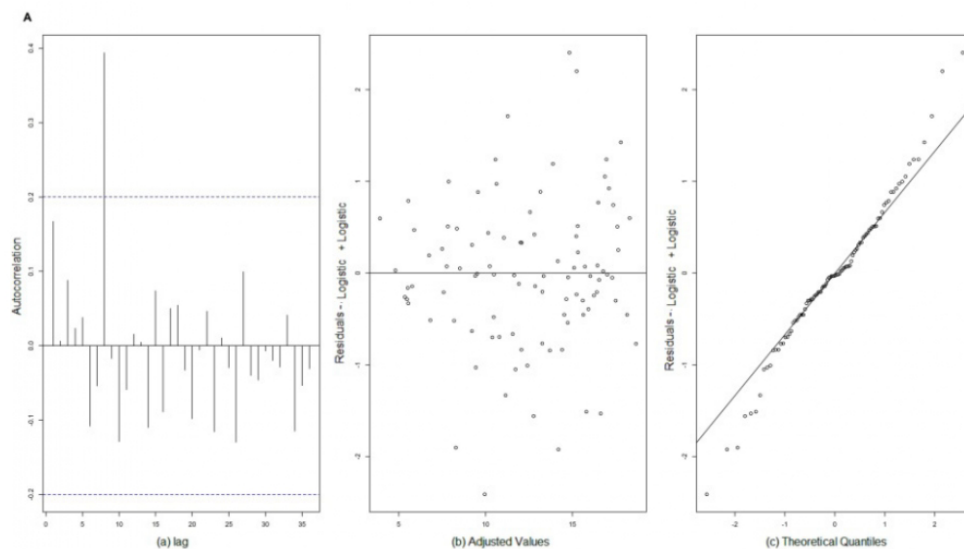


Figure 9A. Residual analysis for length growth, where (a) represents the autocorrelation graph by lags. Residual analysis for length growth, where (b) represents the adjusted values in relation to residuals (c) residual values in relation to the theoretical quantiles for Logistic + Logistic models with random effect on β_1 and β_4 with varPower.

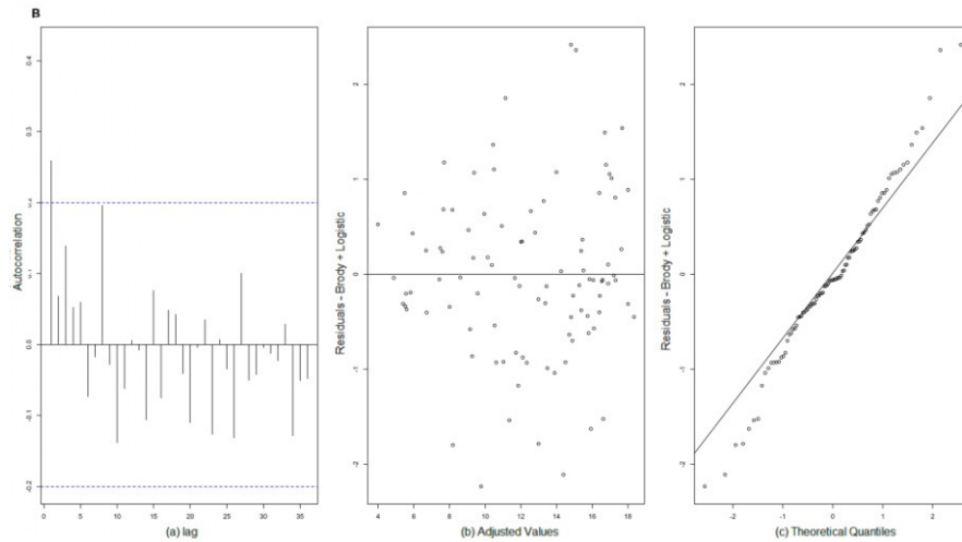


Figure 9B. Residual analysis for length growth, where (a) represents the autocorrelation graph by lags. Residual analysis for length growth, where (b) represents the adjusted values in relation to residuals (c) residual values in relation to the theoretical quantiles for Brody + Logistic models with random effect on β_1 and β_4 with varPower.

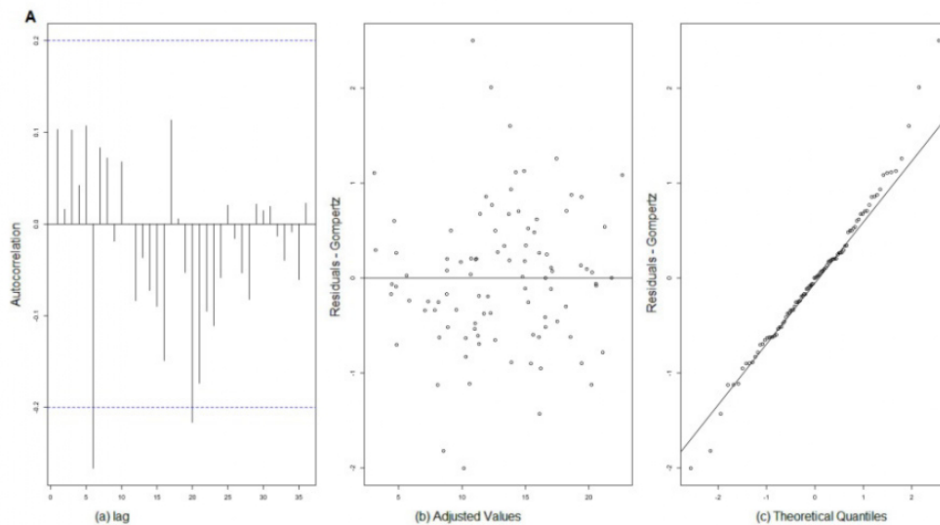


Figure 10A. Residual analysis for length growth, where (a) represents the autocorrelation graph by lags. Residual analysis for length growth, where (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for Gompertz models with random effect on β_1 and β_2 with varPower.

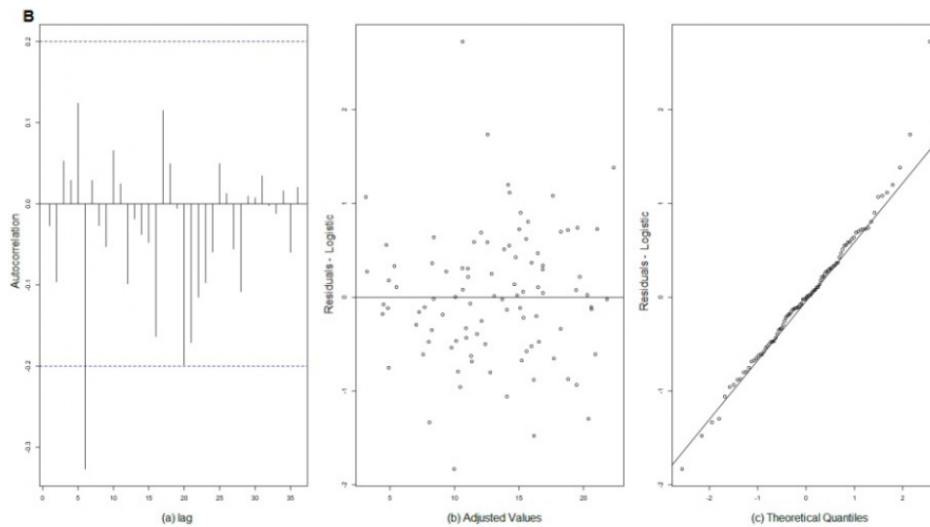


Figure 10B. Residual analysis for length growth, where (a) represents the autocorrelation graph by lags. Residual analysis for length growth, where (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for Logistic models with random effect on β_1 and β_2 with varPower.

As could be observed, for the two best models, for length and diameter, the graphic analysis of lags (a), adjusted values vs residuals (b) and the graph of theoretical quantiles vs residuals (c) that the inclusion in the weight matrix (varPower) and first-order residual autocorrelation were efficient to make the vector of residuals meet the initial assumptions of the statistical modeling and thus be able to have more reliable inferences of the estimated parameters (FRÜHALF A. C. et. al., 2020; JANE et al., 2020; PRADO et al., 2020).

Table 5 shows the estimates of adjusted models with the best fit quality values for the length and diameter of blackberry fruits. The maximum expected diameter growth was 17.24 mm, a statistically similar value in relation to that found by Silva et al. (2020), who found value of 17.11 mm, thus showing that the maximum expected diameter growth of 'Choctaw' blackberry fruits is approximately 17.00 mm, with the inflection point occurring before the expected half of 7.125 days.

Table 5 - Estimates of parameters for Brody + Logistic, Logistic + Logistic, Gompertz and Logistic models applied to the growth data of 'Choctaw' blackberry fruits in relation to diameter and length.

Characteristic	Adjusted model	Parameters	β_1	β_2	β_3	β_4	β_5	β_6
Length	Brody + Logistic	a and a1 (varPower)	17.24*	-2.75	0.07*	20.16*	29.24*	0.66*
	Logistic + Logistic	a and a1 (varPower)	14.69*	4.91*	0.21*	19.54*	28.69*	0.43*
Diameter	Gompertz	a and b (varPower)	17.87*	3.26*	0.08*	---	---	---
	Logistic	a and b (varPower)	17.11*	7.15*	0.12*	---	---	---

*Significant at 5% significance.

As for length, the asymptotic value of the first growth phase occurred when fruits had average of 17.24 mm and maximum expected growth of 20.16 mm. The parameter

linked to the growth rate in the first phase was greater than in the second, which was also observed by Silva et al. (2020) and Fernandes et al (2022).

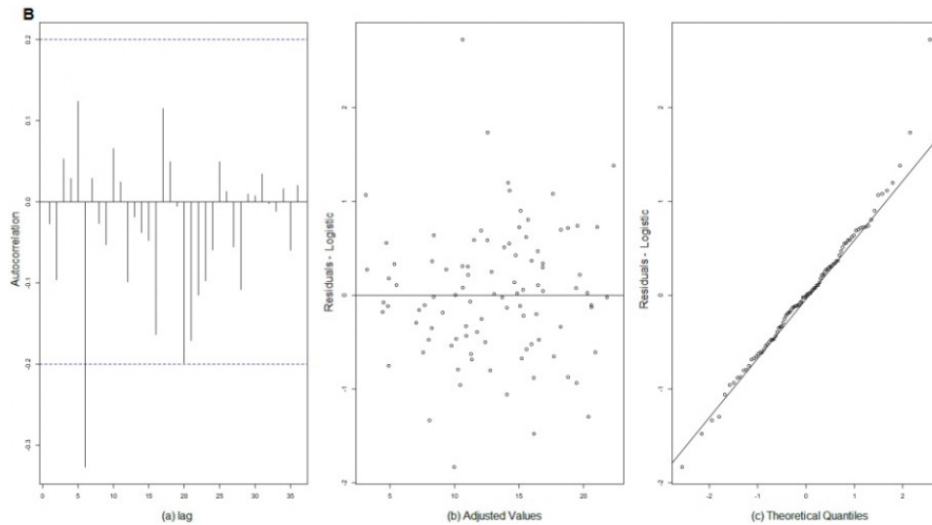


Figure 10B. Residual analysis for length growth, where (a) represents the autocorrelation graph by lags. Residual analysis for length growth, where (b) represents the adjusted values in relation to residuals and (c) residual values in relation to the theoretical quantiles for Logistic models with random effect on β_1 and β_2 with varPower.

As could be observed, for the two best models, for length and diameter, the graphic analysis of lags (a), adjusted values vs residuals (b) and the graph of theoretical quantiles vs residuals (c) that the inclusion in the weight matrix (varPower) and first-order residual autocorrelation were efficient to make the vector of residuals meet the initial assumptions of the statistical modeling and thus be able to have more reliable inferences of the estimated parameters (FRÜHALF A. C. et. al., 2020; JANE et al., 2020; PRADO et al., 2020).

Table 5 shows the estimates of adjusted models with the best fit quality values for the length and diameter of blackberry fruits. The maximum expected diameter growth was 17.24 mm, a statistically similar value in relation to that found by Silva et al. (2020), who found value of 17.11 mm, thus showing that the maximum expected diameter growth of ‘Choctaw’ blackberry fruits is approximately 17.00 mm, with the inflection point occurring before the expected half of 7.125 days.

Table 5 - Estimates of parameters for Brody + Logistic, Logistic + Logistic, Gompertz and Logistic models applied to the growth data of ‘Choctaw’ blackberry fruits in relation to diameter and length.

Characteristic	Adjusted model	Parameters	β_1	β_2	β_3	β_4	β_5	β_6
Length	Brody + Logistic	a and a1 (varPower)	17.24*	-2.75	0.07*	20.16*	29.24*	0.66*
	Logistic + Logistic	a and a1 (varPower)	14.69*	4.91*	0.21*	19.54*	28.69*	0.43*
Diameter	Gompertz	a and b (varPower)	17.87*	3.26*	0.08*	---	---	---
	Logistic	a and b (varPower)	17.11*	7.15*	0.12*	---	---	---

*Significant at 5% significance.

As for length, the asymptotic value of the first growth phase occurred when fruits had average of 17.24 mm and maximum expected growth of 20.16 mm. The parameter

linked to the growth rate in the first phase was greater than in the second, which was also observed by Silva et al. (2020) and Fernandes et al (2022).

Analyzing results found in Table 5, it was possible to observe that all parameters were significant to explain the non-linear relationship between variable time and fruit diameter and length. According to Silva et al. (2020) and Fernandes et al. (2022), both models are indicated to describe such relationship, that is, parameters were significant, with vector of residuals met.

With adjustments of the best model for length and diameter, Figure 11 shows how, over the days, fruit growth occurred individually and in the random and fixed-effect models for Logistic + Logistic (length) and Logistic (diameter) models. As demonstrated in steps so far, in this work, the adjustment of models, adding random effects to parameters, better explained the existing data variance.

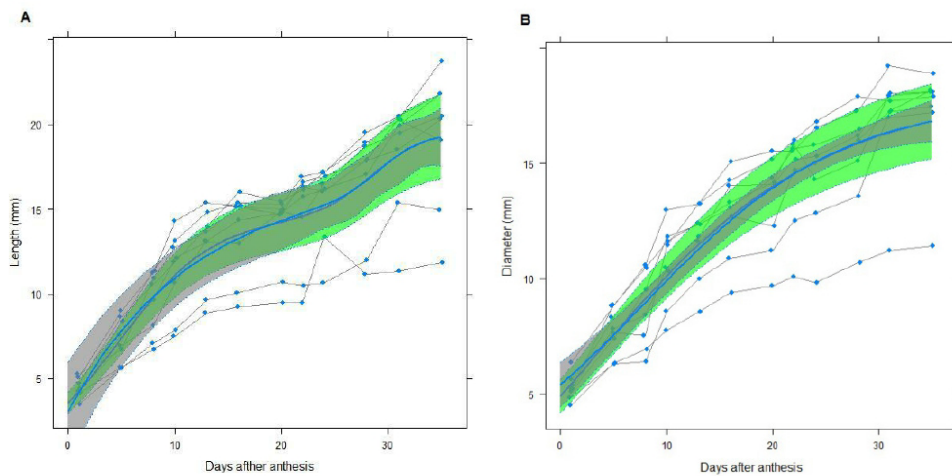


Figure 11 - Adjusted curves of fixed-effect and random-effect models for the Logistic + Logistic model (A) and the Logistic model (B) referring to 'Choctaw' blackberry fruits.

Analyzing both graphs, it was possible to observe that the range of the curve estimated for random-effect models better shows the variation in the growth of individuals, both at the beginning and at the end of the productive cycle (Figure 11). Analyzing graphs, it is clear that the confidence interval for the adjustment of random effects (in green) does not remain constant over time, being, at the beginning, more concentrated and, at the end, more dispersed, corroborating results shown in Figures 2, 3 and Table 2; on the other hand, the fixed-effect model remains constant over time.

Analyzing the present results, in Figure 11, considering some parameters with random effects for the growth of blackberry fruits (diameter and length) is necessary due to

the results found so far and, according to Pinheiro and Bates (2004), ignoring this finding can lead to mistaken estimates of inferences in relation to models.

Figure 12 shows graphs of growth rates for length (on the left) and diameter (on the right) of Logistic + Logistic and Simple Logistic models, respectively. The abscissa and ordinate estimates of the maximum acceleration point (pma), inflection point (pi), maximum deceleration point (pmd) and asymptotic deceleration point (pda) were taken from studies by Silva et al. (2021b) and correspond to critical points. However, Silva et al. (2021b) estimated the points for simple sigmoid models (Gompertz and Logistic); however, as double models are the sum of models, estimates can be used both to esti-

Analyzing results found in Table 5, it was possible to observe that all parameters were significant to explain the non-linear relationship between variable time and fruit diameter and length. According to Silva et al. (2020) and Fernandes et al. (2022), both models are indicated to describe such relationship, that is, parameters were significant, with vector of residuals met.

With adjustments of the best model for length and diameter, Figure 11 shows how, over the days, fruit growth occurred individually and in the random and fixed-effect models for Logistic + Logistic (length) and Logistic (diameter) models. As demonstrated in steps so far, in this work, the adjustment of models, adding random effects to parameters, better explained the existing data variance.

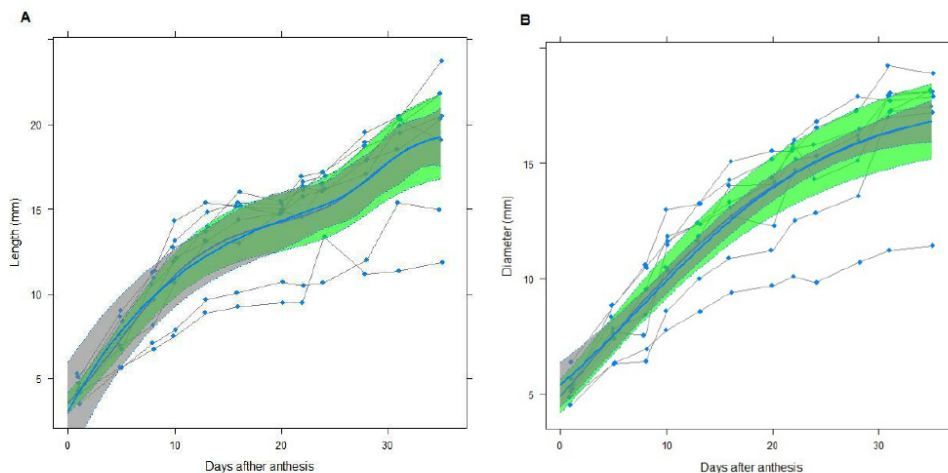


Figure 11 - Adjusted curves of fixed-effect and random-effect models for the Logistic + Logistic model (A) and the Logistic model (B) referring to 'Choctaw' blackberry fruits.

Analyzing both graphs, it was possible to observe that the range of the curve estimated for random-effect models better shows the variation in the growth of individuals, both at the beginning and at the end of the productive cycle (Figure 11). Analyzing graphs, it is clear that the confidence interval for the adjustment of random effects (in green) does not remain constant over time, being, at the beginning, more concentrated and, at the end, more dispersed, corroborating results shown in Figures 2, 3 and Table 2; on the other hand, the fixed-effect model remains constant over time.

Analyzing the present results, in Figure 11, considering some parameters with random effects for the growth of blackberry fruits (diameter and length) is necessary due to

the results found so far and, according to Pinheiro and Bates (2004), ignoring this finding can lead to mistaken estimates of inferences in relation to models.

Figure 12 shows graphs of growth rates for length (on the left) and diameter (on the right) of Logistic + Logistic and Simple Logistic models, respectively. The abscissa and ordinate estimates of the maximum acceleration point (pma), inflection point (pi), maximum deceleration point (pmd) and asymptotic deceleration point (pda) were taken from studies by Silva et al. (2021b) and correspond to critical points. However, Silva et al. (2021b) estimated the points for simple sigmoid models (Gompertz and Logistic); however, as double models are the sum of models, estimates can be used both to esti-

mate a simple development cycle and a double development cycle.

When analyzing results in Figure 12, it was observed that the inflection point occurred exactly at the peak of each curve, that is, for the simple Logistic model with 8.55 mm and at 7.15 DAA, for the double Logistic model, for the first growth phase was 7.35 mm and

4.91 DAA, while in the second phase, 9.77 mm and 28.69 DAA. The asymptotic deceleration point was 15.54 and 14.69 for diameter and length, respectively, and Michan and Pinho (2014) reported that this value corresponds to approximately 90% of the maximum expected fruit growth and can therefore be used as a parameter for fruit harvest.

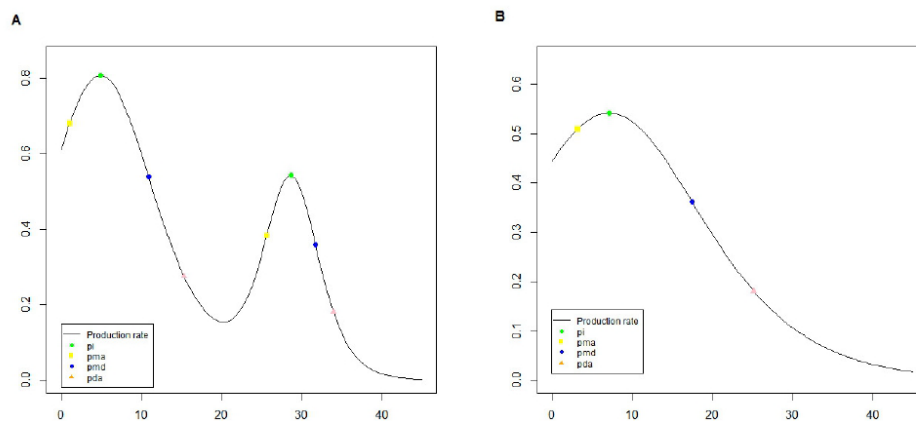


Figure 12 - Growth rate curves for Logistic + Logistic (A) and Logistic (B) models with maximum acceleration points (PMA), inflection point (PI), maximum deceleration point (PMD) and asymptotic deceleration point (PDA) for both length development stages and for the single diameter development stage.

Conclusions

By the fit quality evaluators, the choice of models that best adjusted to data were the double Logistic and the simple Logistic, respectively, for length and diameter. In all models, it was necessary to incorporate the varPower function and the first-order residual autocorrelation. The incorporation of the

random effect was indicated for the Double Logistic model in parameters β_1 and β_4 and in the Simple Logistic model in parameters β_1 and β_2 .

Acknowledgments

The authors would like to thank CAPES and CNPq for financial support.

References

- 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.
- ALLAIRE, J.J. *Interactive plots for RStudio*. R package version 1.0.1. 2014. Disponível em: <https://CRAN.R-project.org/package=manipulate>.
- BREUSCH, T.S.; PAGAN, A.R. A simple test for heterocedasticity and random coeficiente variation. *Econometrica*, New York, v.47, n.5, p.1287-94, 1979.

- CAVALINI, F. C. et al. Maturity indexes for 'kumagai' and 'paluma' guavas. **Revista Brasileira de Fruticultura**, Jaboticabal, v.28, n.2, p.176-9, 2006.
- DIEL, M.I.; SARI, B.G.; KRYSZCZUN, D.K.; OLIVORO, T.; PINEHIRO, M.V.M.; MEIRA, D.; SCHMIDT, D.; LUCIO, A.D. Nonlinear regression for description of strawberry (*Fragaria x ananassa*) production. **The Journal of Horticultural Science and Biotechnology**, Warnick, v.94, n.2, p.259-73, 2019.
- DURBIN, J.; WATSON, G.S. Testing for serial correlation in last squares regression. **Biometrika**, London, v.38, n.1/2, p.159- 77, 1951.
- FACHINELLO, J.C.; PASA, M.S.; SCHMITZ, J.D.; BETEMPS, D.L. Situação e perspectivas da fruticultura de clima temperado no Brasil. **Revista Brasileira de Fruticultura**, Jaboticabal, v.33, p.109-20, 2011.
- FERNANDES, J.G.; SILVA, E.M. da; RIBEIRO, T.D.; SILVA, E.M.; FERNANDES, T.J.; MUNIZ, J.A. Description of the peach fruit growth curve by diphasic sigmoidal non linear models. **Revista Brasileira de Fruticultura**, Jaboticabal, v.44, n.3, p.e-875, 2022.
- FERNANDES, T.J.; MUNIZ, F.A.; PEREIRA, A.A.; REZENDE MUNIZ, F.; MUIANGA, C.A.. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum**, Maringá, v.37, n.4, p.397, 2015.
- FERNANDES, T.J.; PEREIRA, A.A.; MUNIZ, J.A. Double sigmoidal models describing the growth of coffee berries. **Ciência Rural**, Santa Maria, v. 47, n. 8, 2017.
- 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.207-15, 2014.
- FRÜHAUF, A.C.; PEREIRA, G.A.; BARBOSA, A.C.M.C.; FERNANDES, T.J.; MUNIZ, J.A. Nonlinear models in the study of the cedar diametric growth in a seasonally dry tropical forest. **Revista Brasileira de Ciências Agrárias**, Fortaleza, v.15, p.e8558, 2020.
- FRUHAUF, A.C.; SILVA, E.M. da; GRANATO-SOUZA, D.; SILVA, E.M.; MUNIZ, J.A.; FERNANDES, T.J. Description of height growth of hybrid eucalyptus clones in semi-arid region using non-linear models. **Brazilian Journal of Biometrics**, Lavras, v.40, p.138-51, 2022.
- 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 non linear models. **Ciência Rural**, Santa Maria, v. 50, n.33, p.e20190408, 2020.
- LINDSTROM, M. J.; BATES, D. M. Nonlinear mixed effects models for repeated measures data. **Biometrics**, Malden, p.673-87, 1990.
- MAZZINI, A.R.A.; MUNIZ, J.A.; AQUINO, L.H.; SILVA, F.F. Análise da curva de crescimento de machos hereford. **Ciência e Agrotecnologia**, Lavras, v.27, n.5, p.1105-12, 2003.
- MEDEIROS, S.D.S.; LIMA, C.G.; SAVIAN, T.V.; MALHEIROS, E.B.; WERNER, S.S.. Mixed nonlinear models in ruminal in situ degradability trials. **Ciência Animal Brasileira**, Goiânia, v.21, 2020.
- MIRANDA, L.F.; FRÜHAUF, A.C.; LIMA, K.P.; SILVA, E.M.; MUNIZ, J.A. Non linear models to describe the growth of *Jatropha* plant (*Jatropha curcas* L.). **Revista Ciência Agronômica**, Fortaleza, v. 52, p. e20207602, 2021.
- MISCHAN, M.M.; PINHO, S.Z. **Modelos não lineares: funções assintóticas de crescimento**. São Paulo: Cultura Acadêmica, 2014.
- MUIANGA, C.A.; MUNIZ, J.A.; NASCIMENTO, M.S.; FERNANDES, T.J.; SAVIAN, T.V. Descrição da curva de crescimento de frutos do cajueiro por modelos não lineares. **Revista Brasileira de Fruticultura**, Jaboticabal, v.38, p.22-32, 2016.

- OLIVEIRA, A.C.R.; CECON, P.R.; PUIATTI, G.A.; GUIMARÃES, M.E.S.; CRUZ, C.D.; FINGER, F.L.; NASCIMENTO, M.; PUIATTI, M.; LACERDA, M.S. Nonlinear models based on quantiles in the fitting of growth curves of pepper genotypes. **Brazilian Journal of Biometrics**, Lavras, v.39, n.3, 2021.
- PINHEIRO, C.J.; BATES, D. **nlme**: linear and nonlinear mixed effects models. R package version 3.1-153. 2021 Disponível em: <https://CRAN.R-project.org/package=nlme>.
- PINHEIRO, C.J.; BATES, D.M. **Mixed-effects models in S and S-PLUS**. New York: Springer-Verlag, 2004. 528 p.
- PRADO, T.K.L.; SAVIAN, T.V.; FERNANDES, T.J.; MUNIZ, J.A. Study on the growth curve of the internal cavity of Dwarf green coconut fruits. **Revista Ciência Agronômica**, Fortaleza, v.51, p.e20154591, 2020.
- R Core Team. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2022. Disponível em: <https://www.R-project.org/>.
- RASEIRA, M.C.B.; FRANZON, R.C. Melhoramento genético e cultivares de amora-preta e mirtilo. **Informe Agropecuário**, Belo Horizonte, v.33, n.268, p.11–20, 2012.
- RIBEIRO, T.D.; MATTOS, R.W.P.; MORAIS, A.R.; MUNIZ, J.A. Description of the growth of pequi fruits by nonlinear models. **Revista Brasileira de Fruticultura**, Jaboticabal, v.40, p.e-949, 2018a.
- RIBEIRO, T.D.; SAVIAN, T.V.; FERNANDES, T.J.; MUNIZ, J.A. The use of the nonlinear models in the growth of pears of 'Shinseiki' cultivar. **Ciência Rural**, Santa Maria, v.48, n.1, p.1-7, 2018b.
- SARI, B.G.; LÚCIO, A.D.; SANTANA, C.S.; OLIVOTO, T.; DIEL, M.I.; KRYSCZUN, D.K. Nonlinear growth models: An alternative to ANOVA in tomato trials evaluation. **European Journal of Agronomy**, Amsterdam, v.104, p.21-36, 2019b.
- SARI, B.G.; LÚCIO, A.D.; SANTANA, C.S.; SAVIAN, T.V. Describing tomato plant production using growth models. **Scientia Horticulturae**, New York, v.246, p.146-54, 2019a.
- SARI, B.G.; OLIVOTO, T.; DIEL, M.I.; KRYSCZUN, D.K.; LUCIO, A.D.C.; SAVIAN, T.V. Nonlinear modeling for analyzing data from multiple harvest crops. **Agronomy Journal**, Madison, v. 110, n. 6, p. 2331-2342, 2018.
- SHAPIRO, S.S.; WILK, M.B. An analysis of variance test for normality (complete samples). **Biometrika**, New York, v.52, n.3/4, p.591-611, 1965.
- SILVA, E.M.; FRUHAUF, A.C.; SILVA, E.M.; MUNIZ, J.A.; FERNANDES, T.J.; SILVA, V.F. Evaluation of the critical points of the most adequate nonlinear model in adjusting growth data of 'green dwarf' coconut fruits. **Revista Brasileira de Fruticultura**, Jaboticabal, v.43, n.1, p.e-726, 2021b.
- SILVA, E.M.; TADEU, M.H.; SILVA, V.F.; PIO, R.; FERNANDES, T.J.; MUNIZ, J.A. Description of blackberry fruit growth by nonlinear regression models. **Revista Brasileira de Fruticultura**, Jaboticabal, v.42, p.e-177, 2020.
- SILVA, W.S.; FERNANDES, F.A.; MUNIZ, F.R.; MUNIZ, J.A.; FERNANDES, T.J. Eucalyptus grandis X Eucalyptus urophylla growth curve in different site classifications, considering residual autocorrelation. **Revista Brasileira de Biometria**, Lavras, v.39, p.122-38, 2021a.
- TADEU, M.H.; SOUZA, F.B.M.; PIO, R.; VALLE, M.H.R.; LOCATELLI, G.; GUIMARÃES, G.F.; SILVA, B.E.C. Poda drástica de verão e produção de cultivares de amoreira-preta em região subtropical. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v.50, n.2, p.132-40, 2015.
- WICKHAM, H. **ggplot2**: elegant graphics for data analysis. New York: Springer-Verlag, 2016.
- WYZYKOWSKI, J.; CUSTÓDIO, A.A.P.; CUSTÓDIO, A.A.P.; GOMES, N.M.; MORAIS, A.R. Análise do diâmetro de copa do cafeeiro recepado utilizando um modelo não linear misto. **Revista Brasileira de Biometria**, São Paulo, v.33, n.3, p.243-56, 2015.

**ARTIGO 3 - DEMONSTRATION OF CRITICAL POINTS IN NONLINEAR
GROWTH MODELS LOGISTIC AND GOMPERTZ, AND PROVISION OF R
STATISTICAL SOFTWARE SCRIPTS**

(VERSÃO PRELIMINAR)

Redigido para futura submissão na Brazilian Journal of Biometrics



ARTICLE

Demonstration of critical points in nonlinear growth models Logistic and Gompertz, and provision of R statistical software scripts.

Édipo Menezes da Silva,^{*1} Nataniele Cristina Chaves,^{*2} Victor Ferreira da Silva,^{*1} Joel Augusto Muniz,^{*1} and Tales Jesus Fernandes^{*1}

¹Department of Statistics, Federal University of Lavras, Lavras, Brazil

²Department of Matematics, Universidade de Uberlândia, Uberlândia, Brazil

*Corresponding author. Email: ediposvm01@gmail.com; natanielechaves@gmail.com; victor.est.ufop@gmail.com; joamuniz@ufla.br; tales.jfernandes@ufla.br

(Received: January 1,2022; Revised: January 1,2022; Accepted: January 1,2022; Published: January 1,2022)

Abstract

This article underscores the importance of nonlinear growth models, particularly those from the Richards family, in depicting the growth characteristics of living organisms over time. Models such as the Logistic and Gompertz are renowned for their ability to generate "S"-shaped curves, offering valuable insights into growth rates, acceleration points, and asymptotes. Critical points, defined as estimates of acceleration and asymptotic points obtained through derivatives from first to fourth order with respect to time, are highlighted as essential tools for detailed growth analysis. While many researchers have employed these models, not all have fully explored critical points. Recent studies have successfully demonstrated the localization and interpretation of these points in simple sigmoid models like the Logistic and suggest that these analyses can be extrapolated to double sigmoid models. The primary objective of this work is to conduct a comprehensive analysis of derivatives and critical points in Logistic and Gompertz models, providing practical tools for estimating and visualizing these points through routines in R. As the main outcome of this study, critical points have been successfully demonstrated analytically for the simple sigmoid and extrapolated to the double sigmoid.

Keywords: Partial Derivatives; Biological Interpretation; Simple Sigmoid; Double Sigmoid.

1. Introduction

Growth curves represent the evolution of growth characteristics of living beings, whether animals or plants, over time. To model this relationship between time and growth characteristics, such as weight, height, and diameter, for example, there are various statistical models available. However,

nonlinear growth models, particularly those from the Richards family, including the Brody model, the Logistic model, the Gompertz model, and the Von Bertalanffy model, are the most widely adopted to describe this relationship. One of the main advantages of these models over others is the ability to interpret their parameters in a practical and straightforward manner (SILVA et al., 2020; FRUHAUF et al., 2022; FERNANDES et al., 2022).

In the literature, nonlinear statistical growth models stand out due to their feature of generating "S"-shaped curves, which provides significant *insights*. Moreover, these models offer valuable insights into the growth rate, acceleration points, and asymptotes, whose values can be interpreted practically and relevantly from a biological perspective (ALLAMAN; JELIHOVSCHI, 2022).

The use of adjusted critical points in simple and double nonlinear models can greatly assist professionals in a deeper understanding of fruit or animal development (DIEL et al., 2020; TEIXEIRA et al., 2021). These points estimated in classical nonlinear models provide researchers with the ability to examine the "acceleration" and "asymptote" of a given growth process, which can be fundamental for a more detailed and enlightening analysis (ALLAMAN; JELIHOVSCHI, 2022).

It is interesting to note that several researchers have satisfactorily used these growth models in their studies. However, it is important to highlight that these authors often have not explored important features that these models offer, such as critical points (MUIANGA et al. 2016, RIBEIRO et al. 2018, FERNANDES et al., 2020, FRUHAUF et al., 2022). Critical points, as defined by Michan et al. (2014), are estimates of the abscissa and ordinate of the models obtained through derivatives from first to fourth order concerning time variation. This approach can be valuable for a more detailed and comprehensive analysis of growth in future studies.

Bem et al. (2020) and Teixeira et al. (2020) conducted studies in which they characterized critical points through graphical representations. They visually demonstrated the location of each critical point in nonlinear models applied to data from sunn hemp and Mangalarga Marchador horse breeds, respectively. On the other hand, Diel et al. (2020) explored the growth of biquinho pepper data using a nonlinear Logistic model and extracted practical interpretations of critical points.

Additionally, Allaman and Jelihovschi (2022) conducted a detailed analysis of critical points in a non-epidemiological Logistic model, identifying the maximum acceleration point, inflection point, maximum deceleration point, and asymptotic acceleration point. Their estimates provided a comprehensive understanding of the model's characteristics. The main purpose of this study was to develop a comprehensive analytical analysis of the first four derivatives of Logistic and Gompertz nonlinear models. This process involved determining the coordinates (abscissas and ordinates) of critical points representing maximum acceleration (PAM), inflection point (PI), maximum deceleration point (PMD), and maximum asymptotic deceleration point (PDA) concerning simple sigmoid nonlinear models.

The research extended to the application of these analyses to Logistic and Gompertz nonlinear models with two growth phases. As part of the study, routines in R were made available for estimating these critical points and creating graphs representing the growth rate, including visual identification of the aforementioned points. In summary, this work aimed primarily for a comprehensive analysis of derivatives and critical points in Logistic and Gompertz models, as well as to provide practical tools for estimating and visualizing these points through routines in R.

2. Materials and Methods

The simple sigmoid models were applied to the dataset from the study conducted by Benassi (2007), which addresses the longitudinal outer diameter growth (LOD) of Green Dwarf Coconut fruits. As for the double sigmoid models, the data used originates from the research conducted by Tadeu et al. (2015), which focused on the study of fruit diameter of Blackberry.

For the determination of critical points in Gompertz and Logistic nonlinear models, the parameterization adopted was based on the methodology proposed by Fernandes et al. (2015). The

methodological approach established in this study served as a basis for the analysis and interpretation of critical points, providing a systematic approach to understanding growth in relation to simple and double sigmoid models.

$$y_i = \beta_1 e^{-e^{\beta_3(\beta_2-t_i)}} + \epsilon_i \quad (1)$$

$$y_i = \frac{\beta_1}{1 + e^{\beta_3(\beta_2-t_i)}} + \epsilon_i \quad (2)$$

$$y_i = \beta_1 e^{-e^{\beta_3(\beta_2-t_i)}} + (\beta_4 - \beta_1) e^{-e^{\beta_6(\beta_5-t_i)}} + \epsilon_i \quad (3)$$

$$y_i = \frac{\beta_1}{1 + e^{\beta_3(\beta_2-t_i)}} + \frac{(\beta_4 - \beta_1)}{1 + e^{\beta_6(\beta_5-t_i)}} + \epsilon_i \quad (4)$$

The models 1, 2, 3, and 4 represent, respectively, the simple Gompertz, simple Logistic, double Gompertz, and double Logistic models. In the mentioned models, t_i denotes the measurement times, y_i is the response at time i , β_1 and β_4 represent the maximum expected growth in each growth curve, β_2 and β_5 the inflection point in the respective curves, β_3 and β_6 are the parameters related to the growth rate in each curve, and ϵ_i is the random error associated with each model, with $\epsilon_i \sim N(0, \sigma^2)$.

In the previously mentioned simple sigmoid models, the first to fourth-order derivatives were calculated to determine the coordinates (abscissas and ordinates) of each of these derivatives. It is relevant to note that the coordinates of the first and fourth-order derivatives, as emphasized by Michan and Pinho (2014), are commonly termed critical points.

Critical points a fundamental role in growth analysis, allowing the identification of crucial characteristics of growth curves, such as inflection points, maximum acceleration, maximum deceleration, among other relevant aspects. These points provide essential information for a more complete understanding of the growth processes modeled by the models.

In the analysis of the first derivative of the model 1 the rules of derivation were applied, and the result can be expressed as follows:

$$\frac{dy_i}{dt} = \beta_3 \beta_1 e^{-e^{\beta_3(\beta_2-t_i)}} e^{\beta_3(\beta_2-t_i)}$$

$$\frac{d^2 y_i}{dt^2} = \beta_1 \beta_3^2 e^{-e^{\beta_3(\beta_2-t_i)}} e^{(\beta_3(\beta_2-t_i))} (\beta_3 e^{\beta_3(\beta_2-t_i)} - \beta_3)$$

With the second derivative, it is possible to find the abscissa and ordinate of the inflection point by setting it to zero. To facilitate calculations, let $W = e^{(\beta_3(\beta_2-t_i))}$, rearranging the second derivative:

$$\beta_1 \beta_3^2 e^{-W + \beta_3(\beta_2-t_i)} (W - 1) = 0$$

Performing distribution and isolation of W , obtain $W = 1$. Reintroducing W into the expression, have $e^{\beta_3(\beta_2-t_i)} = 1$. Applying the natural logarithm to $\ln(e^{\beta_3(\beta_2-t_i)}) = \ln(1)$, resulting in $\beta_3(\beta_2 - t_i) = 0$. Isolating t_i , the abscissa of the inflection point is $t_i = \beta_2$. To calculate the ordinate, use $y = \beta_1 e^{(-W)}$. Since $W = 1$, the ordinate of the inflection point of the model is $y = \frac{\beta_1}{e}$.

Deriving again, lies the third derivative and the points of maximum acceleration (PMA) and maximum deceleration (PMD).

$$\frac{d^3 y_i}{dt^3} = \beta_1 \beta_3^3 (e^{-e^{\beta_3(\beta_2-t_i)+3\beta_3(\beta_2-t_i)}} - 3e^{-e^{\beta_3(\beta_2-t_i)+2\beta_3(\beta_2-t_i)}} + e^{-e^{\beta_3(\beta_2-t_i)+\beta_3(\beta_2-t_i)}})$$

Again substituting W , get:

$$\beta_1 \beta_3^3 e^{-W} (W^3 - 3W^2 + W) = 0$$

Using the same rationale as for the second derivative and applying the *Bhaskara's* formula, the results obtained were: $W_{PAM} = \frac{3+\sqrt{5}}{2}$ and $W_{PDM} = \frac{3-\sqrt{5}}{2}$.

For $W_{PAM} = \frac{3+\sqrt{5}}{2}$,

$$\ln(W_{PAM}) = \ln\left(\frac{3+\sqrt{5}}{2}\right) \rightarrow \ln(e^{\beta_3(\beta_2-t_i)}) = \frac{3+\sqrt{5}}{2} \rightarrow t_{PAM} = \frac{\beta_2\beta_3 - 0.962}{\beta_3}$$

Using the result found for t_{PAM} to find the ordinate:

$$y_{PAM} = \beta_1 e^{-\left(\frac{3+\sqrt{5}}{2}\right)}$$

To find the same values of the abscissa and ordinate for PMD, the above steps are repeated and the following results are found, $t_{PDM} = \frac{\beta_2\beta_3+0.962}{\beta_3}$ and $y_{PDM} = \beta_1 e^{-\left(\frac{3-\sqrt{5}}{2}\right)}$.

Finally, to find the critical points of model [1](#), the fourth derivative was computed using the same rationale as above, and the following results were obtained:

$$\beta_1 \beta_3^4 (e^{-e^{\beta_3(\beta_2-t_i)+4\beta_3(\beta_2-t_i)}} - 6e^{-e^{\beta_3(\beta_2-t_i)+3\beta_3(\beta_2-t_i)}} + 7e^{-e^{\beta_3(\beta_2-t_i)+2\beta_3(\beta_2-t_i)}} - e^{-e^{\beta_3(\beta_2-t_i)+\beta_3(\beta_2-t_i)}}) = 0$$

Finding the values of W and using the value $W = 0.1657$, then:

$$\ln(W) = e^{\beta_3(\beta_2-t_i)} \rightarrow \ln(0.1657) = \beta_3(\beta_2-t_i) \rightarrow t_{PDA} = \frac{\beta_3\beta_2 + 1.7975}{\beta_3}$$

With the value of t_{PAM} , find the value of $y_{PDA} = \frac{\beta_1}{e^{0.1657}}$

Using the same approach adopted in the [1](#) model for the derivatives, the results obtained for the [2](#) model were the following:

$$\frac{dy_i}{dt} = \frac{\beta_1 \beta_3 e^{\beta_3(\beta_2-t_i)}}{(1 + e^{\beta_3(\beta_2-t_i)})^2}$$

$$\frac{dy_i^2}{dt} = \frac{\beta_1 \beta_3^2 e^{\beta_3(\beta_2-t_i)} (e^{\beta_3(\beta_2-t_i)} - 1)}{(1 + e^{\beta_3(\beta_2-t_i)})^3}$$

$$\frac{dy_i^3}{dt} = \frac{\beta_1 \beta_3^3 e^{\beta_3(\beta_2-t_i)} (e^{2\beta_3(\beta_2-t_i)} - 4e^{\beta_3(\beta_2-t_i)} - 1)}{(1 + e^{\beta_3(\beta_2-t_i)})^4}$$

$$\frac{d^4 y_i}{dt} = \frac{\beta_1 \beta_3^4 e^{\beta_3(\beta_2 - t_i)} (e^{\beta_3(\beta_2 - t_i)} + 11e^{\beta_3(\beta_2 - t_i)} - 11e^{2\beta_3(\beta_2 - t_i)} - 1)}{(1 + e^{\beta_3(\beta_2 - t_i)})^5}$$

The results of the first four derivatives of model 2 indicate that, when considering $W = e^{\beta_3(\beta_2 - t_i)}$, the abscissa and ordinate for the second-order derivative are:

$$\frac{d^2 y_i}{dt} = \frac{\beta_1 \beta_3^2 W(W - 1)}{1 + e^{\beta_3(\beta_2 - t_i)}} = 0 \rightarrow W(W - 1) \rightarrow W = 1 \rightarrow e^{\beta_3(\beta_2 - t_i)} = 1 \rightarrow \beta_3(\beta_2 - t_i) = \ln(1) \rightarrow t_{PI} = \beta_2$$

Com os resultados da terceira derivada,

$$\frac{d^3 y_i}{dt} = \frac{\beta_1 \beta_3^3 W(W^2 - 4W + 1)}{(1 + e^{\beta_3(\beta_2 - t_i)})^2} = 0 \rightarrow W^2 - 4W + 1 = 0$$

The result of the quadratic equation above yields the values of W as 3.73 and 0.27. Using the value of $W=3.73$, $\rightarrow e^{\beta_3(\beta_2 - t_i)} = 3.73 \rightarrow \beta_3(\beta_2 - t_i) = \ln(3.73) \rightarrow t_{PAM} = \frac{\beta_2 \beta_3 - 1.319}{\beta_3}$. For $W = 0.27$ $\rightarrow e^{\beta_3(\beta_2 - t_i)} = 0.27 \rightarrow \beta_3(\beta_2 - t_i) = \ln(0.27) \rightarrow t_{PDM} = \frac{\beta_2 \beta_3 + 1.319}{\beta_3}$.

$$\frac{d^4 y_i}{dt} = \frac{\beta_1 \beta_3^4 W(W^3 + 11W - 11W^2 + 1)}{(1 + e^{\beta_3(\beta_2 - t_i)})} = 0 \rightarrow W^3 + 11W - 11W^2 + 1 = 0$$

As it is a third-degree equation, three values were found, but the only one of interest is the $W = 0.101 \rightarrow e^{\beta_3(\beta_2 - t_i)} = 1 \rightarrow \beta_3(\beta_2 - t_i) = \ln(1) \rightarrow t_{PDA} = \frac{\beta_2 \beta_3 + 2.2924}{\beta_3}$

With the values of the abscissas of the critical points, the ordinates are a consequence of these values:

$$y_{PI} = \frac{\beta_1}{1 + W} \rightarrow \frac{\beta_1}{1 + 1} \rightarrow \frac{\beta_1}{2}$$

$$y_{PAM} = \frac{\beta_1}{1 + W} \rightarrow \frac{\beta_1}{1 + 3.73} \rightarrow \frac{\beta_1}{3 + \sqrt{3}}$$

$$y_{PDM} = \frac{\beta_1}{1 + W} \rightarrow \frac{\beta_1}{1 - 3.73} \rightarrow \frac{\beta_1}{3 - \sqrt{3}}$$

$$y_{PDA} = \frac{\beta_1}{1 + W} \rightarrow \frac{\beta_1}{1 + 0.101} \rightarrow \frac{\beta_1}{1.101}$$

With the above results, Table 1 can be constructed, as shown by Silva et al. (2021).

Table 1. Expressions to find abscissas and ordinates of the maximum acceleration point (PAM), point of inflection (PI), maximum deceleration point (PDM), and asymptotic deceleration point (PDA) for Logistic and Gompertz models using the parameterization of this work from the first sigmoid.

Model		PAM	PI	PDM	PDA
Logistic	Abscissa	$\frac{\beta_2\beta_3-1.319}{\beta_3}$	β_2	$\frac{\beta_2\beta_3+1.319}{\beta_3}$	$\frac{\beta_2\beta_3+2.2924}{\beta_3}$
	Ordinate	$\frac{\beta_1}{3+\sqrt{3}}$	$\frac{\beta_1}{2}$	$\frac{\beta_1}{3-\sqrt{3}}$	$\frac{\beta_1}{1.101}$
Gompertz	Abscissa	$\frac{\beta_2\beta_3-0.962}{\beta_3}$	β_2	$\frac{\beta_2\beta_3+0.962}{\beta_3}$	$\frac{\beta_2\beta_3+1.7975}{\beta_3}$
	Ordinate	$\beta_1 e^{-\left(\frac{3+\sqrt{3}}{2}\right)}$	$\frac{\beta_1}{e}$	$\beta_1 e^{-\left(\frac{3-\sqrt{3}}{2}\right)}$	$\frac{\beta_1}{e^{1.1657}}$

The aforementioned results can be applied to estimate the critical points in Table 2, considering the nature of the double sigmoid models (models 3 and 4). Due to the structure of these models, which consists of the sum of two simple models, it is possible to use the critical points from Table 1. In this context, the parameters β_1 , β_2 , and β_3 can be replaced by the respective parameters β_4 , β_5 , and β_6 , as presented in Table 2.

Table 2. Expressions to find abscissas and ordinates of the maximum acceleration point (PAM), point of inflection (PI), maximum deceleration point (PDM), and asymptotic deceleration point (PDA) for Logistic and Gompertz models using the parameterization of this work from the second sigmoid.

Model		PAM	PI	PDM	PDA
Logistic	Abscissa	$\frac{\beta_5\beta_6-1.319}{\beta_6}$	β_5	$\frac{\beta_5\beta_6+1.319}{\beta_6}$	$\frac{\beta_5\beta_6+2.2924}{\beta_6}$
	Ordinate	$\frac{\beta_4}{3+\sqrt{3}}$	$\frac{\beta_4}{2}$	$\frac{\beta_4}{3-\sqrt{3}}$	$\frac{\beta_4}{1.101}$
Gompertz	Abscissa	$\frac{\beta_5\beta_6-0.962}{\beta_6}$	β_5	$\frac{\beta_5\beta_6+0.962}{\beta_6}$	$\frac{\beta_5\beta_6+1.7975}{\beta_6}$
	Ordinate	$\beta_4 e^{-\left(\frac{3+\sqrt{3}}{2}\right)}$	$\frac{\beta_4}{e}$	$\beta_4 e^{-\left(\frac{3-\sqrt{3}}{2}\right)}$	$\frac{\beta_4}{e^{1.1657}}$

To ensure the validity of the inferences derived from the models and the effectiveness of the calculated critical points, it is crucial that the regression models satisfy certain prerequisites, including normality, homoscedasticity, and independence, in addition to the significance of the investigated parameters. To assess the assumptions of the residuals vector, the following tests were employed: Shapiro-Wilk normality test (1965), Breusch-Pagan homoscedasticity test (1979), Durbin-Watson independence test (1951), and significance through the Student's t-test. In the hypothesis test, rejecting the hypothesis that the test value is different from zero indicates the significance of the parameter. To evaluate the quality of the fits, the coefficient of determination (R^2) was used.

In addition to estimating the critical points in the two growth phases, a graph was generated with the growth rate (first derivative of the models), using the abscissas and ordinates of the critical points. All analyses were conducted in the statistical software R (R CORE TEAM, 2022) with the assistance of the nlme, qpcR, car, and lmtest packages. The significance level adopted for the proposed hypothesis tests was 1%.

3. Results and Discussion

The initial fittings were performed on the models, and the residual analysis is presented in Table 3. As observed in this table, for the results of the single sigmoid models, a violation of the assumption of independence of residuals was identified. In contrast, in the case of the double sigmoid models, no violations of the assumptions related to the residual vector were observed. As done in previous works, such as Silva et al. (2020) and Cassiano and Sáfiadi (2015), in models where a violation of independence was identified, this dependence was modeled by including the autoregressive term of order 1 (AR1).

Table 3. P-value of Shapiro-Wilk (SW), Breusch-Pagan (BP), and Durbin-Watson (DW) tests for the Logistic, Gompertz, Logistic + Logistic, and Gompertz + Gompertz models.

Fruit	Variable	Model	SW	BP	DW
Coconut	DEL	Logistic	0.6102	0.0245	0.0045
Coconut	DEL	Gompertz	0.2598	0.0120	0.0000
Blackberry	Diameter	Double Logistic	0.0670	0.066	0.7937
Blackberry	Diameter	Double Gompertz	0.3654	0.2541	0.1112

The aim of this work was to analytically construct critical points and provide an R routine for generating graphs, without addressing the impacts of residual analysis of the fitted models. Emphasis was placed on ensuring that these assumptions were met, aiming to ensure reliable results in parameter estimates and, consequently, in critical points.

In addition to residual analysis, another crucial step before making inferences about model parameters was to check the significance of these parameters (Silva et al., 2020; Fruhauf et al., 2022). When analyzing the model parameters, it was found that all of them were statistically significant. Finally, before constructing graphs with critical points and presenting R routines, the coefficient of determination (R^2) was analyzed, with all values being above 95%.

From the results of the model estimates, as presented in Table 4, it was possible to estimate critical points and construct the growth rate curve, identifying the locations of each point. For each model related to each fruit and characteristic, the corresponding graphs were generated, and the R routines were made available.

Table 4. Parameter estimates for the Logistic, Gompertz, Logistic + Logistic, and Gompertz + Gompertz models applied to coconut and blackberry growth data.

Fruit	Variable	Model	β_1	β_2	β_3	β_4	β_5	β_6
Coconut	DEL	Logistic	21.40	152.24	0.01	---	---	---
Coconut	DEL	Gompertz	23.61	133.04	0.00834	---	---	---
Blackberry	Diameter	Double Logistic	11.71	5.72	0.18	14.47	28.50	0.82
Blackberry	Diameter	Double Gompertz	12.16	3.23	0.13	14.58	28.14	1.09

```
###DOUBLE GOMPERTZ GROWTH RATE ADJUSTMENT#####
windows()
par(mfrow=c(1,2))
plot(DAAI,DEL,pch=16, xlab="Days after inflorescence opening",
ylab="DEL (mm)",ylim=c(0,21),xlim=c(3,350),axes=F)
lines(DAAI,fitted(log_del1),col="gray")
axis(1,at=c(0,50,100,150,200,250,300,350,400,450), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,3,6,9,12,15,20,25), cex.axis=1,las=1,outer=F)
```

```
###LOGISTI GROWTH RATE ADJUSTMENT#####
dy=function(x,b1,b2,b3){
(b1*b3*exp(b3*(b2-x)))/((1+exp(b3*(b2-x)))^2)
}
```

8 *Brazilian Journal of Biometrics*

```

b1=coef(n0)[1];b1
b2=coef(n0)[2];b2
b3=coef(n0)[3];b3

# Plotting the first derivative curve

curve(dy(x,b1=b1,
b2=b2,
b3=b3),
xlim=range(0,350),ylim=c(0,0.1),
xlab="Days after inflorescence opening",
ylab="Acceleration curve",main="",axes=F)
axis(1,at=c(0,50,100,150,200,250,300,350,400,450), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,0.03,0.06,0.09,0.15), cex.axis=1,las=1,outer=F)
legend(0,0.018,c("Logistic","pi","pam","pdm","pda"),lty=c(1,NA,NA,NA,NA),
pch=c(NA,19,15,16,17),col=c("gray","green","yellow","blue","orange"),cex=0.75)

#INFLECTION POINT
points(b2,dy(b2,
b1=b1,
b2=b2,
b3=b3),pch=19,col="green")

pam = (b3*b2-1.319)/b3
pam=abs(pam)
#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3),pch=15,col="yellow")
pdm = (b3*b2+1.319)/b3
#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,
b1=b1,
b2=b2,
b3=b3),pch=16,col="blue")

pda = ((b3*b2)+2.2924)/b3
#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,
b3=b3),pch=17,col="pink")

###GOMPERTZ GROWTH RATE ADJUSTMENT#####

```

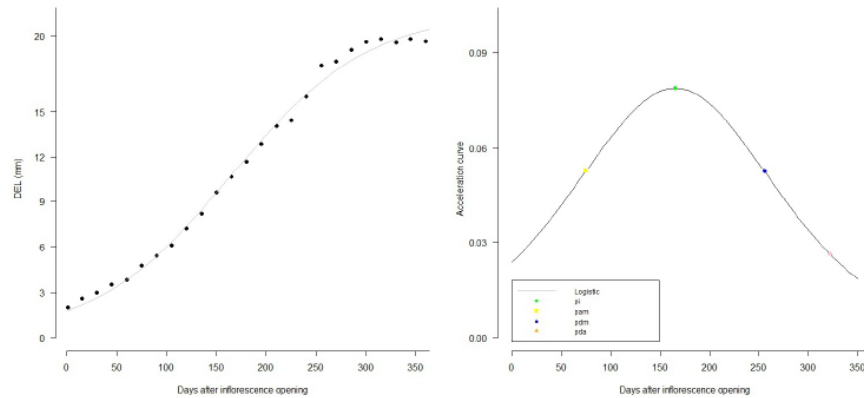


Figure 1. Graphs of the adjusted data and growth rate of coconut with respect to the Logistic model, with indication of critical points.

```

windows()
par(mfrow=c(1,2))
plot(DAAI,DEL,pch=16, xlab="Days after inflorescence opening",
ylab="DEL (mm)",ylim=c(0,21),xlim=c(3,400),axes=F)
lines(DAAI,fitted(n0),col="gray")
axis(1,at=c(0,50,100,150,200,250,300,350,400,450), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,3,6,9,12,15,20,25), cex.axis=1,las=1,outer=F)

```

```

###DOUBLE GOMPERTZ GROWTH RATE ADJUSTMENT#####

```

```

dy = function(x,b1,b2,b3,b4,b5,b6){
b1*b3*exp(-exp(b3*(b2-x)))+b3*(b2-x)

```

```

}

```

```

b1=coef(n0)[1];b1
b2=coef(n0)[2];b2
b3=coef(n0)[3];b3

```

```

curve(dy(x,b1=b1,
b2=b2,
b3=b3),
xlim=range(0,400),ylim=c(0,0.1),
xlab="Days after inflorescence opening",
ylab="Acceleration curve",main="",axes=F)
axis(1,at=c(0,50,100,150,200,250,300,350,400), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,0.03,0.06,0.09,0.15), cex.axis=1,las=1,outer=F)
legend(0,0.018,c("Gompertz","pi","pam","pdm","pda"),lty=c(1,NA,NA,NA,NA),

```

10 *Brazilian Journal of Biometrics*

```
pch=c(NA,19,15,16,17),col=c("gray","green","yellow","blue","orange"),cex=0.75)

#INFLECTION POINT
points(b2,dy(b2,
b1=b1,
b2=b2,
b3=b3),pch=19,col="green")

pam = (b3*b2-0.962)/b3
pam
pam_y = b1*exp(-((3+sqrt(5))/2))

#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3),pch=15,col="yellow")
pdm = (b3*b2+0.962)/b3
#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,
b1=b1,
b2=b2,
b3=b3),pch=16,col="blue")

pda = ((b3*b2)+1.7975)/b3

#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,
b3=b3),pch=17,col="pink")

###DOUBLE LOGISTIC GROWTH RATE ADJUSTMENT#####
windows()
par(mfrow=c(1,2))

plot(tempo1,comp1,pch=16, xlab="Days after anthesis", ylab="Length (mm)",ylim=c(0,20),:
lines(tempo1,fitted(n0),col="gray")
mtext("", side = 4)
axis(1,at=c(0,5,10,
15,20,25,30,35,40,45), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,3,6,9,12,15,18), cex.axis=1,las=1,outer=F)
```

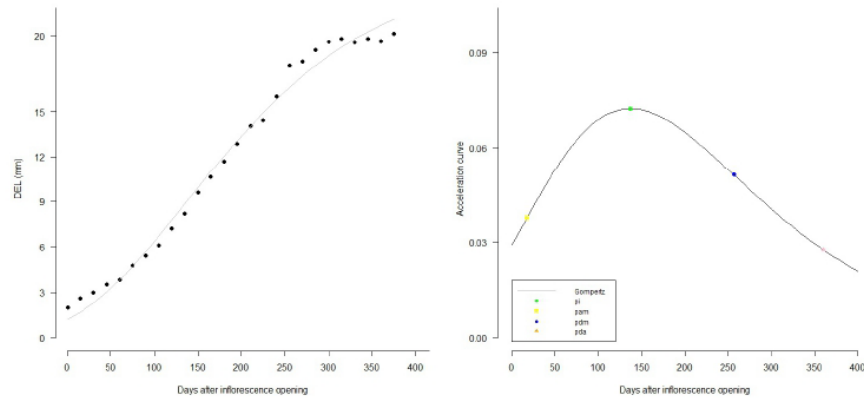


Figure 2. Graphs of the fitted data and growth rate of coconut with respect to the Gompertz model, with indication of critical points.

```
dy = function(x,b1,b2,b3,b4,b5,b6){
  b1*b3*exp(b3*(b2-x))/(1+exp(b3*(b2-x)))^2 + (b4-b1)*b6*exp(b6*(b5-x))/(1+exp(b6*(b5-
```

```
})

b1=coef(n0)[1];b1
b2=coef(n0)[2];b2
b3=coef(n0)[3];b3
b4=coef(n0)[4];b4
b5=coef(n0)[5];b5
b6=coef(n0)[6];b6
```

```
curve(dy(x,b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),
xlim=c(0,40),ylim=c(0,1.2),
xlab="",
ylab="Days after anthesis",
ylab="Acceleration curve",axes=F)
legend(0,0.8,c("Logistic","pi","pam","pdm","pda"),lty=c(1,NA,NA,NA,NA),
pch=c(NA,19,15,16,17),col=c("green","green","yellow","blue","orange"),cex=0.75)
mtext("", side = 4)
axis(1,at=c(0,5,10,15,20,25,30,35,40,45), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,.2,.4,0.6,.8,1.0,1.2), cex.axis=1,las=1,outer=F)
#INFLECTION POINT
points(b2,dy(b2,
b1=b1,
```

12 *Brazilian Journal of Biometrics*

```
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=19,col="green")
```

```
pam = (b3*b2-1.319)/b3
pam=abs(pam)
#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=15,col="yellow")
pdm = (b3*b2+1.319)/b3
#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=16,col="blue")
```

```
pda = ((b3*b2)+2.2924)/b3
#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=17,col="pink")
```

```
#CRITICAL POINTS OF THE SECOND PHASE OF GROWTH
#INFLECTION POINT
points(b5,dy(b5,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=19,col="green")
```

```
pam = (b6*b5-1.319)/b6
```

```
pam=abs(pam)
#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=15,col="yellow")
pdm = (b6*b5+1.319)/b6
#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=16,col="blue")

pda = ((b6*b5)+2.2924)/b6
#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=17,col="pink")
```

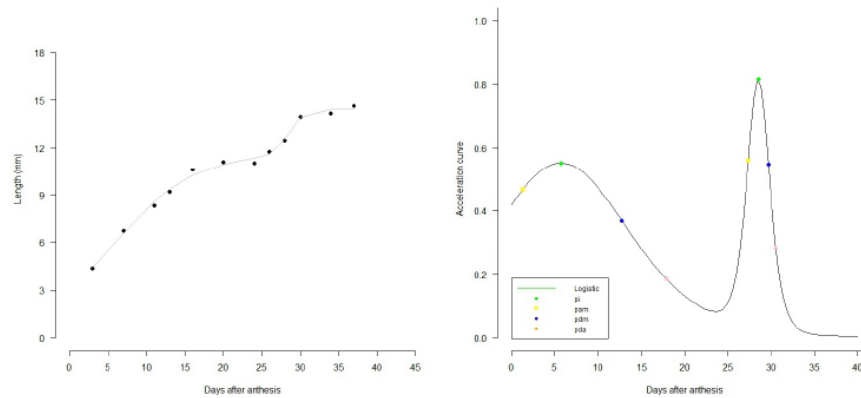


Figure 3. Graphs of the adjusted data and growth rate of blackberry with respect to the Logistic+Logistic model, with indication of critical points for both growth phases.

```

####DOUBLE GOMPERTZ GROWTH RATE ADJUSTMENT#####
windows()
par(mfrow=c(1,2))
plot(tempo1,comp1,pch=16, xlab="Days after anthesis",
ylab="Length (mm)",ylim=c(3,15),xlim=c(0,38),axes=F)
lines(tempo1,fitted(n0),col="gray")
legend(220,5,c("IC","Gompertz","Productionrate","pam","pdm","pda"),
lty=c(2,1,1,NA,NA,NA),pch=c(NA,NA,19,15,16,17),
col=c("red","green","black","yellow","blue","orange"),cex=0.75)
mtext("", side = 4)
axis(1,at=c(0,5,10,15,20,25,30,35,40,45), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,3,6,9,12,15), cex.axis=1,las=1,outer=F)

####DOUBLE GOMPERTZ GROWTH RATE ADJUSTMENT#####
dy = function(x,b1,b2,b3,b4,b5,b6){
b1*b3*exp(-exp(b3*(b2-x))+b3*(b2-x)) + (b4-b1)*b6*exp(-exp(b6*(b5-x))+b6*(b5-x))
}

b1=coef(n0)[1];b1
b2=coef(n0)[2];b2
b3=coef(n0)[3];b3
b4=coef(n0)[4];b4
b5=coef(n0)[5];b5
b6=coef(n0)[6];b6
# Plotting the first derivative curve
curve(dy(x,b1=

```

```

b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),
xlim=range(0,45),ylim=c(0,1.28),
xlab="Days after anthesis",
ylab="Acceleration curve",main="",axes=F)
axis(1,at=c(0,5,10,15,20,25,30,35,40,45), cex.axis=1,las=1,outer=F)
axis(2,at=c(0,.2,.4,0.6,.8,
1.0,1.2), cex.axis=1,las=1,outer=F)

legend(0,0.9,c("Gompertz","pi","pam","pdm","pda"),lty=c(1,NA,NA,NA,NA),
pch=c(NA,19,15,16,17),col=c("gray","green","yellow","blue","orange"),cex=0.75)

#INFLECTION POINT
points(b2,dy(b2,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=19,col="green")

pam = (b3*b2-0.962)/b3
pam =
pam_y = b1*exp(-((3+sqrt(5))/2))

#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=15,col="yellow")
pdm = (b3*b2+0.962)/b3
#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=16,col="blue")

```

16 *Brazilian Journal of Biometrics*

```

pda = ((b3*b2)+1.7975)/b3

#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=17,col="pink")

#CRITICAL POINTS OF THE SECOND PHASE OF GROWTH
#INFLECTION POINT
points(b5,dy(b5,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=19,col="green")

pam = (b6*b5-0.962)/b6

#POINT OF MAXIMUM ACCELERATION
points(pam,dy(pam,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=15,col="yellow")
pdm = (b6*b5+0.962)/b6

#POINT OF MAXIMUM DECELERATION
points(pdm,dy(pdm,
b1=b1,
b2=b2,
b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=16,col="blue")

pda = ((b6*b5)+1.7975)/b6
#ASYMPTOTIC DECELERATION POINT
points(pda,dy(pda,
b1=b1,
b2=b2,

```

```

b3=b3,
b4=b4,
b5=b5,
b6=b6),pch=17,col="pink")

```

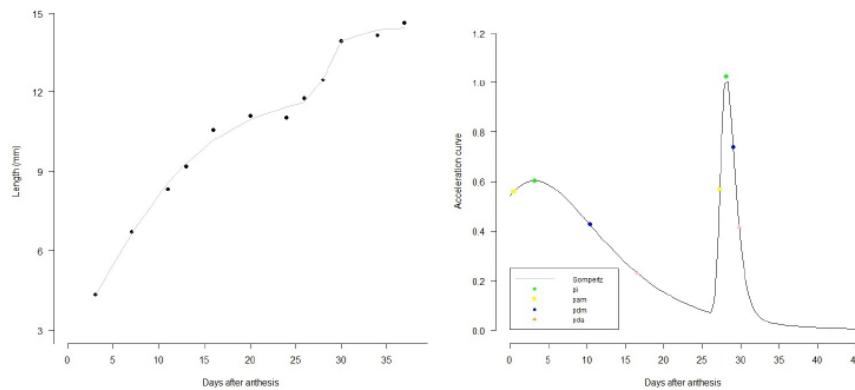


Figure 4. Graphs of the fitted data and growth rate of blackberry with respect to the Gompertz+Gompertz model, with indication of critical points for both growth phases.

In Figures 1, 2, 3, and 4, the graphs of raw data, fitted curves, and their respective growth rates are depicted. When analyzing the results of critical points, Michan and Pinho (2014) emphasize that the inflection point occurs for the Gompertz model at approximately 1/3 of the total growth and for the Logistic model at half of the growth. This behavior was also observed in estimating the critical points in this study, as for the coconut data, these points were 133.04 and 152.24 DAAI, respectively, for the Gompertz and Logistic models.

The analysis of blackberry data regarding the double sigmoid revealed that the Gompertz + Gompertz model presented inflection points (PI) at 3.23 and 28.14 DAA. The first value corresponded to the first phase of growth, while the second referred to the second phase of growth. It was observed that these points occurred at approximately 1/3 of the total expected growth in both phases. For the Logistic + Logistic model, representative values of the midpoint of each development cycle were identified as 5.72 and 28.50 DAA, representing the first and second PI, respectively.

Another relevant critical point with biological interpretation was the asymptotic deceleration point (pda). The formula to estimate the pda for the Gompertz model is $\frac{\beta_1}{e^{0.1657}}$, where the value of $e^{0.1657}$ is equal to 1.18, and for the Logistic model, it is $\frac{\beta_1}{1.101}$. The pda of the Gompertz model in each development cycle estimates 84% of the expected growth, while the Logistic model estimates 90%, corroborating the results of Michan and Pinho (2014).

Overall, estimating critical points represents an advantage when using nonlinear growth models because, according to Allan and Elihovschi (2022), these critical points may have biological interpretation. Teixeira et al. (2021) used the pda to estimate the equilibrium point and highlighted that this point is extremely important as it is closer to the asymptote, considered the stabilization point

3.1 References

- ALLAMAN, I. B.; JELIHOVSKI, E. G. Estimation of the critical points of an epidemic by means of a logistic growth model. **Brazilian Journal of Biometrics**, Lavras, v. 40, n. 2, 2022.
- BEM, C. M. et al. Univariate and multivariate nonlinear models in productive traits of the sunn hemp. **Revista Ciência Agronômica**, Fortaleza, v. 51, 2020.
- BENASSI, A. C. et al. Caracterização biométrica de frutos de coqueiro, *Vocos nucifera* L. variedade anã-verde, em diferentes estádios de desenvolvimento. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 29, p. 302-307, 2007.
- BREUSCH, T.S.; PAGAN, A.R. A simple test for heterocedasticity and random coeficiente variation. **Econometrica**, New York, v.47, n.5, 1287-1294, 1979.
- CASSIANO, F. R.; SÁFADI, T. Modelos de crescimento animal para tempos irregulares. **Pesquisa Agropecuária Brasileira**, Brasília, v. 50, p. 1114-1119, 2015.
- DIEL, M. I. et al. Production of biquinho pepper in different growing seasons characterized by the logistic model and its critical points. **Ciência Rural**, Santa Maria, v. 50, 2020.
- DURBIN, J.; WATSON, G.S. Testing for serial correlation in last squares regression. **Biometrika**, London, v.38, n.1/2, p.159-177, 1951.
- FERNANDES, T. J. et al. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum Technology**, Maringá, v. 37, n. 4, p. 397-402, 2015.
- FERNANDES, F A. et al. Parameterizations of the von Bertalanffy model for description of growth curves. **Brazilian Journal of Biometrics**, Lavras, v. 38, n. 3, p. 369-384, 2020.
- FERNANDES, J. G. et al. Description of the peach fruit growth curve by diphasic sigmoidal non-linear models. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 44, 2022.
- FRUHAUF, A. C. et al. Description of Height Growth of Hybrid Eucalyptus Clones in Semi- Arid Region Using Non-Linear Models. **Brazilian Journal of Biometrics**, Lavras, v. 40, p. 138-151, 2022.
- FRUHAUF, A. C. et al. Predicting height growth in bean plants using non-linear and polynomial models. **Revista Agrogeoambiental**, Pouso Alegre, v. 13, n. 3, p. 488-497, 2022.
- MUIANGA, C. A. et al. Descrição da curva de crescimento de frutos do cajueiro por modelos não lineares. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 38, p. 22-32, 2016.
- MISCHAN, M. M.; PINHO, S. Z. **Modelos não lineares: funções assintóticas de crescimento**. São Paulo: Cultura Acadêmica, 2014.
- MISCHAN, M. M. et al. Inflection and stability points of diphasic logistic analysis of growth. **Scientia Agricola**, v. 72, p. 215-220, 2015.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation

for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

RIBEIRO, T. D. et al. Description of the growth of pequi fruits by nonlinear models. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 40, p.e-949, 2018.

SILVA, E. M. da et al. Description of blackberry fruit growth by nonlinear regression models. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 42, p.e-177, 2020.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, Biometrika Trust, v. 52, n. 3-4, p. 591-611, 1965.

TADEU, M. H. et al. Poda drástica de verão e produção de cultivares de amoreira-preta em região subtropical. **Pesquisa Agropecuária Brasileira**, Brasília, v. 50, n. 2, p. 132-140, 2015.

TEIXEIRA, G. L. et al. Growth curves of campolina horses using nonlinear models. **Livestock Science**, v. 251, p. 104631, 2021.

4. Conclusions

The estimation of critical points was satisfactorily performed for the simple sigmoid and extrapolated to the double sigmoid in both Logistic and Gompertz models, considering the data analyzed in this study. Providing R scripts can facilitate researchers in estimating these points more efficiently and accurately.

Acknowledgments

The authors would like to thank the funding agencies CAPES and CNPq for their financial support for this research.

Conflicts of Interest

The authors declare no conflict of interest.

TERCEIRA PARTE

1 CONSIDERAÇÕES FINAIS

As técnicas propostas, nesta tese: pontos críticos e modelos mistos, apresentaram as seguintes conclusões:

O primeiro artigo utilizou os resultados analíticos dos pontos críticos de forma satisfatória. Ao se analisar os modelos propostos encontraram-se estimativas de pontos críticos que podem ser utilizados para escolher o melhor momento da colheita dos frutos, garantindo assim, maior validade ao fruto.

Os modelos não lineares efeitos mistos foram avaliados no artigo 2. Houve resultados satisfatórios, com os melhores ajustes para o modelo Logístico e duplo Logístico, com a inclusão do efeito aleatório aos parâmetros β_1 e β_2 para o simples e β_1 e β_4 para o duplo.

De forma analítica, foram encontradas as estimativas dos pontos críticos dos modelos sigmoide simples e generalizados esses resultados, para os modelos duplos, porém, vale salientar que, para que essa generalização seja feita, é necessário ter cautela, pois não foi possível demonstrar os seus resultados, somente a utilização nos dados do segundo artigo desta tese.

Para trabalhos futuros ficam possibilidades de demonstrar os pontos críticos da segunda fase do crescimento analiticamente ou via simulação e a inclusão de efeitos aleatórios nesses pontos.