



Natália Albieri Koritiaki

**Crescimento de tilápias-do-Nilo considerando
diferentes curvas, estruturas de variância e
distribuições para os erros: Uma abordagem
Bayesiana**

Maringá - Paraná
Dezembro de 2025

Natália Albieri Koritiaki

Crescimento de tilápias-do-Nilo considerando diferentes curvas, estruturas de variância e distribuições para os erros: Uma abordagem Bayesiana

Dissertação apresentada ao Programa de Pós-graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá como requisito parcial à obtenção do título de Mestre em Bioestatística.

Universidade Estadual de Maringá
Programa de Pós-Graduação em Bioestatística

Orientador: Prof. Dr. Robson Marcelo Rossi

Maringá - Paraná
Dezembro de 2025

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá - PR, Brasil)

K84

Koritiaki, Natália Albieri

Crescimento de tilápias-do-Nilo considerando diferentes curvas, estruturas de variância e distribuição para os erros : uma abordagem Bayesiana / Natália Albieri Koritiaki. -- Maringá, PR, 2026.

59 f. : il. color., figs., tabs.

Orientador: Prof. Dr. Robson Marcelo Rossi.

Dissertação (mestrado) - Universidade Estadual de Maringá, Centro de Ciências Exatas, Departamento de Estatística, Programa de Pós-Graduação em Bioestatística, 2026.

1. Bioestatística. 2. Modelos não lineares. 3. Curvas de crescimento. 4. Stan. 5. Distribuições assimétricas. I. Rossi, Robson Marcelo, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Estatística. Programa de Pós-Graduação em Bioestatística. III. Título.

CDD 23.ed. 519.5

NATÁLIA ALBIERI KORITIAKI

**Crescimento de tilápias-do-Nilo considerando diferentes curvas, estruturas de
variância e distribuições para os erros: Uma abordagem Bayesiana**

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

BANCA EXAMINADORA

Documento assinado digitalmente
ROBSON MARCELO ROSSI
Data: 05/01/2026 14:52:52-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Robson Marcelo Rossi
Universidade Estadual de Maringá – PBE/UEM

Documento assinado digitalmente
VANDERLY JANEIRO
Data: 05/01/2026 07:11:16-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Vanderly Janeiro
Universidade Estadual de Maringá – PBE/UEM

Documento assinado digitalmente
CARLOS ANTONIO ZARZAR
Data: 26/12/2025 10:12:45-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Carlos Antônio Zarzar
Universidade Federal do Oeste do Pará - UFOPA

Maringá, 17 de dezembro de 2025.

Agradecimentos

A Deus,

Ao meu marido, aos meus filhos e meus pais,

Ao meu orientador Prof. Dr. Robson M. Rossi,

Aos membros da banca de qualificação e defesa que aceitaram contribuir para a melhoria desse trabalho,

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo auxílio financeiro,

A todos que direta ou indiretamente contribuíram de maneira acadêmica, profissional ou pessoal para a realização deste trabalho, obrigada!

Resumo

A piscicultura é uma atividade de relevância no Brasil, com a tilápia-do-Nilo (*Oreochromis niloticus*) respondendo por grande parte da produção nacional. A precisão na modelagem do crescimento somático dessa espécie é importante para otimizar o manejo zootécnico, prever o ponto ótimo de abate e orientar programas de melhoramento genético. Contudo, a aplicação de Modelos de Crescimento Não Lineares tradicionais (como Brody, Gompertz, Logístico e Von Bertalanffy) baseia-se em pressupostos clássicos de homocedasticidade (variância do erro constante) e simetria (distribuição Normal para os resíduos). Em dados longitudinais reais de crescimento biológico, observa-se frequentemente uma violação dessas premissas, manifestada pela heterocedasticidade (usualmente a variabilidade residual aumenta com a idade e o peso) e pela presença de caudas pesadas ou assimetria nos resíduos. A não consideração dessas características pode levar a estimativas de parâmetros de crescimento viesadas, como o peso assintótico e a taxa de crescimento, comprometendo a robustez das inferências biológicas. A estrutura de crescimento da tilápia, sujeita a intensas variações, sugere que o crescimento não segue um padrão de taxa constante. Nesse contexto, o modelo de Von Bertalanffy Segmentado (VBS) permite que a taxa de crescimento mude em um ponto de transição não fixo, capturando fases distintas de aceleração e desaceleração. Para lidar com as estruturas de erro, o presente trabalho utilizou distribuições adicionais à Normal de erro robustas: distribuições simétricas flexíveis, *t*-Student, capaz de modelar caudas pesadas (curtose) por meio do parâmetro de graus de liberdade; e distribuições assimétricas: Skew-Normal e Skew-*t*, que incluem um parâmetro adicional para modelar a assimetria. A Heterocedasticidade foi explicitamente modelada, definindo a variância residual (σ_i^2) como uma função da idade (t), através das seguintes estruturas multiplicativas e aditivas, controladas pelo parâmetro de heterocedasticidade (ρ): potência multiplicativa: $\sigma_i^2 = \sigma^2 t^\rho$, exponencial multiplicativa: $\sigma_i^2 = \sigma^2 \exp(\rho t)$, e aditiva (Linear): $\sigma_i^2 = \sigma^2 + \rho t$. O estudo utilizou dados longitudinais de biometrias de tilápia-do-Nilo (peso) de machos e fêmeas, pertencentes a um programa de melhoramento genético da Universidade Estadual de Maringá. A análise exploratória revelou um dimorfismo sexual acentuado e um aumento na dispersão do peso com a idade, justificando a separação por sexo e a modelagem da heterocedasticidade. Foi realizado o ajuste dos modelos clássicos e do modelo VBS, combinados com as quatro distribuições de erro (Normal, *t*-Student, Skew-Normal, Skew-*t*) em versões homocedásticas. O modelo de melhor ajuste foi então testado com as três estruturas heterocedásticas. A estimação foi realizada via Inferência Bayesiana no ambiente computacional *Stan*, utilizando o algoritmo MCMC (NUTS/HMC). Foram utilizadas distribuições *a priori* fracamente informativas. O desempenho preditivo dos modelos foi avaliado pelo Critério de Informação LOO-CV (LOOIC), que penaliza modelos complexos pela complexidade efetiva e estima a capacidade preditiva fora da amostra, sendo o modelo com o menor LOOIC o preferido. O modelo VBS apresentou o melhor desempenho preditivo em todas as comparações

(Menor LOOIC), demonstrando ser o mais adequado para representar as fases de transição do crescimento da tilápia. Dentre as distribuições homocedásticas, a distribuição t -Student foi a que resultou no menor LOOIC (Machos: 1.407; Fêmeas: 2.007) para o VBS, superando as distribuições Normal e Skew. O valor baixo do parâmetro de graus de liberdade ($\nu \approx 1,7$ a $1,9$) confirmou a presença de caudas pesadas nos resíduos, conferindo maior robustez ao modelo frente a observações discrepantes. O ganho preditivo do t -Student sobre o Skew- t sugeriu que a curtose é a principal violação da Normalidade, sendo a assimetria menos relevante para o ajuste preditivo. A inspeção visual dos resíduos e os resultados da modelagem de variância confirmaram a heterocedasticidade nos dados. O modelo potência multiplicativa ($\sigma_i^2 = \sigma^2 t^\rho$) apresentou o melhor ajuste preditivo entre as estruturas testadas. O parâmetro ρ para esta estrutura foi significativamente diferente de zero e próximo de 2,0 (machos $\rho \approx 1,85$; fêmeas $\rho \approx 1,78$). Este resultado indica uma heterocedasticidade forte. Esse comportamento reflete o padrão observado nos dados disponíveis, nos quais o peso tende a aumentar com a idade; contudo, ressalta-se que as observações não abrangem todo o período de crescimento dos peixes, restringindo a interpretação desse padrão ao intervalo etário amostrado. O modelo final, Von Bertalanffy Segmentado com erro t -Student e variância heterocedástica potência multiplicativa forneceu estimativas relevantes para os parâmetros de maior interesse biológico. O peso assintótico (β_1) apresentou valores claramente, significamente distintos entre os sexos, sendo de 2.610 g (ICr 95%: 2.120–3.090 g) para os machos e 1.730 g (ICr 95%: 1.390–2.050 g) para as fêmeas, evidenciando um dimorfismo sexual acentuado. No modelo von Bertalanffy segmentado, o parâmetro τ representou o ponto de transição entre dois regimes de crescimento, refletindo uma mudança na taxa de ganho de peso dentro do intervalo etário observado. Esse ponto de transição (τ) ocorreu aos 234 dias para os machos e aos 247 dias para as fêmeas. Então, foi possível concluir que a modelagem Bayesiana é uma ferramenta essencial para a análise de crescimento em tilápias-do-Nilo, permitindo uma flexibilidade na estrutura do erro que não é viável em abordagens clássicas. A combinação do modelo VBS (para a dinâmica de crescimento) com a distribuição t -Student heterocedástica (para a estrutura de erro) demonstrou ser a abordagem mais robusta e biologicamente coerente. As principais contribuições deste trabalho incluem a validação de uma metodologia robusta para a modelagem de crescimento em aquicultura, fornecendo estimativas de parâmetros com maior precisão e interpretabilidade. Para estudos futuros, recomenda-se: a incorporação de efeitos hierárquicos (por exemplo, diferenças entre linhagens) e covariáveis ambientais (por exemplo, temperatura, qualidade da água) para explicar a variabilidade não residual; o uso de informações *a priori* informativas, quando disponíveis, para refinar a estimação; e a ampliação do número de observações em idades mais avançadas, a fim de reduzir a incerteza preditiva e refinar a estimativa do peso assintótico.

Palavras-chave: Aquicultura; Distribuições assimétricas; Modelos não lineares; Stan.

Abstract

Fish farming is a significant activity in Brazil, with Nile tilapia (*Oreochromis niloticus*) accounting for a large portion of national production. Accuracy in modeling the somatic growth of this species is essential for optimizing zootechnical management, predicting the optimal slaughter point, and guiding genetic improvement programs. However, the application of traditional Nonlinear Growth Models (such as Brody, Gompertz, Logistic, and Von Bertalanffy) is based on classical assumptions of homoscedasticity (constant error variance) and symmetry (Normal distribution for residuals). In real longitudinal biological growth data, a violation of these premises is frequently observed, manifested by heteroscedasticity (residual variability usually increases with age and weight) and the presence of heavy tails or asymmetry in the residuals. Failing to account for these characteristics can lead to biased growth parameter estimates, such as asymptotic weight and growth rate, compromising the robustness of biological inferences. The growth structure of tilapia, subject to intense variations, suggests that growth does not follow a constant rate pattern. In this context, the Segmented Von Bertalanffy (SVB) model allows the growth rate to change at a non-fixed transition point, capturing distinct phases of acceleration and deceleration. To handle error structures, this study utilized robust error distributions in addition to the Normal: flexible symmetric distributions, such as Student's t , capable of modeling heavy tails (kurtosis) through the degrees of freedom parameter (ν); and asymmetric distributions, such as Skew-Normal and Skew- t , which include an additional parameter to model skewness. Heteroscedasticity was explicitly modeled by defining the residual variance (σ_i^2) as a function of age (t) through the following multiplicative and additive structures, controlled by the heteroscedasticity parameter (ρ): multiplicative power: $\sigma_i^2 = \sigma^2 t^\rho$; multiplicative exponential: $\sigma_i^2 = \sigma^2 \exp(\rho t)$; and additive (Linear): $\sigma_i^2 = \sigma^2 + \rho t$. The study used longitudinal biometric data (weight) from male and female Nile tilapia from a genetic improvement program at the State University of Maringá (UEM). Exploratory analysis revealed pronounced sexual dimorphism and an increase in weight dispersion with age, justifying sex-specific analysis and heteroscedasticity modeling. Classical models and the SVB model were fitted in combination with four error distributions (Normal, Student's t , Skew-Normal, Skew- t) in homoscedastic versions. The best-fitting model was then tested with the three heteroscedastic structures. Estimation was performed via Bayesian Inference in the *Stan* computational environment using the MCMC algorithm (NUTS/HMC). Weakly informative prior distributions were used. Predictive performance was evaluated by the LOO-CV Information Criterion (LOOIC). The SVB model showed the best predictive performance in all comparisons (lowest LOOIC). Among the homoscedastic distributions, the Student's t distribution resulted in the lowest LOOIC (Males: 1,407; Females: 2,007) for the SVB. The low value of the degrees of freedom parameter ($\nu \approx 1.7$ to 1.9) confirmed the presence of heavy tails in the residuals. The multiplicative power model ($\sigma_i^2 = \sigma^2 t^\rho$) presented the best predictive fit. The parameter ρ for this structure was

significantly different from zero and close to 2.0 (males $\rho \approx 1.85$; females $\rho \approx 1.78$), indicating strong heteroscedasticity. The final model (SVB with Student's t error and multiplicative power variance) provided relevant estimates: the asymptotic weight (β_1) was 2,610 g (95% CrI: 2,120–3,090 g) for males and 1,730 g (95% CrI: 1,390–2,050 g) for females. The transition point (τ) occurred at 234 days for males and 247 days for females. Bayesian modeling proved to be an essential tool, offering flexibility in error structure and biological coherence. The main contributions include the validation of a robust methodology for aquaculture growth modeling, providing parameter estimates with greater precision and interpretability.

Keywords: Aquaculture; Asymmetric distributions; Nonlinear models; Stan.

Lista de ilustrações

Figura 1 – Gráficos de caixa do peso por grupos decis de idade de tilápias, por sexo. .	34
Figura 2 – Curvas de crescimento ajustadas pelos modelos homocedásticos <i>t</i> -Student para tilápias, por sexo.	35
Figura 3 – Curvas de crescimento do Modelo de Von Bertalanffy Segmentado heterocedástico multiplicativo potência, por sexo, com extrapolação até 600 dias.	40

Lista de tabelas

Tabela 1 – Modelos não lineares aplicados às curvas de crescimento animal.	18
Tabela 2 – Modelos não lineares aplicados às curvas de crescimento animal.	27
Tabela 3 – Distribuições <i>a priori</i> utilizadas na modelagem dos parâmetros de crescimento e estrutura de erro.	31
Tabela 4 – Medidas descritivas para as os pesos em gramas de tilápias por grupo decil de idade e sexo.	33
Tabela 5 – Valores do LOOIC para os modelos homocedásticos segundo a distribuição dos erros e o sexo dos animais.	34
Tabela 6 – Estimativas dos parâmetros do modelo de Von Bertalanffy Segmentado homocedástico para tilápias, por sexo, sob diferentes distribuições de erro.	37
Tabela 7 – Comparação preditiva das estruturas de heterocedasticidade aplicadas ao modelo Von Bertalanffy Segmentado <i>t</i> -Student, por sexo.	39
Tabela 8 – Estimativas <i>a posteriori</i> dos parâmetros do modelo de Von Bertalanffy Segmentado <i>t</i> -Student heterocedástico multiplicativo potência, por sexo.	40

Sumário

1	Introdução	12
2	Fundamentação teórica	14
2.1	Estudo do Crescimento de Peixes	14
2.2	Modelagem Não Linear e Abordagem Bayesiana	16
2.3	Inferência Bayesiana com <i>Stan</i> em Modelos de Crescimento	22
3	Aplicação (Artigo)	24
3.1	Introdução	25
3.2	Materiais e métodos	26
3.2.1	Dados	26
3.2.2	Modelagem Estatística	26
3.2.3	Erros Assimétricos	28
3.2.3.1	Distribuições dos Erros	29
3.2.4	Heterocedasticidade	29
3.2.5	Estimação Bayesiana	29
3.2.5.1	Especificação das Distribuições <i>a Priori</i> e da Estrutura de Erro.	30
3.2.5.2	Procedimento de Amostragem (MCMC)	31
3.2.5.3	Diagnóstico de Convergência	32
3.2.5.4	Comparação dos Modelos	32
3.3	Resultados e discussão	32
3.4	Conclusão	41
4	Considerações finais	42
	Referências	44
	Apêndices	48
APÊNDICE A	Análise exploratória dos dados	49
APÊNDICE B	Modelo Von Bertalanffy <i>t-Student Homocedástico - Stan</i>	52
APÊNDICE C	Modelo Von Bertalanffy <i>t-Student Heterocedástico Multiplicativo Potência - Stan</i>	55
APÊNDICE D	Estrutura do R para ajuste e análise do modelo Von Bertalanffy Segmentado	58

Capítulo 1

Introdução

A produção e a comercialização de tilápias-do-Nilo (*Oreochromis niloticus*) vêm crescendo mundialmente, e o Brasil figura entre os principais produtores dessa espécie, respondendo por mais da metade da produção total da piscicultura nacional (ZARDIN et al., 2019). Nesse contexto, compreender o crescimento dos peixes é fundamental para o aprimoramento dos sistemas produtivos, uma vez que o desenvolvimento corporal está diretamente relacionado ao ganho de peso e, conseqüentemente, à produtividade.

Um ponto de interesse é o tempo necessário até o animal atingir o peso ou comprimento assintótico, ou seja, quando os animais, em média, param de crescer. Esse fato é interessante para otimização da produção comercial, pois animais que atingem o peso assintótico em menos tempo, podem ser selecionados em programas de melhoramento genético.

Assim sendo, os modelos não lineares de curvas de crescimento são frequentemente usados para descrever o crescimento somático de animais, já que esse crescimento geralmente não é linear ao longo do tempo (KAPS; LAMBERSON, 2004).

Os modelos clássicos não lineares de curvas de crescimento são: Brody, Gompertz, Logístico, Richards e Von Bertalanffy e cada um desses modelos tem diferentes parâmetros, que podem ser ajustados aos dados de crescimento observados, para prever o desenvolvimento de um animal ao longo do tempo (ROSSI et al., 2017). A escolha do modelo depende das características específicas do crescimento do animal em estudo.

No contexto da modelagem somática, o modelo de Von Bertalanffy é um dos mais empregados, caracterizando o crescimento por meio de um número reduzido de parâmetros de fácil interpretação biológica, como o valor assintótico e a taxa de crescimento (BERTALANFFY, 1938; BEVERTON; HOLT, 1957). Contudo, a estrutura matemática desses modelos clássicos assume implicitamente que a taxa de crescimento no início da vida é constante ou segue um padrão rígido e homogêneo. Para a tilápia, espécie sujeita a intenso manejo e variações climáticas,

essa restrição de homogeneidade na taxa de crescimento pode não ser biologicamente adequada, indicando a necessidade de modelos que permitam a segmentação da curva para capturar com maior precisão os períodos distintos de aceleração e desaceleração do desenvolvimento.

Tradicionalmente, os modelos de regressão não lineares são aplicados sob pressuposições de resíduos normais e homocedásticos. Em muitas situações biológicas, contudo, a falta de simetria e a presença de *outliers* (valores extremos) violam essas pressuposições. Dessa forma, a Inferência Bayesiana proporciona a flexibilidade para utilizar distribuições mais robustas e adaptáveis que modelam o parâmetro de assimetria e a curtose, especialmente quando é necessário lidar com dados que apresentam caudas mais longas do que as da distribuição normal (FACCIN; ROSSI, 2024), como as distribuições *t*-Student, Skew-normal (AZZALINI, 1985) e Skew-*t* (GUPTA; XU, 2003).

Adicionalmente, a análise exploratória de dados longitudinais de crescimento frequentemente revela um comportamento heterocedástico, onde a variância aumenta em função da idade. A modelagem explícita dessas estruturas de erro (aditiva ou multiplicativa) pode levar a inferências mais precisas dos parâmetros de crescimento.

O objetivo deste trabalho foi comparar e avaliar o comportamento do crescimento de tilápias-do-Nilo por meio da Inferência Bayesiana, ajustando os modelos não lineares clássicos de crescimento (Brody, Gompertz, Logístico, Von Bertalanffy) e a extensão Von Bertalanffy Segmentado. Adicionalmente, o estudo visou identificar a distribuição do erro mais adequada para o modelo de melhor ajuste, considerando estruturas homocedásticas (Normal, *t*-Student, Skew-normal e Skew-*t*) e heterocedásticas (aditiva, multiplicativa exponencial e multiplicativa tipo potencia), e investigar o dimorfismo sexual no crescimento da espécie.

Capítulo 2

Fundamentação teórica

Neste capítulo são apresentados os principais conceitos e abordagens metodológicas que orientam o desenvolvimento deste trabalho. Discutem-se inicialmente os fundamentos biológicos e zootécnicos relacionados ao crescimento de peixes, com ênfase na espécie tilápia-do-Nilo (*Oreochromis niloticus*), amplamente cultivada no Brasil. Em seguida, são exploradas as técnicas de modelagem não linear aplicadas à descrição de curvas de crescimento, incluindo os modelos clássicos utilizados na literatura e suas propriedades. Por fim, são abordadas as contribuições da Inferência Bayesiana para a estimação de modelos complexos, com destaque para o uso de distribuições de erros Simétricos e Assimétricos, Homocedásticos e Heterocedásticos, implementados no ambiente computacional *Stan*.

2.1 Estudo do Crescimento de Peixes

A piscicultura é uma atividade em plena expansão no Brasil, destacando-se pela sua relevância econômica, social e ambiental. Conforme o Anuário da Peixe BR ([PEIXE BR, 2024](#)), a produção nacional de peixes de cultivo atingiu 887.029 toneladas em 2023, representando um crescimento de 2,5% em relação ao ano anterior. A espécie com maior representatividade é a tilápia-do-Nilo, responsável por 66,4% do total produzido, correspondendo a 589.712 toneladas.

A grande expressão da tilápia na piscicultura brasileira se deve à sua rusticidade, rápido crescimento, boa conversão alimentar, além da aceitação no mercado interno e externo. Estados como Paraná, São Paulo e Minas Gerais concentram as maiores produções da espécie, impulsionados por programas de melhoramento genético, tecnologias de cultivo intensivo e políticas públicas voltadas ao desenvolvimento do setor ([ZARDIN et al., 2019](#)).

Dessa forma, compreender e modelar o crescimento somático de tilápias é essencial para aprimorar o manejo e a produtividade. Nesse contexto, a análise de curvas não linear de crescimento nos permite compreender o desenvolvimento corporal dos peixes ao longo do

tempo, prever o desempenho produtivo dos animais, planejar o momento ideal de abate e orientar programas de melhoramento genético. Tais curvas descrevem o crescimento por meio de funções matemáticas que resumem, em poucos parâmetros, a relação entre a idade e o peso ou comprimento corporal (ZARDIN et al., 2019).

Dentre os modelos não lineares comumente utilizados na literatura para descrever curvas de crescimento, o Modelo de Von Bertalanffy é historicamente um dos mais empregados por sua base fisiológica, que relaciona o anabolismo e o catabolismo do animal (TRINDADE; PRATA, 2024; BEVERTON; HOLT, 1957). Outros modelos importantes, como Brody, Gompertz e Logístico, também se destacam. O Gompertz, por exemplo, é valorizado pela capacidade de representar o crescimento assimétrico observado em diversas espécies de peixes (FIALHO, 1999; SANTOS et al., 2007). No entanto, a escolha do modelo ideal depende das características específicas do crescimento do animal em estudo.

Zardin et al. (2019) investigaram o crescimento em diferentes grupos genéticos da tilápia-do-Nilo, utilizando o modelo de Gompertz. Os autores observaram que os peixes classificados como pertencentes ao grupo genético superior, com maiores valores genéticos estimados para ganho de peso diário, apresentaram maiores pesos assintóticos e taxas de crescimento, evidenciando o impacto do melhoramento genético sobre o desempenho zootécnico dos animais. Além disso, identificaram diferenças significativas entre sexos, com os machos apresentando crescimento superior às fêmeas, principalmente após o início da maturação sexual. O estudo destacou também a importância de considerar a interação entre sexo e grupo genético na modelagem do crescimento, uma vez que essa interação pode influenciar diretamente na forma e nos parâmetros da curva. Assim, os autores ajustaram modelos diferenciados para machos e fêmeas dentro de cada grupo genético, utilizando restrições nos parâmetros da curva para avaliar a similaridade entre os grupos e sexos. Os melhores ajustes foram obtidos em modelos que permitiam diferentes valores assintóticos para o peso de machos e fêmeas.

A modelagem adequada das curvas permite identificar precocidade e estimar o ponto ótimo de colheita, impactando diretamente na rentabilidade da atividade. Então, a modelagem do crescimento de organismos cultivados é uma ferramenta essencial de gestão na aquicultura, uma vez que os lucros do setor são diretamente relacionados ao peso comercial dos animais. Zarzar (2023) destaca que, em fazendas comerciais, é comum que os dados de crescimento sejam limitados a observações abaixo do ponto de inflexão da curva sigmoide, em razão de decisões econômicas ou exigências de mercado. Tal limitação pode acarretar viés nas estimativas dos parâmetros dos modelos de crescimento. Para reduzir esse problema, o Zarzar (2023), em pesquisa com camarões, propõe o uso de modelos não lineares com estrutura hierárquica Bayesiana, que permitem capturar com maior precisão a variabilidade entre unidades produtivas, resultando em inferências mais robustas e sensíveis. Essa abordagem é particularmente relevante para sistemas de produção animal, como o cultivo de tilápias, por reforçar a relevância de considerar tanto a estrutura hierárquica quanto as limitações do conjunto de dados em estudos

de crescimento animal.

A criação de tilápias está sujeita a variações ambientais, e a suposição de homogeneidade na taxa de crescimento, implícita nos modelos clássicos, pode não representar adequadamente o comportamento biológico da espécie. Essa limitação reforça a necessidade de modelos mais flexíveis, capazes de capturar as mudanças na taxa de crescimento ao longo do ciclo produtivo. Nesse contexto, o modelo não linear segmentados apresenta-se como uma alternativa promissora, pois sua estrutura matemática permite identificar fases distintas de aceleração e desaceleração do desenvolvimento.

Após compreender os fundamentos biológicos do crescimento, torna-se necessário descrever esse processo por meio de modelos adequados. Assim, a próxima seção apresenta os principais modelos não lineares e suas propriedades sob a ótica da inferência Bayesiana.

2.2 Modelagem Não Linear e Abordagem Bayesiana

Modelos de regressão não linear têm sido amplamente utilizados para descrever fenômenos cuja evolução ao longo do tempo não segue um padrão linear, como ocorre com frequência em estudos de crescimento biológico. Diferentemente dos modelos lineares, esses modelos permitem representar o comportamento sigmoide típico de curvas de crescimento, utilizando um número reduzido de parâmetros com interpretação biológica clara, como o valor assintótico, a taxa de crescimento e o ponto de inflexão (CAMPOS; FILHO, 2011).

A utilização de curvas de crescimento com abordagem longitudinal e parâmetros de fácil interpretação biológica é um aspecto valorizado em estudos zootécnicos (FACCIN; ROSSI, 2024). Podem ser aplicáveis à modelagem de crescimento somático, como no presente estudo, com tilápias.

Assim sendo, o estudo e a modelagem do crescimento em organismos animais frequentemente se baseiam na estrutura de um modelo de regressão não linear geral. Este modelo é fundamental para descrever a relação entre uma variável de resposta, como o peso corporal (y), e uma variável preditora, tipicamente o tempo ou idade (t), quando essa relação não pode ser adequadamente representada por uma linha reta.

A forma geral deste modelo é expressa por:

$$y_{ij} = f(\beta, t_i) + \epsilon_{ij}$$

Nesta formulação:

- y_{ij} representa a observação da característica de interesse (por exemplo, o peso) para o j -ésimo indivíduo;
- t é a variável preditora (tempo ou idade) associada à observação y_{ij} ;
- $f(\beta, t_i)$ é o preditor não linear que contém a função ou curva de regressão que descreve o crescimento;
- ϵ_{ij} é o erro aleatório associado à observação, representando a variação estocástica não explicada pela função $f(\cdot)$.

De modo geral, assume-se que estes erros são independentes e identicamente distribuídos, seguindo uma distribuição normal, $\epsilon_{ij} \sim N(0, \sigma^2)$, o que implica em média zero e variância constante σ^2 . Tal suposição é a base para os modelos clássicos de regressão não linear, implicando em homocedasticidade. Entretanto, em estudos biológicos e zootécnicos, particularmente em dados longitudinais de crescimento, é comum observar o fenômeno da heterocedasticidade (PINHEIRO; BATES, 2000), que ocorre quando a variabilidade residual aumenta (ou diminui) conforme o crescimento do indivíduo (por exemplo, indivíduos maiores tendem a apresentar maior dispersão no peso em comparação com os juvenis). A não consideração da heterocedasticidade, quando presente, pode levar a estimativas de erro padrão enviesadas e, conseqüentemente, a inferências estatísticas incorretas sobre os parâmetros do modelo.

Modelos clássicos como o de Brody, Gompertz, Logístico e Von Bertalanffy (Tabela 1) são, na verdade, especificações particulares dessa função genérica $f(\cdot)$. Cada um desses modelos define uma forma funcional específica, o que lhes confere características distintas em relação ao formato da curva e à localização do ponto de inflexão. Essa particularização da função $f(\cdot)$ é o que influencia diretamente na escolha do modelo mais apropriado para descrever o padrão de crescimento observado em diferentes contextos experimentais (SEBER; WILD, 1989).

Uma crítica aos modelos clássicos é impor uma taxa de crescimento constante (curvatura) que pode não ser biologicamente precisa, especialmente em peixes, no qual as condições ambientais iniciais causam variabilidade (LEE et al., 2020). Essa limitação justifica a inclusão de modelos mais flexíveis.

Tabela 1 – Modelos não lineares aplicados às curvas de crescimento animal.

Modelo	Fórmula (y_{ij})	Referência
Brody	$\beta_1 (1 - e^{-\beta_2(t_{ij}-\beta_3)})$	Brody (1945)
Gompertz	$\beta_1 e^{-e^{-\beta_2(t_{ij}-\beta_3)}}$	Gompertz (1825)
Logístico (Verhulst)	$\frac{\beta_1}{1 + e^{-\beta_2(t_{ij}-\beta_3)}}$	Verhulst (1838)
Von Bertalanffy	$\beta_1 (1 - e^{-\beta_2(t_{ij}-\beta_3)})^3$	Bertalanffy (1938)

y_{ij} : Peso corporal predito (g) i no tempo j ;

β_1 : Peso assintótico teórico (g);

β_2 : Taxa de crescimento (velocidade com que o peso se aproxima de β_1);

β_3 : Parâmetro de localização temporal (idade teórica ao peso zero para Brody e Von Bertalanffy; idade no ponto de inflexão para Gompertz e Logístico);

t_{ij} : Idade do animal (dias) i no tempo j .

A extensão segmentada do modelo de Von Bertalanffy é fundamental, pois permite que os parâmetros de taxa de crescimento (β_2 e β_4) mudem em um ponto de inflexão não forçado. Ele aborda a heterogeneidade da taxa de crescimento ao longo da vida do animal, sendo mais capaz de capturar períodos distintos de aceleração e desaceleração do desenvolvimento. Essa formulação permite identificar mudanças fisiológicas no metabolismo ao longo do crescimento, refletindo, por exemplo, transições entre fases de adaptação, crescimento acelerado e maturação. Sua forma empírica para o peso corporal predito (g) é:

$$y_{ij} = \begin{cases} \beta_1(1 - e^{-\beta_2(t_{ij}-\beta_3)})^3 & \text{se } t \leq \tau \\ \mu_\tau + (\beta_1 - \mu_\tau)(1 - e^{-\beta_4(t_{ij}-\tau)})^3 & \text{se } t > \tau \end{cases}$$

no qual $\mu_\tau = \beta_1(1 - e^{-\beta_2(\tau-\beta_3)})^3$ é o peso predito na idade de transição τ .

- β_1 : Peso assintótico teórico (g) (μ_∞);
- β_2 : Taxa de crescimento da 1ª fase;
- β_3 : Idade teórica com peso zero;

- β_4 : Taxa de crescimento da 2ª fase;
- τ : Idade de transição entre as fases de crescimento (estimada pelo modelo).

Apesar de sua versatilidade, os modelos não lineares tradicionalmente assumem que os erros seguem uma distribuição normal com variância constante. No entanto, essa suposição pode não ser realista, especialmente quando os dados apresentam assimetria ou heterocedasticidade, ou seja, a variância não é constante ao longo dos valores da variável explicativa. Nessas situações, a aplicação de distribuições assimétricas se torna necessária para garantir inferências mais robustas e precisas (CAMPOS; FILHO, 2011).

As distribuições Skew-normal (AZZALINI, 1985) e Skew- t (GUPTA; XU, 2003) surgem como alternativas interessantes por incorporarem parâmetros que permitem modelar tanto a assimetria quanto a curtose dos resíduos. A distribuição Skew-normal é uma generalização da normal simétrica, com um parâmetro adicional que define a direção e intensidade da assimetria. Já a Skew- t inclui, além da assimetria, um parâmetro de graus de liberdade, o que a torna mais adequada para modelar dados com caudas pesadas ou presença de *outliers*.

Rossi e Santos (2014) utilizaram uma abordagem Bayesiana para o ajuste de modelos não lineares com erros seguindo distribuições assimétricas, como a Skew-normal e Skew- t . Utilizando dados de peso corporal de codornas, os autores demonstraram que o modelo Gompertz com erros Skew-normal para machos e Skew- t para fêmeas apresentou melhor ajuste segundo o critério de informação de *deviance* Bayesiano (DIC) e maior cobertura dos intervalos de credibilidade. A modelagem permitiu incorporar diferentes estruturas para os erros, resultando em estimativas mais realistas e robustas para os parâmetros de crescimento.

A adequação da estrutura dos erros é uma etapa crucial na modelagem de dados assimétricos e com possível heterocedasticidade. Xie, Wei e Lin (2009) desenvolveram testes de escore para avaliar a homogeneidade da variância e do parâmetro de assimetria em modelos de regressão não linear com erros Skew-normal. Os autores mostraram, por meio de simulações de Monte Carlo, que a suposição de homocedasticidade pode ser inapropriada e comprometer a eficiência dos estimadores. Assim, a detecção da variabilidade na dispersão e na assimetria se torna essencial para justificar o uso de modelos mais flexíveis, como os que assumem heterocedasticidade e assimetria simultaneamente.

Modelos de regressão não linear com suposição de normalidade dos erros podem ser inadequados em situações com presença de assimetria, caudas pesadas ou heterocedasticidade. Lachos, Bandyopadhyay e Garay (2011) propuseram uma classe de modelos conhecida como SMSN-HNLM (*heteroscedastic nonlinear regression models based on scale mixtures of Skew-normal distributions*), que incorpora simultaneamente variabilidade condicional da variância e flexibilidade na forma da distribuição dos erros. Os autores demonstraram, por meio de estudos de simulação e aplicação a dados reais, que os modelos baseados em distribuições como Skew- t e *Skew-slash* mostram-se mais robustos diante de observações discrepantes e melhor ajuste

em relação à modelos normais. Essa abordagem destaca-se como alternativa poderosa na modelagem de dados de crescimento com estrutura complexa de erro, como os observados em experimentos zootécnicos.

Louzada, Ferreira e Diniz (2014) discutem a inadequação das suposições clássicas de normalidade e homocedasticidade em modelos de crescimento, propondo a utilização de distribuições assimétricas, como a Skew-normal, aliadas a estruturas de variância heterocedásticas. Os autores demonstram, por meio de simulações e aplicação a dados reais, que essas abordagens proporcionam estimativas mais realistas e ajustadas aos dados observados, evidenciando a importância de modelos flexíveis na descrição do crescimento biológico.

Adicionalmente, a abordagem Bayesiana se destaca como uma estratégia eficaz para a estimação dos parâmetros em modelos não lineares com erros assimétricos e heterocedásticos. Por meio da inferência baseada em amostragem de Monte Carlo via Cadeias de Markov (MCMC), é possível obter distribuições *a posteriori* dos parâmetros mesmo em estruturas hierárquicas complexas, além de incorporar incertezas associadas às estimativas (CAMPOS; FILHO, 2011; ZARZAR et al., 2022).

Modelos heterocedásticos são essenciais na estrutura Bayesiana, pois permitem que a variância dos resíduos (σ_i^2) não seja constante, mas sim uma função de covariáveis, como a idade dos indivíduos. Tradicionalmente, essa dependência é modelada por meio de uma função multiplicativa m_i , de modo que $\sigma_i^2 = \sigma^2 m_i$. As formas mais comuns para m_i são a exponencial ($m_i = \exp(\rho x_i)$) ou a linear aditiva ($\sigma_i^2 = \sigma^2 + \rho x_i$), as quais permitem que a dispersão dos dados varie de forma crescente ou decrescente ao longo do tempo. Em estudos de crescimento animal, essa heterocedasticidade é esperada e biologicamente justificada: indivíduos mais velhos apresentam maior dispersão de peso, refletindo a crescente diversidade natural entre taxas metabólicas, diferenças fisiológicas e condições ambientais. A inclusão dessas funções garante, portanto, um melhor ajuste e uma representação mais realista da variabilidade observada.

Cheng e Riu (2006) abordam modelos de regressão linear nos quais ambas as variáveis estão sujeitas a erros de medição com variância heterocedástica, destacando que a não consideração dessa estrutura pode levar a estimativas inconsistentes dos parâmetros. A relevância desse tipo de modelagem é particularmente evidente em áreas como química e ciências biológicas, no qual a variabilidade inerente às medições experimentais é frequente. Tais contribuições reforçam a importância de considerar estruturas flexíveis de variância, mesmo em modelos não lineares de crescimento.

A possibilidade de definir a variância como função de covariáveis é um recurso especialmente importante em modelos de crescimento, nos quais a dispersão dos dados frequentemente aumenta com a idade dos indivíduos (DINIZ; LOUZADA-NETO; MORITA, 2012; CAMPOS; FILHO, 2011).

Em estudos aplicados ao crescimento animal, a modelagem da variância residual é fundamen-

tal. Por exemplo, [Mazucheli, Souza e Philippsen \(2011\)](#) demonstraram que, no crescimento de codornas, a suposição de homocedasticidade pode superestimar o peso assintótico, justificando o uso de modelos com variância dependente da idade. Ao ajustarem o modelo de Gompertz com estrutura heterocedástica multiplicativa ([HARVEY, 1976](#)) sob os enfoques clássico e Bayesiano, os autores mostraram que a modelagem da variância na forma $\sigma_i^2 = \sigma^2 t_i^\rho$ resultou em estimativas mais coerentes e em melhor ajuste segundo o critério DIC (*Deviance Information Criterion*). Reforçando essa estratégia, [Diniz, Louzada-Neto e Morita \(2012\)](#) propuseram o modelo de Von Bertalanffy com a mesma estrutura, que é vantajosa por representar uma relação log-linear entre a variância e o tempo. A inferência foi conduzida sob uma abordagem Bayesiana completa com uso de métodos MCMC, e os autores demonstraram que a consideração explícita da heterocedasticidade levou à redução dos intervalos de credibilidade e maior precisão nas estimativas dos parâmetros. Tais resultados sustentam a adoção de estruturas flexíveis de erro e abordagem Bayesiana em modelos de crescimento animal, sendo esta estratégia particularmente relevante para dados de crescimento de tilápias, foco do presente estudo.

[Faccin e Rossi \(2024\)](#) demonstraram a eficácia da modelagem Bayesiana para curvas de crescimento de animais, ao utilizarem o modelo de Gompertz com diferentes distribuições de erro. Os autores identificaram que a versão com erros heterocedásticos e distribuição Skew- t apresentou melhor desempenho segundo o critério WAIC, reforçando a importância de considerar distribuições mais flexíveis e assimétricas para descrever padrões de crescimento em dados zootécnicos reais.

No presente estudo, foram consideradas três formas funcionais para a variância condicional, multiplicativa tipo potencia $\sigma_i^2 = \sigma^2 t_i^\rho$, multiplicativa exponencial $\sigma_i^2 = \sigma^2 \exp(\rho t_i)$ e aditiva $\sigma_i^2 = \sigma^2 + \rho t_i$. Essas formulações permitem avaliar se a variabilidade do peso aumenta de forma linear, exponencial ou proporcional à idade, ajustando-se de maneira mais realista à dinâmica biológica do crescimento de tilápias.

Modelos de regressão não linear que assumem homocedasticidade e simetria dos erros podem ser inadequados em diversas aplicações práticas, especialmente quando há variabilidade na dispersão ou assimetria dos dados. [Ferreira et al. \(2020\)](#) propuseram uma abordagem Bayesiana baseada em misturas de escala de distribuições Skew-normal para modelar simultaneamente heterocedasticidade e assimetria em modelos de regressão não linear. Os autores demonstraram que tal estrutura melhora substancialmente o ajuste dos modelos, permitindo capturar variações na dispersão condicional e caudas assimétricas nos dados. Além disso, foram utilizados critérios como LPML e Pseudo- R^2 para avaliação do desempenho preditivo, evidenciando a superioridade do modelo proposto em comparação com alternativas mais simples.

Portanto, a combinação entre modelos não lineares e distribuições de erro flexíveis, sob a abordagem Bayesiana, constitui uma alternativa metodológica robusta e flexível para a análise de dados biométricos, como os de crescimento de peixes. Este trabalho segue essa linha metodológica, com especial atenção às distribuições Skew-normal e Skew- t heterocedásticas,

para avaliar o crescimento de tilápias em função da idade e do sexo.

A avaliação e a seleção do modelo mais adequado a um conjunto de dados podem ser realizadas por meio de diversos critérios que equilibram a qualidade do ajuste com a penalidade imposta pela complexidade do modelo.

Tradicionalmente, critérios Bayesianos, como o *Deviance Information Criterion* (DIC), o *Expected Akaike Information Criterion* (EAIC) e o *Expected Bayesian Information Criterion* (EBIC) são utilizados, permitindo avaliar a qualidade e a parcimônia dos modelos (SPIEGELHALTER et al., 2002; CAMPOS; FILHO, 2011), entretanto há outros mais sofisticados, principalmente para a comparação entre modelos ser realizada de forma mais robusta, por exemplo por meio do Critério de Informação de *Watanabe-Akaike* (WAIC) e da Validação Cruzada, *Leave-One-Out* (LOO-CV). O WAIC é um critério totalmente Bayesiano que considera a incerteza associada aos parâmetros e penaliza modelos complexos pela complexidade efetiva. Já a LOO-CV, implementada por meio do método de suavização de importância por Pareto (PSIS-LOO), estima a capacidade preditiva do modelo removendo e reestimando cada observação da amostra. O LOOIC (*LOO Information Criterion*), derivado do LOO-CV, é amplamente aceito como a abordagem mais robusta para a avaliação preditiva em modelos Bayesianos, por ser mais estável e ter uma estimativa menos enviesada do poder preditivo (VEHTARI; GELMAN; GABRY, 2017).

2.3 Inferência Bayesiana com *Stan* em Modelos de Crescimento

A Inferência Bayesiana tem se consolidado como uma abordagem robusta para a estimação de modelos estatísticos complexos, sobretudo em situações no qual a modelagem frequentistas enfrenta limitações, como em modelos não lineares com erros assimétricos e heterocedásticos. Nesse contexto, a modelagem em ambiente computacional *Stan* (STAN DEVELOPMENT TEAM, 2025) destaca-se como uma das ferramentas mais modernas e eficientes para a implementação de modelos probabilísticos com estrutura hierárquica e distribuições flexíveis (GELMAN et al., 2020).

O *Stan* é uma linguagem probabilística que permite especificar modelos estatísticos por meio de equações relacionando parâmetros, variáveis latentes e observações. Para realizar a inferência, algoritmos de amostragem são sofisticados, com destaque para o Hamiltonian Monte Carlo (HMC) e sua versão adaptativa *No-U-Turn Sampler* (NUTS), que são particularmente eficazes em espaços de parâmetros de alta dimensão, com correlações complexas entre os parâmetros (STAN DEVELOPMENT TEAM, 2025).

A principal vantagem do HMC/NUTS frente a métodos tradicionais como Gibbs Sampling ou Metropolis-Hastings está na sua capacidade de explorar o espaço *a posteriori* de forma mais eficiente, reduzindo o tempo necessário para convergência e proporcionando amostras menos autocorrelacionadas. Isso é essencial em modelos de crescimento com múltiplos parâmetros

não linearmente relacionados e distribuições de erro que não seguem o padrão normal clássico.

Em particular, o *Stan* facilita a modelagem de variância heterocedástica e de distribuições de erro assimétricas, por permitir a escrita explícita da função de verossimilhança.

Outro aspecto fundamental na utilização do *Stan* é a verificação rigorosa da convergência das cadeias MCMC. O *Stan* fornece diagnósticos automáticos baseados na estatística $\hat{R}$, aprimorada por [Vehtari et al. \(2019\)](#), que utiliza rank-normalização, dobragem e localização para detectar falhas de convergência de forma mais sensível do que o método clássico. Além disso, métricas como a divergência das amostras, o número efetivo de amostras e gráficos de traço auxiliam na avaliação da qualidade da amostragem.

A escolha do *Stan* neste trabalho justifica-se pela sua flexibilidade para lidar com curvas de crescimento não lineares, incorporando estruturas de variância complexas e distribuições de erro não usuais na modelagem, além de oferecer confiabilidade nas estimativas por meio de algoritmos de inferência eficientes e ferramentas avançadas de diagnóstico, para descrever o crescimento de tilápias. A seguir, apresenta-se a aplicação prática dessa abordagem.

Capítulo 3

Aplicação (Artigo)

Resumo

Crescimento de tilápias-do-Nilo considerando diferentes curvas, estruturas de variância e distribuição para os erros: uma abordagem Bayesiana

A tilápia-do-Nilo (*Oreochromis niloticus*) exige modelagem de crescimento somático robusta, pois seus dados longitudinais apresentam alta heterocedasticidade e caudas pesadas, características negligenciadas por modelos não lineares clássicos. O objetivo deste trabalho foi comparar abordagens Bayesianas, incluindo modelos Brody, Gompertz, Logístico, Von Bertalanffy e o Von Bertalanffy Segmentado (VBS), com diferentes distribuições de erro (Normal, t -Student, Skew-normal e Skew- t) e estruturas de variância homocedástica e heterocedástica (aditivo, multiplicativos potência e exponencial). A estimação via Stan/MCMC e a comparação preditiva (LOOIC) foram conduzidas em dados de 572 machos e 499 fêmeas criados em tanques rede. O Modelo de VBS se mostrou superior aos clássicos, com o menor LOOIC, validando a necessidade de segmentar a taxa de crescimento ($\tau \approx 234$ - machos a 247 dias - fêmeas). A distribuição t -Student foi importante para modelar a curtose e garantir a robustez. A heterocedasticidade foi confirmada, visto que seu parâmetro (ρ) indicou um crescimento da variância com a idade ($\rho \approx 1,8$), sendo a estrutura potência multiplicativa a mais adequada indicada via LOOIC. Foi encontrado dimorfismo sexual acentuado, com peso assintótico superior para machos ($\beta_1 \approx 2.610g$) em comparação com fêmeas ($\beta_1 \approx 1.730g$). Sugere-se que a modelagem Bayesiana do VBS com erro t -Student e variância heterocedástica é uma abordagem recomendada para descrever o crescimento de tilápias, fornecendo estimativas precisas para o manejo.

Palavras-chave: Curvas de crescimento; Distribuições assimétricas; Modelos não lineares; Stan.

3.1 Introdução

A tilápia (*Oreochromis niloticus*) é uma das espécies de peixe mais cultivadas globalmente, devido à sua adaptabilidade, crescimento rápido e alta aceitação no mercado. O Brasil está entre os maiores produtores dessa espécie, correspondendo mais da metade da produção total da piscicultura nacional (ZARDIN et al., 2019).

Em tilápias, o dimorfismo sexual influencia significativamente o crescimento, com machos apresentando desempenho superior em parâmetros morfométricos em comparação com fêmeas, assim sendo, considerar essas diferenças, é essencial para o desenvolvimento de estratégias de cultivo que melhorem o rendimento e a saúde dos peixes (ZARDIN et al., 2019).

O estudo de curvas de crescimento, utilizando modelos não lineares, é fundamental para compreender o desenvolvimento de organismos ao longo do tempo (KAPS; LAMBERSON, 2004), sendo amplamente aplicado em diversas áreas. Nas espécies de peixes é importante para otimizar práticas de manejo, permitindo uma produção mais eficiente e sustentável.

Nos estudos de crescimento, um dos principais interesses é o tempo necessário para que o animal atinja o peso de mercado, isto é, o peso no qual, em média, o animal já apresenta crescimento próximo à estabilização e passa a ser economicamente viável para a comercialização (ZARZAR et al., 2022). Esse peso de mercado não corresponde necessariamente ao peso assintótico teórico estimado por modelos de crescimento, mas o reflete de forma indireta, uma vez que está associado à desaceleração do ganho de peso. Animais que atingem o peso de mercado em menor tempo são considerados mais precoces, e essa precocidade pode ser utilizada como característica ou critério de seleção em programas de melhoramento genético animal.

Embora modelos não lineares sejam amplamente empregados na modelagem de curvas de crescimento, a suposição de variância constante ao longo do tempo permanece comum em aplicações práticas. Em dados biológicos, como os de crescimento de peixes, essa suposição raramente é plausível, uma vez que a variabilidade do peso tende a aumentar com a idade. Ignorar essa estrutura pode resultar em inferências viesadas e subestimar a incerteza associada às estimativas, justificando a adoção de modelos que permitam variância dependente do tempo.

Distribuições mais flexíveis para os resíduos, como a *t*-Student, a Skew-normal (AZZALINI, 1985) e a Skew-*t* (GUPTA; XU, 2003), têm sido amplamente empregadas em contextos nos quais a suposição de normalidade é inadequada. Essas distribuições permitem modelar assimetria por meio de parâmetros específicos e acomodar caudas mais pesadas, características frequentemente observadas em dados biológicos e não contempladas pelo modelo normal (FACCIN; ROSSI, 2024). Dessa forma, a modelagem se beneficia de uma estrutura residual mais realista, capaz de lidar com a variabilidade e a assimetria presentes nos dados.

A abordagem Bayesiana oferece vantagens importantes para a estimação de parâmetros em modelos não lineares, uma delas é a sua flexibilidade que permite considerar diferentes

estruturas de erros, incluindo erros homocedásticos e heteroscedásticos, bem como distribuições simétricas e assimétricas. Além disso, destaca-se por incorporar informações prévias, quando disponíveis, na estimação e quantificar incertezas via distribuições *a posteriori*.

Este trabalho tem por objetivo entender o comportamento do crescimento de tilápias-do-Nilo por estimativas Bayesianas dos parâmetros de diferentes modelos de crescimento (Brody, Gompertz, Logístico, Von Bertalanffy e Von Bertalanffy Segmentado) considerando diferentes estruturas de erros homocedásticos, heterocedástico (aditivo, multiplicativo potência e exponencial), simétricos (Normal e *t*-Student) e assimétricos (Skew-normal e Skew-*t*).

3.2 Materiais e métodos

3.2.1 Dados

Os dados referem-se a biometria (peso em gramas e idade em dias) de 572 machos e 499 fêmeas de tilápia-do-Nilo pertencentes ao programa de melhoramento genético da linhagem GIFT, desenvolvido pela Universidade Estadual de Maringá (UEM) (ZARDIN et al., 2019).

O peso corporal, em gramas dos peixes foi registrado em cinco períodos distintos, resultando em uma série temporal de medidas que permitem a modelagem do crescimento individual dos animais ao longo do tempo, em dias.

Embora o protocolo experimental previsse pesagens em idades semelhantes para todos os animais, variações operacionais e logísticas impediram que as medições fossem realizadas exatamente nas mesmas idades para todos os indivíduos. Como consequência, a idade passou a ser observada como uma variável contínua, resultando em dados longitudinais com leve desbalanceamento temporal. Para fins exclusivamente descritivos, optou-se por agrupar as observações em 10 grupos de idade (G1 a G10), definidos pelos decis da idade observada. Essa estratégia permite obter grupos com tamanhos amostrais semelhantes, facilitando a visualização da evolução do peso e de sua dispersão ao longo do período avaliado.

3.2.2 Modelagem Estatística

Os quatro modelos de regressão não lineares utilizados (Tabela 2) foram: Brody, o qual representa o crescimento em estágios iniciais e é frequentemente considerado um modelo sem ponto de inflexão sigmoide claro (BRODY, 1945); Gompertz, que caracteriza-se por uma assimetria, com a fase de crescimento mais rápida ocorrendo tipicamente na primeira parte da vida (GOMPERTZ, 1825); Logístico, o qual apresenta uma curva simétrica, com a taxa de crescimento máxima ocorrendo exatamente na metade do peso assintótico (VERHULST, 1838); e o modelo de Von Bertalanffy, que assume que a taxa de ganho de massa é proporcional à área de superfície (anabolismo) menos a taxa de perda de massa (catabolismo) (BERTALANFFY, 1938).

Tabela 2 – Modelos não lineares aplicados às curvas de crescimento animal.

Modelo	Fórmula (y_{ij})	Referência
Brody	$\beta_1 \left(1 - e^{-\beta_2(t_{ij}-\beta_3)}\right)$	Brody (1945)
Gompertz	$\beta_1 e^{-e^{-\beta_2(t_{ij}-\beta_3)}}$	Gompertz (1825)
Logístico (Verhulst)	$\frac{\beta_1}{1 + e^{-\beta_2(t_{ij}-\beta_3)}}$	Verhulst (1838)
Von Bertalanffy	$\beta_1 \left(1 - e^{-\beta_2(t_{ij}-\beta_3)}\right)^3$	Bertalanffy (1938)

y_{ij} : Peso corporal predito (g) i no tempo j ;

β_1 : Peso assintótico teórico (g);

β_2 : Taxa de crescimento (velocidade com que o peso se aproxima de β_1);

β_3 : Parâmetro de localização temporal (idade teórica ao peso zero para Brody e Von Bertalanffy; idade no ponto de inflexão para Gompertz e Logístico);

t_{ij} : Idade do animal (dias) i no tempo j .

Adicionalmente foi ajustado o modelo de Von Bertalanffy com extensão segmentada, a qual permite que a taxa de crescimento seja descrita por dois parâmetros distintos (β_2 e β_4), associados a diferentes regimes de crescimento antes e após um ponto de quebra não fixo ([GALLANT, 1987](#)). Enquanto no modelo clássico a dinâmica de crescimento é governada por um único parâmetro de taxa ao longo de todo o período observado, a formulação segmentada relaxa essa restrição, possibilitando que o crescimento apresente comportamentos distintos ao longo da vida do animal. Dessa forma, o modelo é capaz de capturar períodos de crescimento relativamente mais acelerado ou mais lento, refletindo mudanças no padrão de desenvolvimento ao longo do tempo. Essas mudanças no regime de crescimento podem estar associadas a transições fisiológicas ao longo do desenvolvimento, embora sua interpretação deva ser restrita ao intervalo etário amostrado.

A forma empírica do Modelo de Von Bertalanffy (1987) Segmentado para o peso corporal predito (g) é:

$$y_{ij} = \begin{cases} \beta_1(1 - e^{-\beta_2(t_{ij}-\beta_3)})^3 & \text{se } t \leq \tau \\ \mu_\tau + (\beta_1 - \mu_\tau)(1 - e^{-\beta_4(t_{ij}-\tau)})^3 & \text{se } t > \tau \end{cases}$$

no qual $\mu_\tau = \beta_1(1 - e^{-\beta_2(\tau-\beta_3)})^3$ é o peso predito na idade de transição τ .

- β_1 : Peso assintótico teórico (g) (μ_∞);
- β_2 : Taxa de crescimento da 1ª fase;
- β_3 : Idade teórica com peso zero;
- β_4 : Taxa de crescimento da 2ª fase;
- τ : Idade de transição entre as fases de crescimento (estimada pelo modelo).

3.2.3 Erros Assimétricos

A distribuição Skew- t constitui uma extensão da distribuição t -Student, que por sua vez a |Normal, sendo aplicada em situações nas quais, além da presença de assimetria nos resíduos, ocorrem também valores extremos, caracterizando dados com estruturas assimétricas e/ou de caudas pesadas. Essa distribuição mostra-se mais eficaz do que a normal assimétrica proposta por [Azzalini \(1985\)](#), uma vez que apresenta maior robustez nesses cenários ([CRUZ; BRANCO, 2009](#); [MANGUEIRA et al., 2016](#)).

De acordo com [Godoi \(2007\)](#), se uma variável aleatória Z é considerada como seguindo a distribuição t -Student assimétrica padrão, ou Skew- t , com parâmetro de assimetria λ e curtose ν , sua função densidade de probabilidade é dada por:

$$f_z(Z) = 2f_\nu(z)T_{\nu+1} \left(\lambda z \sqrt{\frac{1+\nu}{\nu+z^2}} \right),$$

em que f_ν corresponde à função de densidade de probabilidade de uma distribuição t -Student padrão com ν graus de liberdade, $T_{\nu+1}$ é a função de distribuição acumulada de uma t -Student padrão com $\nu + 1$ graus de liberdade, e $-\infty < z < +\infty$.

O parâmetro λ controla a forma da distribuição: quando $\lambda < 0$, há assimetria negativa; quando $\lambda > 0$, ocorre assimetria positiva; e quando $\lambda = 0$, a distribuição é simétrica, reduzindo-se à t -Student padrão e, no limite $\nu \rightarrow \infty$, à distribuição normal ou gaussiana.

3.2.3.1 Distribuições dos Erros

Neste estudo, foram consideradas diferentes estruturas de erros para capturar possíveis variações dos dados. Inicialmente, foram avaliadas distribuições simétricas, incluindo a distribuição Normal e a distribuição t -Student. Ambas assumem que os resíduos são simétricos e homocedásticos, ou seja, apresentam variância constante ao longo do tempo. Todavia a t -Student permite ajustar caudas mais pesadas, tornando-a adequada para dados com possíveis *outliers* ou desvios acentuados da normalidade.

Além das distribuições simétricas, foram consideradas distribuições assimétricas, que proporcionam maior flexibilidade na modelagem dos dados. As distribuições Skew-normal e Skew- t incorporam um parâmetro de assimetria, possibilitando a captura de desvios da simetria perfeita e um ajuste ainda mais adaptável a diferentes padrões de dados.

Considerar distribuições assimétricas é particularmente relevante quando os dados apresentam assimetrias evidentes ou quando os resíduos não seguem um padrão simétrico, o que é comum em estudos de crescimento. A inclusão dessas distribuições aumenta a capacidade do modelo em capturar nuances dos dados, resultando em estimativas mais precisas e adequadas.

3.2.4 Heterocedasticidade

Considerando que em dados biológicos a variabilidade normalmente aumenta com a média, e para evitar a violação do pressuposto de homocedasticidade, a heterocedasticidade residual foi modelada utilizando a abordagem de modelos de regressão heterocedásticos (DAVIDIAN; GILTINAN, 1995). A variância condicional (σ_i^2) para a observação i foi modelada como função de uma covariável (t , a idade), utilizando a estrutura $\sigma_i^2 = \sigma^2 g(t, \rho)$, no qual σ^2 é o parâmetro de escala e ρ é o parâmetro de variabilidade, com $\rho > 0$ (HARVEY, 1976).

A estrutura geral permite avaliar a intensidade da heterocedasticidade por meio da estimativa do parâmetro ρ . Segundo Xie, Wang e Zou (2000), a magnitude de ρ pode ser interpretada da seguinte forma: valores de $\rho \approx 0$ sugerem homocedasticidade, valores de $\rho \approx 0,5$ indicam heterocedasticidade moderada, e valores de $\rho \approx 2,0$ forte.

Para a função $g(t, \rho)$, no presente estudo, foram consideradas três formas funcionais para a variância condicional σ_i^2 : potência multiplicativa $\sigma_i^2 = \sigma^2 t_i^\rho$ (HARVEY, 1976; CAMPOS; FILHO, 2011), exponencial multiplicativa $\sigma_i^2 = \sigma^2 \exp(\rho t_i)$ (XIE; WANG; ZOU, 2000; FERREIRA et al., 2020) e aditiva $\sigma_i^2 = \sigma^2 + \rho t_i$ (CHENG; RUE, 2006).

3.2.5 Estimação Bayesiana

Um dos pilares da inferência Bayesiana é a chamada operação de Bayes, a qual combina a informação disponível nos dados via verossimilhança com o conhecimento prévio sobre os parâmetros do modelo. Essa operação resulta na distribuição *a posteriori*, que contém toda

a informação atualizada sobre o parâmetro de interesse. Formalmente, a relação pode ser expressa como:

$$\pi(\beta | Y) \propto L(\beta | Y) \pi(\beta),$$

em que $\pi(\beta | Y)$ representa a distribuição *a posteriori*, $L(\beta | Y)$ corresponde à função de verossimilhança, e $\pi(\beta)$ é a distribuição *a priori* do parâmetro. Essa formulação evidencia que a análise Bayesiana incorpora tanto os dados observados quanto o conhecimento prévio, quando disponível, permitindo inferências mais flexíveis e robustas, especialmente em contextos com estruturas complexas ou amostras reduzidas.

Tal abordagem foi escolhida para a estimação dos parâmetros dos modelos não lineares devido à sua capacidade de fornecer distribuições de probabilidade posteriores completas para cada parâmetro estimado. Diferente da abordagem frequentistas, que fornece estimativas pontuais, a inferência Bayesiana permite quantificar a incerteza de forma direta, fornecendo não apenas o valor mais provável do parâmetro, mas toda a sua densidade de probabilidade. Isso possibilita a construção de intervalos de credibilidade e a análise detalhada da variabilidade biológica, tornando a metodologia robusta para lidar com diferentes distribuições de erro e estruturas de variância complexas.

3.2.5.1 Especificação das Distribuições *a Priori* e da Estrutura de Erro.

As distribuições de probabilidade *a priori* para os parâmetros estruturais e de erro foram selecionadas como fracamente informativas, garantindo que a inferência *a posteriori* fosse guiada predominantemente pelos dados observados.

Para os parâmetros que exigem restrição de positividade devido à sua natureza biológica e física, como as taxas de crescimento (β_2, β_4) e os parâmetros de escala e forma ($\sigma, \sigma_e, \rho, \nu$), adotaram-se distribuições com suporte nos reais positivos ($\mathbb{R}^+$), especificamente as distribuições Log-Normal, Cauchy⁺ e Gama. Já para parâmetros de maior escala ou que podem assumir valores em toda a reta real, como o peso assintótico (β_1) e o ponto de transição (τ), utilizaram-se distribuições Normais com variâncias configuradas para permitir uma ampla exploração do espaço paramétrico.

O emprego da distribuição *t*-Student para os resíduos (ε_{ij}) conferiu robustez ao modelo frente a possíveis valores extremos nos dados de biometria. As especificações detalhadas de todas as distribuições e seus respectivos hiperparâmetros encontram-se na Tabela 3.

As *prioris* foram definidas com base em uma estratégia de fraca informatividade, utilizando as seguintes considerações de escala e restrição: os parâmetros de crescimento β_1, β_2 (taxa de crescimento 1) e β_4 (taxa de crescimento 2) foram definidos na escala padronizada para melhorar a eficiência das cadeias MCMC; a distribuição Normal(0; 50) ou Log-Normal(0; 1) foi utilizada para garantir a positividade das taxas; os parâmetros de localização física β_3 (t_0) e τ foram

Tabela 3 – Distribuições *a priori* utilizadas na modelagem dos parâmetros de crescimento e estrutura de erro.

Parâmetro	Descrição	Distribuição <i>a priori</i>
β_1	Peso assintótico	Normal(0; 50)
β_2 e β_4	Taxa de crescimento 1 e 2	Log-Normal(0; 1)
β_3	Idade teórica / inflexão	Normal(0; 50)
τ	Ponto de Transição	Normal(μ_{idade} ; dp_{idade})
σ	Desvio-padrão residual - homocedástico	Cauchy ⁺ (0; 1)
σ_e	Desvio-padrão residual - heterocedástico	Gama(1; 0, 5)
ρ	Parâmetro da heterocedasticidade	Gama(1; 0, 5)
ν	Graus de liberdade (<i>t</i> -Student, Skew- <i>t</i>)	Gama(2; 0, 1)
λ	Parâmetro de assimetria (Skew-normal, Skew- <i>t</i>)	Normal(0; 2)

mantidos na escala real (idade em dias), utilizando *prioris* amplas (Normal(0; 50)) ou centrados na média observada. O desvio-padrão residual (σ ou σ_e) e o parâmetro de heterocedasticidade (ρ) utilizam *prioris* de referência (Cauchy⁺ e Gama) para impor as restrições de positividade necessárias. ρ é definido na escala real nos modelos heterocedásticos.

Essas especificações refletem o equilíbrio entre realismo biológico e flexibilidade estatística, permitindo que as estimativas *a posteriori* sejam fortemente guiadas pelos dados, mas mantendo coerência com a teoria do crescimento animal.

3.2.5.2 Procedimento de Amostragem (MCMC)

Os modelos foram ajustados separadamente para machos e fêmeas, considerando idade e no peso uma padronização (centralização e escalonamento) para melhorar a eficiência numérica e a convergência das cadeias, revertendo-se a escala dos parâmetros de interesse na apresentação dos resultados.

As estimativas foram obtidas por Cadeias de Markov via Monte Carlo (MCMC), com o algoritmo Hamiltoniano Monte Carlo (HMC) na implementação *No-U-Turn Sampler* (NUTS), utilizando o programa *Stan* via pacote *rstan* (STAN DEVELOPMENT TEAM, 2025) do R (R CORE TEAM, 2025).

Cada modelo foi ajustado com 4 cadeias independentes, 2.000 iterações por cadeia, sendo as 1.000 iniciais destinadas ao *warm-up* (adaptação), resultando em 4.000 amostras *a posteriori* por parâmetro. Os valores iniciais dos parâmetros foram definidos por funções que ancoram quantidades biológicas plausíveis, obtidas por estimações frequentistas, favorecendo boa exploração inicial do espaço paramétrico.

Foram estimadas quatro famílias de erro (Normal, *t*-Student, Skew-normal e Skew-*t*) com versões homocedásticas e, quando indicado, heterocedásticas; nestas, a variância condicional σ_i^2 foi modelada como função da idade segundo estruturas aditiva, potência multiplicativa e exponencial multiplicativa conforme descrito na subseção anterior.

3.2.5.3 Diagnóstico de Convergência

A qualidade da amostragem e a convergência dos modelos foram rigorosamente avaliadas por diversos diagnósticos: o diagnóstico $\hat{R}$ de Gelman e Rubin (com valor máximo aceitável de $\leq 1,01$) indicou a convergência das cadeias (no qual valores próximos de 1 refletem uma satisfatória *Potential Scale Reduction Factor*) (GELMAN; RUBIN, 1992); os Tamanhos Efetivos de Amostras (ESS) quantificaram a eficiência da amostragem; e o monitoramento das divergências, do *maximum tree depth* e do BMI (*Bayesian Fraction of Missing Information*) assegurou a qualidade da amostragem gerada pelo algoritmo NUTS (*No-U-Turn Sampler*) do Stan. Complementarmente, foram realizadas inspeções gráficas usuais (trajetória das cadeias, autocorrelação e densidades *a posteriori*) para verificar a mistura adequada e estacionariedade. Para a comparação preditiva, calculou-se o LOOIC (*Leave-One-Out Information Criterion*) via PSIS-LOO (*Pareto Smoothed Importance Sampling LOO*) a partir dos valores de log-verossimilhança, utilizando a eficiência relativa para corrigir a dependência entre amostras quando necessário.

3.2.5.4 Comparação dos Modelos

A comparação entre os modelos e distribuições de erro foi realizada com base no *Leave-One-Out Information Criterion* (LOOIC) (VEHTARI; GELMAN; GABRY, 2017), calculado via *Pareto-Smoothed Importance Sampling* (PSIS-LOO). Esse critério é totalmente Bayesiano e estima a capacidade preditiva fora da amostra, penalizando modelos excessivamente complexos. O LOOIC apresenta desempenho superior ao WAIC em termos de estabilidade numérica e diagnóstico de pontos influentes, sendo particularmente indicado para modelos hierárquicos e não lineares ajustados por MCMC.

Os resultados foram interpretados considerando que menores valores de LOOIC indicam melhor ajuste preditivo. As comparações foram realizadas separadamente para machos e fêmeas, identificando o modelo e a distribuição de erro com melhor desempenho para cada sexo.

3.3 Resultados e discussão

A Tabela 4 apresenta os valores médios, desvios-padrão e o número de observações (n) do peso (em gramas) dos animais em diferentes grupos etários (G_1 a G_{10}), definidos com base nos decis, cobrindo o período de 52 a 322 dias de idade. Observa-se uma tendência clara de aumento da média e da variabilidade do peso com o avanço da idade para ambos os sexos, padrão condizente com a fase de crescimento acelerado observada neste estudo. Contudo, ressalta-se que essa relação positiva entre média e variância pode não ser estritamente monótona em estágios de vida mais tardios (ex: peixes acima de 2 kg), os quais não foram contemplados no domínio experimental deste trabalho. À medida que os indivíduos se aproximam do peso

assintótico (β_1), a variabilidade tende a se estabilizar ou sofrer um decréscimo em função do limite biológico de crescimento da espécie.

Tabela 4 – Medidas descritivas para as os pesos em gramas de tilápias por grupo decil de idade e sexo.

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
Idade mín. (dias)	52	78	101	169	180	200	207	283	292	322
Idade máx. (dias)	77	96	165	177	199	206	281	290	315	322
Machos										
Média (g)	10,4	16,0	110,0	160,0	199,0	244,0	391,0	521,0	654,0	745,0
Desvio-padrão	2,6	5,2	42,4	43,8	57,9	61,3	144,0	122,0	153,0	145,0
n	348	230	316	379	274	170	321	367	286	169
Fêmeas										
Média (g)	9,7	14,3	83,0	124,0	154,0	187,0	251,0	345,0	430,0	506,0
Desvio-padrão	3,0	4,4	33,4	37,1	44,2	51,8	89,7	93,8	113,0	117,0
n	287	216	228	366	237	165	230	356	247	163

Na comparação entre sexos, verifica-se que as fêmeas apresentam consistentemente um menor peso que os machos. Esta diferença se acentua nos grupos etários mais avançados, culminando em um peso dos machos (745,0 g em G_{10}) superior ao das fêmeas (506,0 g em G_{10}). Além disso, a variabilidade do peso, representada pelo desvio-padrão, aumenta substancialmente, especialmente a partir do grupo G_6 em ambos os sexos, sugerindo uma forte heterogeneidade na taxa de crescimento dos indivíduos.

Esses padrões são reforçados pelos gráficos de caixa apresentados na Figura 1, que ilustram a distribuição dos pesos por grupo etário. Nota-se o aumento da mediana e, crucialmente, o aumento da dispersão ao longo do tempo, além da presença de assimetria positiva nos grupos mais avançados. Os dados indicam não apenas um crescimento acelerado nos grupos finais, mas também maior variabilidade individual, especialmente entre os machos. Tais evidências apontam para a necessidade de modelos que considerem estruturas de erro heterocedástico (justificado pelo aumento da variância com a idade) e distribuições assimétricas na análise de crescimento.

Zardin et al. (2019) observaram diferenças significativas entre os sexos em tilápias de diferentes grupos genéticos, reforçando a importância de separar os modelos de crescimento por sexo. Os autores justificaram que a presença de dimorfismo sexual está frequentemente associada a fatores hormonais, comportamentais e fisiológicos que impactam a eficiência alimentar e a alocação energética entre crescimento somático e reprodução, aspectos, com implicações diretas no manejo e na seleção genética em piscicultura.

Os valores de LOOIC obtidos para os diferentes modelos homocedásticos de crescimento

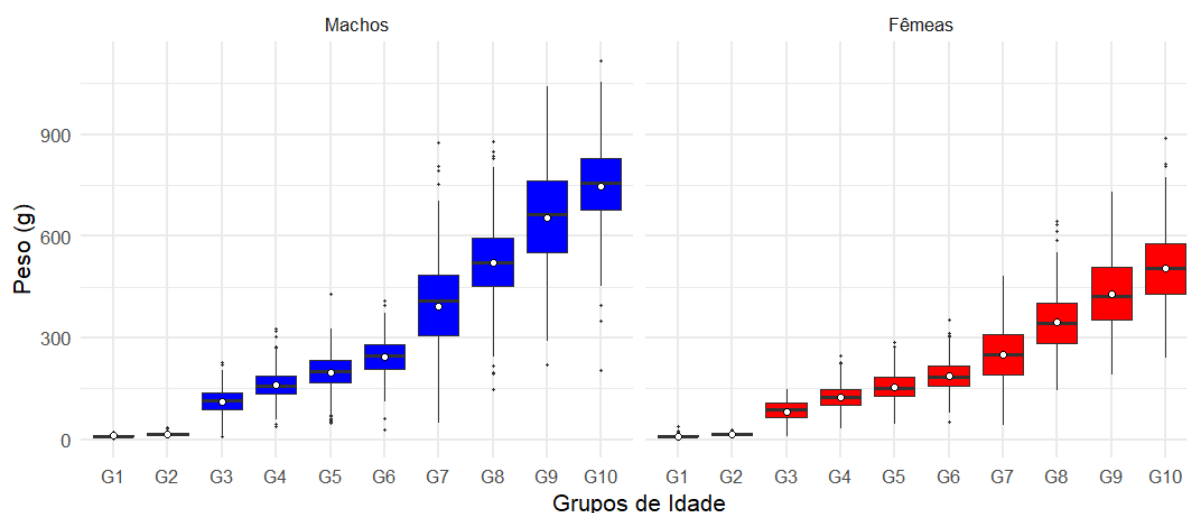


Figura 1 – Gráficos de caixa do peso por grupos decis de idade de tilápias, por sexo.

e distribuições de erro são apresentados na Tabela 5. Observa-se que o modelo de Von Bertalanffy Segmentado (VBS) apresentou os menores valores de LOOIC tanto para machos quanto para fêmeas, indicando o melhor desempenho preditivo entre as estruturas testadas. Essa superioridade sugere que a segmentação da taxa de crescimento permite capturar de forma mais realista as mudanças fisiológicas que ocorrem ao longo do desenvolvimento, sobretudo nas fases de transição entre crescimento acelerado e estabilização.

Tabela 5 – Valores do LOOIC para os modelos homocedásticos segundo a distribuição dos erros e o sexo dos animais.

Modelo	Sexo	Distribuição			
		N	SN	T	ST
Logístico	Macho	5.984	3.750	5.999	3.765
	Fêmea	5.387	3.779	5.400	3.792
Gompertz	Macho	5.974	3.715	5.998	3.730
	Fêmea	5.383	3.764	5.396	3.775
Brody	Macho	3.443	3.443	3.272	—*
	Fêmea	3.297	3.298	2.955	—*
Von Bertalanffy	Macho	3.442	3.367	3.266	3.179
	Fêmea	3.296	3.093	2.916	2.829
Von Bertalanffy Segmentado	Macho	2.210	2.158	1.407	1.495
	Fêmea	2.555	2.478	2.007	2.037

N: Normal, T: *t*-Student, SN: Skew-normal, SN: Skew-*t*;

*O modelo de Brody Skew-*t* não atingiu a convergência.

Além da superioridade do modelo VBS, é fundamental observar a distribuição de erro. Entre as distribuições de erro, a *t*-Student e a Skew-*t* apresentaram os melhores ajustes em comparação com as distribuições simétricas (Normal e Skew-normal). O modelo *t*-Student para erros homocedásticos apresentou o menor LOOIC (1.407 para machos e 2.007 para fêmeas),

indicando o melhor desempenho preditivo entre os modelos avaliados. Essa distribuição se mostrou mais adequada para descrever a variabilidade dos dados de crescimento, especialmente pela sua capacidade de lidar com valores extremos.

De modo geral, as diferenças observadas nos valores de LOOIC indicam que o uso do modelo VBS, aliado a distribuições de erro robustas, aprimora a capacidade preditiva e a representatividade biológica das curvas ajustadas para o crescimento de tilápias.

A Figura 2 apresenta a comparação visual das curvas de crescimento ajustadas pelos diferentes modelos homocedásticos (com erro t -Student) para machos e fêmeas. A inspeção visual confirma que o modelo VBS é o único que se ajusta de forma satisfatória ao longo de todo o ciclo de vida, especialmente capturando a fase inicial de crescimento lento e o rápido crescimento intermediário.

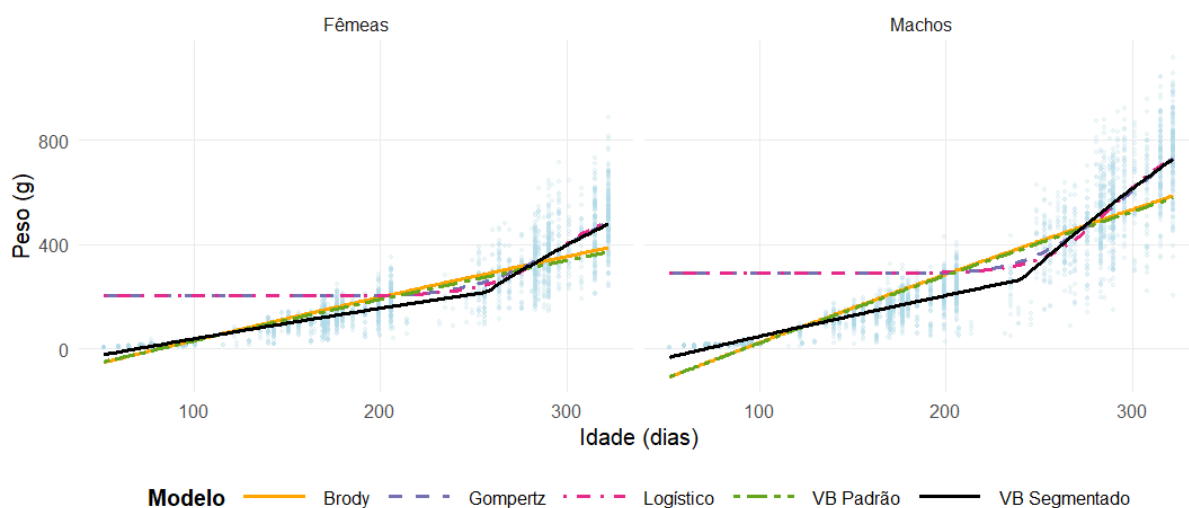


Figura 2 – Curvas de crescimento ajustadas pelos modelos homocedásticos t -Student para tilápias, por sexo.

Para ambos os sexos, o modelo com erro t -Student apresentou os menores valores de LOOIC, indicando melhor desempenho preditivo entre os modelos avaliados. Essa distribuição se mostrou mais adequada para capturar a variabilidade dos dados de crescimento, especialmente pela capacidade de lidar com valores extremos, comuns em experimentos zootécnicos.

Considerando os modelos de Von Bertalanffy Segmentado homocedásticos, para os machos, o modelo t -Student apresentou o menor LOOIC (1.407), seguido do modelo Skew- t (1.495), enquanto os modelos Normal (2.210) e Skew-normal (2.158) exibiram desempenho inferior. Para as fêmeas, observou-se padrão semelhante, com o modelo t -Student apresentando o menor LOOIC (2.007), seguido pelo Skew- t (2.037) Tabela 5.

O modelo Skew-normal apresentou desempenho próximo ao Normal, mas sem ganhos expressivos em termos de predição. Apesar da maior flexibilidade, o modelo Skew- t não superou o desempenho do t -Student, sugerindo que a inclusão simultânea de parâmetros de assimetria e

caudas pesadas pode ter levado a uma penalização maior no LOOIC. Assim, o modelo *t*-Student mostrou-se o mais parcimonioso e eficiente para representar o crescimento de machos e fêmeas de tilápia, conciliando bom ajuste e complexidade moderada.

Resultados semelhantes foram observados por [Faccin e Rossi \(2024\)](#), que utilizaram o modelo de Gompertz com distribuição *t*-Student para descrever o crescimento de codornas e encontraram desempenho superior em termos de WAIC em comparação com modelos normais. A presença de caudas pesadas foi melhor descrita pelo modelo *t*-Student, corroborando os achados do presente estudo.

A superioridade do modelo *t*-Student sobre o Normal indica que os dados de crescimento apresentam caudas mais pesadas do que a distribuição Normal prevê. Isto é, biologicamente, isso significa que há uma frequência maior de indivíduos com padrões de crescimento atípicos, seja muito lento ou muito rápido, do que o esperado sob a suposição de normalidade. Esse comportamento é comum em populações biológicas, em que fatores genéticos e ambientais interagem de forma heterogênea. A inclusão do parâmetro de assimetria nos modelos *Skew* não resultou em melhora no ajuste em relação ao modelo *t*-Student, sugerindo que a principal violação da normalidade nos resíduos está associada à curtose (caudas pesadas), e não à assimetria.

A Tabela 6 apresenta as estimativas dos parâmetros do modelo de Von Bertalanffy Segmentado homocedásticos ajustado para machos e fêmeas de tilápia sob diferentes distribuições de erro. Observa-se variação considerável nas estimativas dos parâmetros conforme a distribuição assumida, especialmente para o peso assintótico (β_1) e desvio-padrão do erro (σ), reforçando a importância de considerar estruturas de erro alternativas na modelagem do crescimento.

Embora o modelo *Skew-t* homocedástico seja teoricamente mais flexível, seu desempenho inferior em termos de LOOIC (Tabela 5) pode estar relacionado à penalização decorrente da maior complexidade do modelo, especialmente em situações com menor evidência de assimetria acentuada. [Ferreira et al. \(2020\)](#) destacam que modelos com muitos parâmetros adicionais devem ser avaliados com cautela quanto ao equilíbrio entre ajuste e parcimônia.

Com base no modelo *t*-Student, que apresentou o melhor desempenho preditivo (menor LOOIC - Tabela 5), é possível interpretar os parâmetros de crescimento estimados. Nos machos, o peso assintótico (β_1) foi estimado em 2.437 g (1.958 e 3.035 g). Para as fêmeas, o valor médio estimado foi de 1.326 g (1.024 a 1.723 g) (Tabela 6).

Esses resultados evidenciam um acentuado dimorfismo sexual no crescimento, com os machos apresentando um potencial de peso assintótico (β_1) superior ao das fêmeas. Além disso, os intervalos de credibilidade não se sobrepõem, o que confere alta confiança estatística a essa diferença. Tal resultado é reforça sob a perspectiva produtiva, que populações exclusivamente masculinas pode resultar em maior rendimento de biomassa.

Os parâmetros de taxa de crescimento inicial (β_2) e final (β_4) apresentaram valores médios

Tabela 6 – Estimativas dos parâmetros do modelo de Von Bertalanffy Segmentado homocedástico para tilápias, por sexo, sob diferentes distribuições de erro.

		Machos		Fêmeas	
		Média (DP)	ICr[2,5%; 97,5%]	Média (DP)	ICr[2,5%; 97,5%]
N	β_1	2.050 (241)	[1.629; 2.548]	1.225 (169)	[952; 1.594]
	β_2	0,0008 (0,0001)	[0,0007; 0,0011]	0,0011 (0,0002)	[0,0008; 0,0015]
	β_4	0,0037 (0,0006)	[0,0027; 0,0050]	0,0049 (0,0010)	[0,0033; 0,0070]
	σ	0,356 (0,047)	[0,347; 0,650]	0,403 (0,058)	[0,392; 0,415]
	τ	241,000 (1,700)	[237,730; 244,320]	256,870 (1,650)	[253,83; 260,45]
T	β_1	2.437 (274)	[1.958; 3.035]	1.326 (178)	[1.024; 1.723]
	β_2	0,0007 (0,0002)	[0,0005; 0,0009]	0,0010 (0,0001)	[0,0007; 0,0012]
	β_4	0,0030 (0,0004)	[0,0023; 0,0040]	0,0043 (0,0005)	[0,0029; 0,0059]
	σ	0,163 (0,058)	[0,152; 0,174]	0,206 (0,008)	[0,191; 0,221]
	ν	1,700 (0,091)	[1,530; 1,890]	1,940 (0,124)	[1,710; 2,190]
SN	β_1	2.182 (248)	[1.746; 2.711]	1.132 (165)	[844; 1.484]
	β_2	0,0008 (0,0001)	[0,0006; 0,0011]	0,0010 (0,0001)	[0,0008; 0,0013]
	β_4	0,0037 (0,0006)	[0,0027; 0,0051]	0,0046 (0,0009)	[0,0031; 0,0067]
	σ	0,467 (0,011)	[0,444; 0,491]	0,554 (0,014)	[0,526; 0,587]
	λ	-1,400 (0,111)	[-1,620; -1,180]	1,720 (0,135)	[1,460; 1,990]
ST	β_1	2.359 (273)	[1.851; 2.933]	1.243 (166)	[963; 1.611]
	β_2	0,0007 (0,0001)	[0,0006; 0,0010]	0,0010 (0,0001)	[0,0008; 0,0013]
	β_4	0,0031 (0,0005)	[0,0021; 0,0041]	0,0047 (0,0009)	[0,0031; 0,0068]
	σ	0,184 (0,032)	[0,172; 0,195]	0,212 (0,026)	[0,236; 0,258]
	λ	-0,058 (0,026)	[-0,103; -0,015]	0,103 (0,024)	[0,054; 0,132]
	ν	3,020 (0,083)	[3,000; 3,080]	3,040 (0,081)	[3,030; 3,090]
	τ	240,212 (1,545)	[237,215; 243,355]	256,879 (1,506)	[254,097; 260,370]

N: Normal, T: *t*-Student, SN: Skew-normal, SN: Skew-*t*;

DP: desvio-padrão, ICr: intervalo de credibilidade;

β_1 : Parâmetro assintótico (peso máximo ou limite superior de crescimento);

β_2 : Taxa de crescimento inicial (segmento 1);

β_4 : Taxa de crescimento final (Segmento 2);

σ : Desvio-padrão residual;

λ : Parâmetro de assimetria;

ν : Graus de liberdade;

τ : Idade de transição.

de 0,0007 e 0,0030 para os machos, e de 0,0010 e 0,0043 para as fêmeas, respectivamente. Esses resultados sugerem que o crescimento inicial é mais lento, com aceleração progressiva até o ponto de transição (τ), após o qual ocorre uma desaceleração moderada, comportamento característico do modelo segmentado. Apesar de pequenas diferenças entre os sexos, o padrão geral indica uma dinâmica de crescimento biologicamente coerente, com as fêmeas alcançando o platô em idades inferiores, com menor peso assintótico.

A relevância biológica do ponto de transição (τ), estimado em ≈ 234 dias para machos e

≈ 247 dias para fêmeas, pode ser diretamente associada à sazonalidade e à temperatura da água. A tilápia-do-Nilo é um organismo pecilotérmico, o que significa que seu metabolismo, incluindo as taxas de anabolismo (ganho de massa) e catabolismo (perda de massa) que governam o modelo de Von Bertalanffy, é diretamente regulado pela temperatura ambiente. O período em torno de 230 a 250 dias de cultivo em sistemas subtropicais, como o local de estudo, frequentemente coincide com a chegada das temperaturas mínimas do inverno. A queda térmica provoca uma desaceleração metabólica e, consequentemente, uma redução na ingestão de ração e na taxa de crescimento. Portanto, o ponto τ não apenas segmenta matematicamente a curva, mas demarca o momento em que a taxa de crescimento (β_4) é afetada por um estressor ambiental significativo. Esse achado corrobora a literatura que destaca a relação entre a temperatura da água e a taxa de crescimento da tilápia, sendo um fator limitante em períodos de frio (AZAZA; AL-JEDDA; AL-HASSANEIN, 2010). A modelagem segmentada, portanto, oferece um instrumento preciso para identificar esse período crítico de manejo.

O parâmetro σ , que representa o desvio-padrão do erro, foi inferior nos modelos com distribuição *t*-Student (0,163 para machos e 0,206 para fêmeas), refletindo maior precisão das estimativas e menor variabilidade residual. O parâmetro ν , que controla o grau de liberdade da distribuição *t*-Student, apresentou valores baixos (1,70 para machos e 1,94 para fêmeas), indicando a presença de caudas pesadas e assimétricas confirmando a robustez do modelo frente a possíveis observações discrepantes. Esse resultado reforça a adequação do uso da distribuição *t*-Student, como também observado por Lachos, Bandyopadhyay e Garay (2011), que destacaram sua resistência à influência de *outliers* e à variabilidade elevada em dados biológicos.

Nos modelos que incorporam assimetria (Skew-normal e Skew-*t*), o parâmetro de assimetria (λ) indicou direções opostas entre os sexos, com valores significativos, negativos para machos ($\lambda = -1,40$, $ICr95\% = [-1,620; -1,180]$) e positivos para fêmeas ($\lambda = 1,72$, $ICr95\% [1,460; 1,990]$). Esses resultados sugerem leve assimetria à esquerda nos resíduos dos machos e à direita nas fêmeas, refletindo padrões distintos de dispersão. No entanto, como o modelo *t*-Student apresentou melhor desempenho baseado no menor LOOIC (Tabela 5), conclui-se que a principal violação da normalidade está relacionada à curtose (caudas pesadas), e não à assimetria, sendo desnecessária a complexidade adicional de modelos assimétricos.

Apesar da superioridade do Von Bertalanffy Segmentado *t*-Student homocedástico, a inspeção gráfica dos resíduos e a Tabela 4, que mostra o aumento do desvio-padrão com a idade, indicam a necessidade de modelar a heterocedasticidade. A próxima etapa da análise concentrou-se, portanto, em comparar as diferentes estruturas de variância aplicadas ao Von Bertalanffy Segmentado *t*-Student, heterocedástica aditiva, multiplicativa potência e exponencial, apresentadas na seção 3.2.4.

A comparação das estruturas de variância, apresentada na Tabela 7, demonstrou que o

modelo multiplicativo potência foi o que alcançou o melhor desempenho preditivo (menor LOOIC). Esta superioridade justifica a escolha desta estrutura para modelar o erro.

Tabela 7 – Comparação preditiva das estruturas de heterocedasticidade aplicadas ao modelo Von Bertalanffy Segmentado t -Student, por sexo.

Função de Variância	Machos		Fêmeas	
	ρ [ICr 95%]	LOOIC	ρ [ICr 95%]	LOOIC
Aditiva ¹	0,27 [0,25–0,28]	31.641	0,22 [0,21–0,23]	26.329
Multiplicativa Potência ²	1,85 [1,79–1,91]	31.028	1,78 [1,71–1,85]	25.840
Multiplicativa Exponencial ³	0,01 [0,01–0,01]	31.147	0,01 [0,01–0,01]	25.934

ICr: intervalo de credibilidade; ρ : parâmetro de heterocedasticidade; t_i : idade.

$$^1 \sigma_i^2 = \sigma^2 + \rho t_i;$$

$$^2 \sigma_i^2 = \sigma^2 t_i^\rho;$$

$$^3 \sigma_i^2 = \sigma^2 \exp(\rho t_i).$$

O parâmetro de heterocedasticidade (ρ) estimados no modelo multiplicativo potência (machos: 1,85; fêmeas: 1,78) foram significativamente diferentes de zero, com intervalos de credibilidade ([1,79; 1,91] e [1,71; 1,85]) que não contêm a unidade ($\rho = 0$), rejeitando-se formalmente a hipótese de homocedasticidade e confirmando a necessidade do termo de variância condicional. Ademais, o valor de ρ próximo de 2,0 (XIE; WANG; ZOU, 2000) indica uma heterocedasticidade forte.

A Tabela 8 apresenta as estimativas *a posteriori* do modelo final, Von Bertalanffy Segmentado t -Student heterocedástico multiplicativo tipo potência. Os resultados confirmam um acentuado dimorfismo sexual, com os machos alcançando um peso assintótico teórico (β_1) estimado em 2.610 g, significativamente superior ao das fêmeas (1.730 g). A transição de fase de crescimento (τ) ocorre em idades semelhantes, mas ligeiramente antes nos machos (234 dias) do que nas fêmeas (247 dias).

Os parâmetros de taxa de crescimento final (β_4) são consistentemente superiores às taxas iniciais (β_2) em ambos os sexos, confirmando que a segmentação capturou a aceleração pós-juvenil. O parâmetro ν (Graus de Liberdade) com valores acima de 10 (18,5 para machos e 12,5 para fêmeas) valida a robustez do modelo t -Student, enquanto o σ_e (Variância Base) estabelece o nível de variação residual na escala real.

A Figura 3 apresenta as curvas médias de crescimento, bem como o Intervalo de Credibilidade de 95% (área sombreada), ajustadas pelo modelo de Von Bertalanffy Segmentado heterocedástico multiplicativo potência para machos e fêmeas, com extrapolação até 600 dias. A inspeção visual reforça a precisão e a necessidade das estruturas modeladas: o dimorfismo sexual é evidente, com os machos (curva azul) atingindo um potencial assintótico significativamente superior ao das fêmeas (curva vermelha), conforme quantificado pelos parâmetros β_1 (peso assintótico). A segmentação é claramente marcada pelas linhas verticais pontilhadas, indicando o ponto de transição (τ) no qual a primeira taxa de crescimento muda para a segunda (machos $\tau \approx 241$ dias; fêmeas $\tau \approx 257$ dias). O comportamento mais importante do modelo,

Tabela 8 – Estimativas *a posteriori* dos parâmetros do modelo de Von Bertalanffy Segmentado *t*-Student heterocedástico multiplicativo potência, por sexo.

Parâmetros	Machos		Fêmeas	
	Média (DP)	ICr [2,5%; 97,5%]	Média (DP)	ICr [2,5%; 97,5%]
β_1	2.610 (275)	[2.120; 3.090]	1.730 (188)	[1.390; 2.050]
β_2	0,0006 (0,0001)	[0,0005; 0,0007]	0,0006 (0,0001)	[0,0005; 0,0008]
β_4	0,0027 (0,0004)	[0,0020; 0,0033]	0,0029 (0,0004)	[0,0021; 0,0036]
τ	234,000 (2,080)	[230,000; 238,000]	247,000 (2,230)	[243,000; 251,000]
σ_e	0,0035 (0,0006)	[0,0026; 0,0045]	0,0037 (0,0007)	[0,0025; 0,0049]
ρ	1,850 (0,030)	[1,790; 1,910]	1,780 (0,036)	[1,710; 1,850]
ν	18,500 (5,970)	[10,800; 26,000]	12,500 (3,720)	[7,970; 17,000]

a heterocedasticidade, é validado pela observação do incremento na amplitude do ICr 95% ao longo do tempo.

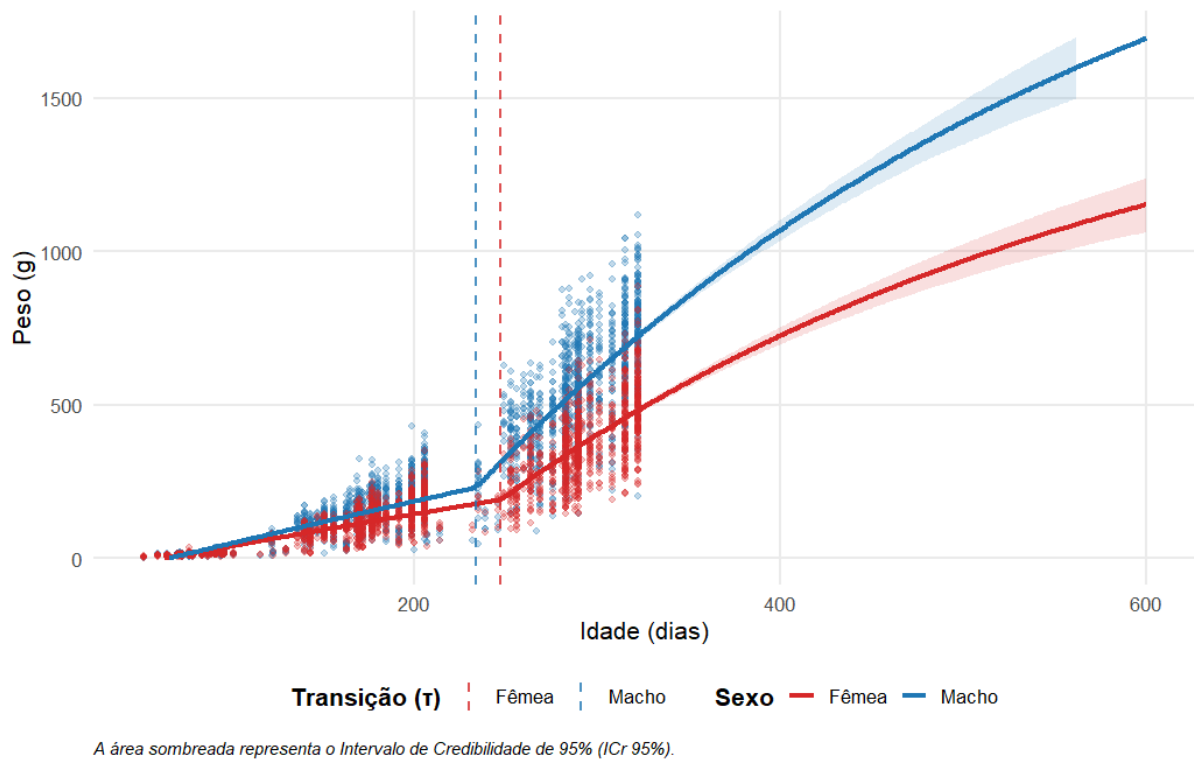


Figura 3 – Curvas de crescimento do Modelo de Von Bertalanffy Segmentado heterocedástico multiplicativo potência, por sexo, com extrapolação até 600 dias.

Recomendamos que estudos futuros em piscicultura e seleção genética adotem essa estrutura para obter inferências mais precisas, especialmente na gestão da uniformidade de lotes e na avaliação do potencial máximo de crescimento, mitigando o viés introduzido pelas suposições simplistas de normalidade e homocedasticidade nos dados biológicos.

3.4 Conclusão

A aplicação do modelo de Von Bertalanffy Segmentado com erros t -Student heterocedástico multiplicativo potência demonstrou ser a ferramenta mais representativa para o estudo do crescimento de tilápias-do-Nilo.

Capítulo 4

Considerações finais

Neste trabalho, avaliou-se o crescimento somático de tilápias-do-Nilo (*Oreochromis niloticus*) por meio do ajuste de modelos não lineares de crescimento: Brody, Gompertz, Logístico, Von Bertalanffy e Von Bertalanffy Segmentado sob a abordagem Bayesiana. Foram consideradas diferentes distribuições de erro, incluindo modelos simétricos (Normal e t -Student) e assimétricos (Skew-normal e Skew- t), em versões homocedásticas e heterocedásticas (aditivo, multiplicativo potência e multiplicativo exponencial), a fim de investigar o desempenho e a adequação de cada especificação na descrição do crescimento dos peixes.

Os resultados indicaram que o modelo de Von Bertalanffy Segmentado apresentou o melhor desempenho preditivo entre os modelos avaliados, sendo o mais capaz de representar as fases de aceleração e desaceleração do crescimento. Dentre as distribuições Homocedástica de erro testadas, a t -Student destacou-se como a mais adequada, evidenciando a presença de caudas pesadas nos resíduos. Essa característica reforça a robustez do modelo frente à ocorrência de observações extremas, comuns em experimentos zootécnicos.

As análises também revelaram um dimorfismo sexual acentuado, com machos apresentando peso assintótico superior ao das fêmeas, além de maior variabilidade residual. Esses resultados estão em consonância com a literatura, que descreve diferenças marcantes no potencial de crescimento entre os sexos na piscicultura.

Por outro lado, a inclusão de estruturas de variância heterocedásticas (aditiva e multiplicativas) resultou em melhorias na qualidade do ajuste. Esse comportamento indica que de fato há heterocedasticidade no fenômeno de crescimento da espécie tilápias-do-Nilo, o que é um achado plausível biologicamente.

A utilização do ambiente computacional *Stan*, mostrou-se eficaz para a estimação dos modelos propostos, permitindo flexibilidade na especificação das estruturas de erro, facilidade na verificação dos diagnósticos de convergência e obtenção de estimativas precisas e interpretáveis.

Conclui-se que a modelagem Bayesiana aplicada à curva de Von Bertalanffy Segmentado com erro *t*-Student heterocedástico multiplicativo potência constitui uma abordagem robusta e biologicamente coerente para descrever o crescimento de tilápias-do-Nilo. Essa combinação oferece um equilíbrio adequado entre ajuste, parcimônia e interpretabilidade, sendo uma ferramenta útil para estudos de crescimento, seleção genética e manejo produtivo na piscicultura.

Diante dos resultados e das discussões apresentadas, recomendam-se para estudos futuros a incorporação de modelos hierárquicos e a inclusão de covariáveis ambientais, visando capturar a variabilidade entre tanques, linhagens e sistemas de manejo, o que permitiria uma decomposição mais precisa da estrutura de variância. Adicionalmente, sugere-se a utilização de distribuições *a priori* informativas, fundamentadas em metanálises ou conhecimentos biológicos consolidados da espécie, para otimizar a convergência e reduzir a amplitude dos intervalos de credibilidade em parâmetros de maior complexidade estimativa. Por fim, propõe-se a ampliação do esforço amostral em idades mais avançadas (acima de 350 dias), possibilitando uma descrição mais fidedigna do platô de crescimento, o que reduziria a incerteza preditiva observada na fase de extrapolação e permitiria refinar a estimativa do peso assintótico (β_1).

Referências

- AZAZA, M. S.; AL-JEDDA, S. A.; AL-HASSANEIN, S. Modeling the effects of water temperature on growth rates, gastric evacuation and the return of appetite in juvenile Nile tilapia, *Oreochromis niloticus* L. *Journal of Agricultural Science*, Canadian Center of Science and Education, v. 2, n. 4, p. 191–201, 2010. Disponível em: <<https://www.ccsenet.org/journal/index.php/jas/article/view/0/43274>>.
- AZZALINI, A. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, Wiley for the Board of the Foundation of the Scandinavian Journal of Statistics, v. 12, n. 2, p. 171–178, 1985.
- BERTALANFFY, L. V. A quantitative theory of organic growth (inquiries on growth laws. ii). *Human Biology*, v. 10, n. 2, p. 181–213, 1938.
- BEVERTON, R. J. H.; HOLT, S. J. *On the dynamics of exploited fish populations*. [S.l.]: H. M. Stationery Office, Ministry of Agriculture, Fisheries and Food, 1957. v. 19. (Fishery investigations, v. 19).
- BRODY, S. *Bioenergetics and Growth*. New York: Reinhold Publishing Corporation, 1945.
- CAMPOS, A. M. de; FILHO, M. G. de A. *Uma abordagem bayesiana para modelos não lineares na presença de assimetria e heteroscedasticidade*. 2011. Workshop de Teses e Dissertações Defendidas, USP/ICMC. Disponível em: <<https://www.icmc.usp.br>>.
- CHENG, C.-L.; RIU, J. On estimating linear relationships when both variables are subject to heteroscedastic measurement errors. *Technometrics*, v. 48, n. 4, p. 511–519, 2006.
- CHENG, J.; RUE, H. The use of generalized additive models for estimation of growth curves. *Journal of the Royal Statistical Society Series C: Applied Statistics*, Wiley Online Library, v. 55, n. 1, p. 1–13, 2006.
- CRUZ, R. D. L.; BRANCO, M. D. Bayesian analysis for nonlinear regression model under skewed errors, with application in growth curves. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, v. 51, n. 4, p. 588–609, 2009.
- DAVIDIAN, M.; GILTINAN, D. M. *Nonlinear Models for Repeated Measurement Data*. [S.l.]: Chapman & Hall, 1995.
- DINIZ, C.; LOUZADA-NETO, F.; MORITA, L. The multiplicative heteroscedastic von bertalanffy model. *Brazilian Journal of Probability and Statistics*, v. 26, n. 1, p. 71–81,

2012. Available online. Disponível em: <<https://www.imstat.org/journals-and-publications/brazilian-journal-of-probability-and-statistics/>>.

FACCIN, G. B.; ROSSI, R. M. Bayesian modeling of the gompertz curve for meat quails growth data considering different error distributions. *Brazilian Journal of Biometrics*, v. 42, n. 1, p. 1–13, 2024. In press.

FERREIRA, L. A. et al. Inference and diagnostics for heteroscedastic nonlinear regression models under skew scale mixtures of normal distributions. *Statistical Papers*, v. 61, p. 777–805, 2020.

FIALHO, F. B. *Modelos matemáticos para descrever o crescimento de animais*. Porto Alegre: Editora Universitária da UFRGS, 1999. Título genérico — confirme na sua fonte se for diferente.

GALLANT, A. R. *Nonlinear Statistical Models*. [S.l.]: John Wiley & Sons, 1987.

GELMAN, A. et al. *Bayesian Data Analysis*. 3rd. ed. [S.l.]: CRC Press, 2020.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science*, v. 7, n. 4, p. 457–472, 1992.

GODOI, L. G. d. *A distribuição t-assimétrica univariada: propriedades e inferência*. Dissertação (MA thesis) — Universidade de São Paulo, São Paulo, 2007.

GOMPERTZ, B. On the nature of the function expressive of the law of human mortality... *Philosophical Transactions of the Royal Society of London*, v. 115, p. 513–585, 1825.

GUPTA, A. K.; XU, H. On the skew t distribution. *Journal of Multivariate Analysis*, v. 84, n. 2, p. 416–430, 2003.

HARVEY, A. C. Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, Wiley, v. 44, n. 3, p. 461–465, 1976.

KAPS, M.; LAMBERSON, W. R. *Biostatistics for Animal Science*. [S.l.]: CABI Publishing, 2004.

LACHOS, V. H.; BANDYOPADHYAY, D.; GARAY, A. M. Heteroscedastic nonlinear regression models based on scale mixtures of skew-normal distributions. *Statistics & Probability Letters*, v. 81, n. 8, p. 1208–1217, 2011.

LEE, S. H. et al. Estimating growth parameters from size-at-age data of fish using a nonlinear mixed-effects model. *Fisheries Research*, v. 222, p. 105436, 2020.

LOUZADA, F.; FERREIRA, P. H.; DINIZ, C. A. R. Skew-normal distribution for growth curve models in presence of a heteroscedasticity structure. *Biometrical Letters*, v. 51, n. 1, p. 59–70, 2014.

MANGUEIRA, R. A. F. et al. Logistic model considering different error distributions applied in maize height data. *Brazilian Journal of Biometrics*, v. 34, n. 2, p. 317–333, 2016. Disponível em: <<http://jbiom.org.br/ojs/index.php/jbiom/article/view/229>>.

MAZUCHELI, J.; SOUZA, R. M. de; PHILIPPSEN, A. S. Modelo de crescimento de gompertz na presença de erros normais heterocedásticos: um estudo de caso. *Revista Brasileira de Biometria*, v. 29, n. 1, p. 91–101, 2011.

- PEIXE BR. *Anuário da Piscicultura 2024*. 2024. <<https://www.peixebr.com.br/anuario>>. Associação Brasileira da Piscicultura.
- PINHEIRO, J. C.; BATES, D. M. *Mixed-effects models in S and S-PLUS*. [S.l.]: Springer Science & Business Media, 2000.
- R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2025. Disponível em: <<https://www.R-project.org/>>.
- ROSSI, R. M. et al. Curvas de crescimento em codornas de corte e postura: uma perspectiva bayesiana. *Semina: Ciências Agrárias*, v. 38, n. 4, suplemento 1, p. 2743–2754, 2017.
- ROSSI, R. M.; SANTOS, L. A. d. Modelagem bayesiana para curvas de crescimento de codornas assumindo assimetria nos erros. *Semina: Ciências Agrárias*, v. 35, n. 3, p. 1637–1648, 2014.
- SANTOS, V. B. d. et al. Aplicação de modelos não lineares ao crescimento de tilápia do nilo. *Revista Brasileira de Zootecnia*, v. 36, n. 4, p. 912–918, 2007.
- SEBER, G. A. F.; WILD, C. J. *Nonlinear Regression*. New York: John Wiley & Sons, 1989.
- SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 64, n. 4, p. 583–639, 2002.
- STAN DEVELOPMENT TEAM. *Stan User's Guide, Version 2.32*. 2025. Disponível em: <<https://mc-stan.org/docs>>.
- TRINDADE, W.; PRATA, R. A. C. Análise dos modelos gompertz, verhulst e de von bertalanffy para o crescimento linear dos machos albinos da tilápia do nilo. *REMAT: Revista Eletrônica da Matemática*, v. 10, n. 1, p. e3008, 2024. Disponível em: <<https://periodicos.ifrs.edu.br/index.php/REMAT/article/view/6762>>.
- VEHTARI, A.; GELMAN, A.; GABRY, J. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, v. 27, n. 5, p. 1413–1432, 2017.
- VEHTARI, A. et al. Rank-normalization, folding, and localization: An improved r-hat for assessing convergence of mcmc. *arXiv preprint arXiv:1903.08008*, 2019. Disponível em: <<https://arxiv.org/abs/1903.08008>>.
- VERHULST, P. F. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique*, v. 10, p. 113–121, 1838.
- XIE, F.-C.; WEI, B.-C.; LIN, J.-G. Homogeneity diagnostics for skew-normal nonlinear regression models. *Statistics and Probability Letters*, v. 79, n. 6, p. 821–827, 2009.
- XIE, G.; WANG, J.; ZOU, G. Analysis of nonlinear growth models with heteroscedasticity. *Acta Agronomica Sinica*, v. 26, n. 4, p. 385–390, 2000.
- ZARDIN, A. M. da S. O. et al. Growth curves by gompertz nonlinear regression model for male and female nile tilapias from different genetic groups. *Aquaculture*, v. 511, p. 734243, 2019.
- ZARZAR, C. A. *Estatística aplicada à aquicultura: desafios e métodos para a modelagem de crescimento do camarão cinza*. Tese (Tese de Doutorado) — Universidade Federal do Ceará, Fortaleza, CE, 2023. Disponível em: <<https://repositorio.ufc.br/handle/riufc/74361>>.

ZARZAR, L. H. et al. Evidence of parameters underestimation from nonlinear growth models for data classified as limited. *Aquaculture Reports*, v. 24, p. 101152, 2022.

Apêndices

APÊNDICE A

Análise exploratória dos dados

```

> library(dplyr)
> library(ggplot2)
> library(readxl)
> library(tidyverse)
>
> dados <- read_tsv("dados5obs.txt", col_types = cols(
+   # Categorias / IDs
+   id = col_factor(), idanimal = col_factor(), animal =
+     ↪ col_factor(),
+   pai = col_factor(), mae = col_factor(), sexo =
+     ↪ col_factor(),
+   tanque = col_factor(), igrupo = col_factor(), familia =
+     ↪ col_factor(),
+   # Numericas
+   idade = col_double(), # x
+   peso = col_double() # y
+ ))
>
> head(dados)
# A tibble: 6      14
id idanimal animal pai mae sexo tanque igrupo familia idade
  ↪ ct cp cc peso
<fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct> <dbl>
  ↪ <dbl> <dbl> <dbl> <dbl>
1 1 1 2456 800002      7000      7000      1 1 1 132  91 10      7.3
  ↪ 2.4  16
2 2 2 2456 800002      7000      7000      1 1 1 132 147 19      16
  ↪ 5    128

```

```

3 3 2456 800002      7000      7000      1 1 1 132 176 22      18
  ↪ 5.3 212
4 4 2456 800002      7000      7000      1 1 1 132 260 30.5 24.5
  ↪ 7.5 597
5 5 2456 800002      7000      7000      1 1 1 132 292 30      25
  ↪ 8.3 909
6 6 2578 800002      7000      7001      1 2 2 146  78  8      6.4
  ↪ 2.2  11
>
> # Filtrar para nao ter NA em idade e peso
> dados <- dados %>%
+   filter(!is.na(idade), !is.na(peso))
>
> # Criar decil de idade geral (mesma divis o para todos)
> dados <- dados %>%
+   mutate(decil_idade = cut(
+     idade,
+     breaks = quantile(idade, probs = seq(0, 1, 0.1), na.rm
  ↪ = TRUE),
+     labels = paste0("G", 1:10),
+     include.lowest = TRUE
+   ))
>
> estatisticas_por_sexo_decil <- dados %>%
+   group_by(sexo, decil_idade) %>%
+   summarise(
+     media_peso = mean(peso, na.rm = TRUE),
+     desvio_padrao = sd(peso, na.rm = TRUE),
+     n = n(),
+     .groups = "drop"
+   )
>
> library(ggplot2)
>
> # Mapeando a cor pelo sexo manualmente
> ggplot(dados, aes(x = decil_idade, y = peso, fill = sexo))
  ↪ +
+   geom_boxplot(outlier.size = 0.5) +
+   stat_summary(fun = mean, geom = "point", shape = 21,
  ↪ color = "black", fill = "white", size = 2) +
+   scale_fill_manual(
+     values = c("1" = "blue",

```

```
+           "2" = "red")
+   ) +
+   facet_wrap(~sexo, labeller = labeller(sexo = c("1" =
+     ↪ "Machos", "2" = "Femeas")))) +
+   labs(
+     title = "",
+     x = "Grupos de Idade",
+     y = "Peso (g)"
+   ) +
+   theme_minimal(base_size = 14) +
+   theme(legend.position = "none") # remove a legenda
```

APÊNDICE B

Modelo Von Bertalanffy *t-Student* *Homocedástico - Stan*

```

data {
  int<lower=0> N;
  vector[N] age;          // Idade PADRONIZADA
  vector[N] weight;       // Peso PADRONIZADO

  real min_age; real max_age;
  real mean_age; real sd_age;
  real mean_weight; real sd_weight;
}

parameters {
  // Parametros
  real<lower=0> beta1; // Peso assintotico (W_inf)
  real<lower=0> beta2; // Taxa de crescimento 1 (k1)
  real<lower=0> beta4; // Taxa de crescimento 2 (k2)
  real beta3;         // Idade teorica com peso zero (t0)
  real<lower=min_age, upper=max_age> tau; // Ponto de Transicao
    ↪ (tau)

  real<lower=0> sigma;
  real<lower=1> nu;    // Graus de liberdade
}

transformed parameters {
  vector[N] mu;
  // Padroniza t0 e tau para a escala do modelo

```

```

real t0_pad = (beta3 - mean_age) / sd_age;
real tau_pad = (tau - mean_age) / sd_age;

// Peso no ponto de transicao (escala padronizada)
real W_tau = beta1 * pow(1 - exp(-beta2 * (tau_pad - t0_pad)),
  ↪ 3);

for (i in 1:N) {
  if (age[i] <= tau_pad) {
    // Segmento 1: Von Bertalanffy elevado ao cubo
    mu[i] = beta1 * pow(1 - exp(-beta2 * (age[i] - t0_pad)),
      ↪ 3);
  } else {
    // Segmento 2: Transicao suave elevada ao cubo
    mu[i] = W_tau + (beta1 - W_tau) * pow(1 - exp(-beta4 *
      ↪ (age[i] - tau_pad)), 3);
  }
}
}

model {
  // Priors conforme Tabela 3
  beta1 ~ normal(0, 2);
  beta2 ~ lognormal(0, 1);
  beta4 ~ lognormal(0, 1);
  beta3 ~ normal(0, 50);
  tau ~ normal(mean_age, sd_age);
  sigma ~ cauchy(0, 1);
  nu ~ gamma(2, 0.1);

  // Likelihood: t-Student conforme seus resultados
  weight ~ student_t(nu, mu, sigma);
}

generated quantities {
  vector[N] log_lik;
  // Reversao de Escala para Parametros Reais
  real W_inf_real = beta1 * sd_weight + mean_weight;
  real k1_real = beta2 / sd_age;
  real k2_real = beta4 / sd_age;

  for (i in 1:N) {

```

```
    log_lik[i] = student_t_lpdf(weight[i] | nu, mu[i], sigma);  
  }  
}
```

APÊNDICE C

Modelo Von Bertalanffy *t-Student* Heterocedástico Multiplicativo Potência - *Stan*

```

data {
  int<lower=0> N;
  vector[N] x;          // Idade REAL (l_ij)
  vector[N] y;          // Peso REAL (y_ij)

  real mean_age; real sd_age;
  real mean_weight; real sd_weight;
  real min_age; real max_age;
}

parameters {
  // Parametros de Crescimento (Conforme Tabela 3)
  real<lower=0> beta1; // W_inf (Padronizado)
  real<lower=0> beta2; // k1 (Padronizado)
  real<lower=0> beta4; // k2 (Padronizado) - Tabela 3 usa beta4
    ↪ para taxa 2
  real beta3;          // t0 (Real) - Tabela 3 usa beta3 para
    ↪ localizacao
  real<lower=min_age, upper=max_age> tau; // Ponto de Transicao
    ↪ (Real)

  // Parametros de Heterocedasticidade e Erro (Escala REAL)
  real<lower=0> sigmae; // Sigma base (sigma_e na Tabela 3)
  real<lower=0> rho;    // Parametro de heterocedasticidade
  real<lower=1> nu;     // Graus de liberdade (Tabela 3 usa nu >
    ↪ 1)

```

```

}

transformed parameters {
  vector[N] mu_pad;
  vector[N] mu;          // Mu na escala REAL
  vector[N] sigma;       // Sigma na escala REAL

  real t0_pad = (beta3 - mean_age) / sd_age;
  real tau_pad = (tau - mean_age) / sd_age;

  // Peso no ponto de transicao com potencia 3
  real W_tau = beta1 * pow(1 - exp(-beta2 * (tau_pad - t0_pad)),
    ↪ 3);

  for (i in 1:N) {
    real age_pad = (x[i] - mean_age) / sd_age;

    // 1. Curva de Crescimento Segmentada com expoente 3
    if (age_pad <= tau_pad) {
      mu_pad[i] = beta1 * pow(1 - exp(-beta2 * (age_pad -
        ↪ t0_pad))), 3);
    } else {
      mu_pad[i] = W_tau + (beta1 - W_tau) * pow(1 - exp(-beta4 *
        ↪ (age_pad - tau_pad))), 3);
    }

    // 2. Estrutura de Variancia Potencia Multiplicativa
    // sigma[i] = sigma_e * idade^rho
    sigma[i] = sigmae * pow(x[i], rho);
  }

  // 3. Despadronizacao para Likelihood
  mu = mu_pad * sd_weight + mean_weight;
}

model {
  // Priors Fracamente Informativos (Conforme Tabela 3)
  beta1 ~ normal(0, 2);
  beta2 ~ lognormal(0, 1);
  beta4 ~ lognormal(0, 1);
  beta3 ~ normal(0, 50);
  tau ~ normal(mean_age, sd_age);
}

```

```
// Priors Heterocedasticidade
sigmae ~ gamma(1, 0.5);
rho     ~ gamma(1, 0.5);
nu      ~ gamma(2, 0.1);

// Likelihood t-Student Robusta
y ~ student_t(nu, mu, sigma);
}

generated quantities {
  vector[N] log_lik;
  real W_inf_real = beta1 * sd_weight + mean_weight;
  real k1_real = beta2 / sd_age;
  real k2_real = beta4 / sd_age;

  for (i in 1:N) {
    log_lik[i] = student_t_lpdf(y[i] | nu, mu[i], sigma[i]);
  }
}
```

APÊNDICE D

Estrutura do R para ajuste e análise do modelo Von Bertalanffy Segmentado

```
library(rstan)
library(tidyverse)
library(loo)

# --- Funcoes Auxiliares de Padronizacao ---
calcular_stats <- function(df) {
  list(
    mean_age = mean(df$idade), sd_age = sd(df$idade),
    mean_weight = mean(df$peso), sd_weight = sd(df$peso)
  )
}

# 1. Lista para MODELOS HOMOCEDASTICOS (Tudo PADRONIZADO -
# ↳ Likelihood Z-score)
criar_stan_lista_homo <- function(df, stats) {
  list(
    N = nrow(df),
    age = (df$idade - stats$mean_age) / stats$sd_age,
    weight = (df$peso - stats$mean_weight) /
      ↳ stats$sd_weight,
    mean_age = stats$mean_age, sd_age = stats$sd_age,
    mean_weight = stats$mean_weight, sd_weight =
      ↳ stats$sd_weight,
    min_age = min(df$idade), max_age = max(df$idade)
  )
}
```

```

# 2. Lista para MODELOS HETEROCEDASTICOS (Tudo REAL -
  ↳ Likelihood Gramas)
criar_stan_lista_het <- function(df, stats) {
  list(
    N = nrow(df),
    x = df$idade,    # Idade REAL para a variancia
    y = df$peso,     # Peso REAL para o likelihood
    mean_age = stats$mean_age, sd_age = stats$sd_age,
    mean_weight = stats$mean_weight, sd_weight =
      ↳ stats$sd_weight,
    min_age = min(df$idade), max_age = max(df$idade)
  )
}

# -- Carregamento e Preparacao dos Dados --
dados <- read_tsv("dados5obs.txt", show_col_types = FALSE)
  ↳ %>% filter(!is.na(idade), !is.na(peso))

dados_machos <- dados %>% filter(sexo == 1)
stats_machos <- calcular_stats(dados_machos)

# Listas Stan
stan_dados_machos_homo <-
  ↳ criar_stan_lista_homo(dados_machos, stats_machos)
stan_dados_machos_het <- criar_stan_lista_het(dados_machos,
  ↳ stats_machos)

# --- Exemplo de Ajuste (Sampling) ---
# modelo_vb_seg_homo <-
  ↳ stan_model("bertalanffy_segmentado_t.stan")
# modelo_vb_seg_het <- stan_model("vb_seg_het_mult1.stan")
#
# fit_m_homo <- sampling(modelo_vb_seg_homo, data =
  ↳ stan_dados_machos_homo, chains = 4, iter = 2000)
# fit_m_het <- sampling(modelo_vb_seg_het, data =
  ↳ stan_dados_machos_het, chains = 4, iter = 2000)

# --- A partir daqui, voce extrai os resultados ---
# looic_homo <- loo(fit_m_homo)
# looic_het <- loo(fit_m_het)

```