



LUCAS DE SOUZA SEGATO

**ANTEVENDO O IMPACTO CLIMÁTICO E DE SANIDADE
FLORESTAL NO POTENCIAL PRODUTIVO DE
PLANTIOS DE EUCALIPTO: UMA ABORDAGEM PARA
ÍNDICE DE SÍTIO BASEADA EM MACHINE LEARNING**

**LAVRAS – MG
2025**

LUCAS DE SOUZA SEGATO

**ANTEVENDO O IMPACTO CLIMÁTICO E DE SANIDADE FLORESTAL NO
POTENCIAL PRODUTIVO DE PLANTIOS DE EUCALIPTO: UMA ABORDAGEM
PARA ÍNDICE DE SÍTIO BASEADA EM MACHINE LEARNING**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Ciências Florestais, para a obtenção do título de Mestre.

Orientador

Prof. Dr. Samuel José Silva Soares da Rocha

Coorientador

Dr. Henrique Ferraço Scolforo

**LAVRAS – MG
2025**

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

Segato, Lucas de Souza.

Antevendo o impacto climático e de sanidade florestal no potencial
produtivo de plantios de eucalipto: uma abordagem para índice de
sítio baseada em machine learning/ Lucas de Souza Segato. - 2025.
71 p.

Orientador(a): Samuel José Silva Soares da Rocha.

Coorientador(a): Henrique Ferraço Scolforo.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2025.

Bibliografia.

1. Modelos Híbridos. 2. Programação Genética. 3. Seleção de
Variáveis. 4. Sensoriamento Remoto. I. Rocha, Samuel José Silva
Soares da. II. Scolforo, Henrique Ferraço. III. Título.

LUCAS DE SOUZA SEGATO

**ANTEVENDO O IMPACTO CLIMÁTICO E DE SANIDADE FLORESTAL NO
POTENCIAL PRODUTIVO DE PLANTIOS DE EUCALIPTO: UMA ABORDAGEM
PARA ÍNDICE DE SÍTIO BASEADA EM MACHINE LEARNING**

**ANTICIPATING CLIMATE AND FOREST HEALTH IMPACT ON THE
PRODUCTIVE POTENTIAL OF EUCALYPTUS PLANTATIONS: A MACHINE
LEARNING-BASED SITE INDEX APPROACH**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Florestal, área de concentração em Ciências Florestais, para a obtenção do título de Mestre.

APROVADA em 26 de fevereiro de 2025.

DSc. Samuel José Silva Soares da Rocha	UFLA
DSc. Lucas Rezende Gomide	UFLA
DSc. Evandro Nunes Miranda	Suzano SA
DSc. Henrique Ferraço Scolforo	Suzano SA
DSc. Talles Hudson Souza Lacerda	Suzano SA

Orientador
Prof. Dr. Samuel José Silva Soares da Rocha

Coorientador
Dr. Henrique Ferraço Scolforo

**LAVRAS – MG
2025**

Para Adriana de Souza,

A mulher que me ensinou que, mesmo com o mundo contra nós, o amor sempre vence.

Você foi, desde sempre, minha fortaleza, minha base e minha inspiração.

Foi você quem me deu asas, mesmo quando a vida te tirava o chão.

*Graças ao seu amor, hoje eu sou capaz de enxergar o mundo com outros olhos — olhos que
você me ensinou a abrir com coragem, sensibilidade e verdade.*

Mais do que mãe, você é a razão pela qual eu luto todos os dias para ser alguém melhor.

E se um dia eu for capaz de carregar uma fração da luz que você irradia, saberei que venci.

Essa conquista também é sua — e para sempre será

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus por estar guiando minha jornada desde o princípio, sem o Senhor eu nada seria. Deixo minhas honras a minha mãe amada e batalhadora que conseguiu me educar, ensinar o que é o amor, que tudo na vida é fruto de muito esforço e muita dedicação, e agradeço imensamente aquele dia que a senhora me deu a opção de estudar, foi a melhor decisão que eu tomei em minha vida, tenho muito orgulho de ser seu filho, você é minha maior inspiração!

Agradeço também a minha namorada Ana, que me apoiou imensamente no final dessa jornada, obrigado pelo apoio emocional, obrigado por ser esse anjo abençoado que Deus colocou no meu caminho para compartilhar a benção do amor.

A toda minha família que me apoiou imensamente em todo o processo, sabemos que a caminhada não foi fácil, mas Deus na frente e tudo torna-se um simples aprendizado, e as dores passam a ser vistas como oportunidades de crescimento. Então obrigado a todos, em especial a meu avô Sidney, meu primo Kauan, meu padrinho Jean, meu outro padrinho Carlo, minha prima Mayra e a todos da família, que eu fico dizendo que é pequena, mas quando começo a agradecer vejo que é bem grande.

Um agradecimento a meus companheiros diários, Cassio, Evandro, Gabriel, Raiana, Julyana, Talles e Marcel, que são minha segunda família. Saibam que eu me torno alguém melhor todos os dias convivendo com vocês.

Deixo um obrigado especial para o irmão que a vida me deu, que dividiu comigo todo esse desafio de conciliar o trabalho com o mestrado. Guilherme, agradeço a Deus por ter você como amigo, essa jornada foi muito mais leve e proveitosa pois estava lá me apoiando em todos os momentos. Conte comigo para sua vida irmão!

Um agradecimento a um pai que vida me deu, meu coorientador Henrique, que tem me apadrinhado desde o início da minha carreira profissional, agradeço por toda instrução, toda paciência, todo carinho e respeito em todos esses anos, e que tenho o prazer de dividir esse trabalho que vos apresento.

Agradeço a meu orientador no início do projeto, professor José Scolforo, agradeço pelos ensinamentos e pela demonstração de paixão pelo manejo florestal e pela modelagem, quisera eu ser 1% do senhor um dia e estarei satisfeito. E agradeço também a meu atual orientador, na etapa final do projeto, Samuel, que me apoiou em diversas frentes, garantindo que eu estivesse aqui hoje entregando esse trabalho.

Agradeço a UFLA pela oportunidade de estar afiando meu machado, aprendendo um pouco mais a cada dia e absorvendo o máximo de conhecimento e experiências possível. Agradeço a Suzano por ter permitido realizar esse projeto de mestrado, conciliando com a rotina empresarial, que embora tenha sido árdua em muitos momentos, hoje é gratificante ter passado por tudo isso e saído ainda mais forte do que entrei.

A minha segunda casa, grande República Pé da Serra, que me recebeu de braços abertos e tornou minha experiência em Lavras menos dolorosa, por conta da saudade da família. Vocês se tornaram parte de mim e sempre voltarei aí para lembrar os velhos e bons momentos.

Por fim, deixo meu agradecimento a duas estrelinhas que estão me vendo e torcendo por mim lá de cima, que sem vocês, eu não seria a pessoa que sou, vocês foram e são muito especiais para mim, amarei vocês para sempre, minhas segundas mães, minha madrinha Selma e minha vizinha Catarina, vocês sempre terão cadeira cativa no meu coração e levarei vocês comigo para onde eu for, amo vocês.

RESUMO GERAL

As mudanças climáticas vêm impondo desafios crescentes ao setor florestal, impactando diretamente a produtividade dos povoamentos devido a eventos como secas prolongadas, aumento da variabilidade climática e alteração nos regimes hídricos. Nesse contexto, esta dissertação buscou aprofundar o entendimento dos efeitos ambientais sobre o crescimento florestal, com foco na projeção da altura dominante, por meio de uma abordagem integrada que abrange revisão de literatura e desenvolvimento de modelos preditivos aprimorados. A primeira parte do trabalho realizou uma análise crítica dos principais avanços na modelagem do crescimento florestal, ressaltando o potencial do uso de variáveis climáticas, dados de sensoriamento remoto e técnicas baseadas em inteligência artificial para melhorar a acurácia das estimativas. Na segunda parte, foi desenvolvido um estudo aplicado que comparou diferentes modelos de projeção da altura dominante em povoamentos de eucalipto, incluindo modelos convencionais, expandidos com variáveis climáticas e híbridos com variáveis orbitais. As variáveis foram selecionadas com apoio da Programação Genética, permitindo a identificação do efeito de carry-over, que representa o impacto de eventos climáticos passados sobre o crescimento atual das árvores. Os resultados demonstraram que a inclusão de variáveis ambientais contribuiu para a redução do viés e aumento do poder explicativo dos modelos, especialmente em cenários de déficit hídrico. A integração de índices orbitais, como o SAVI, também se mostrou promissora para representar a sanidade florestal e capturar variações espaciais da disponibilidade hídrica. A dissertação, portanto, contribui metodologicamente ao propor uma abordagem híbrida, escalável e sensível às condições ambientais, oferecendo subsídios para um manejo florestal mais eficiente, resiliente e adaptável às novas realidades impostas pelas mudanças climáticas.

Palavras chaves: Modelos Híbridos; Programação Genética; Seleção de Variáveis; Sensoriamento Remoto.

GENERAL ABSTRACT

Climate change has increasingly posed significant challenges to the forestry sector, directly affecting forest stand productivity due to prolonged droughts, increased climate variability, and changes in water regimes. In this context, this dissertation aimed to deepen the understanding of environmental effects on forest growth—particularly dominant height—through an integrated approach that combines a comprehensive literature review and the development of improved predictive models. The first part of the study critically reviewed key advances in forest growth modeling, emphasizing the potential of incorporating climatic variables, remote sensing data, and machine learning techniques to enhance estimation accuracy. The second part presented an applied study that compared different dominant height projection models for eucalyptus stands, including conventional models, models expanded with climatic variables, and hybrid models incorporating remote sensing indices. Variable selection was supported by Genetic Programming, enabling the identification of the carry-over effect, which reflects the delayed impact of past climatic events on current tree growth. The results showed that including environmental variables significantly reduced bias and improved model explanatory power, particularly under water deficit scenarios. The integration of remote sensing indices, such as SAVI, also proved valuable for representing forest health and capturing spatial variations in water availability. Therefore, this dissertation contributes methodologically by proposing a hybrid, scalable, and environmentally sensitive modeling approach, offering valuable insights to support more efficient, resilient, and climate-adaptive forest management strategies.

Keywords: Hybrid Models; Genetic Programming; Feature Selection; Remote Sensing.

INDICADORES DE IMPACTO

A dissertação apresenta um modelo híbrido para projeção da altura dominante em povoamentos de eucalipto, incorporando variáveis climáticas históricas e informações de sensoriamento remoto, com ênfase na mitigação dos efeitos de estresse hídrico e variabilidade ambiental. Os resultados obtidos evidenciam impactos em potencial de caráter tecnológico, ambiental, econômico e social. No aspecto tecnológico, o modelo desenvolvido representa um avanço metodológico ao combinar técnicas de programação genética e dados ambientais com alta precisão, superando métodos tradicionais em métricas como R^2 e viés médio, o que possibilita maior acurácia na estimativa de produtividade florestal. Do ponto de vista econômico, esse ganho técnico viabiliza decisões mais eficientes no planejamento silvicultural, com potencial de redução de custos e riscos associados à produção, sobretudo em regiões suscetíveis a déficits hídricos. Ambientalmente, ao considerar a influência do clima em períodos anteriores ao inventário, o modelo contribui para o entendimento dos efeitos das mudanças climáticas sobre o crescimento florestal, auxiliando no desenvolvimento de práticas de manejo adaptativas e sustentáveis. A pesquisa também se aproxima do caráter extensionista ao utilizar dados operacionais de uma empresa florestal brasileira, demonstrando aplicabilidade direta no setor produtivo e promovendo a integração entre universidade e sociedade. O território contemplado inclui regiões de cultivo de eucalipto no sudeste do Brasil, com impactos potenciais sobre técnicos, engenheiros e analistas atuantes no setor florestal, além da comunidade acadêmica envolvida com modelagem, geotecnologias e manejo florestal. Os impactos deste trabalho alinham-se principalmente às áreas temáticas de Meio ambiente e Tecnologia e produção da Política Nacional de Extensão, bem como a quatro Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas: ODS 9 (Indústria, inovação e infraestrutura), ODS 12 (Consumo e produção responsáveis), ODS 13 (Ação contra a mudança global do clima) e ODS 15 (Vida terrestre). Assim, a pesquisa contribui para o avanço técnico-científico da área florestal, com desdobramentos que reforçam a sustentabilidade da produção, a inovação tecnológica no setor e o enfrentamento dos desafios ambientais atuais e futuros.

IMPACT INDICATORS

This dissertation presents a hybrid model for projecting dominant height in eucalyptus stands by incorporating historical climatic variables and remote sensing information, with emphasis on mitigating the effects of water stress and environmental variability. The results reveal potential technological, environmental, economic, and social impacts. From a technological perspective, the developed model represents a methodological advancement by combining genetic programming techniques with environmental data of high precision, outperforming traditional methods in metrics such as R^2 and mean bias, which enables more accurate forest productivity estimates. Economically, this technical improvement allows for more efficient decisions in silvicultural planning, with the potential to reduce costs and risks associated with forest production, especially in regions prone to water deficits. From an environmental standpoint, by considering the influence of past climatic conditions on current growth, the model contributes to understanding the effects of climate change on forest development, supporting the creation of adaptive and sustainable management practices. The research also displays an extensionist character by using operational data from a Brazilian forestry company, demonstrating direct applicability in the productive sector and promoting integration between university and society. The impacted territory includes eucalyptus-growing regions in southeastern Brazil, with potential benefits for forestry professionals such as technicians, engineers, and analysts, as well as the academic community involved with modeling, geotechnologies, and forest management. The impacts of this work align primarily with the thematic areas of Environment and Technology and Production, as defined by the Brazilian National Extension Policy, and with four of the United Nations' Sustainable Development Goals (SDGs): SDG 9 (Industry, Innovation and Infrastructure), SDG 12 (Responsible Consumption and Production), SDG 13 (Climate Action), and SDG 15 (Life on Land). Thus, this research contributes to the technical-scientific advancement of the forestry sector, reinforcing sustainable production, technological innovation, and efforts to address current and future environmental challenges.

LISTA DE FIGURAS

PRIMEIRA PARTE

Figura 1 – Mudança líquida anual global da área florestal, por década, 1990–2020.	17
Figura 2 – Distribuição de área plantada de eucalipto por região em 2023.	18

SEGUNGA PARTE – ARTIGO

Figura 1 – Localização do estado de Mato Grosso do Sul e suas condições edafoclimáticas..	44
Figura 3 – Comportamento sazonal dos índices de vegetação ao longo do crescimento do plantio.	48
Figura 4 – Fluxograma do método proposto e como ele foi executado.	49
Figura 5 – Distribuição de resíduos das estimativas de SAVI por intervalo de projeção.	59
Figura 6 – Distribuição geral dos resíduos da altura dominante projetada.	60
Figura 7 – Distribuição dos resíduos considerando apenas locais que sofreram estresse hídrico severo.....	61
Figura 8 – Distribuição dos resíduos para os modelos avaliados a partir da reamostragem por Bootstrap, sob diferentes regimes hídricos.	62
Figura 9 – Distribuições dos resíduos em diferentes condições de sanidade florestal e de disponibilidade hídrica.	64
Figura 10 – Simulação de cenários de deficiência hídrica e perda de sanidade florestal.....	65

LISTA DE TABELAS

SEGUNDA PARTE - ARTIGO

Tabela 1 – Descrição geral da altura dominante da área de estudo nas diferentes regiões climáticas.	45
Tabela 2 – Descrição geral das variáveis climáticas da área de estudo nas diferentes regiões.	46
Tabela 3 – Índices de vegetação utilizados no estudo e suas fórmulas.	47
Tabela 4 – Descrição geral de alguns índices de vegetação utilizados no estudo.	48
Tabela 5 – Coeficientes ajustados para os três modelos avaliados.	56
Tabela 6 – Parâmetros estimados do modelo de projeção do SAVI.	58

SUMÁRIO

PRIMEIRA PARTE	14
1 INTRODUÇÃO	16
2 REVISÃO DE LITERATURA.....	17
2.1 Setor florestal	17
2.2 Plantios comerciais de <i>Eucalyptus</i>	19
2.3 Crescimento florestal e sua modelagem.....	20
2.3.1 Modelagem de crescimento e produção para eucalipto	21
2.3.2 Variáveis essenciais para modelos de produção	21
2.3.3 Índice de sítio	22
2.3.4 Modelos híbridos.....	23
2.4 Clima e seu impacto no crescimento do eucalipto	24
2.4.1 Impacto das variáveis climáticas.....	25
2.5 Sensoriamento remoto e dados orbitais.....	25
2.5.1 Tipos de satélites e sensores para o monitoramento florestal.....	26
2.5.2 Índices de vegetação e aplicação no manejo do eucalipto	27
2.6 Seleção de variáveis	28
2.6.1 Métodos de seleção.....	28
2.7 Inteligência artificial e aprendizado de máquina	29
2.7.1 Programação genética	30
3 CONSIDERAÇÕES FINAIS.....	31
REFERÊNCIAS.....	31
SEGUNDA PARTE	37
ARTIGO.....	38
1 INTRODUÇÃO	41
2 MATERIAL E MÉTODOS	43
2.1 Caracterização da área de estudo	43
2.2 Base de dados de inventário.....	44
2.3 Base climática e compatibilização com os dados de inventário.....	45
2.4 Base de dados de sensoriamento remoto e compatibilização com dados de inventário.....	46
2.5 Sistema de modelagem	49
2.5.1 Modelo base (convencional) de Chapman-Richards	50

2.5.2	Expansão dos parâmetros via programação genética	51
2.5.2.1	Modelo expandido com variáveis climáticas	51
2.5.2.2	Modelo expandido com variáveis climáticas e de sensoriamento remoto	52
2.5.2.3	Funcionamento da programação genética para seleção e combinação de variáveis	53
2.5.3	Validação dos modelos	55
3	RESULTADOS	56
3.1	Modelos de projeção de altura dominante	56
3.2	Modelo de projeção do índice de vegetação	57
3.3	Análise comparativa média dos modelos de crescimento da altura dominante	59
3.4	Análise comparativa do índice de sítio sob diferentes condições climáticas ..	61
3.5	Análise comparativa do índice de sítio sob diferentes condições climáticas x sanidade florestal	63
3.6	Cenários de variação do crescimento da altura dominante sob variadas condições de clima e sanidade florestal.....	64
4	DISCUSSÃO	66
5	CONCLUSÃO.....	69
	REFERÊNCIAS.....	71
	TERCEIRA PARTE.....	74

PRIMEIRA PARTE

1 INTRODUÇÃO

As florestas plantadas surgiram como uma evolução silvicultural necessária devido a altas taxas de exploração de florestas nativas. Aliado a incentivos fiscais, a expansão dessa atividade iniciou na década de 60, e isso proporcionou o maior uso dessa matéria prima pelos setores industriais, inicialmente empreendimentos de celulose e papel, e posteriormente siderurgias e serrarias (Moreira *et al.*, 2017). Com o objetivo de sustentar a produção dessas indústrias, o Brasil detém aproximadamente 9,94 milhões de hectares de florestas plantadas, sendo 76% (7,6 milhões de hectares) apenas composta por espécies do gênero *Eucalyptus*. Esse setor possui uma grande participação na geração de empregos do país, alcançando, em 2022, a geração de 2,6 milhões de empregos direta e indiretamente. É um setor que recebe investimentos constantes e que está em ampla expansão, dessa forma prevê que esse número tende a aumentar ao longo dos próximos anos (IBÁ, 2024).

Enfrentando o desafio de assegurar um fornecimento contínuo e alinhado ao planejamento estratégico de curto, médio e longo, o setor precisa de ferramentas eficientes e precisas para a gestão de suas operações. O inventário florestal surge como um instrumento crucial, permitindo não só o acompanhamento do plantio, mas também o entendimento da dinâmica de crescimento dos povoamentos florestais, consistindo em coletas de amostras do plantio, possibilitando a estimativa de características qualitativas e quantitativas da floresta (Guedes *et al.*, 2012). Um aspecto chave nesse planejamento é a quantificação precisa do volume de madeira disponível ao longo dos anos, considerando diferentes cenários produtivos. Isso se torna possível através dos modelos de crescimento e produção, que, fundamentados em dados de inventários florestais e históricos de medições, são capazes de prever o volume futuro de povoamentos florestais (Scolforo *et al.*, 2019a).

Os modelos clássicos de crescimento e produção são ajustados a partir de uma base de dados abrangente, entretanto, como eles baseiam-se em regimes produtivos históricos, estes podem não ser tão eficazes em cenários atípicos, como em momentos de secas. Assim, muitos estudos foram realizados buscando tornar os modelos de crescimento e produção híbridos, adicionando em sua formulação variáveis relacionadas ao clima e ao solo local, buscando uma melhoria nas estimativas, além de fornecer informações flexíveis independentemente do cenário climático (Casnati *et al.*, 2016; Ferraz Filho *et al.*, 2011; Maestri, 2003; Salekin *et al.*, 2021; Scolforo *et al.*, 2017).

Recentemente, identificou-se uma nova oportunidade de aprimorar esses modelos, porém atuando por uma nova perspectiva, de modo que seja possível integrar lógicas algébricas

buscando prever quedas produtivas com antecedência, de forma rápida e em curto espaço de tempo, a fim de mitigar impactos na prognose do volume do plantio.

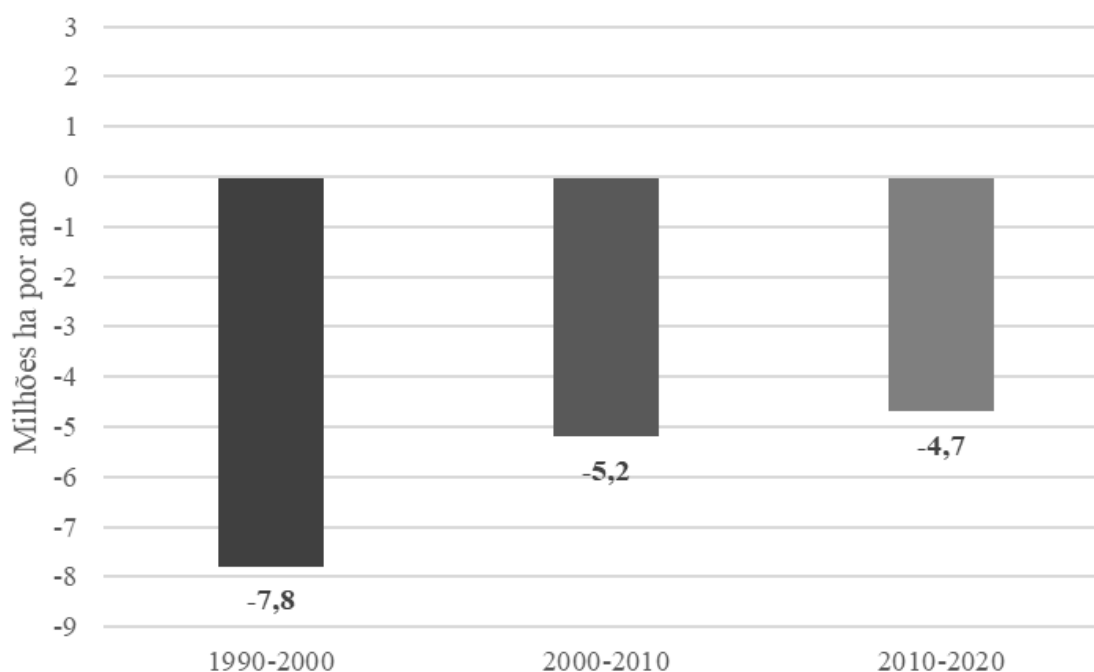
2 REVISÃO DE LITERATURA

2.1 Setor florestal

O levantamento global de recursos florestais realizado pela FAO em 2020 revelou que aproximadamente 4,06 bilhões de hectares do planeta são cobertos por vegetação florestal, o que corresponde a cerca de 31% da área terrestre. A maior parte dessa cobertura florestal (cerca de 54%) está concentrada em cinco países: Rússia, Brasil, Canadá, Estados Unidos e China. Apesar dessa significativa extensão florestal, o desmatamento continua a ser uma preocupação global.

Embora a taxa de perda de cobertura florestal tenha diminuído na última década (Figura 1), a redução contínua de áreas naturais ressalta a importância dos plantios comerciais. Esses plantios desempenham um papel crucial na mitigação da pressão sobre as florestas naturais, uma vez que fornecem madeira de forma sustentável, reduzindo a necessidade de exploração de fragmentos de vegetação nativa.

Figura 1– Mudança líquida anual global da área florestal, por década, 1990–2020.



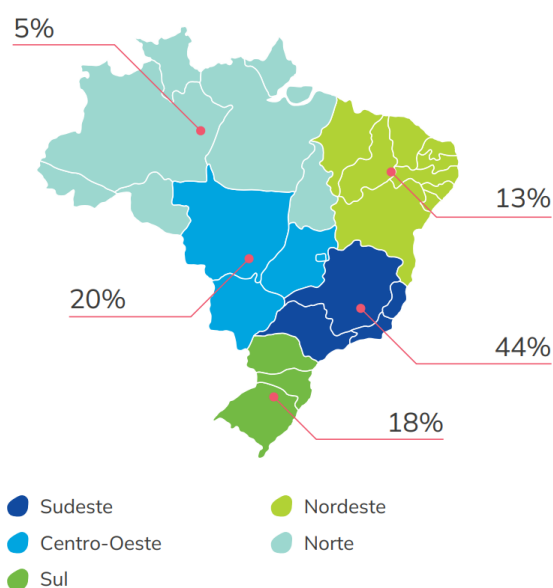
Fonte: Adaptado de Relatório da FAO, *Global Forest Resources Assessment 2020*.

Os plantios florestais desempenham um papel essencial na geração de empregos e no desenvolvimento socioeconômico das comunidades locais. Em 2023, foram criados aproximadamente 33,4 mil novos postos de trabalho, elevando o total para cerca de 2,69 milhões de empregos diretos e indiretos. Além do impacto econômico, esses plantios também promovem importantes benefícios ambientais, como a captura de carbono, a conservação do solo e a regulação do ciclo hidrológico. Nesse mesmo ano, aproximadamente 1,8 milhão de mudas foram plantadas em áreas previamente degradadas, contribuindo para a recuperação ambiental e a ampliação da sustentabilidade no setor (IBÁ, 2024).

No Brasil, o setor florestal tem apresentado uma significativa expansão nos últimos anos. Comparando os dados dos relatórios anuais da Indústria Brasileira de Árvores (IBÁ) de 2014 e 2024, observa-se que, em um intervalo de 10 anos, a área de plantios florestais aumentou de 7,6 para 10,2 milhões de hectares. Segundo o mesmo relatório, em 2023, cerca de 7,8 milhões de hectares desses plantios eram compostos exclusivamente por espécies de eucalipto, o que evidencia a importância dessa cultura no contexto florestal brasileiro.

Dentre as regiões do país, destaca-se o Sudeste, que concentra aproximadamente 44% dessa área plantada, seguida pela região Centro-Oeste, responsável por 20% do total (Figura 2). Esse crescimento pode ser atribuído a fatores como avanços no manejo florestal, melhorias genéticas e a alta demanda por madeira de rápido crescimento para suprir diferentes indústrias, como papel e celulose, energia renovável e biomassa.

Figura 2 – Distribuição de área plantada de eucalipto por região em 2023.



Fonte: Extraído de Relatório Anual IBÁ 2024 – Canopy.

Além de sua expressiva relevância em área, a cadeia produtiva florestal tem um grande impacto no setor econômico brasileiro, representando aproximadamente 0,9% do PIB nacional e movimentando cerca de US\$ 12,7 bilhões em exportações. Esse desempenho se destaca mesmo em um cenário global desafiador, caracterizado por flutuações nos preços de commodities, ajustes de estoques no período pós-pandemia e uma crescente volatilidade das demandas globais. Adicionalmente, o Brasil mantém a liderança como maior exportador de celulose no mundo, com um valor exportado de aproximadamente US\$ 7,9 bilhões, o que equivale a 63% do total das exportações do setor (IBÁ, 2024).

2.2 Plantios comerciais de *Eucalyptus*

Atender à crescente demanda por madeira requer uma seleção criteriosa das espécies a serem cultivadas, levando em consideração não apenas a qualidade do produto final, mas também a sustentabilidade econômica e ambiental do negócio. Entre as diversas opções disponíveis, o gênero *Eucalyptus* se sobressai como a principal escolha. O Brasil, por sua vez, é amplamente reconhecido como o país mais eficiente na produção de madeira de *Eucalyptus* proveniente de plantios comerciais, consolidando sua liderança global nesse segmento. Essa eficiência é resultado de avanços no manejo florestal, melhorias genéticas e condições climáticas favoráveis, que tornam o país referência mundial na produção sustentável de madeira (Machado *et al.*, 2014).

Além de seu notável crescimento rápido e da conveniência de ser implantada em larga escala, abrangendo vastas extensões territoriais, o eucalipto apresenta uma notável capacidade de adaptação a uma variedade de condições climáticas. Assim, pode ser encontrada em todo o território nacional, desde áreas frias e secas até regiões quentes e úmidas, o que ressalta sua versatilidade e seu potencial para contribuir significativamente para diversos setores da economia brasileira (Coelho *et al.*, 2020).

A adaptabilidade do eucalipto, predominantemente cultivado através de plantios clonais, reflete os avanços no melhoramento genético realizados ao longo de décadas. Esses esforços visam otimizar as interações entre o genótipo e o ambiente em que são cultivadas. Esse processo contínuo de aprimoramento genético tem sido fundamental para garantir que as plantações de eucalipto possam prosperar em diferentes condições ambientais, contribuindo assim para a robustez e a sustentabilidade desse importante setor (Stape *et al.*, 2010).

A ampla escala de plantios de eucalipto no Brasil destaca a importância de processos precisos de mensuração e monitoramento desse ativo biológico. A avaliação detalhada, tanto

do volume quanto da qualidade da madeira, é fundamental para garantir um controle eficiente e a sustentabilidade da produção. Isso ocorre porque há uma relação direta e determinante entre as características da matéria-prima e a qualidade do produto final obtido nas etapas industriais, independentemente do destino ou aplicação, seja na fabricação de celulose, papel, biomassa ou outros produtos derivados. Além disso, a mensuração adequada permite identificar variações na produtividade, avaliar o impacto de práticas silviculturais e assegurar a viabilidade econômica e ambiental do negócio, contribuindo para a competitividade do setor em mercados cada vez mais exigentes e globalizados (Fantuzzi Neto, 2012).

Diversos métodos de mensuração podem ser aplicados em campo, variando conforme o estágio de desenvolvimento do plantio e os objetivos específicos de cada avaliação. Nos empreendimentos florestais que têm o eucalipto como principal espécie plantada, é comum a realização de medições qualitativas nos estágios iniciais do cultivo. Essas avaliações iniciais têm como objetivo principal verificar a adequação dos tratamentos silviculturais realizados, garantindo que as práticas adotadas estejam alinhadas com o manejo planejado.

Posteriormente, quando os plantios atingem cerca de dois anos de idade, inicia-se a realização de inventários quantitativos, que têm como foco a mensuração precisa de variáveis como volume, altura e diâmetro das árvores. Esses inventários são periodicamente repetidos ao longo do ciclo de desenvolvimento da floresta até a idade de corte, em um processo conhecido como Inventário Florestal Contínuo (IFC).

Os dados coletados por meio do Inventário Florestal Contínuo (IFC) permitem acompanhar a dinâmica de crescimento do plantio ao longo do tempo. Com base nesses dados, é possível utilizar sistemas de crescimento e produção para realizar projeções da produtividade futura (Hall *et al.*, 2020). Essas projeções, que indicam o volume estimado em idades específicas, são fundamentais para o planejamento florestal das empresas. Essa estimativa futura serve como base para garantir um abastecimento contínuo e eficiente das unidades fabris, atendendo às demandas futuras de produção. Dessa forma, as empresas podem planejar de maneira estratégica a utilização dos recursos florestais, otimizando a gestão do ativo biológico e assegurando a sustentabilidade de sua cadeia produtiva.

2.3 Crescimento florestal e sua modelagem

O crescimento de um plantio florestal ocorre por meio de diferentes processos, como o alongamento das árvores (incremento em altura), o aumento do diâmetro do tronco, bem como o crescimento dos galhos e das raízes. Esse crescimento reflete o acúmulo de biomassa ao longo

do tempo e pode ser expresso de diferentes formas, dependendo da escala de análise considerada. Em intervalos de tempo mais curtos, o crescimento pode ser avaliado por incrementos periódicos, como o incremento médio anual em altura ou volume. Já em escalas mais amplas, observa-se o desenvolvimento total do povoamento, incluindo a produção acumulada de madeira e mudanças na estrutura do dossel ao longo do ciclo produtivo.

A modelagem do crescimento e da produção florestal é uma ferramenta essencial para traduzir esses processos biológicos em projeções quantitativas. Ao trabalharmos com esses modelos, avaliamos precisamente as variáveis que se alteram ao longo do tempo, como altura, diâmetro e volume, com o objetivo de inferir a quantidade de madeira disponível no futuro. Essas projeções são indispensáveis para o planejamento estratégico e operacional de empresas do setor florestal, permitindo uma gestão mais eficiente dos recursos e o alinhamento da produção às demandas do mercado.

2.3.1 Modelagem de crescimento e produção para eucalipto

Na modelagem voltada especificamente para o eucalipto, é essencial considerar as características únicas dessa espécie, como seu rápido crescimento, alta plasticidade fenotípica e resposta significativa às práticas de manejo silvicultural. Entre os principais modelos utilizados, destacam-se os modelos empíricos, que se baseiam em dados históricos para gerar projeções, e os modelos de processo, que buscam simular os mecanismos biológicos subjacentes ao desenvolvimento das árvores.

Além disso, os modelos podem ser aplicados em diferentes escalas de ajuste. Em nível de povoamento, destacam-se os modelos de Clutter (Casas *et al.*, 2022; Clutter, 1963) e Scolforo (Scolforo *et al.*, 2019b), amplamente utilizados no planejamento florestal. Já em nível de classes diamétricas e árvores individuais, o modelo proposto por Scolforo *et al.* (2019a) é particularmente relevante por compatibilizar as projeções realizadas em nível de povoamento com aquelas realizadas para árvores individuais, proporcionando maior precisão nas estimativas e um entendimento mais detalhado da dinâmica florestal.

2.3.2 Variáveis essenciais para modelos de produção

Ao trabalharmos com modelos de crescimento e produção, avaliamos precisamente essas variáveis que se alteram ao longo do tempo, como altura, diâmetro e volume, com o objetivo de inferir a quantidade de madeira disponível no futuro. Essas projeções são

fundamentais para o planejamento estratégico e operacional de empresas do setor florestal, permitindo uma gestão mais eficiente dos recursos e o alinhamento da produção às demandas de mercado.

Algumas variáveis são fundamentais para os modelos de produção, sendo indispensáveis para garantir a precisão e robustez das projeções. A idade do povoamento é a variável principal, pois incorpora a dimensão temporal, essencial para compreender a evolução do crescimento ao longo do ciclo produtivo. A segunda variável independente tradicionalmente utilizada é a qualidade de sítio, que representa a capacidade de determinada área de sustentar e promover o crescimento das árvores. De forma prática, a qualidade de sítio pode ser definida como o conjunto de fatores ecológicos, como clima, solo e topografia, que influenciam diretamente o crescimento dos povoamentos em uma localização específica. Em plantios equiâneos, como ocorre na maioria dos plantios comerciais de eucalipto, essa variável é particularmente eficaz para determinar a capacidade produtiva do local.

A terceira variável essencial é a densidade do povoamento, que pode ser interpretada, segundo Clutter *et al.* (1983), como a taxa de ocupação do sítio pelo povoamento e, consequentemente, como um indicador do potencial de crescimento presente. A densidade pode ser mensurada de diferentes formas, como o número de árvores por hectare, a área basal por hectare, índices de densidade de povoamento ou de espaçamento relativo, entre outros. Cada abordagem fornece insights complementares sobre a ocupação e a competição intraespecífica no povoamento.

Por fim, mas não menos importante, temos a sobrevivência do plantio, uma variável crítica para a precisão das predições futuras. A mortalidade, que afeta diretamente a densidade, pode ocorrer de maneira regular, de forma esporádica no espaço e no tempo, ou de maneira irregular, como resultado de fatores exógenos. Exemplos de fatores que causam mortalidade irregular incluem catástrofes ambientais, ataques de pragas, doenças e outras adversidades que podem comprometer significativamente o desenvolvimento do povoamento. A incorporação adequada da sobrevivência nos modelos permite uma estimativa mais confiável da produtividade futura, essencial para o planejamento florestal.

2.3.3 Índice de sítio

A qualidade do sítio é um fator determinante para estimar o potencial produtivo dos plantios florestais, sendo essencial para o manejo e planejamento estratégico da produção. Entre as métricas utilizadas para essa avaliação, o índice de sítio destaca-se como uma das principais

ferramentas de quantificação da capacidade produtiva de uma área florestal (Burkhardt e Tomé, 2012).

Além de indicar o potencial produtivo de uma área, o índice de sítio possui diversas aplicações no setor florestal. Ele pode ser utilizado como critério na avaliação e aquisição de novas áreas, auxiliando na seleção de locais mais propícios para o cultivo de determinadas espécies. No âmbito do manejo florestal, permite ajustar práticas silviculturais, como espaçamentos, desbastes e adubações, de forma a otimizar a produtividade. Também desempenha um papel fundamental no planejamento estratégico, contribuindo para a projeção de produção futura, a alocação eficiente de recursos e a tomada de decisão baseada em dados confiáveis.

O índice de sítio é definido como a altura dominante projetada de um povoamento florestal para uma idade de referência previamente estabelecida. Essa métrica pode ser determinada por meio de métodos indiretos, que utilizam variáveis como tipo de solo, microclima ou profundidade do lençol freático, ou por métodos diretos, que se baseiam em dados históricos do plantio. Entre os métodos diretos, a estimativa do índice de sítio pode ser obtida a partir do volume e da área basal, de medições de altura ao longo do tempo ou, mais comumente, pela modelagem do crescimento periódico da altura dominante, o que permite projeções mais precisas por meio de modelos estatísticos (Clutter *et al.*, 1983).

Os métodos diretos apresentam vantagens significativas, pois o índice de sítio não sofre influência direta do espaçamento do plantio e reflete, de forma integrada, os efeitos edafoclimáticos sobre o crescimento das árvores (Iles, 2003). Isso possibilita a comparação da qualidade do sítio entre diferentes áreas, independentemente das densidades de plantio adotadas, tornando-o uma ferramenta robusta para a análise da produtividade florestal.

A determinação do índice de sítio desempenha um papel essencial na modelagem do crescimento e produção florestal, sendo amplamente utilizada tanto em pesquisas acadêmicas quanto na gestão de plantações comerciais. Sua aplicação permite o refinamento contínuo das práticas de manejo, contribuindo para uma produção florestal mais eficiente e sustentável, além de embasar a tomada de decisões estratégicas no setor florestal.

2.3.4 Modelos híbridos

Os modelos híbridos têm surgido como uma abordagem promissora para a previsão e projeção do crescimento de plantações de eucalipto e outras espécies florestais. Esses modelos combinam informações de diversas fontes, incluindo variáveis climáticas, orbitais e

biométricas, para melhorar a precisão das estimativas e fornecer insights mais abrangentes sobre o comportamento e o desenvolvimento das florestas (Scolforo *et al.*, 2017).

Nos últimos anos, tem-se observado um avanço significativo na modelagem do crescimento de povoamentos florestais, especialmente com o surgimento da abordagem híbrida que combina elementos da modelagem tradicional com covariáveis climáticas. O pioneirismo de Maestri (2003) foi crucial nesse sentido, ao propor a inclusão de fatores ambientais na modelagem do crescimento da altura dominante das árvores.

Seguindo essa linha de pesquisa, Ferraz Filho *et al.* (2011) expandiram ainda mais os horizontes ao integrar dados detalhados sobre a precipitação anual e a radiação solar na já estabelecida equação de crescimento de Chapman-Richards. Esse esforço resultou em uma notável redução no erro da estimativa, evidenciando a importância da consideração dos fatores climáticos na predição do crescimento florestal.

Em paralelo, Scolforo *et al.* (2017) trouxeram novas contribuições ao relatar que a inclusão de covariáveis relacionadas à disponibilidade de água no solo pode aumentar significativamente o poder explicativo dos modelos de crescimento da altura dominante, destacando a complexa interação entre fatores ambientais e o desenvolvimento das florestas.

Adicionalmente, a utilização de informações provenientes do sensoriamento remoto tem se mostrado uma ferramenta valiosa na modelagem híbrida, permitindo correlacionar dados climáticos e índices de vegetação com a produtividade dos povoamentos florestais. Como destacado por Li e Roy (2017), essa capacidade de monitoramento contínuo abre novas perspectivas para o aprimoramento dos modelos de crescimento florestal, permitindo uma melhor compreensão das dinâmicas ambientais e sua influência sobre a produtividade das florestas.

2.4 Clima e seu impacto no crescimento do eucalipto

O clima desempenha um papel fundamental no crescimento e desenvolvimento do eucalipto. Diferentes condições climáticas, desde a temperatura até a precipitação, influenciam diretamente aspectos como taxa de crescimento, saúde das árvores e produtividade da madeira. O entendimento de como o clima afeta o eucalipto é crucial para a otimização das práticas de silvicultura e para garantir a sustentabilidade e rentabilidade das plantações (Silva *et al.*, 2022). Este segmento da dissertação explora a relação intrínseca entre o clima e o cultivo do eucalipto, cobrindo desde as variáveis climáticas básicas até as metodologias de coleta de dados

climáticos, impacto dessas variáveis no crescimento do eucalipto, e as estratégias de adaptação e manejo em resposta às condições climáticas.

2.4.1 Impacto das variáveis climáticas

O clima é definido como o padrão médio do tempo em uma região ao longo de um longo período. Este padrão é determinado por variáveis cruciais como temperatura, precipitação, umidade relativa, velocidade do vento e radiação solar. Cada uma dessas variáveis tem um papel significativo no desenvolvimento das plantas (Lorençone *et al.*, 2020).

A temperatura afeta processos fisiológicos das árvores, como a fotossíntese e a transpiração. As faixas de temperatura ideais variam entre diferentes espécies e fases de crescimento do eucalipto, tendo uma relação direta com a fotossíntese da planta, de forma que condições extremas podem afetar negativamente o desenvolvimento da planta (Battaglia *et al.*, 1996). A precipitação é igualmente crítica, pois fornece a água necessária para o crescimento das plantas. Variações na quantidade e na distribuição da chuva podem levar a estresses hídricos ou à saturação do solo, ambos impactando negativamente o crescimento do eucalipto (Silva *et al.*, 2022). A umidade relativa influencia a evapotranspiração, que, quando combinadas a altas taxas de temperatura, tendem a ocasionar a redução do potencial de crescimento do eucalipto (Assad *et al.*, 2021). A velocidade do vento pode afetar a taxa de transpiração e até causar danos físicos às árvores em condições extremas. Por fim, a radiação solar é fundamental para a fotossíntese, portanto, a quantidade e a qualidade da luz solar recebida podem afetar diretamente a taxa de crescimento do eucalipto (Nardini *et al.*, 2020).

Estudos demonstraram correlações entre essas variáveis climáticas e o crescimento do eucalipto, destacando a necessidade de compreender e monitorar o clima para otimizar as práticas de manejo florestal e garantir a sustentabilidade das plantações de eucalipto (Campoe *et al.*, 2016; Queiroz *et al.*, 2023; Scolforo *et al.*, 2019c). Compreender essas variáveis e sua interação é essencial para avaliar como o clima afeta o cultivo do eucalipto, permitindo previsões mais precisas sobre o crescimento das árvores e a produtividade das florestas.

2.5 Sensoriamento remoto e dados orbitais

O sensoriamento remoto é uma técnica essencial para o manejo florestal, pois permite a coleta de informações detalhadas sobre grandes áreas de floresta de forma eficiente e econômica. Utilizando sensores em satélites ou aeronaves, gestores florestais e pesquisadores

podem monitorar a cobertura florestal, a saúde das árvores, e detectar mudanças devido a fatores naturais ou atividades humanas sem necessidade de contato direto. Essa capacidade de observação abrangente e contínua facilita o planejamento sustentável, a implementação de práticas de manejo eficazes, e a resposta rápida a potenciais problemas, desempenhando um papel crucial na conservação dos ecossistemas florestais e na otimização da produção florestal.

2.5.1 Tipos de satélites e sensores para o monitoramento florestal

Diversos tipos de imagens têm sido utilizados ao longo dos anos, para estimativas de biomassa e acompanhamento da dinâmica de crescimento do povoamento florestal. Embora a interferometria por radar e o Lidar (*Light Detection and Ranging*) sejam técnicas promissoras na estimativa da biomassa florestal, a necessidade imediata dessas estimativas e suas mudanças em longos períodos demonstra necessidade de alternativas, e, em obtenção de informações em grandes escalas e com alta recorrência, destaca-se a relevância das imagens ópticas. Estas últimas são essenciais devido à limitação das técnicas de sensoriamento remoto ativo em atender a tais demandas de forma abrangente. Assim, apesar de uma precisão potencialmente inferior, as imagens ópticas mantêm um papel crucial na monitorização da biomassa florestal (Maire *et al.*, 2011).

Satélites ópticos capturam imagens usando a luz visível e infravermelha, similar à fotografia, mas em uma escala muito maior. Eles são excelentes para capturar imagens de alta resolução da cobertura terrestre em condições claras, portanto necessitam da presença de luz solar (Liu *et al.*, 2021). Entre os sensores mais conhecidos estão o *Landsat* e o *Sentinel*. O programa *Landsat*, uma colaboração entre a NASA e o USGS, tem fornecido dados de alta resolução desde os anos 1970, sendo crucial para o monitoramento de mudanças ambientais a longo prazo, passando por diversas melhorias ao longo das décadas, inicialmente fornecia imagens com resolução de 60 metros a partir do *Landsat-1*, alcançando uma resolução de 30 metros com o *Landsat-8*, além do aumento do número de bandas captadas pelo sensor (Hemati *et al.*, 2021).

O programa *Copernicus*, sob a Agência Espacial Europeia (ESA), lançou o primeiro satélite *Sentinel* - *Sentinel-1A*, marcando o início de várias missões de satélite, incluindo *Sentinels-1, 2, 3 e 5*. Os satélites *Sentinel-2*, com seus instrumentos multispectrais, destacam-se por fornecer dados de alta resolução para monitoramento de cobertura e uso do solo, mudanças climáticas e monitoramento de desastres. Esses dados são essenciais para diversas aplicações, como agricultura, desenvolvimento urbano e monitoramento florestal,

complementando outros programas globais de satélite, como *Landsat*, para garantir a continuidade no monitoramento das dinâmicas na superfície da Terra (Phiri *et al.*, 2020).

2.5.2 Índices de vegetação e aplicação no manejo do eucalipto

Índices de vegetação são medidas quantitativas derivadas de dados de sensoriamento remoto, usadas para estimar a densidade e a saúde da vegetação em uma área. São obtidos a partir da combinação entre duas ou mais bandas presentes nas imagens capturadas pelos sensores ópticos mencionados (Huang *et al.*, 2021). Diversos estudos correlacionam índices de vegetação com a estimativa de biomassa do povoamento florestal (Deng *et al.*, 2020; Silva *et al.*, 2021; Villacrés *et al.*, 2021; Wongchai *et al.*, 2020), e, entre os mais conhecidos índices estão o NDVI (*Normalized Difference Vegetation Index*), EVI (*Enhanced Vegetation Index*) e SAVI (*Soil Adjusted Vegetation Index*).

Os índices de vegetação têm aplicações vitais no manejo do eucalipto. Primeiramente, permite a detecção precoce de estresse hídrico e doenças, possibilitando intervenções rápidas para proteger as plantações. Além disso, é uma ferramenta poderosa para a avaliação de biomassa e crescimento florestal, fornecendo informações detalhadas sobre a saúde e a produtividade das árvores (Maire *et al.*, 2011; Reis *et al.*, 2018; Silva *et al.*, 2021). Por fim, o monitoramento de mudanças na cobertura florestal é essencial para o manejo sustentável, permitindo o acompanhamento de práticas de reflorestamento e o impacto de eventos naturais ou atividades humanas nas florestas de eucalipto.

A coleta de dados de satélite pode ser realizada através de diversas plataformas, tanto públicas quanto privadas. Plataformas públicas como o USGS Earth Explorer e o ESA's Copernicus Open Access Hub oferecem acesso gratuito a uma vasta gama de dados de sensoriamento remoto. Por outro lado, empresas privadas fornecem dados especializados, muitas vezes com maior resolução e capacidades específicas.

O processamento destes dados envolve várias etapas, incluindo a correção atmosférica, para remover distorções causadas pela atmosfera, e a calibração espectral, para garantir que as medidas sejam consistentes e passíveis de comparação ao longo do tempo e entre diferentes sensores (Francisco *et al.*, 2020). Estas etapas são cruciais para garantir a precisão e a utilidade dos dados no monitoramento florestal.

2.6 Seleção de variáveis

A seleção de variáveis é um processo crítico na modelagem estatística e no aprendizado de máquina, especialmente relevante para o manejo florestal e a análise de dados de sensoriamento remoto e climáticos. Esse processo envolve identificar as variáveis mais informativas e relevantes para serem incluídas nos modelos preditivos, de modo a criar modelos mais eficientes, eliminando dados redundantes ou irrelevantes que podem complicar a modelagem e interpretação dos resultados. Ao identificar as variáveis mais significativas, os pesquisadores podem melhorar a precisão das previsões e reduzir a complexidade dos modelos, tornando-os mais fáceis de entender e mais rápidos para serem executados (Rostami; Berahmand; Forouzandeh, 2021).

No contexto do manejo florestal, a seleção eficaz de variáveis é essencial para entender e prever fenômenos complexos, como o crescimento e a saúde das florestas, onde modelos simplificados, mas precisos, podem significar a diferença entre prever corretamente o impacto das mudanças climáticas em povoamentos de eucalipto ou falhar em antecipar problemas críticos de saúde e crescimento das florestas, otimizando as decisões baseadas em dados precisos e relevantes (Li *et al.*, 2020).

2.6.1 Métodos de seleção

Existem vários métodos para a seleção de variáveis, cada um com suas vantagens e contextos ideais de aplicação. Os métodos de filtragem (*filters*), por exemplo, utilizam análises de correlação, variância e outras estatísticas para identificar e remover variáveis com pouco ou nenhum poder preditivo antes mesmo da modelagem. Em contrapartida, os métodos de embrulho (*wrapper*) atribuem um valor de importância para cada variável em relação a variável predita, modelos também conhecidos por serem “caixa-preta”. Já os métodos de incorporação (*embedded*), como a regularização LASSO, realizam a seleção de variáveis como parte integrante do processo de modelagem, penalizando os coeficientes de variáveis menos relevantes até que seu impacto no modelo seja minimizado ou eliminado (Roffo *et al.*, 2020).

Além do LASSO, outros métodos de incorporação, como *Ridge Regression* e *Elastic Net*, oferecem formas de penalizar a complexidade do modelo ao ajustar os coeficientes das variáveis. Enquanto o LASSO tende a zerar os coeficientes de variáveis menos importantes, possibilitando sua exclusão direta (Frejeiro-González; Febrero-Bande; González-Manteiga, 2022), o *Ridge Regression* penaliza os coeficientes de forma que variáveis menos significativas

tenham seus valores reduzidos, mas não necessariamente zerados (Schreiber-Gregory, 2018). O *Elastic Net* combina as propriedades do LASSO e do *Ridge*, oferecendo um compromisso entre a seleção de variáveis e a retenção de informações importantes quando há variáveis correlacionadas (Xu *et al.*, 2021).

Outra possibilidade é a seleção *stepwise*, uma técnica iterativa que avalia a importância das variáveis adicionando ou removendo-as uma a uma, baseando-se em critérios estatísticos específicos, como o p-valor, para determinar sua contribuição ao modelo. Existem duas abordagens principais dentro do *stepwise*: *forward selection*, que começa com nenhum preditor e adiciona variáveis uma a uma, e *backward elimination*, que começa com todas as variáveis candidatas e remove a menos significativa em cada etapa. Essa técnica é útil para reduzir a complexidade do modelo e identificar as variáveis mais impactantes, mas é criticada por algumas de suas propriedades estatísticas, como o aumento do risco de *overfitting* e a dependência da ordem de entrada das variáveis (XU; ZHANG, 2001).

Métodos baseados em árvores também são aplicados, como *Random Forest* e *Gradient Boosting*, oferecem uma medida intrínseca da importância das variáveis, baseada na contribuição de cada variável para a redução da impureza nos nós das árvores ou no aumento do ganho de informação. Essa abordagem fornece insights valiosos sobre quais variáveis são mais importantes para prever o resultado, permitindo uma seleção natural e intuitiva das variáveis mais relevantes para o modelo (Desai; Ouarda, 2021; Velthoen *et al.*, 2023).

As heurísticas, como os algoritmos genéticos, oferecem uma abordagem alternativa para a seleção de variáveis, simulando processos de evolução natural para explorar o espaço de busca de variáveis de forma mais eficiente. Esses métodos aplicam operações genéticas, como seleção, cruzamento e mutação, a uma população de soluções potenciais (conjuntos de variáveis), procurando otimizar um critério de desempenho específico. Essa abordagem é vantajosa por sua capacidade de escapar de ótimos locais e explorar grandes espaços de busca, tornando-a adequada para problemas de seleção de variáveis com muitas variáveis candidatas e relações complexas (Abdel-Basset; Abdel-Fatah; Sangaiah, 2018).

2.7 Inteligência artificial e aprendizado de máquina

A inteligência artificial (IA) tem ganhado destaque na engenharia florestal devido à sua superioridade em relação aos modelos tradicionais de regressão, especialmente pela capacidade de lidar com dados complexos e não-lineares. Haykin (2001) destacou sua eficiência na resolução de problemas complexos, enquanto Binoti, M., Binoti, D. e Leite (2013) ressaltaram

sua menor necessidade de amostras para obter estimativas precisas. Além disso, a IA não depende de premissas estatísticas específicas, proporcionando maior flexibilidade e tornando-se uma alternativa viável para a otimização de recursos no setor florestal.

Diante do crescimento expressivo das bases de dados e da complexidade dos problemas modernos, o aprendizado de máquina tem se consolidado como uma ferramenta essencial para identificar padrões e realizar previsões (Lidberg; Nilsson; Ågren, 2019). Esse campo abrange diversas técnicas inspiradas em fenômenos naturais, como redes neurais artificiais, computação evolutiva e sistemas *fuzzy* (Konar, 2005).

Embora as metaheurísticas sejam eficazes na busca de soluções, seus resultados podem variar conforme a semente aleatória inicial, exigindo processos iterativos intensos (Chumney; Simpson, 2006). Essas técnicas ajustam parâmetros até atingir um erro mínimo (Payandeh, 1983), tornando-se valiosas na exploração de espaços de busca complexos.

2.7.1 Programação genética

Dentre as metaheurísticas, destacam-se o algoritmo genético (AG) e a programação genética (PG), amplamente utilizados em problemas de otimização (Jin *et al.*, 2016). Esses algoritmos, baseados na seleção natural, geram múltiplas soluções por iteração, permitindo sua evolução até uma resposta ótima global (Li; Lucasius; Kateman, 1992).

O AG, proposto por Holland na década de 1990, é amplamente utilizado em diversas áreas da engenharia devido à sua capacidade de encontrar soluções ótimas em contextos complexos (Garg, 2016). Sua eficiência é influenciada pela codificação dos cromossomos, que representam soluções potenciais, e pelo tamanho da população inicial, que afeta o desempenho do algoritmo (Michalewicz, 1996).

A função de avaliação (*fitness function*) seleciona os cromossomos mais adequados para gerar novas soluções. O operador de crossover permite combinações genéticas, aumentando a diversidade das soluções, enquanto a mutação introduz variações aleatórias para evitar convergência prematura (Balieiro *et al.*, 2014). Para garantir eficiência, critérios de parada devem ser estabelecidos, como limite de iterações ou estabilidade dos resultados (Gomide, 2009).

A PG, proposta por Koza (1990), compartilha fundamentos com o AG, mas difere na abordagem: enquanto o AG requer um modelo predefinido, a PG, também chamada de regressão simbólica, gera automaticamente a estrutura do modelo e seleciona variáveis e coeficientes conforme o problema. Essa capacidade torna a PG uma ferramenta poderosa para

problemas de otimização complexos, permitindo maior exploração e busca por soluções robustas em diferentes áreas da engenharia.

3 CONSIDERAÇÕES FINAIS

Eventos climáticos adversos estão se tornando cada vez mais frequentes, impactando diretamente as estimativas de crescimento de povoamentos florestais, especialmente quando são utilizados modelos empíricos em sua aplicação. Para garantir um planejamento florestal eficiente, é essencial trabalhar com informações precisas e constantemente atualizadas, possibilitando ajustes estratégicos e a mitigação de impactos decorrentes da menor disponibilidade de madeira ao final do ciclo produtivo.

Nesse contexto, a utilização de estratégias híbridas de modelagem, que incorporam variáveis climáticas e ambientais, surge como uma abordagem promissora. A integração dessas variáveis a imagens de sensoriamento remoto permite a extração de índices de vegetação, viabilizando o mapeamento da sanidade do povoamento em escala reduzida. Dessa forma, torna-se possível refinar as estimativas em nível talhão a talhão, aumentando a precisão das projeções.

A Programação Genética (PG) para a seleção de variáveis mais adequadas para captar os efeitos climáticos e ambientais no crescimento florestal se destaca como uma ferramenta robusta. Essa metodologia permite explorar um amplo espaço de busca, considerando um grande número de variáveis, suas transformações matemáticas e interações, algo inviável por meio de métodos tradicionais. Além disso, o avanço computacional facilita a geração de modelos mais complexos e a integração de variáveis de forma automatizada, otimizando o processo de modelagem e melhorando a capacidade preditiva das estimativas florestais.

REFERÊNCIAS

ABDEL-BASSET, M.; ABDEL-FATAH, L.; SANGAIAH, A. Metaheuristic algorithms: A comprehensive review. **Computational intelligence for multimedia big data on the cloud with engineering applications**, p. 185-231, 2018.

ASSAD, E. D. *et al.* Mudanças do clima e a cultura do eucalipto. **O eucalipto e a Embrapa: quatro décadas de pesquisa e desenvolvimento**, Brasília, 2021.

BALIEIRO, A. *et al.* A multi-objective genetic optimization for spectrum sensing in cognitive radio. **Expert Systems with Applications**, Oxford, v. 41, n. 8, p. 3640–3650, 2014.

BATTAGLIA, M. *et al.* Photosynthetic temperature responses of *Eucalyptus globulus* and *Eucalyptus nitens*. **Tree physiology**, v. 16, n. 1-2, p. 81-89, 1996.

BINOTI, M. L. M. da S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v. 37, n. 4, p. 639-645, 2013.

BURKHART, H. E.; TOME, M. **Modeling forest trees and stands**. Springer, New York, 2012.

CAMPOE, O. C. *et al.* Meteorological seasonality affecting individual tree growth in forest plantations in Brazil. **Forest Ecology and Management**, v. 380, p. 149-160, 2016.

CASAS, G. G. *et al.* Improving Yield Projections from Early Ages in Eucalypt Plantations with the Clutter Model and Artificial Neural Networks. **Pertanika Journal of Science & Technology**, v. 30, n. 2, 2022.

CHUMNEY, E. C. G.; SIMPSON, K. N. **Methods and designs for outcomes research**. ASHP, 2006.

CLUTTER, J. L. Compatible growth and yield models for loblolly pine. **Forest Science**, v. 9, n. 3, p. 354-371, 1963.

CLUTTER, J. L. *et al.* **Timber management; a quantitative approach**. New York: John Wiley & Sons, 333p. 1983.

COELHO, J. C. F. *et al.* Wood grain angles variations in *Eucalyptus* and their relationships to physical-mechanical properties. **Holzforschung**, v. 74, n. 12, p. 1089-1097, 2020.

DENG, X. *et al.* Identification of short-rotation eucalyptus plantation at large scale using multi-satellite imageries and cloud computing platform. **Remote Sensing**, v. 12, n. 13, p. 2153, 2020.

DESAI, S.; OUARDA, T. B. M. J. Regional hydrological frequency analysis at ungauged sites with random forest regression. **Journal of Hydrology**, v. 594, p. 125861, 2021.

FANTUZZI NETO, H. **Qualidade da madeira de eucalipto para produção de celulose Kraft**. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, 2012.

FAO - Food and Agriculture Organization of the United Nations. **Global Forest Resources Assessment - 2020 - Key Findings**. Disponível em: <<https://www.fao.org/3/ca9825en/ca9825en.pdf>>. (Acesso em 07 de janeiro de 2025).

FERRAZ FILHO, A. C. *et al.* Dominant height projection model with the addition of environmental variables. **Cerne**, v. 17 (3), p. 427-433, 2011.

FRANCISCO, C. N. *et al.* Análise do impacto da correção atmosférica no cálculo do Índice de Vegetação por Diferença Normalizada a partir de Imagem Landsat 8/OLI. **Revista Brasileira de Geografia Física**, v. 13, n. 1, p. 76-86, 2020.

- FREIJEIRO-GONZÁLEZ, L.; FEBRERO-BANDE, M.; GONZÁLEZ-MANTEIGA, W. A critical review of LASSO and its derivatives for variable selection under dependence among covariates. **International Statistical Review**, v. 90, n. 1, p. 118-145, 2022.
- GARG, H. A hybrid GSA-GA algorithm for constrained optimization problems. **Information Sciences**, New York, v. 478, p. 292–305, 2016.
- GUEDES, I. C. L. *et al.* Geostatistical techniques and spatial interpolators in the stratification of Eucalyptus sp. Stands. **Ciência Florestal**, v. 22, n. 3, p. 541-550, 2012.
- GOMIDE, L. R. **Planejamento florestal espacial**. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2009.
- HALL, K. B. *et al.* A growth and yield model for Eucalyptus benthamii in the southeastern United States. **Forest Science**, v. 66, n. 1, p. 25-37, 2020.
- HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900p
- HEMATI, M. A. *et al.* A systematic review of landsat data for change detection applications: 50 years of monitoring the earth. **Remote Sensing**, v. 13, n. 15, p. 2869, 2021.
- HUANG, S. *et al.* A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. **Journal of Forestry Research**, v. 32, n. 1, p. 1-6, 2021.
- IBÁ - Indústria Brasileira de Árvores. **Relatório Anual Ibá 2024**, ano base 2023 Disponível em: <<https://www.iba.org/datafiles/publicacoes/relatorios/relatorio-anual-iba2024-compactado.pdf>> . (Acesso em 07 de janeiro de 2025).
- ILES, K. **A sampler of inventory topics**. Kim Iles & Associates Ltd., Nanimo, 2003.
- JIN, Y. *et al.* Selection of sand models and identification of parameters using an enhanced genetic algorithm. **International Journal for Numerical and Analytical Methods in Geomechanics**, v. 40, n. 8, p. 1219-1240, 2016.
- KONAR, A. **Computational Intelligence: Principles, Techniques and Applications**. New York, Springer-Verlag Berlin Heidelberg, 2005.
- KOZA, J. R. **Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems**. Stanford University, Department of Computer Science, Stanford, 1990.
- LI, Y. *et al.* Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. **Scientific reports**, v. 10, n. 1, p. 9952, 2020.
- LI, T.; LUCASIUS, C. B.; KATEMAN, G. Optimization of calibration data with the dynamic genetic algorithm. **Analytica Chimica Acta**, v. 268, n. 1, p. 123-134, 1992.
- LI, J.; ROY, D. P. A global analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring. **Remote Sensing**, v. 9, n. 9, p. 902, 2017.

LIDBERG, W.; NILSSON, M.; ÅGREN, A. Using machine learning to generate high-resolution wet area maps for planning forest management: A study in a boreal forest landscape. **Ambio**, Stockholm, p. 1-12, 2019.

LIU, X. *et al.* Integration of Sentinel-1 and ALOS/PALSAR-2 SAR datasets for mapping active landslides along the Jinsha River corridor. **Engineering Geology**, China, v. 284, p. 106033, 2021.

LORENÇONE, P. A. *et al.* Implementação do modelo de balanço hídrico climatológico de Thornthwaite e Mather (1995) no aplicativo IFCClima. **Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN)**, v. 4, n. 1, 2020.

MACHADO, C. M. *et al.* **Eucaliptocultura no Brasil, silvicultura, manejo e ambiência**. Viçosa: SIF, p. 284-302, 2014.

MAESTRI, R. **Modelo de crescimento e produção para povoamentos clonais de Eucalyptus grandis considerando variáveis ambientais**. Tese de Doutorado. Universidade Federal do Paraná, Curitiba, 143 p. , 2003.

MAIRE, G. *et al.* MODIS NDVI time-series allow the monitoring of Eucalyptus plantation biomass. **Remote sensing of Environment**, v. 115, n. 10, p. 2613-2625, 2011.

MICHALEWICZ, Z. **Genetic Algorithms + Data Structures = Evolution Programs**. 3. ed. New York: Springer-Verlag Berlin Heidelberg, 1996.

MOREIRA, J. M. M. A. P. *et al.* Importância e desempenho das florestas plantadas no contexto do agronegócio brasileiro. **Floresta**, v. 47, n. 1, p. 85-94, 2017.

NARDINI, C. *et al.* Biomass and radiation use efficiency in Eucalyptus plantations as affected by spacing of planting. **Scientia Forestalis**, v. 48, n. 128, p. 1-14, 2020.

PAYANDEH, B. Some applications of nonlinear regression models in forestry research. **The Forestry Chronicle**, v. 59, n. 5, p. 244-248, 1983.

PHIRI, D. *et al.* Sentinel-2 data for land cover/use mapping: A review. **Remote Sensing**, v. 12, n. 14, p. 2291, 2020.

QUEIROZ, T. B. *et al.* A New Method to Calibrate Cardinal Temperatures for Eucalyptus Plantation. **Forests**, v. 14, n. 8, p. 1631, 2023.

REIS, A. A. *et al.* Classification of Eucalyptus plantation Site Index (SI) and Mean Annual Increment (MAI) prediction using DEM-based geomorphometric and climatic variables in Brazil. **Geocarto International**, v. 37, n. 5, p. 1256-1273, 2022.

ROFFO, G. *et al.* Infinite feature selection: a graph-based feature filtering approach. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 43, n. 12, p. 4396-4410, 2020.

ROSTAMI, M.; BERAHMAND, K.; FOROUZANDEH, S. A novel community detection based genetic algorithm for feature selection. **Journal of Big Data**, v. 8, n. 1, p. 1-27, 2021.

SCHREIBER-GREGORY, D. N. Ridge Regression and multicollinearity: An in-depth review. **Model Assisted Statistics and Applications**, v. 13, n. 4, p. 359-365, 2018.

SALEKIN, S. *et al.* Hybrid height growth and survival model for juvenile *Eucalyptus globoides* (Blakely) and *E. bosistoana* (F. Muell) in New Zealand. **Forest Ecology and Management**, v. 490, p. 119074, 2021.

SCOLFORO, H. F. *et al.* Incorporating rainfall data to better plan *Eucalyptus* clones deployment in eastern Brazil. **Forest Ecology and Management**, v. 391, p. 145–153, 2017.

SCOLFORO, H. F. *et al.* *Eucalyptus* growth and yield system: Linking individual-tree and stand-level growth models in clonal *Eucalypt* plantations in Brazil. **Forest Ecology and Management**, v. 432, p. 1-16, 2019a.

SCOLFORO, H. F. *et al.* Stand-level growth and yield model system for clonal *eucalypt* plantations in Brazil that accounts for water availability. **Forest Ecology and Management**, v. 448, p. 22-33, 2019b.

SCOLFORO, H. F. *et al.* Modeling whole-stand survival in clonal *eucalypt* stands in Brazil as a function of water availability. **Forest Ecology and Management**, v. 432, p. 1002-1012, 2019c.

SILVA, A. K. V. *et al.* Predicting *eucalyptus* diameter at breast height and total height with UAV-based spectral indices and machine learning. **Forests**, v. 12, n. 5, p. 582, 2021.

SILVA, V. E. *et al.* Influência do clima na produtividade e resposta à seca do eucalipto e uma proposta de maximização da produtividade de madeira em função dos atributos do solo no Brasil. **Ciência Florestal**, v. 32, p. 523-547, 2022.

STAPE, J. L. *et al.* The Brazil *Eucalyptus* Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production. **Forest Ecology and Management**, Amsterdam, v. 259, n. 9, p. 1684-1694, 2010.

VELTHOEN, J. *et al.* Gradient boosting for extreme quantile regression. **Extremes**, v. 26, n. 4, p. 639-667, 2023.

VILLACRÉS, J. *et al.* Retrieval of vegetation indices related to leaf water content from a single index: A case study of *Eucalyptus globulus* (Labill.) and *Pinus radiata* (D. Don.). **Plants**, v. 10, n. 4, p. 697, 2021.

WONGCHAI, W. *et al.* Above-ground biomass estimation of *Eucalyptus* plantation using remotely sensed data and field measurements. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020. p. 012042.

XU, L.; ZHANG, W. Comparison of different methods for variable selection. **Analytica Chimica Acta**, v. 446, n. 1-2, p. 475-481, 2001.

XU, Q. F. *et al.* An elastic-net penalized expectile regression with applications. **Journal of Applied Statistics**, v. 48, n. 12, p. 2205-2230, 2021.

CASNATI, A. C. R. **Hybrid mensurational-physiological models for *Pinus taeda* and *Eucalyptus grandis* in Uruguay**. 2016. 208 p. PhD Dissertation. University of Canterbury, Canterbury, New Zealand, 2016.

REIS, A. A. *et al.* Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. **New Zealand Journal of Forestry Science**, v. 48, n. 1, p. 1-17, 2018.

SEGUNDA PARTE

ARTIGO

Antevendo o impacto climático e de sanidade florestal no potencial produtivo de plantios de eucalipto: uma abordagem para índice de sítio baseada em machine learning

Anticipating climate and forest health impacts on the productive potential of eucalyptus plantations: a machine learning-based site index approach

RESUMO

As mudanças climáticas e seus efeitos, como períodos prolongados de seca, impactam significativamente a produtividade dos povoamentos florestais. Diante desse desafio, este estudo propõe uma abordagem inovadora para aprimorar modelos convencionais de projeção da altura dominante, incorporando variáveis climáticas e de sensoriamento remoto. A metodologia desenvolvida busca aumentar a precisão das estimativas, permitindo projeções mais robustas em diferentes condições ambientais e escalas espaciais. Foram comparados três modelos: um modelo convencional (M_O), um modelo expandido com variáveis climáticas (M_C) e um modelo híbrido com variáveis climáticas e orbitais (M_C_SR). Os resultados demonstraram que a incorporação de variáveis ambientais reduziu significativamente o viés das projeções, especialmente em cenários de déficit hídrico, onde o erro sistemático do modelo base (-0,5 m) foi reduzido para -0,1 m nos modelos expandidos. Além disso, o coeficiente de determinação (R^2) aumentou de 82,3% em M_O para 84,6% em M_C_SR, evidenciando maior capacidade explicativa. A análise também revelou a importância de considerar o *carry-over effect*, ou seja, o atraso na resposta do crescimento das árvores a eventos climáticos passados. Esse efeito foi identificado pelo uso da Programação Genética (PG) na seleção de variáveis a serem inseridas nos modelos expandidos, por meio seleção recorrente de variáveis climáticas com atraso, reforçando sua relevância na previsão da produtividade florestal. Adicionalmente, a utilização de índices de vegetação derivados de sensoriamento remoto, como o SAVI, permitiu uma melhor representação da sanidade florestal, tornando o modelo mais sensível às variações espaciais da disponibilidade hídrica. Embora a projeção do índice orbital tenha introduzido um erro acumulado, os resultados indicam que sua inclusão melhora a acurácia das previsões e amplia a aplicabilidade dos modelos em diferentes cenários operacionais. A integração de técnicas híbridas baseadas em aprendizado de máquina com modelagem convencional representa um avanço metodológico promissor, possibilitando um manejo florestal mais eficiente, adaptável e resiliente às mudanças climáticas.

Palavras-chave: Modelos Híbridos; Programação Genética; Seleção de Variáveis; Sensoriamento Remoto.

ABSTRACT

Climate change and its effects, such as prolonged drought periods, significantly impact the productivity of forest stands. To address this challenge, this study proposes an innovative approach to improving conventional dominant height projection models by incorporating climatic and remote sensing variables. The developed methodology aims to enhance estimation accuracy, enabling more robust projections across different environmental conditions and spatial scales. Three models were compared: a conventional model (M_O), an expanded model incorporating climatic variables (M_C), and a hybrid model integrating both climatic and orbital variables (M_C_SR). The results demonstrated that incorporating environmental variables significantly reduced projection bias, particularly in water deficit scenarios, where the systematic error of the base model (-0.5 m) decreased to -0.1 m in the expanded models. Additionally, the coefficient of determination (R^2) increased from 82.3% in M_O to 84.6% in M_C_SR, indicating improved explanatory power. The analysis also highlighted the importance of considering the carry-over effect, which refers to the delayed response of tree growth to past climatic events. This effect was identified through the use of Genetic Programming (GP) in variable selection for the expanded models, where climatic variables with a time lag were recurrently selected, reinforcing their relevance in predicting forest productivity. Furthermore, the inclusion of remote sensing-derived vegetation indices, such as SAVI, improved the representation of forest health, making the model more sensitive to spatial variations in water availability. Although the projection of orbital indices introduced accumulated error, the results indicate that their inclusion enhances prediction accuracy and broadens the applicability of the models across different operational scenarios. The integration of hybrid techniques based on machine learning with conventional modeling represents a promising methodological advancement, enabling more efficient, adaptable, and resilient forest management in response to climate change.

Keywords: Hibrid Models; Genetic Programming, Feature Selection; Remote Sensing.

1 INTRODUÇÃO

A eucaliptocultura tem se expandido rapidamente, especialmente na última década, impulsionada pelo aumento da área cultivada e pelo crescimento das exportações de produtos madeireiros. Atualmente, o Brasil é o maior exportador global de celulose e o segundo maior produtor mundial. Com uma área de plantio que ultrapassa 7,6 milhões de hectares, é fundamental a execução de um planejamento sustentável, independentemente da destinação da madeira (IBÁ, 2024).

A qualidade de sítio é uma informação fundamental para determinar o potencial produtivo dos plantios florestais, sendo o índice de sítio uma das métricas mais utilizadas na composição do planejamento florestal. Esse índice representa a altura dominante do plantio projetada para uma idade de referência pré-definida, trazendo a vantagem de não ser influenciado pelo espaçamento do plantio e refletindo o histórico dos efeitos edafoclimáticos sobre a produtividade da floresta (Iles, 2003).

Os modelos convencionais utilizados para a estimativa do índice de sítio baseiam-se em séries históricas de medições, traçando tendências de crescimento da altura dominante a partir do comportamento observado ao longo do tempo, ou seja, esses modelos, apresentam elevada eficiência para o planejamento de longo prazo. Na eucaliptocultura, onde o crescimento das árvores é sensível às variações climáticas (Stape; Binkley; Ryan, 2004), os modelos convencionais de altura dominante podem ter sua precisão comprometida especialmente no curto e médio prazo dado a drástica alteração no ritmo de crescimento médio do eucalipto no caso de qualquer evento climático atípico no período (Scolforo *et al.*, 2019).

Nesse contexto, diversos estudos têm explorado abordagens híbridas buscando integrar variáveis ambientais aos modelos convencionais de maneira a permitir uma representação mais realista do crescimento das árvores de acordo com a condição climática vigente (Koirala; Montes; Bullock, 2021).

Maestri (2003) foi pioneiro em estudos dessa natureza na eucaliptocultura através da introdução de variáveis ambientais em modelos de crescimento da altura dominante, abordagem posteriormente replicada por Ferraz Filho *et al.* (2011), que incorporaram precipitação e radiação solar ao modelo de Chapman-Richards. Da mesma forma, Scolforo *et al.* (2016) exploraram a decomposição do parâmetro de inclinação (em função da precipitação do período) do modelo de Chapman-Richards demonstrando o quanto qualquer oscilação no regime hídrico tende a afetar significativamente a variação do crescimento das florestas de eucalipto. Koirala, Montes e Bullock (2021) demonstraram que a inclusão de variáveis climáticas relacionadas ao

balanço hídrico pode aprimorar a explicação do crescimento do Pinus. De forma semelhante, Kinane *et al.* (2022) utilizaram essas variáveis para estimar a área foliar do Pinus, evidenciando que o impacto do clima se manifesta em diferentes espécies, embora os períodos para sua influência no crescimento possam variar.

Vale ressaltar, no entanto, que oscilações climáticas, por si só, podem não ser suficiente para explicar plenamente o potencial produtivo de um determinado clone, especialmente considerando a ampla variabilidade genética existente dentro do gênero *Eucalyptus*. Assim, o sensoriamento remoto emerge como uma ferramenta crítica ao processo, permitindo a obtenção de informações acerca da condição fisiológica (sanidade) da floresta (Reis *et al.*, 2018). Os índices de vegetação derivados de imagens orbitais fornecem uma medida indireta do índice de área foliar (IAF), que, por sua vez, está diretamente relacionado à capacidade fotossintética do povoamento e, conseqüentemente, ao seu potencial de crescimento (Kinane *et al.*, 2022).

Diversos estudos têm buscado estabelecer relações quantitativas entre índices de vegetação obtidos por sensoriamento remoto e estimativas de biomassa em povoamentos florestais. Wongchai *et al.* (2020) utilizaram o índice TNDVI para estimar a biomassa de plantios de *Eucalyptus* na Tailândia, enquanto Reis *et al.* (2018) empregaram dados do satélite Landsat para estimar a área basal e o volume dessas florestas. Essa abordagem explora o potencial dos dados orbitais para fornecer informações em larga escala e com alta frequência temporal, permitindo um monitoramento mais dinâmico e preciso da produtividade florestal no nível de talhão (Deng *et al.*, 2020; Villacrés *et al.*, 2021).

Além disso, diversos estudos mostram que a integração de índices espectrais com variáveis dendrométricas melhora a acurácia das estimativas de biomassa. Maire *et al.* (2011) demonstraram essa relação ao estimar o índice de área foliar (IAF) com séries temporais de NDVI do satélite MODIS. Da mesma forma, Silva *et al.* (2021) previram o diâmetro à altura do peito (DAP) e a altura total de povoamentos de *Eucalyptus* com base em índices espectrais. Essas técnicas reduzem a dependência de medições manuais em campo, tornando a gestão florestal mais eficiente e adaptável a diferentes condições ambientais.

Incertezas climáticas têm se tornado cada vez mais um desafio crescente para a modelagem do crescimento da altura dominante de florestas de eucalipto no Brasil. Os modelos convencionalmente não incorporam de maneira eficaz a influência do clima na taxa de crescimento dos plantios. Além disso, esses modelos não consideram informações satelitais, que oferecem uma visão detalhada acerca da condição de sanidade das florestas. Enquanto as variáveis climáticas fornecem uma perspectiva macro, refletindo tendências regionais, os dados de sensoriamento remoto permitem uma análise micro, capturando variações locais talhão a

talhão, uma vez que o efeito do clima no crescimento das florestas depende da sanidade das mesmas.

Há um grande desafio na construção de modelos híbridos de índice de sítio que contemplem a inclusão de variáveis de clima e sensoriamento remoto dado a variedade de informações disponíveis. Nesse contexto, estudos voltados ao aprendizado de máquina têm sido cada vez mais explorados, oferecendo diferentes abordagens para a seleção das variáveis mais responsivas quanto a estimativa da variável de interesse (Miranda, 2022). Métodos de seleção de variável vão desde técnicas amplamente conhecidas como *stepwise* até abordagens de simulação, como por exemplo, a programação genética. A programação genética se destaca por sua capacidade não apenas de selecionar variáveis, mas também de construir combinações ótimas que melhor representam a interação entre o plantio e o ambiente, captando de forma mais precisa os efeitos das condições ambientais e de sanidade sobre o desenvolvimento futuro esperado para as florestas (Cunha, 2023).

Nesse contexto, este estudo objetivou o desenvolvimento de um novo modelo de crescimento da altura dominante, capaz de gerar projeções de curto e médio prazo com alta precisão tanto em nível regional quanto para aplicações em nível de talhão. Para isso, um modelo de altura dominante teve seus parâmetros decompostos para inclusão da melhor combinação de informações climáticas e/ou satelitais através do uso da programação genética. Os objetivos específicos do estudo foram: 1) identificar a melhor seleção e combinação de variáveis por meio de programação genética; 2) construir um modelo baseado na expansão das melhores variáveis climáticas para aplicação global; 3) construir um modelo que combine variáveis climáticas e dados de satélite para atualização do inventário florestal com máxima precisão.

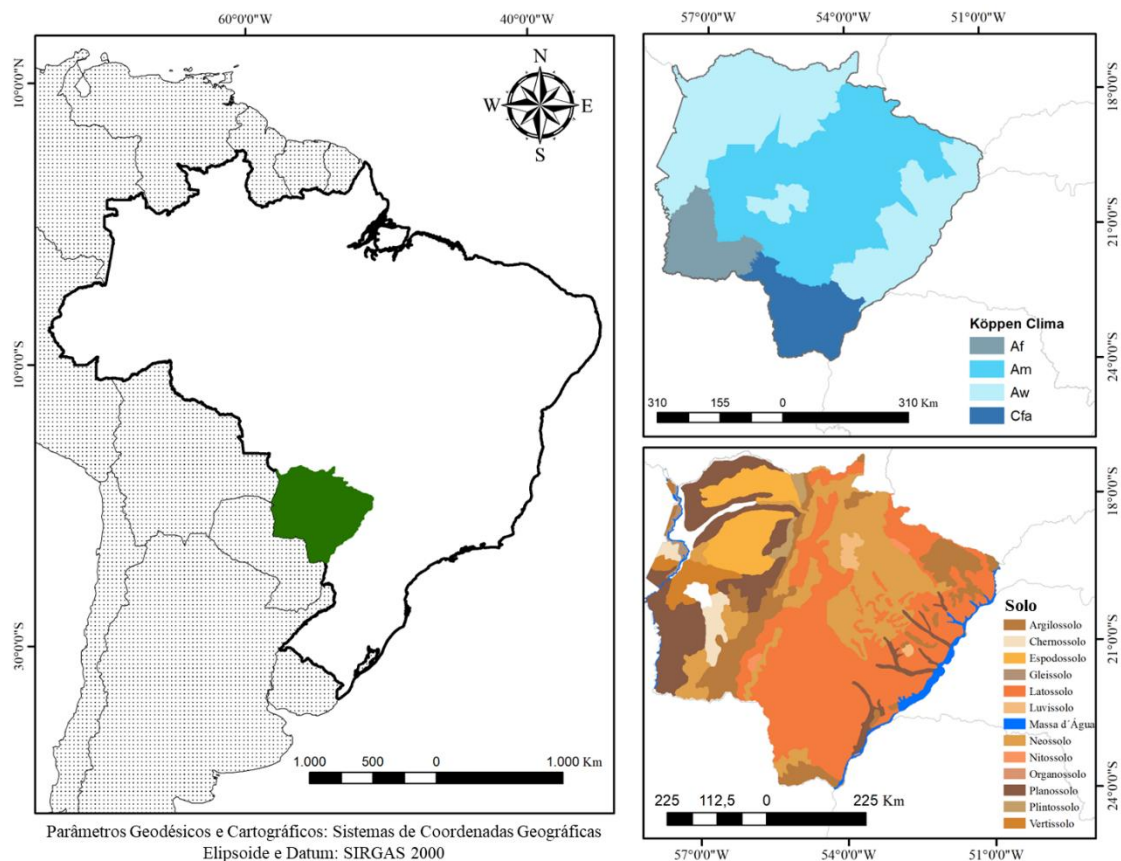
2 MATERIAL E MÉTODOS

2.1 Caracterização da área de estudo

O estudo foi conduzido no Mato Grosso do Sul abrangendo quatro regiões climáticas, segundo a classificação de Köppen (Alvares *et al.*, 2013). Essas regiões abrangem desde o clima Af (clima quente, sem estação seca), Aw (clima quente, com verões chuvosos), Am (clima de monção) até o Cfa (clima temperado, sem estação seca e verão quente). Nas regiões de estudo também há ampla diversidade de solos, havendo predominância de Latossolos, Neossolos e Argissolos, conforme a classificação da EMBRAPA (2013). Por fim, as altitudes das regiões

de estudo também apresentam variações significativas, oscilando entre 260 e 760 metros acima do nível do mar.

Figura 1 – Localização do estado de Mato Grosso do Sul e suas condições edafoclimáticas.



Fonte: Do autor (2025).

2.2 Base de dados de inventário

Os dados do estudo foram obtidos a partir de inventários anuais em plantios comerciais do gênero *Eucalyptus* com variação de idade das florestas entre 2 e 7 anos. As informações foram coletadas em parcelas circulares de 400 m². Em cada parcela, foi medido o diâmetro à 1,30 metros acima do solo (DAP) de todas as árvores, enquanto a altura total foi mensurada para as 10 árvores situadas na fileira central da parcela. A altura dominante foi mensurada para as 4 árvores mais grossas da parcela, conforme o conceito de Assmann (1970).

As coletas das parcelas tiveram início em 2010 e foram concluídas em 2024, abrangendo um período de 14 anos e permitindo a captura de uma ampla variedade de variações climáticas. Um total de 248 parcelas foram mensuradas no período englobando diferentes regimes de manejo (talhadia e alto fuste), diferentes níveis de produtividade e variados clones. Essa

diversidade possibilita uma análise mais abrangente dos fatores que influenciam o crescimento e a dinâmica da floresta ao longo do tempo. A altura dominante apresenta amplitude nas regiões de estudo, conforme demonstrado na Tabela 1.

Tabela 1 – Descrição geral da altura dominante da área de estudo nas diferentes regiões climáticas.

Região Climática	Alt. Dominante		
	Mín.	Méd.	Máx.
Clima 1	12,3	25,4	35,3
Clima 2	12,2	24,2	37,1
Clima 3	12,0	23,5	36,7
Clima 4	12,7	23,7	32,1

Fonte: Do autor (2025).

2.3 Base climática e compatibilização com os dados de inventário

Os dados climáticos utilizados neste estudo foram obtidos a partir do sistema *Mesoscale Convective Systems Rainfall Estimation* (MERGE), que integra estimativas de precipitação provenientes de satélites com informações de observação da superfície (INPE, 2020). A partir dessa combinação, é gerada uma grade espacial com resolução de 10 x 10 km, permitindo dados contínuos com frequências horária ou diária. A aplicação dessa metodologia é particularmente útil para estimar precipitação em locais onde não há dados de estações de monitoramento.

A disponibilidade de informações climáticas em escala diária permitiu a compatibilização dessa base de dados com o inventário florestal, inclusive, para identificação de qual a melhor combinação possível entre clima e resposta em crescimento florestal. Assim, os dados climáticos foram extraídos considerando desde o período de plantio até a medição de cada parcela.

As variáveis climáticas utilizadas neste estudo foram aquelas diretamente relacionadas à disponibilidade hídrica para o plantio. A Tabela 2 apresenta um panorama dessas variáveis, incluindo valores médios e acumulados ao longo do período analisado, considerando as 248 parcelas inventariadas. Foram consideradas as variáveis: déficit hídrico (DEF), excedente hídrico (EXC), precipitação (PPT), evapotranspiração potencial (ETP) e evapotranspiração real (ETR), permitindo uma caracterização abrangente das condições hídricas e sua influência na produtividade florestal.

Tabela 2 – Descrição geral das variáveis climáticas da área de estudo nas diferentes regiões.

Variável Climática	Média do ciclo (mm/ano)			Acumulado no ciclo (mm)		
	Mín.	Méd.	Máx.	Mín.	Méd.	Máx.
Precipitação	723	1.130	1.778	674	3.322	6.732
Déficit hídrico	-288	-731	-1.128	-230	-2.222	-4.697
Excedente hídrico	0	73	444	0	216	677
Evapotranspiração real	674	1.057	1.571	672	3.108	6.178
Evapotranspiração potencial	1.610	1.787	2.022	1.158	5.327	9.729

Fonte: Do autor (2025).

2.4 Base de dados de sensoriamento remoto e compatibilização com dados de inventário

Os dados de sensoriamento remoto utilizados neste estudo foram obtidos a partir do satélite Landsat 7, um satélite artificial norte-americano operado em parceria pela *National Aeronautics and Space Administration* (NASA) e pela *United States Geological Survey*. Esse satélite fornece imagens nas bandas RGB e infravermelho com resolução espacial de 30 metros e um período de revisitação de 16 dias (USGS, 2024). As imagens foram coletadas de 2010 a 2024, mesmo período em que as parcelas de campo foram mensuradas. A obtenção das imagens foi realizada por meio da plataforma *Google Earth Engine* (GEE) (GOOGLE, 2025), em conjunto com a linguagem R e o ambiente de desenvolvimento RStudio.

A partir das imagens adquiridas, foram calculados 19 índices de vegetação, cujas nomenclaturas, abreviações e fórmulas de cálculo estão descritas na Tabela 3. Cada índice possui um significado específico, refletindo distintos aspectos do desenvolvimento do plantio e permitindo diferentes formas de interpretação. O Índice de Vegetação por Diferença Normalizada (*Normalized Difference Vegetation Index* – NDVI), por exemplo, é amplamente empregado na avaliação da saúde da vegetação (MAIRE *et al.*, 2011), enquanto o Índice de Umidade por Diferença Normalizada (*Normalized Difference Moisture Index* – NDMI) apresenta forte correlação com o teor de umidade do plantio (Alonso *et al.*, 2024). A seleção e o cálculo foram realizados com o intuito de serem posteriormente analisados quanto às suas correlações com o crescimento da altura dominante.

Tabela 3 – Índices de vegetação utilizados no estudo e suas fórmulas.

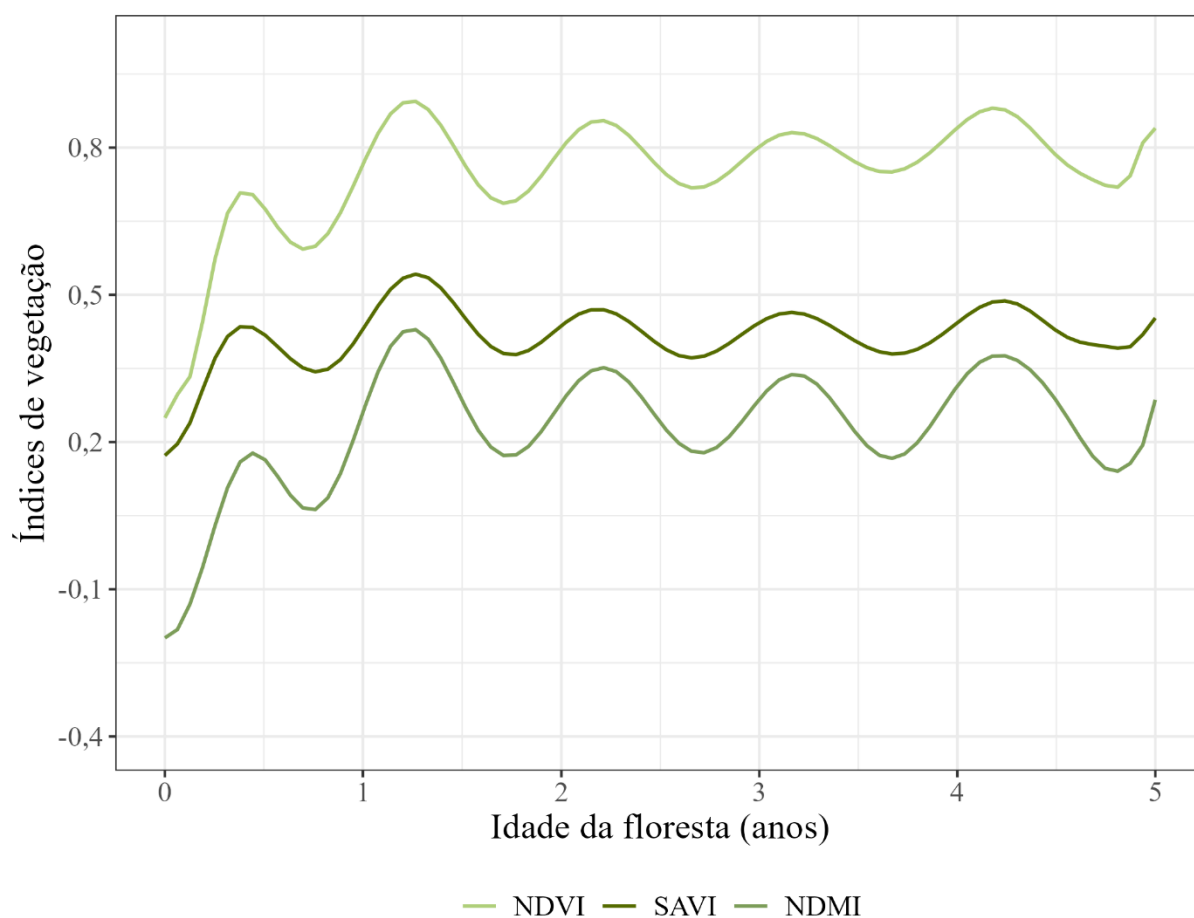
Índice de vegetação	Abreviação	Fórmula
Índice de Vegetação Resistente à Atmosfera	ARVI	$ARVI = \frac{(NIR - 2 * RED + BLUE)}{(NIR + 2 * RED + BLUE)}$
Índice de Estresse Hídrico por Doença	DSWI	$DSWI = \frac{RED\ EDGE\ 4 + GREEN}{SWIR\ 1 + RED}$
Índice de Vegetação Melhorado	EVI	$EVI = 2,5 * \frac{NIR - RED}{NIR + 6 * RED - 7,5 * BLUE + 1}$
Índice de Vegetação Melhorado 2	EVI2	$EVI2 = 2,5 * \frac{NIR - RED}{NIR + 2,4 * RED + 1}$
Índice de Vegetação da Proporção Verde	GRVI	$GRVI = \frac{GREEN - RED}{GREEN + RED}$
Índice de Umidade	MSI	$MSI = \frac{SWIR\ 2}{NIR}$
Razão Simples Modificada	MSR	$MSR = \frac{NIR - RED - 1}{\sqrt{1 + NIR - RED}}$
Índice de Queima Normalizada	NBR	$NBR = \frac{NIR - SWIR\ 2}{NIR + SWIR\ 2}$
Índice de Queima Normalizada 2	NBR2	$NBR2 = \frac{SWIR\ 2 - SWIR\ 1}{SWIR\ 2 + SWIR\ 1}$
Índice de Diferença Normalizada de Umidade	NDMI	$NDMI = \frac{NIR - SWIR\ 1}{NIR + SWIR\ 1}$
Índice de Água por Diferença Normalizada	NDWI	$NDWI = \frac{GREEN - NIR}{GREEN + NIR}$
Índice de Vegetação por Diferença Normalizada	NDVI	$NDVI = \frac{NIR - RED}{NIR + RED}$
Índice de Vegetação Ajustado ao Solo Otimizado	OSAVI	$OSAVI = \frac{NIR - RED}{NIR + RED + 0,16}$
Índice de Vegetação por Diferença Renormalizada	RDVI	$RDVI = \frac{RED\ EDGE\ 4 - RED}{\sqrt{RED\ EDGE\ 4 + RED}}$
Índice de Vigor Relativo	RVI	$RVI = \frac{RED}{NIR}$
Índice de Vegetação Ajustado ao Solo	SAVI	$SAVI = \frac{NIR - RED}{NIR + RED + 0,5} * 1,5$
Índice de Pigmento Insensível à Estrutura	SIPI	$SIPI = \frac{NIR - BLUE}{NIR - RED}$
Índice de Simples Razão	SR	$SR = \frac{NIR}{RED}$
Índice de Vegetação Transformado	TVI	$TVI = \sqrt{NDVI + 0,5}$

Fonte: Do autor (2025).

O intervalo de coleta para cada parcela considerou os 12 meses anteriores à medição, correspondendo ao último ano de crescimento do plantio, com o objetivo de avaliar a sanidade da floresta. A escolha desse período deve-se ao fato de que, ao longo de um ano, o plantio atravessa diferentes estações, caracterizadas por variações nos regimes hídricos, o que influencia diretamente a interpretação dos índices de vegetação. Ao adotar um intervalo de 12 meses, busca-se capturar integralmente os efeitos sazonais, reduzindo o viés associado a

variações temporais e garantindo uma análise mais representativa das condições ambientais ao longo do ciclo anual. Após a fase inicial de estabelecimento do povoamento, o índice de vegetação se estabiliza, apresentando variações sazonais com comportamento senoidal (Figura 3). Os dados de sensoriamento remoto utilizados podem ser observados na Tabela 4, utilizando como exemplo NDVI, NDMI e Índice de Vegetação Ajustado ao Solo (*Soil Adjusted Vegetation Index* – SAVI).

Figura 2 – Comportamento sazonal dos índices de vegetação ao longo do crescimento do plantio.



Fonte: Do autor (2025).

Tabela 4 – Descrição geral de alguns índices de vegetação utilizados no estudo.

Índice de vegetação	Valor Anual (considerando 12 meses de dados)		
	Mín.	Méd.	Máx.
NDVI	0,154	0,799	0,942
NDMI	-0,248	0,344	0,792
SAVI	0,105	0,453	0,676

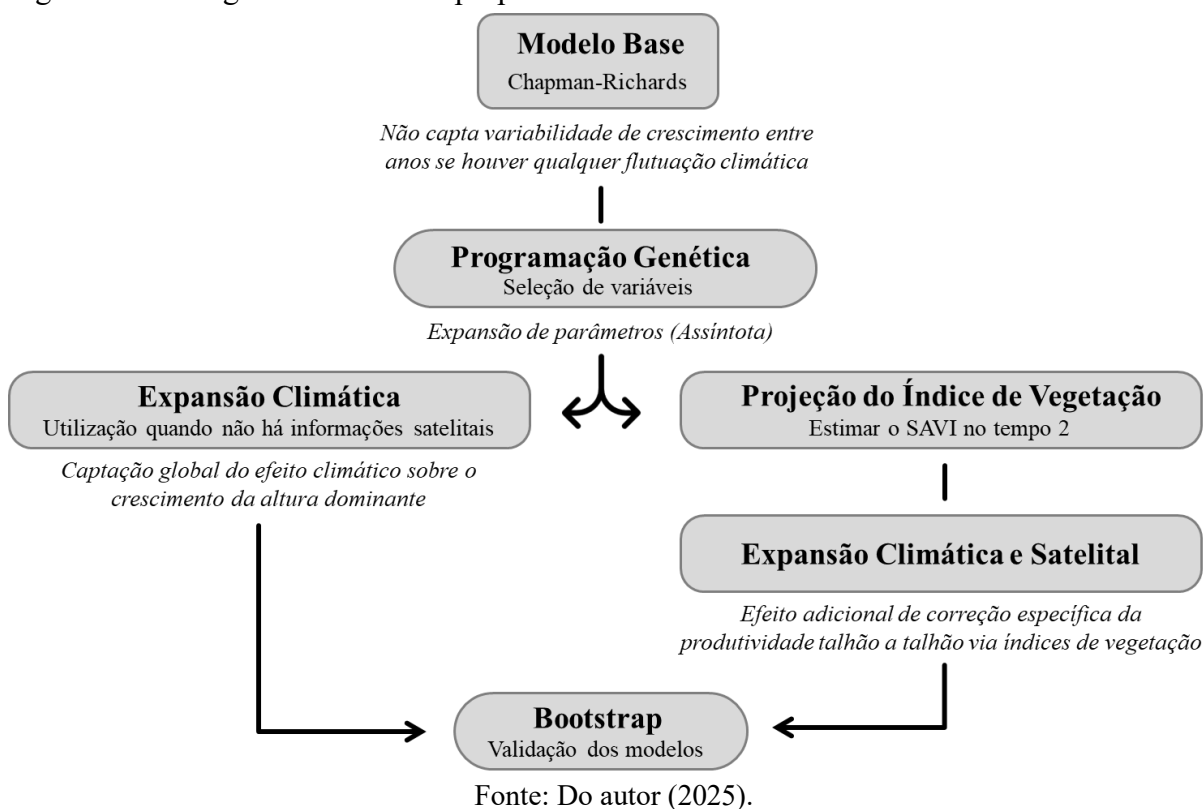
Fonte: Do autor (2025).

2.5 Sistema de modelagem

Diversos modelos podem ser aplicados para representar o crescimento da altura dominante em povoamentos de eucalipto. Com base no estudo realizado por Scolforo *et al.* (2019), optou-se por utilizar o modelo Chapman-Richards como referência, devido à sua ampla aplicação na modelagem do crescimento florestal e à sua flexibilidade para descrever diferentes padrões de crescimento ao longo do tempo. Esse modelo permite a representação do crescimento sigmoidal das árvores, sendo amplamente utilizado por sua capacidade de capturar variações na dinâmica de crescimento sob diferentes condições ambientais.

O foco principal do estudo foi realizar a expansão de parâmetros do modelo convencional, integrando variáveis climáticas e de sensoriamento remoto, para melhor representar a taxa de crescimento da altura dominante do eucalipto em função das oscilações ambientais. No entanto, devido à alta dimensionalidade do problema, considerando a grande número de variáveis disponíveis e suas interações, tornou-se essencial adotar uma abordagem de aprendizado de máquina (programação genética) para garantir que apenas as variáveis mais relevantes fossem incorporadas ao modelo (Figura 4).

Figura 3 – Fluxograma do método proposto e como ele foi executado.



Toda programação foi realizada no software R (R CORE TEAM, 2024). Os coeficientes do modelo original foram ajustados por meio da função `nls` do pacote `stats` (R CORE TEAM, 2024). Já os modelos híbridos foram ajustados via programação genética por meio do pacote `gramEvol` (Noorian; De Silva; Leong, 2016).

O processamento e a organização dos dados foram realizados com os pacotes `dplyr` (Wickham, 2023) e `tidyr` (Wickham; Vaughan; Girlich, 2024), enquanto a validação dos modelos contou com o uso do pacote `purrr` (Wickham; Henry, 2023). Para a visualização dos dados, foi utilizado o pacote `ggplot2` (Wickham, 2016).

2.5.1 Modelo base (convencional) de Chapman-Richards

O modelo de Chapman-Richards (denominado no estudo como M_O , por ser o modelo original) (Equação 1) é amplamente utilizado para descrever o crescimento da altura dominante em plantios de eucalipto. Esse modelo sigmoidal é definido por três principais parâmetros: a taxa de crescimento, que determina a velocidade de incremento da altura; o ponto de inflexão, representando a idade em que o crescimento relativo é máximo; e a assíntota, que corresponde ao valor máximo teórico alcançado pela variável ao longo do tempo.

$$Hd = a (1 - \exp(bt))^{-c} \quad (1)$$

em que, Hd é a altura dominante em metros; t é a idade em anos; a , b e c são parâmetros a serem estimados.

A construção das curvas de índice de sítio foi realizada por meio do método da Diferença Algébrica (ADA), primeiramente descrito por Bailey e Clutter (1974), que permite a derivação de equações a partir de qualquer modelo que relacione a altura dominante das árvores com a idade do povoamento. Uma das principais vantagens desse método é a flexibilidade na obtenção de curvas de índice de sítio polimórficas ou anamórficas, independentemente da idade de referência escolhida.

O parâmetro de referência de sítio selecionado foi o ponto de inflexão da curva de crescimento. A transformação do modelo via método da Diferença Algébrica (ADA) (Equação

2) resultou em curvas polimórficas, nas quais diferentes classes produtivas apresentam diferentes trajetórias de crescimento ao longo do tempo.

$$Hd_2 = a \left(1 - \exp(bt_2) \right)^{\frac{\ln\left(\frac{Hd_1}{a}\right)}{\ln(1 - \exp(bt_1))}} \quad (2)$$

em que, Hd_1 e Hd_2 são as alturas dominantes em metros no tempo 1 e tempo 2, respectivamente; t_1 e t_2 são as idades em anos no tempo 1 e tempo 2, respectivamente; a e b são parâmetros a serem estimados.

2.5.2 Expansão dos parâmetros via programação genética

2.5.2.1 Modelo expandido com variáveis climáticas

O primeiro modelo desenvolvido neste estudo incorpora exclusivamente variáveis climáticas na assíntota do modelo base (M_O) e foi denominado Modelo com Expansão Climática (M_C). A principal motivação para a construção desse modelo é avaliar o impacto médio das variações no regime hídrico sobre a taxa de crescimento da altura dominante, possibilitando a correção de possíveis vieses nas estimativas e capturando a influência dos padrões sazonais e eventos climáticos extremos no desenvolvimento das árvores.

A seleção inicial de variáveis restringiu-se a informações climáticas, com o objetivo central de aprimorar a projeção da altura dominante e mitigar distorções induzidas por eventos climáticos atípicos, como por exemplo secas prolongadas. A hipótese fundamental subjacente a essa abordagem é que a incorporação dessas variáveis permite ao modelo ajustar-se dinamicamente às variações ambientais, refletindo de maneira mais realista a resposta do crescimento florestal às condições climáticas.

Para a construção do modelo M_C, a Programação Genética (PG) foi utilizada como ferramenta para explorar o espaço de busca das variáveis climáticas previamente selecionadas. O algoritmo teve a capacidade de avaliar individualmente cada variável e, simultaneamente, criar combinações e transformações matemáticas que melhor representassem a relação entre os fatores climáticos e o crescimento das árvores. Esse processo permitiu a modelagem de relações não lineares complexas, possibilitando uma representação mais fiel da influência de variáveis como precipitação acumulada, evapotranspiração real e potencial, déficit e excedente hídrico na dinâmica de crescimento da altura dominante.

2.5.2.2 Modelo expandido com variáveis climáticas e de sensoriamento remoto

A segunda estrutura do modelo expandido, denominada Modelo com Expansão Climática e Satelital (M_C_SR) incorporou simultaneamente variáveis climáticas e de sensoriamento remoto, permitindo uma análise mais abrangente e detalhada do crescimento da altura dominante. Enquanto as variáveis climáticas possibilitam capturar de forma precisa o impacto médio das variações sazonais e eventos extremos, os índices de vegetação derivados do sensoriamento remoto fornecem informações adicionais sobre a saúde e o vigor do plantio em diferentes escalas espaciais e temporais. Essa abordagem possibilita não apenas uma modelagem mais refinada do crescimento florestal, mas também um monitoramento contínuo e atualizado das condições do povoamento.

Para essa estrutura de modelo, foram testados diferentes índices espectrais, como NDVI, SAVI, NDMI, entre outros indicadores relacionados à biomassa e ao estresse hídrico. A inclusão dessas variáveis permitiu avaliar a resposta do plantio às condições ambientais de forma mais granular, capturando padrões de crescimento que podem não ser detectáveis apenas com dados climáticos. Além de aprimorar a capacidade do modelo de representar a influência do clima sobre o desenvolvimento florestal, essa abordagem pode aumentar a precisão na escala média e uma boa acurácia ao nível de talhão, tornando viável sua aplicação para atualizações recorrentes do inventário florestal e suporte a decisões operacionais.

No entanto, para tornar a aplicação desse modelo viável, foi necessário desenvolver um modelo secundário para projetar o índice de vegetação em idades futuras. Optou-se por utilizar um modelo linear generalizado para estimar a projeção do índice espectral, permitindo assim obter uma estimativa do índice de vegetação no futuro. A Equação 3 apresenta uma estrutura de modelo de projeção semelhante à adotada neste estudo, utilizando o SAVI como variável orbital de referência, a precipitação como variável climática de referência e a PVAR50 como variável biométrica de referência. Esta última representa o percentual de biomassa das 50% menores árvores em relação à biomassa total do povoamento, servindo como um indicador da heterogeneidade do plantio. Assim, um valor de PVAR50 igual a 50 indica um plantio completamente homogêneo, enquanto valores inferiores evidenciam maior desigualdade na distribuição da biomassa entre as árvores.

$$SAVI_2 \sim f(SAVI_1 + \text{Clima} + \text{Variáveis Florestais} + e) \quad (3)$$

em que, $SAVI_2$ representa a variável orbital no tempo 2; $SAVI_1$ representa a variável orbital no tempo 1; $Precipitação_1$ representa alguma variável climática que possa ter relação com a variável orbital no futuro; $PVAR50_1$ representa alguma variável do povoamento que possa ter relação com a variável orbital no futuro; a_0 , a_1 , a_2 e a_3 são parâmetros a serem estimados; e é o erro do modelo.

2.5.2.3 Funcionamento da programação genética para seleção e combinação de variáveis

A Programação Genética (PG), é uma técnica de aprendizado de máquina e otimização inspirada na evolução biológica. Desenvolvida por Koza (1990), a PG é uma extensão dos Algoritmos Genéticos (AG) e se destaca pela capacidade de gerar modelos matemáticos otimizados sem a necessidade de impor previamente uma estrutura funcional rígida. Sua aplicação é especialmente útil em problemas de otimização com espaços de busca complexos, onde a relação entre variáveis não é trivial.

A Programação Genética (PG) segue uma sequência estruturada de operações evolutivas para encontrar soluções otimizadas para um determinado problema. Inicialmente, as soluções candidatas, denominadas cromossomos, são codificadas em estruturas hierárquicas, geralmente representadas por árvores ou grades. Nessa configuração, os nós correspondem a operadores ou funções matemáticas, enquanto as arestas representam os argumentos ou variáveis associadas. Essa estrutura possibilita a construção de diferentes expressões que serão avaliadas ao longo do processo evolutivo.

A primeira etapa do processo consiste na geração de uma população inicial, composta por um conjunto de soluções candidatas distribuídas aleatoriamente. Cada indivíduo dentro dessa população apresenta uma combinação única de operadores e variáveis, que definem possíveis soluções para o problema em questão. No presente estudo, adotou-se uma população inicial de 100 indivíduos, garantindo diversidade suficiente para o processo evolutivo.

Em seguida, cada indivíduo passa por uma função de avaliação (fitness), cujo objetivo é quantificar a adequação da solução ao problema. Para este estudo, a métrica utilizada foi a raiz do erro quadrático médio (RMSE), derivado de um modelo de regressão não linear múltiplo. O algoritmo foi ajustado de forma iterativa até que os coeficientes do modelo respeitassem as correlações estabelecidas entre a altura dominante e as variáveis independentes.

Com base na avaliação de desempenho, ocorre a seleção dos indivíduos mais aptos para a próxima geração. Esse processo segue um critério probabilístico em que soluções com melhor

desempenho possuem maior chance de serem escolhidas, mas soluções menos aptas também podem ser selecionadas para manter a diversidade genética da população. Neste estudo, adotou-se o método de seleção por torneio, no qual indivíduos são comparados entre si em confrontos aleatórios. Esse método não elitista reduz a probabilidade de o algoritmo ficar preso em ótimos locais, favorecendo a busca por soluções globais mais robustas.

A recombinação genética (crossover) é a principal operação evolutiva da PG, permitindo a geração de novas soluções a partir da combinação de indivíduos selecionados. Esse processo ocorre por meio da troca de variáveis entre cromossomos, promovendo a criação de descendentes que herdam características distintas dos pais, potencialmente melhorando o desempenho da população ao longo das gerações.

Além do cruzamento, o algoritmo também incorpora a mutação, que introduz variações aleatórias nos indivíduos da população. Esse operador desempenha um papel essencial na prevenção da convergência prematura para soluções subótimas, garantindo que a busca continue explorando o espaço de soluções. No presente estudo, adotou-se um critério de mutação aleatória aplicada a 5% da população, com uma probabilidade de ocorrência de 50% para cada indivíduo selecionado.

Esse processo evolutivo se repete ao longo de múltiplas gerações, refinando gradativamente as soluções por meio da interação entre os operadores genéticos. No caso deste estudo, a execução do algoritmo foi conduzida por 200 iterações, definidas como critério de parada. Esse número foi estabelecido com base na análise da convergência do modelo, assegurando que a solução final fosse representativa da relação entre as variáveis climáticas e orbitais e a dinâmica de crescimento da altura dominante.

No presente estudo, as funções matemáticas consideradas na construção dos modelos incluíram o inverso $\left(\frac{1}{x}\right)$, o logaritmo natural $(\ln(x))$ e a função exponencial (e^x) . Os operadores aritméticos utilizados para conectar essas funções e estruturar as equações foram adição, subtração, multiplicação e divisão. Essa abordagem permitiu que a PG explorasse diferentes combinações de variáveis e estruturas funcionais, resultando em modelos com maior capacidade preditiva e melhor ajuste às condições específicas dos dados observados.

É importante destacar que os modelos de altura dominante apresentam natureza não linear, o que torna seu ajuste mais complexo e exige a definição de parâmetros iniciais coerentes para viabilizar sua conversão. Em contraste, no modelo de projeção do índice espectral, optou-se por manter uma estrutura linear. A Programação Genética (PG) pode ser aplicada de diferentes maneiras, seja selecionando variáveis para modelos não lineares e expandindo seus

parâmetros, seja construindo modelos do zero, formulando autonomamente a estrutura mais adequada, e, ambas as aplicações foram utilizadas neste estudo.

2.5.3 Validação dos modelos

Após o ajuste dos diferentes modelos, foi conduzida uma análise comparativa detalhada entre estes. Essa análise avaliou o impacto da inclusão de variáveis climáticas isoladamente e em combinação com variáveis de sensoriamento remoto, buscando identificar o modelo mais responsivo as variações ambientais. A performance dos modelos foi avaliada por meio de estatísticas clássicas de avaliação, incluindo coeficiente de determinação (R^2), erro médio absoluto (MAE), raiz do erro quadrático médio (RMSE) e viés do modelo (BIAS), garantindo uma análise abrangente da acurácia e precisão das estimativas.

$$R^2 = 1 - \frac{\sum_{i=1}^n (Hd_i - \widehat{Hd}_i)^2}{\sum_{i=1}^n (Hd_i - \overline{Hd})^2} \quad (4)$$

$$MAE (m) = \frac{\sum_{i=1}^n |Hd_i - \widehat{Hd}_i|}{n} \quad (5)$$

$$RMSE (m) = \sqrt{\frac{\sum_{i=1}^n (Hd_i - \widehat{Hd}_i)^2}{n}} \quad (6)$$

$$BIAS (m) = \frac{\sum_{i=1}^n (Hd_i - \widehat{Hd}_i)}{n} \quad (7)$$

em que, Hd_i é a altura dominante observada do indivíduo i ; $\widehat{Hd}_i$ é a altura dominante estimada do indivíduo i ; $\overline{Hd}$ é a altura média observada da população; n é o número indivíduos da população.

Além dessas métricas, foi realizada uma validação por reamostragem via Bootstrap com reposição, permitindo testar a robustez do modelo em diferentes cenários gerados a partir de novas amostras randômicas. Para o presente estudo, foram conduzidas 500 repetições, possibilitando a construção de histogramas que ilustram a distribuição dos erros e o desempenho do modelo sob condições adversas. Essa abordagem forneceu uma visão mais ampla da estabilidade do modelo e de sua capacidade preditiva em diferentes contextos.

Todas as validações foram realizadas a partir de uma abordagem comparativa entre diferentes cenários climáticos, com ênfase em condições adversas. Para a definição de cenários com regime hídrico seco, considerou-se como critério a ocorrência de precipitação média anual

inferior a 90% da normal climatológica histórica. Assim, tendo como valor de referência, para exemplificação, uma precipitação histórica média anual de 1.200 mm, os cenários secos seriam definidos como aqueles com precipitação média inferior a 1.080 mm. Por fim, é válido ressaltar que toda a validação do modelo expandido com dados satelitais e climáticos norteou-se no erro agregado via aplicação inicial do modelo de projeção do índice espectral e sua utilização na aplicação do modelo de altura dominante.

3 RESULTADOS

3.1 Modelos de projeção de altura dominante

Os parâmetros ajustados para os modelos foram estatisticamente significativos ($\alpha = 0,05$) e apresentaram sinais e valores biologicamente coerentes (Tabela 5).

Tabela 5 – Coeficientes ajustados para os três modelos avaliados.

Modelo	a <i>Assíntota</i>	a1 $\text{Ln}\left(\frac{1}{\text{Precipitação acum. com delay}}\right)$	a2 <i>SAVI</i>	b <i>Taxa de crescimento</i>
M_O	41,3585	-	-	-0,1428
M_C	-18,3170	-5,6786	-	-0,1176
M_C_SR	-23,7188	-5,9349	5,9437	-0,1219

Fonte: Do autor (2025).

O parâmetro a do modelo M_O indica que a assíntota da altura dominante projetada pode alcançar, no máximo, 38,3 metros. No entanto, em condições que o regime hídrico estiver abaixo do esperado, o potencial de crescimento do plantio não será o mesmo.

Após a exploração exaustiva do espaço de busca envolvendo todas as variáveis disponíveis, a Programação Genética (PG) identificou a variável $\text{Ln}\left(\frac{1}{\text{Precipitação acumulada com delay}}\right)$ como a que apresentou o melhor desempenho em termos de ajuste estatístico e capacidade preditiva para o modelo M_C. Essa variável demonstrou a melhor função de fitness, indicando que foi a mais eficaz na captura das oscilações climáticas e seu impacto na estimativa da altura dominante (Equação 8).

$$Hd_2 = \left(-18,32 - 5,68 * \ln\left(\frac{1}{ppt_2}\right) \right) (1 - \exp(-0,12 * t_2)) \frac{\ln\left(\frac{Hd_1}{\left(-18,32 - 5,68 * \ln\left(\frac{1}{ppt_1}\right)\right)}\right)}{\ln(1 - \exp(-0,12 * t_1))} \quad (8)$$

em que, ppt é a precipitação acumulada com delay de um ano.

Já no modelo M_C_SR foi realizada a busca para a seleção de um indicador orbital que representasse a sanidade do povoamento. O Índice de Vegetação Ajustado ao Solo (SAVI) foi a métrica espectral selecionada, por apresentar forte correlação com a sanidade da vegetação e, consequentemente, com sua condição fisiológica e crescimento (Equação 9).

$$Hd_2 = \left(-23,72 - 5,93 * \ln\left(\frac{1}{ppt_2}\right) + 5,94 * SAVI_2 \right) (1 - \exp(-0,12 * t_2)) \frac{\ln\left(\frac{Hd_1}{\left(-23,72 - 5,93 * \ln\left(\frac{1}{ppt_1}\right) + 5,94 * SAVI_1\right)}\right)}{\ln(1 - \exp(-0,12 * t_1))} \quad (9)$$

em que, ppt é a precipitação acumulada com delay de um ano e $SAVI$ é o índice de vegetação ajustado ao solo dos últimos doze meses.

3.2 Modelo de projeção do índice de vegetação

A projeção de índices de vegetação para o futuro representa um desafio, pois esses índices refletem as condições atuais da vegetação, capturando características como sanidade, biomassa e vigor do plantio. Eles são derivados da reflectância espectral e estão fortemente correlacionados a métricas estruturais da floresta, como o Índice de Área Foliar (IAF) e a biomassa aérea. Assim, a projeção dessas variáveis exige abordagens metodológicas robustas, capazes de lidar com a complexidade das interações entre clima, crescimento vegetal e respostas espectrais da vegetação.

Neste estudo, adotou-se um modelo linear generalizado (GLM) para projetar as variáveis orbitais no futuro. Essa abordagem possibilita a incorporação de variáveis explicativas, como fatores climáticos e fenológicos, ajustando os parâmetros do modelo para minimizar erros preditivos. A formulação do modelo (Equação 10), demonstra como essas variáveis são combinadas para capturar a dinâmica da vegetação e suas interações com o ambiente. Assim, a estratégia adotada viabiliza a estimativa futura dos índices de vegetação, proporcionando maior aderência às condições ambientais e estruturais da floresta.

$$SAVI_2 = a_0 + a_1 * SAVI_1 + a_2 * PVAR50_1 + a_3 * Idade_2 + a_4 * \frac{1}{Def.acum.1} \quad (10)$$

em que, $SAVI_1$, valor do SAVI na idade 1 da medição, refletindo a condição da vegetação; $PVAR50_1$, percentual de biomassa representado pelas 50% menores árvores, indicando a heterogeneidade do plantio; $Idade_2$: referente ao tempo 2, logo a idade para a qual a projeção estará sendo realizada; $\frac{1}{Def.acum.1}$, inverso do déficit hídrico acumulado até a idade 1, refletindo as condições hídricas do solo.

Os parâmetros ajustados para o modelo de projeção do SAVI foram estatisticamente significativos ($\alpha = 0,05$) e apresentaram sinais e valores biologicamente coerentes (Tabela 6).

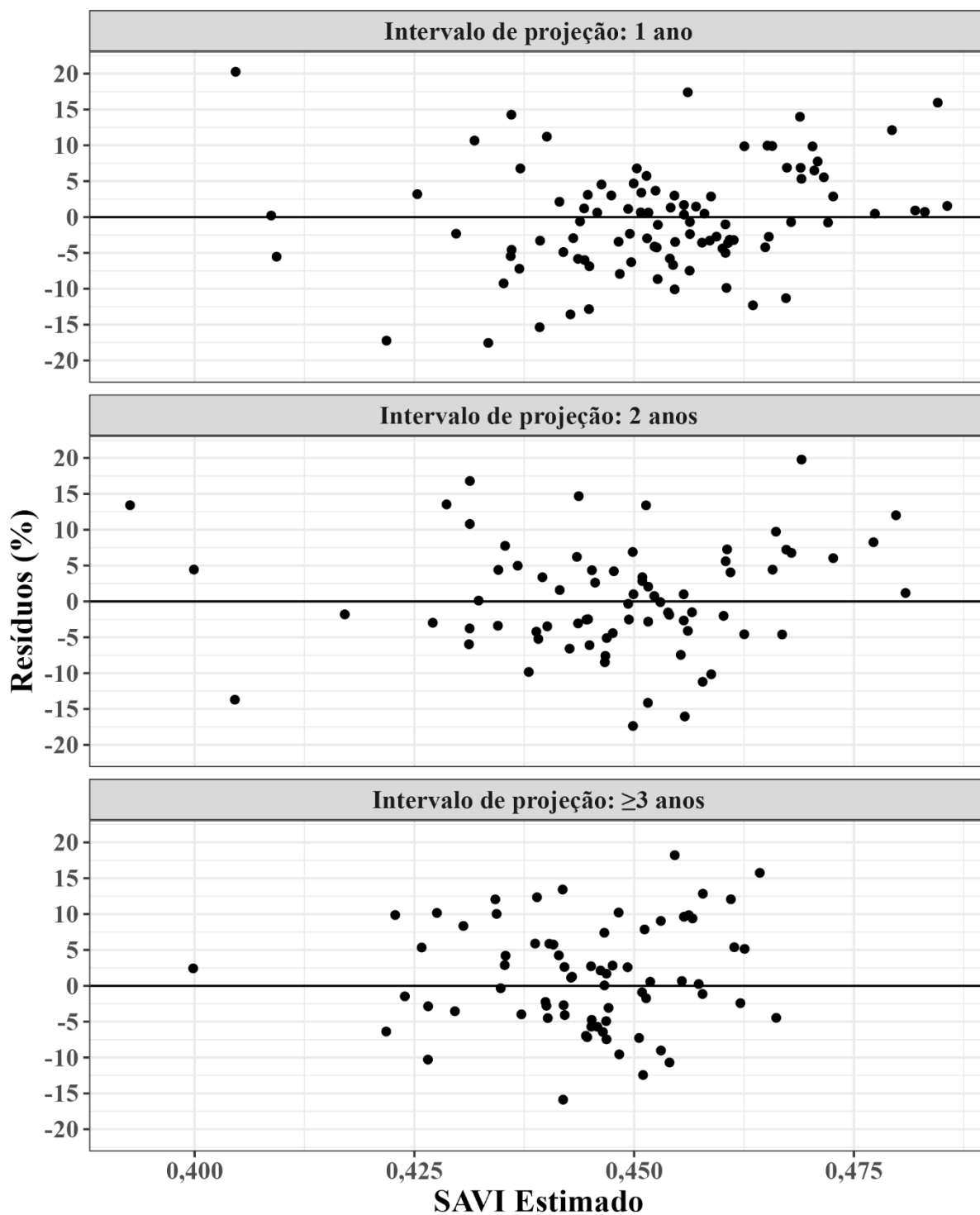
Tabela 6 – Parâmetros estimados do modelo de projeção do SAVI.

Variáveis	Parâmetro	Coefficiente ajustado
Assíntota	a_0	0,5370
$SAVI_1$	a_1	0,2745
$PVAR50_1$	a_2	-0,0025
$Idade_2$	a_3	-0,0047
$\frac{1}{Def.acum.1}$	a_4	32,5108

Fonte: Do autor (2025).

O modelo possui a capacidade de acertar, em média, os valores de SAVI na idade futura, com BIAS sempre tendendo a 0 (Figura 5).

Figura 4 – Distribuição de resíduos das estimativas de SAVI por intervalo de projeção.



Fonte: Do autor (2025).

3.3 Análise comparativa média dos modelos de crescimento da altura dominante

De forma geral, os modelos M_C_SR e M_C demonstraram maior capacidade explicativa do crescimento da altura dominante, conforme evidenciado pelo aumento do

coeficiente de determinação (R^2), redução do RMSE e BIAS (Tabela 7). Esse resultado indica que a inclusão de variáveis adicionais contribuiu para um melhor ajuste do modelo em termos médios de explicação da variação observada nos dados.

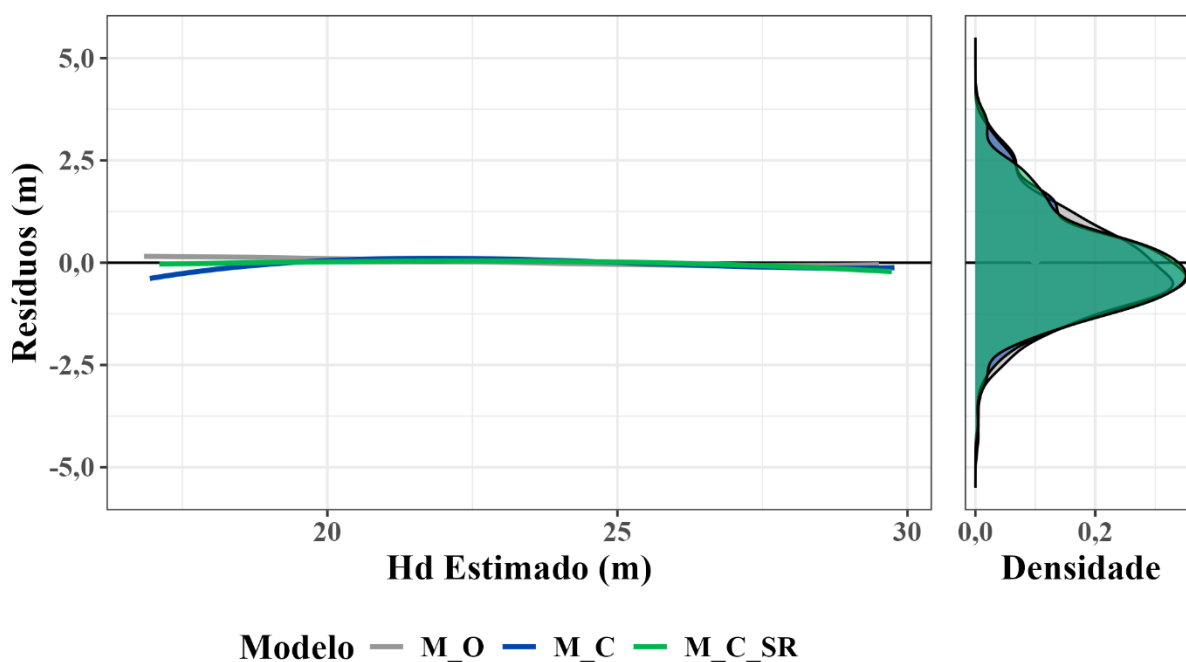
Tabela 7 – Estatísticas de precisão gerais dos modelos.

Modelo	R^2 (%)	MAE (m)	RMSE (m)	BIAS (m)
M_O	82,3	1,0	1,3	0,0
M_C	84,3	1,0	1,2	0,0
M_C_SR	84,6	0,9	1,2	0,0

Fonte: Do autor (2025).

Os modelos M_C e M_C_SR apresentaram uma menor dispersão dos resíduos, ou seja, uma maior concentração perto de 0, controlando as variações, de forma global (Figura 6).

Figura 5 – Distribuição geral dos resíduos da altura dominante projetada.



Fonte: Do autor (2025).

Entretanto, ao considerar apenas locais sujeitos a estresse hídrico, ou seja, com menor disponibilidade de água, os modelos M_C e M_C_SR demonstraram maior sensibilidade à queda no crescimento. Isso resultou em estimativas mais ajustadas e uma redução significativa da superestimação. O BIAS diminuiu de -0,5 metros para próximo de -0,1 metros, acompanhado por um aumento nos coeficientes de determinação e uma redução do RMSE, indicando um melhor desempenho preditivo (Tabela 8).

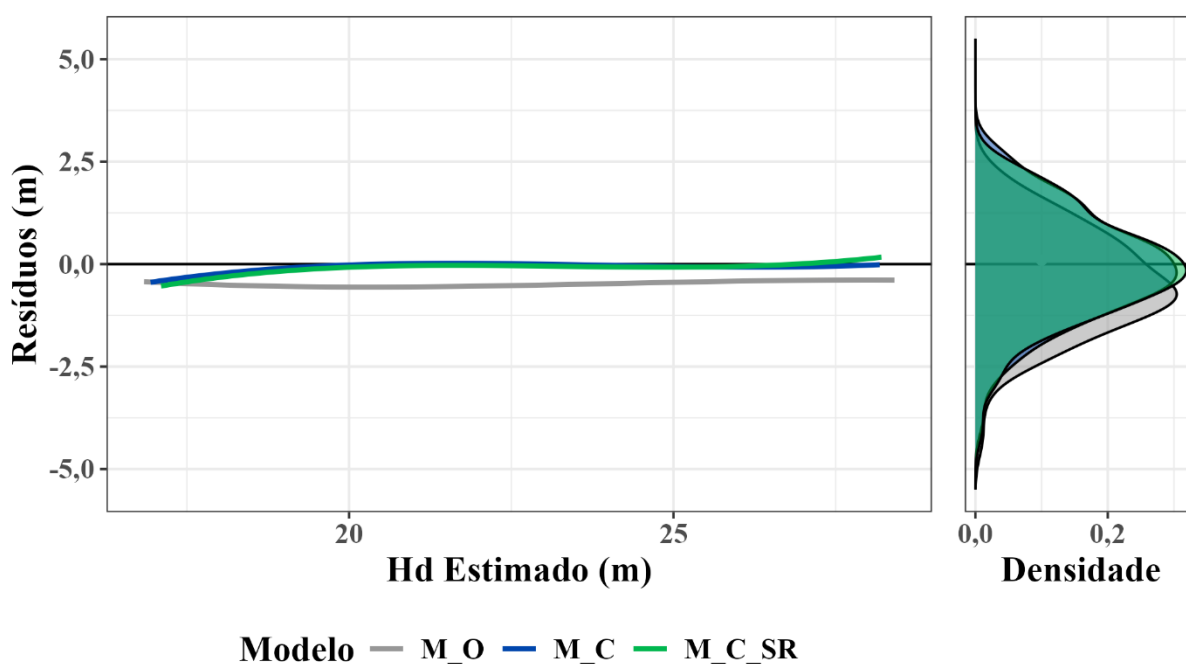
Tabela 8 – Estatísticas de precisão dos modelos considerando apenas locais que sofreram estresse hídrico severo.

Modelo	R ² (%)	MAE (m)	RMSE (m)	BIAS (m)
M_O	77,1	1,0	1,4	-0,5
M_C	78,7	1,0	1,2	-0,1
M_C_SR	79,6	0,9	1,2	-0,1

Fonte: Do autor (2025).

A distribuição de resíduos também é alterada em cenários de estresse hídrico (Figura 7), onde ocorre uma forte tendência de superestimativa da altura dominante do modelo M_O, tendência essa neutralizada com os modelos M_C e M_C_SR.

Figura 6 – Distribuição dos resíduos considerando apenas locais que sofreram estresse hídrico severo.



Fonte: Do autor (2025).

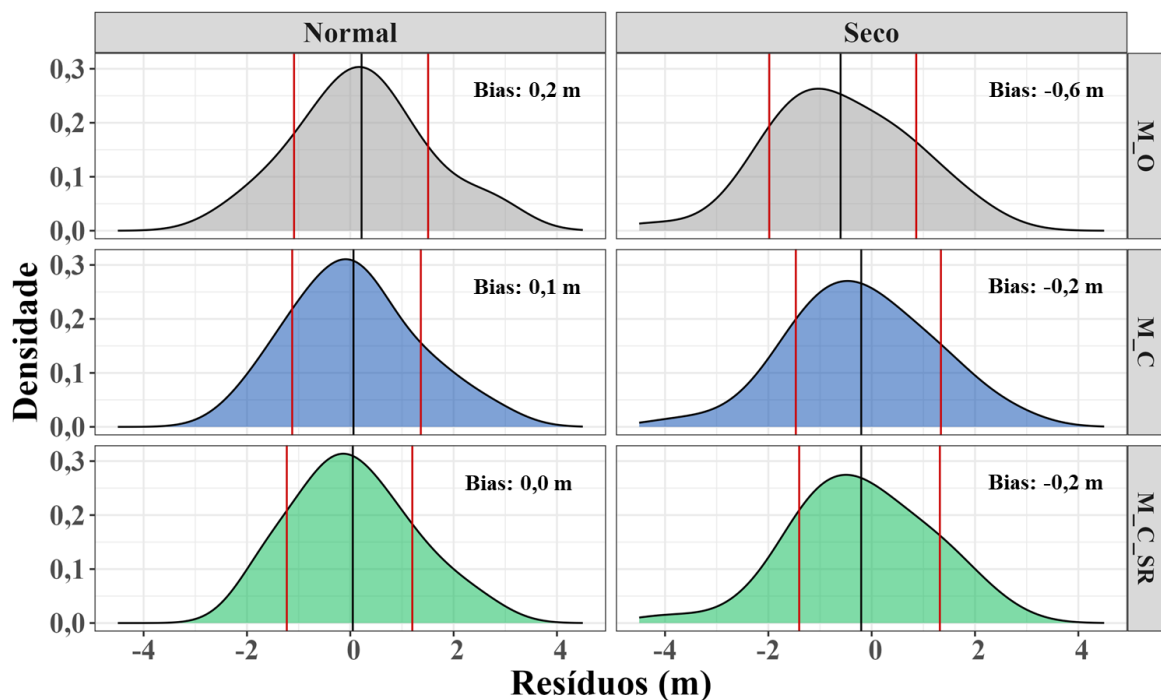
3.4 Análise comparativa do índice de sítio sob diferentes condições climáticas

Dado que o índice de sítio desempenha um papel central no planejamento operacional, as validações foram concentradas em estimativas para a idade de referência de 6 anos, permitindo uma avaliação mais clara do impacto do ajuste da altura dominante na determinação do potencial produtivo dos diferentes locais.

Para avaliar a estabilidade e a robustez dos modelos gerados, foi empregada a técnica de reamostragem Bootstrap com reposição, gerando 1000 novas amostras a partir da população inicial. É válido ressaltar que toda a validação do modelo M_C_R norteou-se no erro agregado via aplicação inicial do modelo de projeção do índice espectral e sua utilização na aplicação do modelo de altura dominante.

O modelo M_O apresentou, em média, estimativas satisfatórias da altura dominante. No entanto, o mesmo gera superestimativas para momentos em que o plantio enfrenta um período de escassez hídrica (precipitação média do ciclo atingindo até 80% da média global, que já estava aproximadamente 100 mm abaixo da normal histórica). Por outro lado, os modelos M_C e M_C_SR demonstraram capacidade generalista acerca das variações climáticas, ou seja, mantêm estimativas precisas em cenários hídricos destoantes da média histórica (Figura 8).

Figura 7 – Distribuição dos resíduos para os modelos avaliados a partir da reamostragem por Bootstrap, sob diferentes regimes hídricos.



| Percentis 15/85 | Média

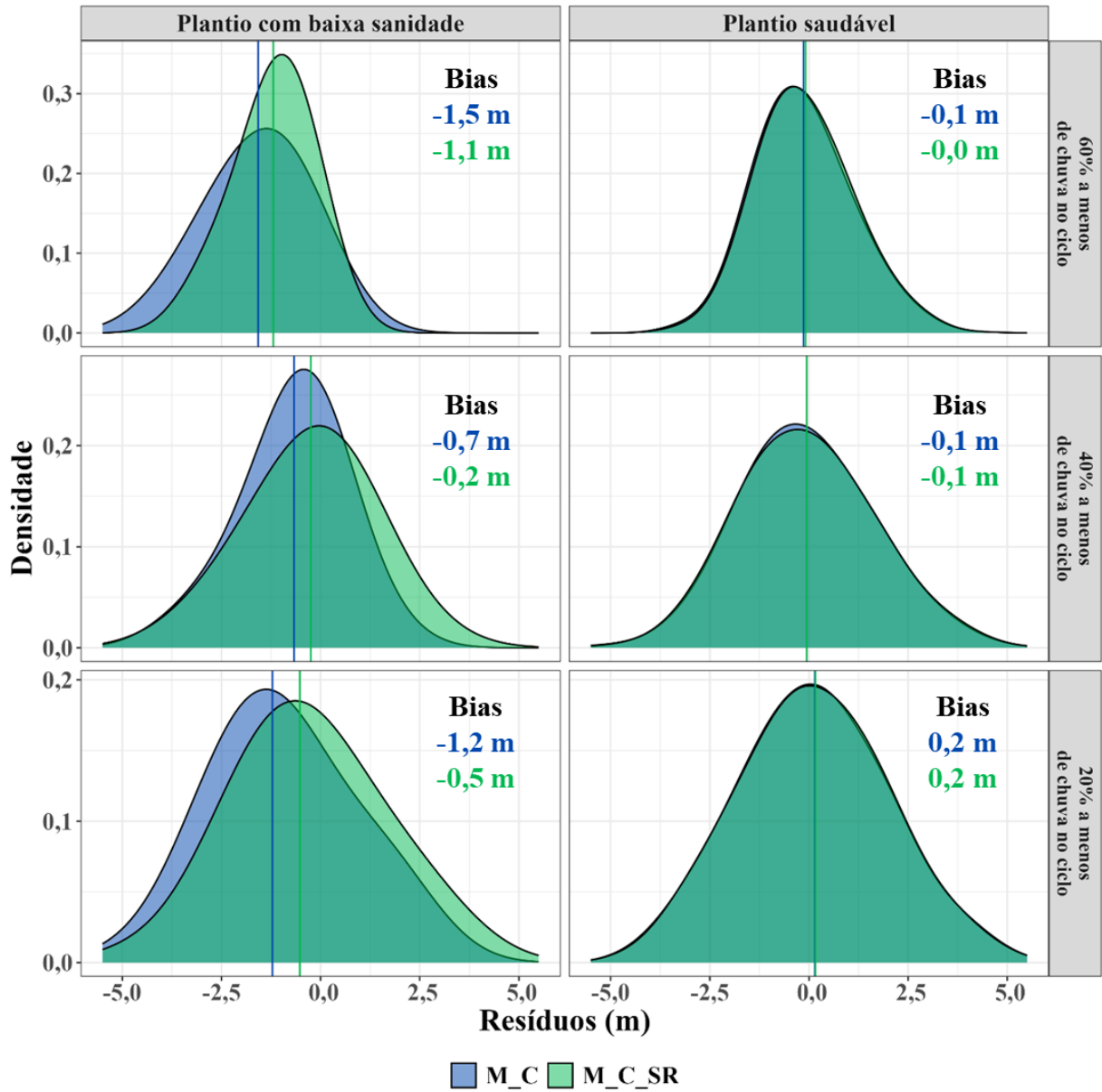
Fonte: Do autor (2025).

3.5 Análise comparativa do índice de sítio sob diferentes condições climáticas x sanidade florestal

O vigor de um plantio está diretamente relacionado à sua saúde, sendo a sanidade florestal um fator determinante para seu potencial produtivo. Neste estudo, a sanidade foi avaliada por meio do índice SAVI, que reflete a capacidade fotossintética das plantas e, consequentemente, sua eficiência no crescimento.

Ao comparar os modelos M_C e M_C_SR, observou-se que, embora a inclusão da variável climática melhore a tendência geral das estimativas de crescimento, ela introduz vieses na previsão do índice de sítio quando há variações na sanidade florestal, representadas pelo Índice de Área Foliar (IAF). Nesses casos, a incorporação da variável orbital se mostrou essencial para aprimorar as estimativas. O modelo M_C_SR apresentou um desempenho superior em áreas de baixa sanidade, associadas a valores reduzidos de SAVI, indicando que o uso de dados orbitais pode corrigir distorções e melhorar a precisão das projeções em condições adversas (Figura 9).

Figura 8 – Distribuições dos resíduos em diferentes condições de sanidade florestal e de disponibilidade hídrica.



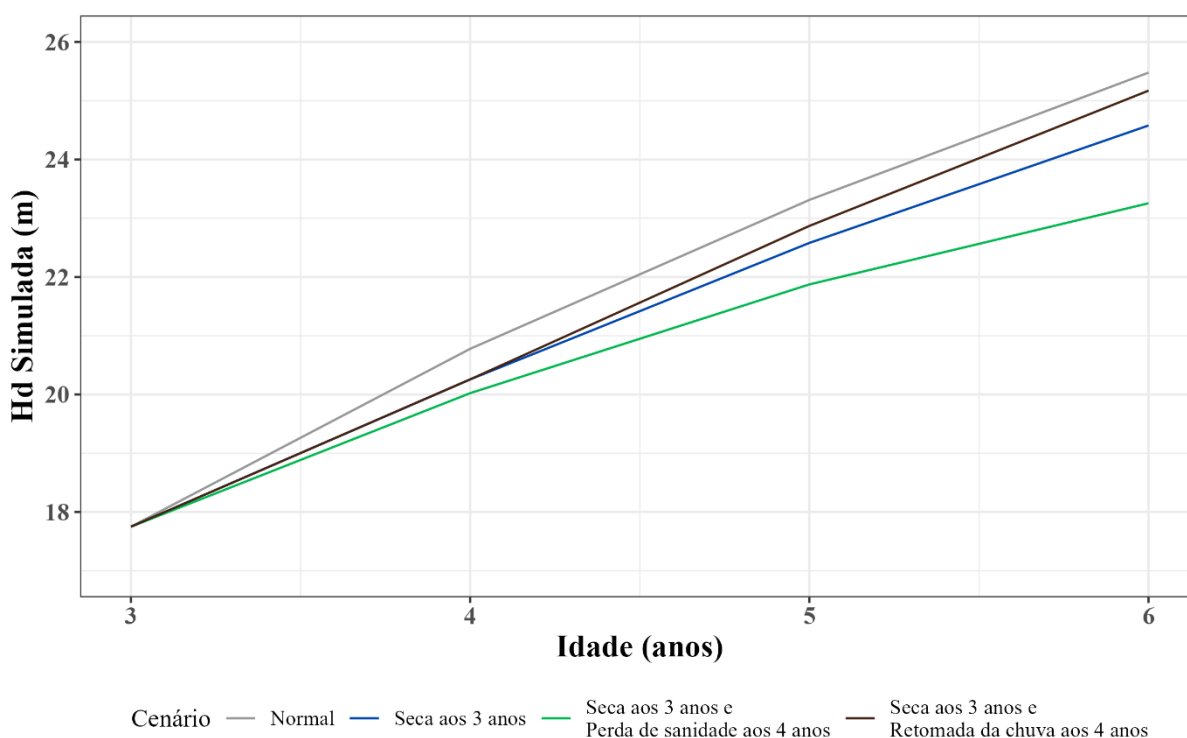
Fonte: Do autor (2025).

3.6 Cenários de variação do crescimento da altura dominante sob variadas condições de clima e sanidade florestal

Ao longo do ciclo, florestas podem ser afetadas por secas e/ou perdas de sanidade devido a diversos fatores ambientais. Para avaliar esse impacto no crescimento, foram simulados quatro cenários a partir de uma mesma altura dominante inicial, projetando sua evolução, a partir do modelo M_C_SR (considera variável climática e satelital) ao longo dos anos. No primeiro cenário (controle), considerou-se um regime hídrico normal, com chuvas dentro do esperado e

manutenção da sanidade florestal, resultando em um índice de sítio de 25,5 metros. No segundo cenário, o plantio foi submetido a uma seca severa que se iniciou aos três anos e manteve até o final do ciclo, reduzindo seu potencial produtivo e levando a uma queda no índice de sítio para 24,6 metros. No terceiro cenário, além da seca iniciando aos três anos, houve uma perda progressiva da sanidade florestal a partir do quarto ano, impactando ainda mais o crescimento e reduzindo o índice de sítio para 23,3 metros. Por fim, o quarto cenário o plantio sofreu a seca aos 3 anos, porém a partir dos 4 anos ocorreu a retomada de chuva dentro dos padrões esperados, refletindo a queda da taxa de crescimento inicial, porém retomada produtiva a partir dos 4 anos, alcançando índice de sítio de 25,2 metros (Figura 10).

Figura 9 – Simulação de cenários de deficiência hídrica e perda de sanidade florestal.



Fonte: Do autor (2025).

Essas simulações demonstram a influência combinada do estresse hídrico e da sanidade na produtividade florestal, ressaltando a necessidade de modelos capazes de captar e refletir os efeitos ambientais sobre o crescimento do plantio. A antecipação desses impactos permite prever variações produtivas com maior precisão, possibilitando tomadas de decisão mais rápidas e estratégicas para a gestão sustentável da floresta. O modelo M_C e M_C_SR são capazes de refletir o impacto das oscilações climáticas e da sanidade florestal de forma biologicamente coerente.

4 DISCUSSÃO

Os impactos das mudanças climáticas têm se tornado cada vez mais evidentes, caracterizados por eventos extremos, como secas prolongadas (Mokhov, 2022), que comprometem significativamente a produtividade dos povoamentos florestais. Diante desse cenário, é essencial que modelos de crescimento e produção capturem e representem adequadamente essas variações ambientais, garantindo estimativas mais precisas e confiáveis para o manejo florestal. Neste contexto, o presente estudo propôs o desenvolvimento de uma metodologia inovadora para aprimorar modelos empíricos de projeção da altura dominante, incorporando variáveis climáticas e de sensoriamento remoto. Essa abordagem visa não apenas capturar os efeitos globais das oscilações climáticas sobre o crescimento das árvores, mas também refinar as projeções em nível mais granular, permitindo estimativas mais precisas a nível do talhão. Os resultados demonstraram que os modelos expandidos M_C e M_C_SR foram eficazes na redução do viés das estimativas, especialmente em cenários de seca severa, garantindo maior aderência às dinâmicas observadas no campo. Além disso, as simulações indicaram que ambas as versões do modelo mantiveram coerência biológica, respondendo de forma adequada às variações na precipitação e às oscilações dos índices de vegetação, reforçando sua aplicabilidade para um planejamento florestal mais resiliente e adaptativo.

Dentre os diversos fatores climáticos que afetam o crescimento florestal, o estresse hídrico se destaca como um dos principais limitantes para a produtividade dos povoamentos. Stape *et al.* (2010) demonstraram que a restrição hídrica pode resultar em perdas significativas no incremento de madeira, com reduções de até 35% em regiões mais áridas. Esses impactos ressaltam a necessidade de aprimorar os modelos de crescimento, incorporando variáveis que representem de forma mais realista as flutuações climáticas e seus efeitos sobre o desenvolvimento florestal.

Além disso, durante a elaboração da metodologia, observou-se que a taxa de crescimento da altura dominante apresentava um efeito retardado em relação às variações das variáveis climáticas. Embora ainda existam poucos estudos que investiguem esse fenômeno, conhecido como *carry-over effect*, na produtividade florestal, Chen *et al.* (2023) evidenciam sua ocorrência, especialmente em períodos de seca. Isso indica que há um intervalo temporal entre a oscilação climática e a resposta efetiva do crescimento do povoamento. Esse comportamento reforça a necessidade de considerar esse atraso temporal na inclusão de

variáveis climáticas nos modelos empíricos, garantindo projeções mais precisas e uma representação mais fiel da relação entre o ambiente e o crescimento da floresta.

Corroborando fortemente as observações metodológicas e os achados de Chen *et al.* (2023), o *carry-over effect* foi capturado durante a construção do modelo expandido por meio de programação genética. Apesar da ampla gama de variáveis disponíveis para seleção, a abordagem computacional selecionou, de forma recorrente, variáveis climáticas que incorporavam esse atraso temporal, evidenciando sua relevância na melhoria da capacidade preditiva do modelo. Dessa forma, a metodologia proposta não apenas reduz os erros de estimativa, mas também aprimora a compreensão da relação entre o regime hídrico e o crescimento florestal, tornando as projeções mais realistas e aplicáveis a cenários operacionais.

Os resultados obtidos com a inclusão de variáveis climáticas com delay temporal e variáveis satelitais, baseadas no espectro analisado dos 12 meses anteriores à medição, confirmaram a relevância dessas variáveis na busca por estimativas mais precisas e acuradas, especialmente em cenários de déficit hídrico. Nesse contexto, a expansão do modelo base (M_O), inicialmente com a incorporação de variáveis climáticas (M_C) e, posteriormente, com a adição de indicadores satelitais (M_C_SR), resultou em uma melhora significativa na capacidade explicativa das projeções de altura dominante. O coeficiente de determinação (R^2) aumentou de 82,3% no modelo base para 84,3% no M_C e 84,6% no M_C_SR, além da redução do RMSE e controle do BIAS, evidenciando que a abordagem híbrida aprimora a acurácia das estimativas.

Além da melhoria no ajuste global dos modelos, a principal vantagem observada foi a redução do viés (bias) em cenários de restrição hídrica, foco principal do estudo. Enquanto o modelo M_O apresentava um viés sistemático de -0,5 metros na estimativa da altura dominante, ambos os modelos M_C e M_C_SR reduziram esse valor para -0,1 metros. Esses achados corroboram estudos prévios, como o de Scolforo *et al.* (2019), que expandiu o modelo de Chapman-Richards ao incluir a variável de Déficit Hídrico no Solo, reduzindo significativamente os erros das estimativas de altura dominante. Assim como descrito por esses autores, a utilização de modelos híbridos permite projeções mais confiáveis em diferentes horizontes temporais (curto, médio e longo prazo), garantindo uma melhor representação do comportamento biológico do povoamento.

A robustez dos modelos expandidos reforça a importância da inclusão de variáveis ambientais para melhorar a acurácia das projeções de crescimento florestal, especialmente em cenários de restrição hídrica. No presente estudo, a programação genética, dentre as variáveis climáticas disponíveis, selecionou a precipitação como variável explicativa em detrimento do

déficit hídrico no solo, utilizado por Scolforo *et al.* (2019). Essa escolha evidencia que diferentes abordagens podem capturar de maneira eficaz os impactos do balanço hídrico sobre o crescimento florestal, desde que considerem a dinâmica temporal e os efeitos acumulados das oscilações climáticas.

A relevância da precipitação como variável preditora já havia sido destacada por Chauke, Mwanbi e Kotze (2022), que, embora tenham adotado uma abordagem distinta – utilizando modelos não lineares de efeito misto –, demonstraram que a inclusão dessa variável resultou em um aumento de 37% na capacidade explicativa das projeções de altura dominante. Esse achado, alinhado aos resultados obtidos no presente estudo, sugere que a precipitação pode atuar como um indicador eficiente das condições hídricas do solo, refletindo de forma indireta a disponibilidade de água para o desenvolvimento do povoamento. Dessa forma, a escolha dessa variável pela programação genética reforça sua relevância para a modelagem do crescimento, corroborando tanto estudos anteriores quanto as evidências obtidas na presente análise.

A projeção da variável orbital introduz um erro acumulado na estimativa final. No presente estudo, utilizou-se um modelo de natureza linear para realizar as estimativas futuras do índice SAVI, embora essa abordagem não seja a mais adequada para a projeção desse tipo de dado. O modelo construído apresentou tendências de subestimativa em valores de SAVI superiores a 0,450, entretanto independente do horizonte de projeção, o BIAS do modelo ficou próximo de 0, indicando que o modelo é aderente e de modo geral não apresenta viés em suas estimativas. No entanto, seria mais apropriado utilizar um modelo autorregressivo de série temporal, a partir de correções das imagens e lapsos temporais que não há informações disponíveis, como proposto por Li *et al.* (2021), para captar de forma mais precisa a tendência futura da variável orbital, especialmente em projeções de longo prazo. Ainda assim, a abordagem adotada teve como principal objetivo demonstrar a importância de projetar essa variável para o futuro e apresentar uma técnica simples para essa finalidade, servindo como ponto de partida para estudos futuros que busquem aprimorar essa estimativa.

Os resultados deste estudo evidenciam a importância de incorporar variáveis climáticas em modelos de projeção de crescimento florestal, tornando as estimativas mais precisas e reduzindo o viés, especialmente em cenários de restrição hídrica. A abordagem utilizada demonstrou que considerar o atraso temporal entre os eventos climáticos e a resposta do povoamento melhora a capacidade preditiva do modelo, contribuindo para um planejamento florestal mais eficiente. Além disso, a redução de erros nas estimativas de altura dominante tem impactos diretos na previsão de volume e no manejo da produção, minimizando riscos e otimizando recursos. O aprimoramento contínuo dessas metodologias, aliado à integração de

novas técnicas e dados ambientais, como variáveis moduladoras do crescimento – aplicadas em modelos como o 3-PG e o PULSE – ou avanços na projeção de variáveis satelitais por meio da inserção de modelos de séries temporais para estimativas futuras de índices de vegetação, pode fortalecer ainda mais a capacidade de prever a produtividade florestal sob diferentes condições climáticas, auxiliando na tomada de decisão e na adaptação a cenários futuros.

5 CONCLUSÃO

As oscilações climáticas, cada vez mais frequentes, tornam essencial o uso de novas abordagens para a estimativa do crescimento florestal, permitindo minimizar perdas produtivas e viabilizar um planejamento estratégico baseado em informações sempre atualizadas.

A aplicação da Programação Genética (PG) na seleção de variáveis e construção de modelos híbridos apresenta um grande potencial para aprimorar as estimativas de crescimento e produtividade florestal. Neste estudo, a incorporação de variáveis climáticas e orbitais demonstrou ser uma alternativa viável para aumentar a precisão das projeções da altura dominante, possibilitando uma modelagem mais adaptável a diferentes cenários ambientais. Além de melhorar as estimativas em nível global, a abordagem proposta permitiu um refinamento significativo em escala talhão a talhão, fator essencial para a atualização dinâmica do inventário florestal.

A expansão da assíntota do modelo original de Chapman-Richards, por meio da inclusão de variáveis climáticas, possibilitou uma melhor captura dos impactos de eventos climáticos extremos, como períodos de déficit hídrico, um fator crítico para a produtividade florestal. Além disso, a integração de variáveis de sensoriamento remoto, como o SAVI, contribuiu para uma correção mais específica das projeções da altura dominante em situações em que a sanidade do plantio está debilitada, além de tornar o modelo mais sensível às variações espaciais da disponibilidade de água.

Diante da carência de estudos que combinam modelagem empírica com abordagens híbridas baseadas em técnicas avançadas de aprendizado de máquina, faz-se necessário o aprofundamento dessas metodologias. A incorporação de variáveis ambientais e orbitais, aliada a métodos de otimização como a Programação Genética, demonstra ser um caminho promissor para aumentar a capacidade preditiva dos modelos florestais e sua aplicação no manejo estratégico de plantações. Essa evolução metodológica não apenas aprimora a acurácia das

estimativas, mas também viabiliza uma tomada de decisão mais eficiente, alinhando tecnologia e sustentabilidade na gestão dos recursos florestais.

REFERÊNCIAS

- ALONSO, L. *et al.* Automatic differentiation of Eucalyptus species through Sentinel-2 images, Worldview-3 images and LiDAR data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 207, p. 264-281, 2024.
- ALVARES, C. A. *et al.* Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**. Berlin, v. 22, n. 6, p. 711–728, 2013.
- ASSMANN, E. **The principles of forest yield study**. 1970.
- ALMEIDA, A. C. *et al.* Needs and opportunities for using a process-based productivity model as a practical tool in Eucalyptus plantations. **Forest Ecology Management**. 193 (1), 167–177, 2004.
- BAILEY, R. L.; CLUTTER, J. L. Base-age invariant polymorphic site curves. **Forest Science**. 20 (2), 155–159, 1974.
- CASNATI, A. C. R. **Hybrid mensurational-physiological models for Pinus taeda and Eucalyptus grandis in Uruguay**. 2016. 208 p. PhD Dissertation. University of Canterbury, Canterbury, New Zealand, 2016.
- CHAUKE, M.; MWAMBI, H.; KOTZE, H. Modelling dominant height growth including a rainfall effect using the algebraic difference approach. **Cerne**, v. 28, 2022.
- CHEN, Liang *et al.* Immediate and carry-over effects of late-spring frost and growing season drought on forest gross primary productivity capacity in the Northern Hemisphere. **Global Change Biology**, v. 29, n. 14, p. 3924-3940, 2023.
- CUNHA, G. T. da. **Uma nova abordagem para inventário de biomassa em plantações de eucalipto no Brasil**. 2023. 67 p. Dissertação (Mestrado em Engenharia Florestal)– Universidade Federal de Lavras, Lavras, 2023.
- DENG, X. *et al.* Identification of short-rotation eucalyptus plantation at large scale using multi-satellite imageries and cloud computing platform. **Remote Sensing**, v. 12, n. 13, p. 2153, 2020.
- EMBRAPA, *et al.* Sistema brasileiro de classificação de solos. **Centro Nacional de Pesquisa de Solos - Empresa Brasileira de Pesquisa Agropecuária**, Rio de Janeiro, v. 3, 2013.
- FERRAZ FILHO, A. C. *et al.* Dominant height projection model with the addition of environmental variables. **Cerne** 17 (3), 427-433, 2011.
- GOOGLE. **Google Earth Engine**. 2025. Disponível em: <https://earthengine.google.com/>. Acesso em: 19 fev. 2025.
- IBÁ - Indústria Brasileira de Árvores. **Relatório Anual Ibá 2024**, ano base 2023. Disponível em: <<https://www.iba.org/datafiles/publicacoes/relatorios/relatorio-anual-iba2024-compactado.pdf>> . (Acesso em 07 de janeiro de 2025).

ILES, K. A sampler of inventory topics. **Kim Iles & Associates Ltd.**, Nanimo, BC, 2003.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Produto de precipitação MERGE**. São José dos Campos: INPE, 2020. Disponível em: <<http://ftp.cptec.inpe.br/modelos/tempo/MERGE/>>. (Acesso em 07 de janeiro de 2025).

KINANE, S. M. *et al.* Influence of environmental variables on leaf area index in loblolly pine plantations. **Forest Ecology and Management**, v. 523, p. 120445, 2022.

KOIRALA, A.; MONTES, C. R.; BULLOCK, B. P. Modeling dominant height using stand and water balance variables for loblolly pine in the Western Gulf, US. **Forest Ecology and Management**, v. 479, p. 118610, 2021.

KOZA, J. R. **Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems**. Stanford University, Department of Computer Science, Stanford, 1990.

LI, S. *et al.* High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. **International Journal of Applied Earth Observation and Geoinformation**, v. 105, p. 102640, 2021.

MAESTRI, R. **Modelo de crescimento e produção para povoamentos clonais de Eucalyptus grandis considerando variáveis ambientais**. 2003. 143 p. Tese (Doutorado em Engenharia Florestal). Universidade Federal do Paraná, Curitiba, Brazil, 2003.

MAIRE, G. *et al.* MODIS NDVI time-series allow the monitoring of Eucalyptus plantation biomass. **Remote sensing of Environment**, v. 115, n. 10, p. 2613-2625, 2011.

MIRANDA, E. N. *et al.* Variable selection for estimating individual tree height using genetic algorithm and random forest. **Forest Ecology and Management**, Amsterdam, v. 504, p. 119828, 2022.

MOKHOV, I. I. Changes in the frequency of phase transitions of different types of El Niño phenomena in recent decades. *Izvestiya*, **Atmospheric and Oceanic Physics**, v. 58, n. 1, p. 1-6, 2022.

NOORIAN, F.; DE SILVA, A. M.; LEONG, P. H. W. gramEvol: Grammatical evolution in R. **Journal of Statistical Software**, v. 71, p. 1-26, 2016.

R CORE TEAM. R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, Vienna. Disponível em: <<http://www.R-project.org>> (Acesso em: 07 de janeiro de 2025), 2024.

REIS, A. A. *et al.* Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. **New Zealand Journal of Forestry Science**, v. 48, n. 1, p. 1-17, 2018.

REIS, A. A. *et al.* Classification of Eucalyptus plantation Site Index (SI) and Mean Annual Increment (MAI) prediction using DEM-based geomorphometric and climatic variables in Brazil. **Geocarto International**, v. 37, n. 5, p. 1256-1273, 2022.

SCOLFORO, H. F. *et al.* Modeling dominant height growth of eucalyptus plantations with parameters conditioned to climatic variations. **Forest Ecology and Management** 380, 182–195, 2016.

SCOLFORO, H. F. *et al.* Site index estimation for clonal eucalypt plantations in Brazil: A modeling approach refined by environmental variables. **Forest Ecology and Management** 366, 1–12, 2019.

SILVA, A. K. V. *et al.* Predicting eucalyptus diameter at breast height and total height with UAV-based spectral indices and machine learning. **Forests**, v. 12, n. 5, p. 582, 2021.

STAPE, J. L., BINKLEY, D., RYAN, M. G. Eucalyptus production and the supply, use and efficiency of use of water, light and nitrogen across a geographic gradient in Brazil. **Forest Ecology and Management**. 193 (1), 17–31, 2004.

U.S. GEOLOGICAL SURVEY. **Landsat 7 Data Users Handbook**. Version 3.0. Sioux Falls, SD: USGS, 2024. Disponível em: <<https://www.usgs.gov/landsat-missions/product-information>>. (Acesso em 15 de janeiro de 2025).

VILLACRÉS, J. *et al.* Retrieval of vegetation indices related to leaf water content from a single index: A case study of *Eucalyptus globulus* (Labill.) and *Pinus radiata* (D. Don.). **Plants**, v. 10, n. 4, p. 697, 2021.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag, New York, 2016.

WICKHAM H. *et al.* **dplyr: A Grammar of Data Manipulation**. Disponível em: <<https://CRAN.R-project.org/package=dplyr>>. (Acesso em: 07 de janeiro de 2025), 2023.

WICKHAM H.; HENRY L. **purrr: Functional Programming Tools**. Disponível em: <<https://CRAN.R-project.org/package=purrr>>. (Acesso em: 07 de janeiro de 2025), 2023.

WICKHAM H.; VAUGHAN D.; GIRLICH M. **tidyr: Tidy Messy Data**. Disponível em: <<https://CRAN.R-project.org/package=tidyr>>. (Acesso em: 07 de janeiro de 2025), 2024.

WONGCHAI, W. *et al.* **Above-ground biomass estimation of Eucalyptus plantation using remotely sensed data and field measurements**. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012042 , 2020.

STAPE, J. L. *et al.* The Brazil *Eucalyptus* Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production. **Forest Ecology and Management**, Amsterdam, v. 259, n. 9, p. 1684-1694, 2010.

TERCEIRA PARTE

CONSIDERAÇÕES FINAIS

Este estudo contribui significativamente para o avanço da modelagem do crescimento florestal ao propor a incorporação de variáveis climáticas e de sensoriamento remoto nos modelos convencionais de projeção da altura dominante. A metodologia desenvolvida, baseada na decomposição de parâmetros e na aplicação da Programação Genética para seleção e combinação de variáveis, demonstrou-se eficaz na melhoria da precisão das estimativas, especialmente em cenários de restrição hídrica. Os resultados reforçam a relevância da precipitação e da sanidade florestal, captada por índices espectrais, na previsão do crescimento das florestas, destacando a importância de abordagens híbridas para refinar projeções em diferentes escalas espaciais e temporais.

Além da inovação metodológica, este trabalho fortalece o conhecimento acadêmico ao evidenciar a existência do *carry-over effect* na modelagem da altura dominante e seu impacto na dinâmica de crescimento dos povoamentos florestais. A recorrente seleção de variáveis climáticas com atraso temporal pela Programação Genética ressalta a necessidade de considerar a inércia do crescimento frente a oscilações ambientais, garantindo projeções mais coerentes e biologicamente realistas. Esses achados ampliam a compreensão da interação entre clima e produtividade florestal, contribuindo para o aprimoramento de modelos empíricos e orientando futuras pesquisas na área de crescimento e produção.

Por fim, este estudo destaca a aplicabilidade das técnicas propostas para o manejo florestal, proporcionando um ferramental mais preciso para o monitoramento e a projeção da produtividade em diferentes cenários operacionais. A integração de dados climáticos e satelitais nos modelos de crescimento oferece uma alternativa viável para reduzir incertezas e otimizar a tomada de decisão, especialmente em um contexto de mudanças climáticas. A continuidade desse trabalho, com a incorporação de modelos autorregressivos para projeção de variáveis orbitais e a validação da abordagem em diferentes regiões e espécies, poderá expandir ainda mais a aplicabilidade dos métodos desenvolvidos, consolidando um novo paradigma na modelagem do crescimento florestal.