



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Botucatu



JOÃO MARCOS REBESSI DE SOUSA

**MODELAGEM DA BIOMASSA DE PLANTAÇÕES FLORESTAIS A PARTIR
DA TECNOLOGIA LIDAR EMBARCADA EM DRONE**

Botucatu

2023

JOÃO MARCOS REBESSI DE SOUSA

**MODELAGEM DA BIOMASSA DE PLANTAÇÕES FLORESTAIS A PARTIR DA
TECNOLOGIA LIDAR EMBARCADA EM DRONE**

Dissertação apresentada à Faculdade de
Ciências Agronômicas da Unesp Câmpus
de Botucatu, para obtenção do título de
Mestre em Engenharia Agrícola.

Orientador: Prof. Dr. Saulo Philipe Sebastião Guerra
Coorientador: Prof. Dr. Humberto de Jesus Eufrade-
Junior

Botucatu
2023

S725m

Sousa, João Marcos Rebessi de

Modelagem da biomassa de plantações florestais a partir da tecnologia LiDAR embarcada em drone / João Marcos Rebessi de Sousa. -- Botucatu, 2024

57 p. : il., tabs., fotos, mapas

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrônômicas, Botucatu

Orientador: Prof. Dr. Saulo Philipe Sebastião Guerra

Coorientador: Prof. Dr. Humberto de Jesus Eufrade-Junior

1. Sensorimaneto Remoto. 2. LiDAR. 3. Mensuração Florestal. 4. Biomassa. 5. Machine Learning. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrônômicas, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

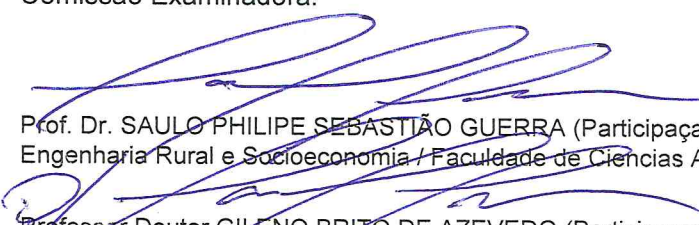
TÍTULO DA DISSERTAÇÃO: MODELAGEM DA BIOMASSA DE PLANTAÇÕES FLORESTAIS A PARTIR DA TECNOLOGIA LIDAR EMBARCADA EM DRONE

AUTOR: JOÃO MARCOS REBESSI DE SOUSA

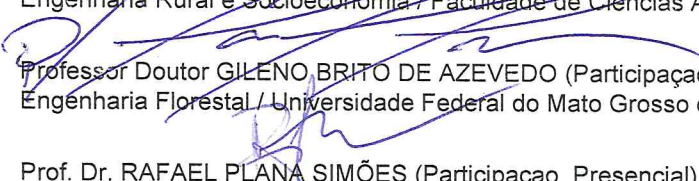
ORIENTADOR: SAULO PHILIPPE SEBASTIÃO GUERRA

COORIENTADOR: HUMBERTO DE JESUS EUFRADE JUNIOR

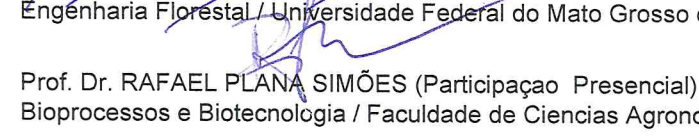
Aprovado como parte das exigências para obtenção do Título de Mestre em Engenharia Agrícola, pela Comissão Examinadora:



Prof. Dr. SAULO PHILIPPE SEBASTIÃO GUERRA (Participação Presencial)
Engenharia Rural e Socioeconomia / Faculdade de Ciências Agrômicas de Botucatu UNESP



Professor Doutor GILENO BRITO DE AZEVEDO (Participação Virtual)
Engenharia Florestal / Universidade Federal do Mato Grosso do Sul



Prof. Dr. RAFAEL PLANA SIMÕES (Participação Presencial)
Bioprocessos e Biotecnologia / Faculdade de Ciências Agrômicas de Botucatu - UNESP

Botucatu, 11 de dezembro de 2023

AGRADECIMENTOS

A Deus por me dar sabedoria em todas as etapas.

Agradeço aos meus familiares por todo apoio em minha formação e trajetória em busca de conhecimentos e formação profissional.

Agradeço a Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) e a Faculdade de Ciência Agronômicas (FCA) campus de Botucatu local onde me formei e vivenciei momentos de muito aprendizado.

Agradeço aos professores Dr. Saulo Philipe Sebastião Guerra e Dr. Humberto de Jesus Eufrade-Junior por todos ensinamentos e pela oportunidade.

Meus sinceros agradecimentos a toda equipe do Laboratório Agroflorestal de Biomassa e Bioenergia (LABB), os quais foram fundamentais em diversas etapas no decorrer deste trabalho.

E por fim agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil – CAPES, o presente trabalho foi realizado com apoio da mesma, Código de financiamento 3300406-4.

RESUMO

A quantificação da biomassa de plantações florestais com o uso de dados LiDAR (*Light Detection And Ranging*) embarcado em drones (ou *unmanned aerial vehicle* – UAV-LS) tem apresentado grande potencial, principalmente pelo elevado nível de detalhamento e agilidade em relação as metodologias tradicionais de inventário florestal. Este trabalho teve como objetivo principal aplicar técnicas de aprendizado de máquinas para prever a biomassa acima do solo com base em dados LiDAR. Para isso, foram obtidos dados UAV-LS e o inventário censitário de uma plantação de um híbrido de *E. torelliana* × *C. citriodora* de 7 anos de 1,25 ha localizado na Fazenda Experimental Lageado da Faculdade de Ciências Agrônômicas (FCA/UNESP) em Botucatu. Após amostragem e seleção das árvores-amostra, houve a determinação da densidade básica da madeira ponderada pelo volume, e determinação indireta da biomassa seca de fuste e ajuste do modelo alométrico de Schumacher-Hall. A partir do georreferenciamento dos indivíduos arbóreos e tratamento dos dados da nuvem de pontos LiDAR no programa computacional R, houve a extração das métricas de altura a nível de árvore individual, que ainda foram selecionadas pela análise de componentes principais (PCA). Por último, os modelos de aprendizado de máquina foram testados em ambiente Python, ranqueados e comparados pelo coeficiente de determinação ajustado e erro quadrático médio. A melhor janela para individualização das árvores foi a de 2 m, com um erro relativo de 11,9 %. Os modelos de aprendizado de máquina baseados na lógica de “Árvores de decisão” foram os que apresentaram melhor ajuste ($R^2_{aj} > 0,85$ e $RMSE < 15\%$) para a população de treinamento, sendo possível obter modelos com elevada performance de ajuste a nível de árvore individual para as condições presentes nesse estudo.

Palavras-chave: mensuração florestal; sensoriamento remoto; aprendizado de máquinas; nuvem de pontos.

ABSTRACT

Quantifying the biomass of forestry plantations using LiDAR (Light Detection And Ranging) data embedded in drones (or unmanned aerial vehicles – UAV-LS) has shown great potential, mainly due to the high level of detail and agility in relation to methodologies traditional forest inventory. The main objective of this work was to apply machine learning techniques to predict above-ground biomass based on LiDAR data. For this, UAV-LS data and the census inventory were obtained from a 7-year-old plantation of a 7-year-old *E. torelliana* × *C. citriodora* hybrid of 1.25 ha located at the Lageado Experimental Farm of the Faculty of Agricultural Sciences (FCA/UNESP) in Botucatu. After sampling and selection of sample trees, the basic density of wood weighted by volume was determined, and the dry biomass of the bole was indirectly determined and the Schumacher-Hall allometric model was adjusted. From the georeferencing of tree individuals and processing of LiDAR point cloud data in the computational program R, height measurements at the individual tree level were removed, which were still selected by principal component analysis (PCA). Finally, the machine learning models were tested in the Python environment, classified and compared by the adjusted coefficient of determination and mean squared error. The best window for individualizing trees was 2 m, with a relative error of 11.9%. The machine learning models based on “Decision Trees” logic were those that obtained the best fit ($R^2_{aj} > 0.85$ and $RMSE < 15\%$) for the training population, making it possible to obtain models with high adjustment performance at the level of individual tree for the conditions present in this study.

Keywords: forest measurement; remote sensing; machine learning; point cloud.

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO BIBLIOGRÁFICA	15
2.1	Biomassa de plantações florestais	15
2.2	Tecnologia LiDAR no setor florestal	17
2.3	Modelagem de dados UAV-LS com uso de técnicas de aprendizado de máquinas.....	22
2.4	A biblioteca Lazy predict em Python.....	24
3	MATERIAL E MÉTODOS	26
3.1	Área de estudo	26
3.2	Coleta de dados no campo.....	26
3.3	UAV-LS e processamento da nuvem de pontos LiDAR	31
3.4	Espaço amostral e Georreferenciamento das árvores em campo.....	34
3.5	Modelagem da biomassa seca com casca de fuste	37
3.6	Estatística para verificação do ajuste dos modelos e variáveis.....	37
4	RESULTADOS E DISCUSSÃO	38
4.1	Produtos obtidos pelo processamento da nuvem de pontos.....	39
4.2	Modelagem preditiva por aprendizado de máquinas.....	46
5	CONCLUSÕES	49
	REFERÊNCIAS.....	50

1 INTRODUÇÃO

As florestas plantadas ocupam cerca de 131 milhões de hectares no mundo, sendo que o continente com a maior área plantada é a América do Sul, com destaque para o cultivo de espécies exóticas (FAO, 2020). Estabelecidas para fins produtivos, são caracterizadas por um manejo intensivo, compostas de uma ou duas espécies com a mesma idade, plantadas com um espaçamento regular. No Brasil, são 9,93 milhões de hectares plantados, sendo que 75,8% é representado pelo cultivo de eucalipto e 19,4% de pinus (Ibá, 2022). As outras espécies, possuem uma área de aproximadamente 475 mil hectares. Estima-se que no ano de 2020, este setor promoveu mais de 2 milhões de empregos de forma direta e indireta no país, bem como foram gerados R\$ 18,8 bilhões em produtos de origem madeireira para atender os segmentos industriais de celulose e papel, painéis de madeira, lenha e carvão vegetal (IBGE, 2021).

No âmbito de plantações florestais, são diversos fatores que afetam a produtividade florestal, dentre eles estão as alterações climáticas, principalmente no que se refere à elevação da temperatura e a redução dos regimes pluviométricos. Neste cenário, estimula-se a obtenção de informações precisas e ágeis, que vinculem o uso de novas tecnologias e técnicas, com destaque para o uso do sensoriamento remoto e da inteligência artificial na aquisição e no processamento dos dados (Zou *et al.*, 2019; Elli *et al.*, 2020).

O conceito de Florestas 4.0 emergiu na era da digitalização e automação dos dados para a administração dos recursos florestais norteando-se pela quarta revolução industrial, algo que corrobora para o uso de técnicas que conciliem velocidade e acurácia na captação de informações (Müller *et al.*, 2019; Sahal *et al.*, 2021). Em contrapartida, os processos de mensuração tradicional são lentos na aquisição de dados se comparado com as novas tecnologias que estão surgindo no setor florestal, devido ao caminhar e medição manual de dados nos plantios florestais, características que tornam o processo demorado, muito oneroso e limitado as condições da área florestal (Soares *et al.*, 2012; Batista *et al.*, 2014; Maltamo *et al.*, 2018).

Contudo, o levantamento das características quantitativas de árvores e povoamentos são imprescindíveis para o planejamento e manejo florestal (Burkhar; Tomé, 2012). Neste contexto, o LiDAR (Light Detection And Ranging) destaca-se como uma

promissora tecnologia, que vem sendo utilizada na quantificação de recursos florestais (Baldin, 2018). Diversos autores têm reportado a adoção da tecnologia de sensoriamento remoto ativo de varredura a laser para a quantificação dos recursos florestais (Corte *et al.*, 2020a; Rudge *et al.*, 2021; Xu *et al.*, 2021, Melo *et al.*, 2021).

Quando embarcado em veículo aéreo não tripulado (Unmanned Aerial Vehicle Laser Scanning – UAV-LS), possibilita a aquisição de uma nuvem de pontos tridimensionais (3D) de alta densidade (1.000 a 20.000 pontos m^{-2}) e georreferenciadas, as varreduras do laser são mais amplas do que plataformas aéreas tradicionais, isso gera medições mais precisas (Kellner *et al.*, 2019). O sensor tem apresentado sucesso na medição da estrutura de florestas e na predição da biomassa acima do solo, algo que impulsiona diversas pesquisas na área de modelagem da biomassa dos plantios (Silva *et al.*, 2017; Gleason *et al.*, 2012; Neuville *et al.*, 2021; Pourshamsi *et al.*, 2021). A medida que penetram o dossel até o solo, permitem o registro dos componentes da estrutura da floresta, incluindo copas, troncos, galhos e vegetação do sub-bosque (Beland *et al.*, 2019; Puliti *et al.*, 2020; Mohan *et al.*, 2021; Yu *et al.*, 2021).

Uma abordagem utilizada para a predição da biomassa acima do solo é a elaboração de modelos de regressão complexos, com uso das métricas LiDAR para as estimativas de biomassa alcançadas via inventário florestal tradicional (Schuh, 2019; Leitold *et al.*, 2015). Neste cenário de inovações, o uso da inteligência artificial se tornou uma realidade no setor de mensuração florestal, principalmente as técnicas de aprendizado de máquinas, as quais possibilitam a criação de modelos mais robustos, que por sua vez, conseguem modelar grande quantidade de dados de forma automatizada, com grande capacidade de adaptabilidade, ao lidar melhor com dados complexos do que os modelos tradicionais de regressão (Montaño, 2016).

Sendo assim, se faz necessário mais estudos que impulsionem o uso do LiDAR para a mensuração florestal, de forma que, futuramente a tecnologia possa substituir ou complementar os processos de inventário tradicional. Assim, este estudo teve como objetivo principal aplicar técnicas de aprendizado de máquinas para prever a biomassa acima do solo com uso de dados obtidos por tecnologia LiDAR embarcado em drone.

2 REVISÃO BIBLIOGRÁFICA

2.1 Biomassa de plantações florestais

As plantações florestais são cultivadas principalmente para propósitos produtivos associados ao uso da madeira na produção de papel e celulose, carvão vegetal, painéis de madeira e construção civil. Para isso são intensamente manejadas, e podem ser compostas por uma ou duas espécies da mesma idade, e plantadas em um espaçamento regular (Oliveira, 2021; FAO, 2020)

A quantificação da produção da biomassa nesses plantios é importante tanto no planejamento e coordenação das atividades florestais, bem como nos estudos do balanço global de carbono e questões ligadas às mudanças climáticas discutidas recentemente (Nakai, 2016). Entende-se biomassa como a massa orgânica produzida por unidade de área, tanto acima ou abaixo do solo, viva ou morta, podendo ser expressa em peso de matéria seca ou úmida, ou então em peso de carbono (FAO, 2004, Odum, 1986).

A biomassa florestal está diretamente correlacionada com a estrutura florestal que é compreendida pelas variáveis quantificáveis em um povoamento. Dentre elas estão o diâmetro, altura total e volume das árvores, que são características que podem ser afetadas pelo arranjo espacial do povoamento florestal, espécie, capacidade de sítio e condições climáticas. Com esse viés, mensurar essas variáveis nos remete maior compreensão do potencial produtivo da biomassa em plantios florestais (Batista *et al.*, 2014).

Diante dos diversos fatores que afetam a produtividade florestal, a diversidade de espécies e tamanhos das árvores, a estrutura espacial e a competição entre os indivíduos na floresta plantada ganham maior importância ao longo do desenvolvimento do povoamento (Zhang *et al.*, 2019; Duncan, 1991). As árvores estão dispostas em um ambiente de elevada competitividade, ao passo que quando as copas se fecham, as árvores começam a competir significativamente entre si por fatores abióticos, e pode fazer com que alguns indivíduos se atrasem no crescimento em relação aos outros, suas copas perdem vigor, resultando eventualmente em morte. Como forma de monitorar a dinâmica e a estrutura das florestas plantadas, são

levantadas características quantitativas das árvores isoladas que posteriormente são estendidas para todo o povoamento (Burkhar; Tomé, 2012; Souza; Soares, 2013).

A variável em uma floresta de mais fácil obtenção é o diâmetro das árvores e com ela é possível gerar a distribuição de diâmetros, que por sua vez é uma ferramenta que caracteriza o inventário florestal (Batista *et al.*, 2014). Por meio da coleta dos diâmetros da floresta é verificado I. a amplitude total de tamanhos de indivíduos no povoamento; II. a distribuição das árvores por classe de tamanho (identificando as árvores dominantes e dominadas); III. o tipo de regeneração que se processa no povoamento (regeneração por ciclos ou regeneração contínua); 4. se existe algum problema no processo de regeneração e na passagem de indivíduos de uma classe de tamanho para outra. O diâmetro à altura do peito (DAP) das árvores tem relação com outras características como a altura, volume e biomassa acima do solo, que são medidos indiretamente nos levantamentos florestais, além de ser um indicador da estrutura do estoque de crescimento e base dos cálculos de sortimento florestal (Loetsch *et al.*, 1973; Campos; Leite, 2017).

As alturas das árvores, também, são informações importantes na quantificação da biomassa florestal, caracterizada por ser uma variável de difícil obtenção. É um atributo correlacionado com o diâmetro à altura do peito, já demonstrados na literatura em estudos de modelagem hipsométrica e são utilizadas na determinação do volume das árvores (Batista *et al.*, 2001; Soares; Tomé, 2002; Batista *et al.*, 2012). Em florestas homogêneas e com a mesma idade, as diferenças nas alturas dos indivíduos que possuem valores próximos de diâmetro na floresta, sugerem diferenças na qualidade do sítio florestal, sendo a altura utilizada na classificação de sítio (Skovsgaard; Vanclay, 2008). Neste último caso, a medida do potencial produtivo do sítio (conhecido como índice de sítio) tem como indicador o crescimento em altura das árvores dominantes (altura dominante). O índice de sítio é muito difundido nos estudos para mensurar unidades com maior potencial produtivo de biomassa florestal (Pretzsch, 2010; Weiskittel *et al.*; 2011).

Diretamente relacionado com a produção de madeira, o volume das árvores e a densidade das plantações sempre nortearam os estudos de manejo e crescimento das florestas e são balizadores no planejamento florestal (Clutter *et al.*, 1983). A curva de crescimento das árvores e as respectivas taxas de crescimento, somada aos critérios econômicos e ambientais, são utilizados para a tomada de decisão relativos ao desenvolvimento dos povoamentos, como a definição do período de rotação

florestal ótimo, a regulação e redistribuição do potencial produtivo das árvores (condução dos desbastes), estoque e sortimento florestal, e também utilizados no balanço de água, nutrientes e carbono (Rubilar *et al.*; 2018; Toda *et al.*, 2010). Particularmente, a densidade do povoamento, que pode ser expressa pela área basal, número de árvores por hectare e densidade de copa, tem uma forte relação com o crescimento, produtividade e as dimensões que a árvore pode atingir durante seu desenvolvimento (Burkhar; Tomé, 2012).

Desta forma, o diâmetro à altura do peito (DAP), a altura, a densidade de plantio e o volume são parâmetros imprescindíveis no manejo dos povoamentos florestais. Com essas variáveis florestais e na busca por automatização e maior precisão, nos últimos anos, as técnicas computacionais de aprendizado de máquinas vêm despertando cada vez mais o interesse dos pesquisadores da área de modelagem para quantificação da biomassa florestal, como o realizado por Azevedo *et al.*, (2020); Costa-Filho *et al.* (2019); Batista (2020) e Bueno e Costa (2020). Embora possam ser medidos (diretamente ou não) com uma precisão aceitável no campo, os altos custos envolvidos na amostragem desses recursos, bem como a aquisição lenta dos dados e a expansão da silvicultura na última década, culminaram na procura e adoção de novas metodologias para a extração das informações no campo de forma não destrutivo, como a realização do levantamento por sensores embarcados em veículo aéreo não tripulado (VANT) (Xu *et al.*, 2021; Maltamo *et al.*, 2018).

2.2 Tecnologia LiDAR no setor florestal

O sensoriamento remoto é a aquisição remota de dados, ou seja, inferir a respeito de determinado alvo, sem que haja a necessidade de um contato direto, envolve três elementos fundamentais: I. uma plataforma responsável pela aquisição dos dados, II. o sensor que é o dispositivo no qual realiza as medições e por fim, III. o objeto, que é aquilo que se quer estudar (ex: florestas). Os sensores remotos são dispositivos que podem ser classificados como passivos e ativos. Ao interagir com o alvo ao captar a radiação emitida, o sistema é denominado como ativo, como por exemplo, os lasers. Já os sensores passivos não possuem esse comportamento, ou seja, não possuem fonte de radiação própria (Pires, 2019; Giongo *et al.*, 2010).

Dentre os sensores passivos estão os multiespectrais e hiperespectrais. A utilização desses sensores no sensoriamento remoto tem impulsionando grandes avanços científicos no manejo e monitoramento das florestas, sobretudo com informações extraídas do dossel florestal para estimativa da deficiência nutricional, idade, doenças, ataques de pragas, estresse das árvores ou produtividade (Dash *et al.*, 2018; Housman *et al.*, 2018; Sandino *et al.*, 2018; Liang *et al.*, 2018; Zhou *et al.*, 2020). Para os sensores ativos, destaca-se o uso do sensor LiDAR que tem capacidade de mensurar simultaneamente a estrutura vertical e horizontal da floresta, bem como a morfologia do terreno, com elevado nível de detalhamento e velocidade (Gleason *et al.*, 2012; Sousa *et al.*, 1998).

O uso sensor consiste em uma técnica de sensoriamento remoto que se baseia na propagação de pulsos laser, que interagem com os objetos na superfície, a reflexão para o sensor permite a medição do tempo entre emissão e retorno (Equação 1), isso devido o funcionamento do conjunto, onde a precisa localização e altitude da unidade laser no momento da emissão e também uma exata coordenada são obtidas em cada ponto pela integração dos dados provenientes do sistema de posicionamento global e dos sensores inerciais. Posteriormente, é calculada as distâncias entre o sensor e o alvo em função do tempo entre emissão e retorno de pulsos (Yadav, 2016; Baltsavias, 1999; Christian; Lindenberger, 1999). É uma forma de detecção remota caracterizada pela emissão e recepção de pulsos de laser em diferentes direções e com uma elevada frequência de repetição para o escaneamento da superfície do objeto de interesse. O sistema é composto essencialmente por uma plataforma móvel ou fixa; um sensor LiDAR; um sistema de medição inercial (INS), receptores GNSS (Sistema de Navegação Global por Satélites) e um dispositivo para armazenamento dos dados (Gorgens *et al.*, 2014; Giongo *et al.*, 2010).

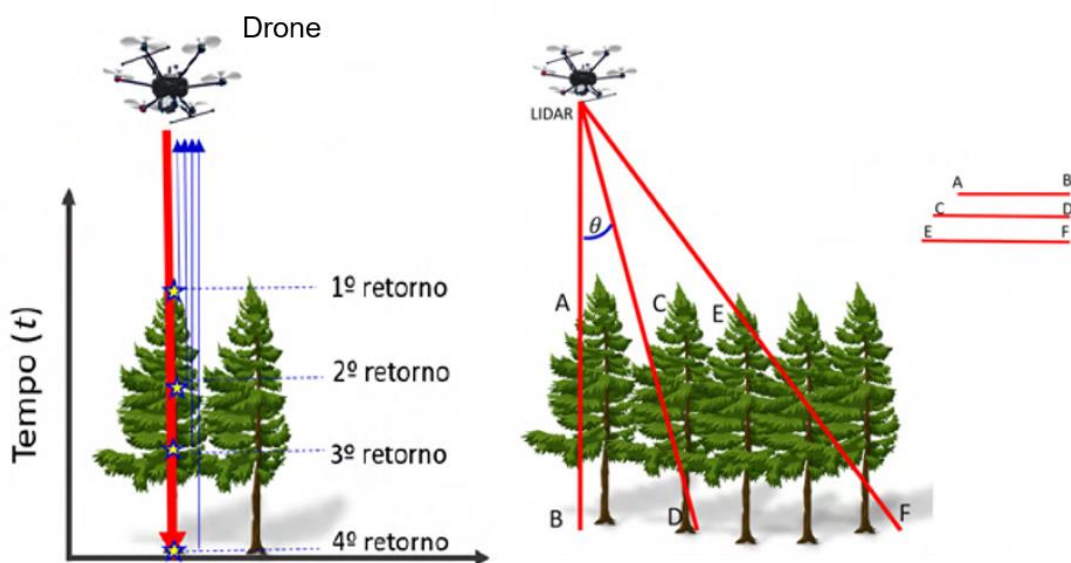
$$d = c \cdot \frac{t}{2} \quad (1)$$

Em que: d = distancia para o objeto (m); c = velocidade do pulso (m s^{-1}); t = tempo entre a luz emitida e o retorno (s).

O sensor LiDAR pode ser operado a partir de satélites, do solo ou então de plataformas aerotransportadas (Airbone Laser Scanning – ALS), como, por exemplo,

em drones (Figura 1). O resultado do escaneamento LiDAR de uma superfície são nuvens de pontos tridimensionais (3D), com grande quantidade de informações, sendo, portanto, densas e georreferenciadas, expressando a reflexão da superfície escaneada (Schuh, 2019; Yadav, 2016).

Figura 1 - Sistema de aquisição de dados LiDAR indicando os diferentes retornos e ângulo de varredura ao longo da emissão dos pulsos em um sobrevoo



Fonte: Adaptado de Corte (2022) e Jensen (2009)

Em âmbito florestal, seu uso possibilita um maior nível de detalhamento no interior da floresta pelo fato de os pulsos penetrarem o dossel e interagirem com o mesmo, isso torna possível captar toda a estrutura vertical da fisionomia florestal. É uma tecnologia com elevado potencial de mensurar variáveis florestais como cobertura e altura do dossel e variáveis dendrométricas. Já a biomassa da parte aérea, área basal e diâmetro podem ser obtidas por técnicas de modelagem (Giongo *et al.*, 2010). Diversos trabalhos utilizam as métricas obtidas dos dados LiDAR para a modelagem da biomassa florestal (Silva *et al.*, 2017; Badin, 2018; Dai, 2018). Segundo Leitold *et al.* (2015) essa é uma abordagem típica para prever essa variável, por meio da construção de modelos que associam as métricas LiDAR às estimativas de inventário florestal.

Para cada sistema de varredura com sensor LiDAR tem-se a obtenção de uma nuvem com diferentes densidades de pontos (pts m⁻²), que implicam no processamento dos modelos digitais e acurácia da derivação das características quantitativas das árvores (Cunha-Neto *et al.*, 2021). Na varredura a laser de alta densidade em VANT é obtido uma nuvem de pontos densas (acima de 1.000 pts m⁻²) que tem se tornado uma alternativa atrativa para a utilização em inventários florestais contínuos frente os métodos de varredura a laser aéreo (ALS) embarcados em aviões ou o de varredura a laser terrestre (TLS) (Maltamo *et al.* 2018; Puliti *et al.*, 2020; Rudge *et al.*, 2021).

Os pesquisadores têm demonstrado que o sistema de sensoriamento remoto ativo de varredura a laser ou LiDAR (Light Detection And Ranging) em veículos aéreos não tripulados é uma saída para estimativa de parâmetros da estrutura da floresta, inclusive para a coleta das variáveis dendrométricas com uma abordagem automática, sem a necessidade de validação em parcelas no campo (Puliti *et al.*, 2020; Hao *et al.*, 2021).

Existem duas abordagens nas estimativas florestais com esse sensor, a primeira com o enfoque por unidade de área e a segunda conhecida como abordagem de árvores individuais. No primeiro caso, as métricas LiDAR são extraídas da nuvem de pontos das unidades amostrais (parcelas), com isso, será possível inferir estimativas para a área total. Já a outra abordagem identifica as árvores uma a uma, de forma é se torna possível inferir estimativas para a área total. A escolha da abordagem a ser utilizada é em função da densidade de pontos da nuvem LiDAR diretamente correlacionado com o modelo LiDAR utilizado e o veículo embarcado. Em nuvens de pontos de baixa densidade 5 - 30 pts.m (normalmente obtidas por sensores aerotransportados), se faz o uso do enfoque por unidade de área (Tabela 1). Em nuvens de pontos de alta densidade, > 100 pts.m (normalmente obtidos por sensores terrestres ou drones), se faz uso da abordagem de árvore individual (Corte *et al.*, 2022). No entanto, o desenvolvimento de aplicações para a obtenção de atributos à nível de árvore individual tem encontrada restrições de resolução espacial e acurácia (Luther *et al.*, 2014; Xu *et al.*, 2021).

Tabela 1 - Potenciais contribuições do uso da tecnologia LiDAR no setor florestal

Plataforma de aquisição	Abordagem	Variável	Autores
ALS	Árvore individual	Diâmetro e Altura	Chasmer <i>et al.</i> (2006); Persson <i>et al.</i> (2002); Vasconcellos <i>et al.</i> (2020)
		Biomassa	Rex <i>et al.</i> (2019); Vasconcellos <i>et al.</i> (2020)
		Deteção de árvores, classificação de espécies, altura de dossel	Jaakkola <i>et al.</i> (2010); Pitkänen <i>et al.</i> (2004); Dalponte <i>et al.</i> (2014)
	Unidade de área	Altura, deteção de árvores	Holopainen <i>et al.</i> (2013); Moe <i>et al.</i> (2020); Næsset <i>et al.</i> (2001); Kolendo <i>et al.</i> (2021)
		Biomassa	Silva <i>et al.</i> (2018); Rex <i>et al.</i> (2020)
		Cobertura do Subdossel	Wing <i>et al.</i> (2012); Almeida <i>et al.</i> (2019)
		Diâmetro	Holopainen <i>et al.</i> (2013); Moe <i>et al.</i> (2020)
		Idade	Racine <i>et al.</i> (2014)
		Carbono	Baccini <i>et al.</i> (2013)
		Volume	Watt <i>et al.</i> (2013); Carvalho <i>et al.</i> (2015)
UAV	Árvore Individual	Altura	Da Cunha <i>et al.</i> (2021)
		Deteção de árvores	Jaakkola <i>et al.</i> (2010)
		Diâmetro, altura	Dalla Corte <i>et al.</i> (2020)
		Diâmetro, altura, volume	Dalla Corte <i>et al.</i> (2020)
	Unidade de área	Área basal	Silva <i>et al.</i> (2017)
		Carbono, área basal, volume	Fekety <i>et al.</i> (2015); Almeida <i>et al.</i> (2020); Almeida <i>et al.</i> (2021)
		Deteção de árvores, diâmetro	Moe <i>et al.</i> (2020)
		Altura, largura de copa	Wallace <i>et al.</i> (2012)

Fonte: Adaptado de CORTE *et al.* (2022)

De forma geral, o processamento de dados LiDAR, a nível de árvore individual se inicia com a classificação e normalização dos pontos, seguido da obtenção do CHM (Canopy Height Model), ou modelo de altura de copa (CHM – Canopy Height Model) resultante da subtração do modelo digital de terreno (Digital Terrain Model – DTM) do modelo digital da superfície (Digital Surface Model – DSM), posteriormente há a

detecção da árvore, e por último, a extração das métricas da nuvem de pontos para modelagem das características de interesse e extrapolação para todo o povoamento (Mohan *et al.*, 2021).

A previsão precisa de atributos individuais de árvores com base nos dados LiDAR dependem dos métodos utilizados na detecção das árvores de geralmente segue três etapas: I. detecção de árvore individual e extração de métricas na nuvem (3D), II. Correspondência de árvore baseada em LiDAR e em inventário campo, e III. modelagem e predição (Silva *et al.*, 2016). A abordagem de individualização de árvores por dados UAV é eficiente para avaliar o número de árvores em florestas de eucalipto, porém se faz necessário que haja a orientação do algoritmo pelo inventário de campo (Cozenza *et al.*, 2018). As métricas LiDAR podem que podem ser extraídas por árvore podem ser de contagem, elevação ou altura, intensidade e topografica, classificadas com subgrupos de posição ou dispersão (Gorgens *et al.*, 2014).

2.3 Modelagem de dados UAV-LS com uso de técnicas de aprendizado de máquinas

A subdivisão da inteligência artificial baseada no aprendizado biológico é chamada de aprendizado de máquina (em inglês, *machine learning*), que inclui técnicas para classificação de elementos, predições de valores, extração de regras e seus resultados muitas vezes são melhores do que técnicas tradicionais (Montaño, 2016). As vantagens destas técnicas estão: a precisão, a automaticidade na explicação de dados, a versatilidade de aplicação, e a melhoria continua dos resultados com aumento do conjunto dos dados. Essas técnicas são fortemente influenciadas pelos hiperparâmetros, cuja a definição se dá antes do treinamento dos modelos (Lary *et al.*, 2016).

Os modelos de machine learning podem ser supervisionadas ou não supervisionados, sendo que os modelos não-supervisionados são recomendados para situações em que não se tem ideia da possível resposta desejada, ou seja, os modelos aprendem mais sobre a estrutura dos dados, geralmente é necessário descobrir estruturas, agrupamentos e padrões nos resultados para a interpretação final. Já os supervisionados são aplicados em situações para fazer predições ou classificar um conjunto de dados em que há uma saída esperada, facilmente rotuladas para que haja comparação entre os resultados (Vendruscolo *et al.*, 2015).

Para a derivação das variáveis dendrométricas a partir da nuvem de pontos LiDAR, tem-se proposto o uso de modelos supervisionados (especificamente de regressão), que são treinados a partir de um conjunto de dados para os quais o rótulo da classe associada é conhecido (ou seja, o banco de dados é coletado em parcelas no campo). Destaque para os algoritmos de redes neurais artificiais (ANN - Artificial Neural Networks); floresta aleatória (RF - Random Forest); e regressão por vetores suporte (SVR - Support Vector Regression) (Lee *et al.*, 2018; Costa-Filho *et al.*, 2019; Marrs; Ni-Meister, 2019; Corte *et al.*, 2021).

O algoritmo de floresta aleatória é baseado no método ensemble que combina o resultado de diversos modelos preditivos para realizar a classificação dos dados. De forma geral, ele possui poucos hiperparâmetros para seu ajuste e o funcionamento consiste em criar várias árvores de decisão a partir da escolha aleatória de variáveis em um conjunto de treinamento, onde o resultado final é obtido pela combinação dos resultados de cada árvore (Breiman, 2001). A quantidade de árvores (Ntree) e de variáveis selecionadas em cada nó (Mtry) a serem utilizadas são definidas previamente no treinamento (Costa-Filho *et al.*, 2019).

São modelos que analogicamente se assemelham a uma estrutura de árvore invertida, com os nós e as folhas que se ramificam, ao simularem um conjunto de decisões e resultados. Dentre os modelos baseados na arquitetura de árvores de decisão estão os: Decision Trees Regressor, Extra Trees, Random Forest e Gradient Boosting Trees. Criados maior robustez e complexidade na modelagem, com o intuito de evitar situações de Overfitting, onde o processo de ajuste indica que o modelo em treino expressa bons resultados, mas quando aplicado em um conjunto de dados não visto, o modelo não é capaz de desempenhar boa performance, pelo fato de absorver os ruídos ou flutuações do conjunto de dados (Domingos, 2012; Liaw *et al.*, 2002).

Nos modelos de regressão em geral, os hiperparâmetros são definidos, com isso, os modelos são definitivamente ajustados no conjunto de treinamento, posteriormente é avaliado seu desempenho no conjunto teste utilizando parâmetros estatísticos. Neste processo, sugere-se que sejam avaliadas estatisticamente as métricas da nuvem de pontos para que sejam utilizadas nos modelos (Corte *et al.*, 2020a).

As métricas extraídas da nuvem de pontos do escaneamento a laser vêm sendo aplicadas em algoritmos de aprendizado de máquina para a predição de variáveis dendrométricas como alternativa à abordagem convencional com modelos estatísticos de regressão (Lee *et al.*, 2018; Marrs; Ni-Meister, 2019; Corte *et al.*, 2020a; Neuville *et al.*, 2021; Yu *et al.*, 2021). As vantagens das técnicas de machine learning incluem a precisão; a automaticidade na explicação de dados, a versatilidade de aplicação; a melhoria contínua dos resultados com aumento do conjunto dos dados; permitem a inclusão de variáveis categóricas; lidam com a não-linearidade entre as variáveis e possuem tolerância as falhas e ruídos (Lary *et al.*, 2016; Haykin, 2009).

2.4 A biblioteca Lazy predict em Python

Programada em Python, a biblioteca Lazy predict desenvolvida em 2019, tem como objetivo avaliar a performance de um conjunto de modelos que podem ser de classificação ou regressão. O Lazy Predict Regressor ajusta de forma automatizada cerca de 46 modelos de regressão, que são avaliados por um conjunto de métricas como: coeficiente de correlação ajustado (R^2_{aj}) e Raiz do Erro Quadrático Médio (RMSE). Como construir e comparar modelos de aprendizado de máquina é um processo complexo, a ferramenta cria então um ranqueamento com base no desempenho do ajuste, algo que norteia a escolha do melhor modelo para determinado conjunto de dados (Shankar, 2022, Silva, 2023).

Por tratar-se de uma biblioteca relativamente recente, são poucos os trabalhos que abordam seu uso, principalmente relacionadas ao setor florestal. Mas pode ser utilizado em diversos conjuntos de dados e problemáticas como por exemplo: na área da saúde, é utilizado com mais frequência, principalmente para avaliação de medicamentos e riscos de doenças (Rached *et al.*, 2023; Gao *et al.*, 2023; Deepa *et al.* 2023). No setor agrícola, existem trabalhos a utilizam para estimar produção de biogás, na avaliação da propriedade física de biocombustíveis e até mesmo na previsão de rendimento e produtividade de plantações de algodão (Severino, 2023; Silva *et al.* 2022; He, 2022).

O uso do Lazy Predict Regressor para trabalhos relacionados a plantações florestais ainda é algo incipiente, porém, não significa que a mesma não possa ser empregada em problemáticas que envolvam mensuração florestal, uma vez que a

biblioteca é uma ferramenta que a priori busca a automatização da modelagem de modelos de aprendizado de máquinas (Silva, 2023).

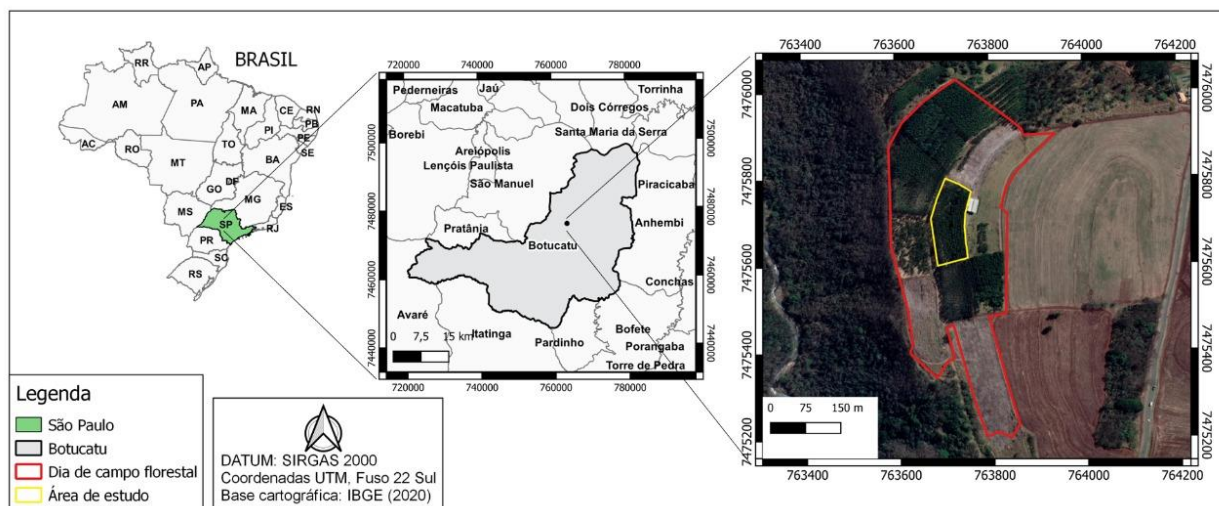
3 MATERIAL E MÉTODOS

3.1 Área de estudo

O estudo foi conduzido na Fazenda Lageado, pertencente a Faculdade de Ciências Agrônômicas (FCA/UNESP), campus de Botucatu (SP). A área experimental, contempla um povoamento clonal de *Eucalyptus toreliodora*, com 7 anos de idade, e 1,25 ha de área total, plantado no espaçamento 3 m × 2 m; a uma longitude 48° 25' 51,9" W e latitude 22° 48' 20,16" S (Figura 2).

Segundo a classificação climática de Köppen-Geiger, é caracterizado por clima subtropical úmido com verão quente - Cfa, possui a temperatura média do mês mais quente superior a 22°C e apresenta no mês mais seco uma precipitação pluviométrica maior de 40 mm (Alvares *et al.*, 2013). O plantio está situado a 727 m de altitude, em região com precipitação pluviométrica média de 1400 mm ano⁻¹ e temperatura do ar média anual de 20°C. O solo do local foi classificado como Nitossolo vermelho com textura argilosa (Embrapa, 2022).

Figura 2 - Área experimental onde foram realizadas as coletas de dados



Fonte: Elaborado pelo autor (2023)

3.2 Coleta de dados no campo

Foi realizado o censo da área pelo qual todas as árvores foram mensuradas com a coleta do diâmetro à altura do peito (DAP) com auxílio de uma suta de alumínio (Haglöf), totalizando 1292 observações. A altura total por meio do clinômetro (Haglöf),

foi coletada de forma sistemática a cada 4 árvores no campo, totalizando 407 observações.

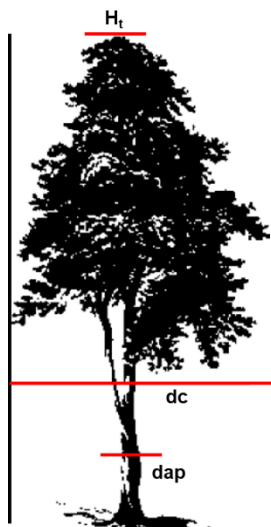
Com base nos dados levantados no campo ainda foram calculadas as estatísticas dendrométricas para o plantio, sendo: diâmetro quadrático médio (D_g) e altura dominante (H_d) (Equação 2). Sendo que o conceito de altura dominante, tem como referência a altura total das 100 árvores com maior crescimento basal, ou seja, os moires valores de DAP considerando uma área fixa de um hectare (Assmann, 1970). A área em estudo obteve um D_g de 15,71 cm e uma altura dominante de 18,71 m.

$$D_g = \sqrt{\frac{\bar{g} \cdot 4}{\pi}} \quad (2)$$

Em que: D_g = diâmetro quadrático médio (cm); G = área basal (m^2); DAP = diâmetro a altura do peito à 1,30 m do solo.

Outra variável fundamental mensurada após a derrubadas das árvores-amostra em campo, foi o diâmetro de copa das árvores (Figura 3) conforme proposto por Wink *et al.*, (2012). Essa informação foi utilizada no algoritmo de segmentação a nível de árvore individual descrito por Silva (2016) presente no Item 3.3.

Figura 3 - Exemplo da obtenção de variáveis dendrométricas no campo.



Fonte: Adaptado Wink *et al.* (2012)

Após coleta do diâmetro à altura do peito - DAP (Tabela 2), as árvores foram subdivididas em classes de diâmetro para seleção dez árvores-amostra que foram abatidas em seis classes de diâmetro (total de 60 árvores), selecionadas a partir do centro de cada classe (Tabela 1) para realizar o procedimento de cubagem rigorosa das árvores pelo método de Smalian para a determinação do volume (Batista *et al.*, 2014; Sanquetta *et al.*, 2002). A partir da base da árvore, foram coletados o diâmetro a cada metro no fuste, a fim de obter o volume individual com casca de cada árvore. Com o uso das equações 3, 4, 5, 6 foi determinado o volume das seções da árvore.

$$V_i = \frac{(g_1 + g_2)}{2} \cdot l \quad (3)$$

$$g_1 = \frac{\pi d_1^2}{4} \quad (4)$$

$$g_2 = \frac{\pi d_2^2}{4} \quad (5)$$

$$V_p = \frac{g_p \cdot l_p}{3} \quad (6)$$

Em que: V_i = volume da seção i (m^3); V_p = volume do ponteiro (m^3); g_1 = área transversal na ponta mais grossa da tora (m^2); g_2 = área transversal na ponta mais fina da tora (m^2); g_p = área transversal da última seção (m^2); d_1 = diâmetro na base da seção (m); d_2 = diâmetro no topo da seção (m); l = comprimento da tora (m); l_p = comprimento da última seção (m).

O volume total do fuste da árvore (V_t) com casca foi calculado pelo somatório dos volumes das seções, descrito pela Equação 7.

$$V_t = V_1 + V_2 + V_3 + \dots + V_n + V_p \quad (7)$$

Tabela 2 - Quantidade de árvores mensuradas e classes estabelecidas para a derrubada das árvores-amostra no processo de cubagem

Classe	CC (cm)	Árvores
1	3,7	103
2	8,7	260
3	13,7	549
4	18,7	288
5	23,7	62
6	28,7	30

Fonte: Elaborado pelo autor (2023)

Em que: CC = centro de classe gerado com base no intervalo superior e inferior (cm).

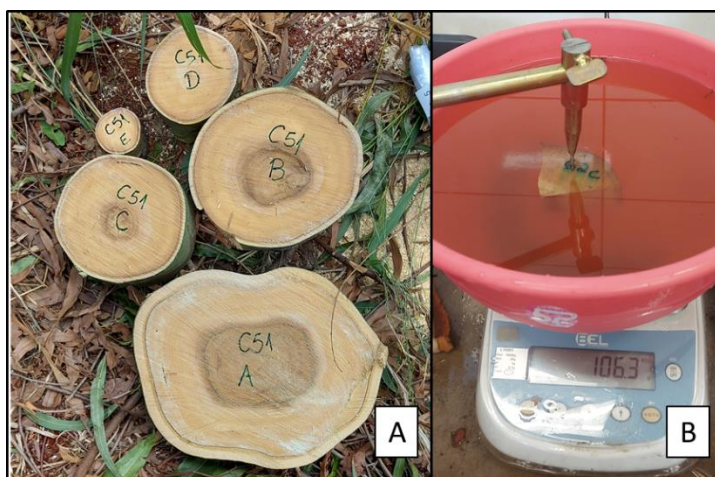
Para quantificação da biomassa, foi obtida a densidade básica da madeira com o uso metodologia descrita (Eufrade-Junior, 2015), que considera a retirada de discos nas posições 0, 25, 50, 75 e 100% da altura comercial (diâmetro comercial de 8 cm) de cada árvore abatida. De acordo com os procedimentos descritos na ABNT NBR 11.941 (2003), os discos de madeira foram submersos em água para a saturação logo após a coleta visando determinar o volume saturado dos discos (Figura 4). Subsequentemente, foram secos em estufa (a $105 \text{ } ^\circ\text{C} \pm 2^\circ\text{C}$) para obtenção da massa seca. Com isso a relação entre massa seca e volume saturado foi utilizada para o cálculo da densidade básica da madeira com casca em cada posição com o uso da

Equação 8. A densidade básica ponderada média com casca foi de 517 (CV = 15%) kg/m³ com casca.

$$D_{b,pond} = \frac{Db_{(0-25\%)} \times V_{(0-25\%)} + \dots + Db_{(75-100\%)} \times V_{(75-100\%)}}{V_{(0-25\%)} + V_{(25-50\%)} + V_{(50-75\%)} + V_{(75-100\%)}} \quad (8)$$

Em que: $D_{b,pond}$ = densidade básica da madeira com casca ponderada pelo volume (kg m⁻³); D_b = densidade básica média entre os discos na posição considerada (kg m⁻³); V = volume da seção do fuste na posição considerada (m³).

Figura 4 - Discos das árvores-amostras abatidas em campo para determinação da densidade básica. A: Discos coletados em campos em diferentes posições. B: Aferição da massa saturada do disco imerso em água



Fonte: Elaborado pelo autor (2023)

Para a determinação da massa seca do fuste com a casca, foi utilizada a metodologia indireta (Blanco; Eufrade-Junior *et al.*, 2023), em função do produto entre o volume dos fustes (calculados pela cubagem rigorosa) e o valor da densidade básica (Equação 9).

$$MS_{fuste} = V_{fuste} \cdot D_{b,pond} \quad (9)$$

Em que: MS_{fuste} = massa seca do fuste com casca (kg árvore^{-1}); V_{fuste} = volume do fuste com casca (m^3); $D_{b,pond}$ = densidade básica da madeira com casca ponderada pelo volume (kg m^{-3}).

A partir da biomassa seca de fuste com casca, foi realizado o ajuste do modelo Schumacher e Hall Logarítmico (1933) para prever a biomassa individual de fuste de todas as árvores do plantio. Este modelo é um dos mais difundidos na área florestal devido às suas propriedades estatísticas, pois quase sempre resulta em estimativas sem viés (Souza; Soares, 2013; Azevedo *et al.*, 2020). A forma linearizada do modelo $\ln(MS_{fuste}) = -4,8830 * \ln DAP + 1,4434 * \ln H + 1,9035 + \epsilon$, apresentou os seguintes parâmetros estatísticos: coeficiente de determinação ajustado (R^2_{aj}) de 0,9787, com um erro padrão de estimativa (S_{yx}) de 5,14 e desvio médio percentual (DMP) de 16,02.

3.3 UAV-LS e processamento da nuvem de pontos LiDAR

Os dados UAV-LiDAR foram coletados no mesmo período do inventário censitário (dezembro/2021), para isso, a nuvem de pontos 3D das árvores foi obtida pelo sensor mdLiDAR1000LR desenvolvido pelas empresas Microdrones / Geo Cue e General Electric (Figura 5). O conjunto consiste em um drone industrial de asas rotativas e com quatro rotores (modelo md4-1000), um sensor LiDAR, um sistema integrado inercial-GNSS da empresa Trimble Applanix modelo APX-15 UAV e um dispositivo de armazenamento de 32GB USB. Foi utilizada uma base Trimble para o pós-processamento para correção do erro inercial. O voo foi realizado com velocidade de 4 m s^{-1} ($14,4 \text{ km h}^{-1}$), com distância de 15 m entre as linhas (para atingir 90% de sobreposição) e altura de 40 m da copa das árvores a fim de promover uma densidade média de pelo menos 245 pontos m^{-2} , a largura da faixa imaginada foi de 180 m com uma densidade média de pulsos de 282 pulsos por faixa.

Figura 5 - A) Drone industrial de asas rotativas quadricoptero (modelo md4-1000); B) Sensor LiDAR mdLiDAR1000LR embarcado em drones



Fonte: Elaborado pelo autor (2023)

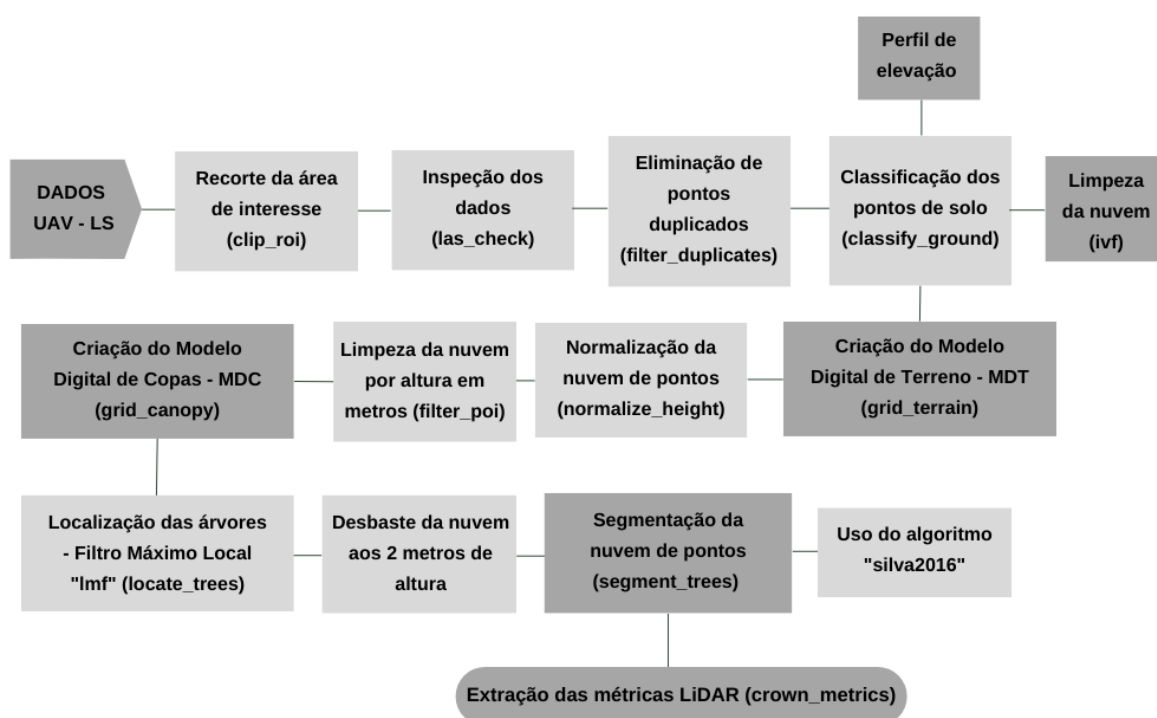
As etapas do processamento estão sintetizadas na Figura 6 com as respectivas funções utilizadas. Os dados foram pré-processados com auxílio do software LP360 (True View Evo). No software, as coordenadas foram corrigidas com uso do processamento “POSPac”, as linhas de voo foram criadas e ajustadas. Após o pré-processamento, foi realizado o processamento em linguagem R (R CORE TEAM, 2019), com auxílio dos pacotes lidR (Roussel; Auty, 2019) e rLiDAR (Silva *et al.*, 2017). Os arquivos tipo raster (MDT e MDC) foram gerados com uma resolução espacial de 1 x 1 m. A nuvem de pontos foi classificada em pontos de solo e pontos de vegetação. Para realizar o processo de segmentação das árvores, ou seja, determinação dos pontos pertencentes a cada árvore na projeção tridimensional, foi utilizado o algoritmo “Silva2016”, onde as entradas são: I. valor máximo de diâmetro de copa; II. exclusão de ruídos; III. o modelo digital de copas (CHM) (Silva *et al.*, 2016).

Um processo crucial para a segmentação das árvores, é a localização espacial das árvores, para isso, foi realizado a comparação entre duas janelas de seleção de pontos de máxima (2 e 3 metros). As janelas do algoritmo de Filtro de Máximo Local (“lmf”) foram avaliadas conforme realizado por Rodrigues (2021) com base no Erro Absoluto (EA) e pelo Erro Relativo (ER) (Equações 10 e 11).

$$EA = N^{\circ} \text{ árvore predita} - N^{\circ} \text{ árvore observada} \quad (10)$$

$$ER = \frac{EA}{N^{\circ} \text{ árvore observada}} \times 100 \quad (11)$$

Figura 6 - Fluxograma do processamento da nuvem de pontos



Fonte: Elaborado pelo autor (2023)

As métricas LiDAR extraídas da função “crown_metrics” estão sumarizadas no Tabela 3.

Tabela 3 - Métricas LiDAR extraídas por árvore

Métrica	Descrição
zmax	Altura máxima
zmean	Média da altura
zsd	Desvio padrão da distribuição da altura
zskew	Assimetria da distribuição da altura
zkurt	Curtose da altura
pzabovzmn	Porcentagem de retornos acima da média
zq (x=5,...,95) range =5	Xth percentil da distribuição de alturas
Zpqum (x=1,...9) range = 1	Percentil acumulativo da distribuição da altura
n points	Número da pontos
convul area	Área projetada de copa

Fonte: Adaptado Corte *et al* (2020)

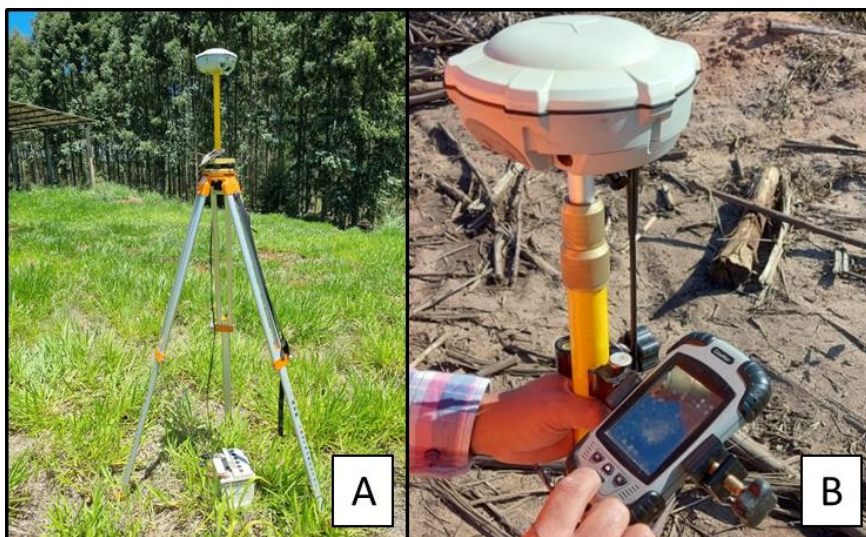
3.4 Espaço amostral e georreferenciamento das árvores em campo

Para o processamento dos dados LiDAR e modelagem da biomassa seca de fuste, a nuvem de pontos foi amostrada em 25% (319 árvores), com uma área 3300 m², realizado para avaliar a identificação individual das árvores e ranquear os modelos em uma população controlada de treinamento. A definição do espaço amostral teve como norteador avaliar a performance do lmf para área de estudo, visto que, é um procedimento crucial para cruzar as informações de campo com as obtidas pela nuvem de pontos. Devido a essa característica, foi realizado uma amostragem de 25% da nuvem de pontos devidamente demarcado em campo, para controle das árvores avaliadas.

Foram coletados coordenadas no campo pela predefinição de 10 pontos de controle na área experimental em posições de elevada captação de sinal satelital em posições que facilitassem a obtenção das coordenadas arbitrárias. O equipamento utilizado é da marca COMNAV, receptores com tecnologia GNSS RTK, 256 canais com localização interligada com sinais de satélites GPS: L1 C/A, L1/L2 P, L5 - BeiDou: B1, B2, B3 - GLONASS: L1/L2 - SBAS: WASS, EGNOS, MASAS. Os dados brutos foram disponibilizados em arquivo DWG e em texto dos pontos processados pelo IBGE em coordenadas geográficas locais. Ambos os equipamentos, Receptor GNSS T300 + Coletor de dados R100, estão ilustrados pela Figura 7.

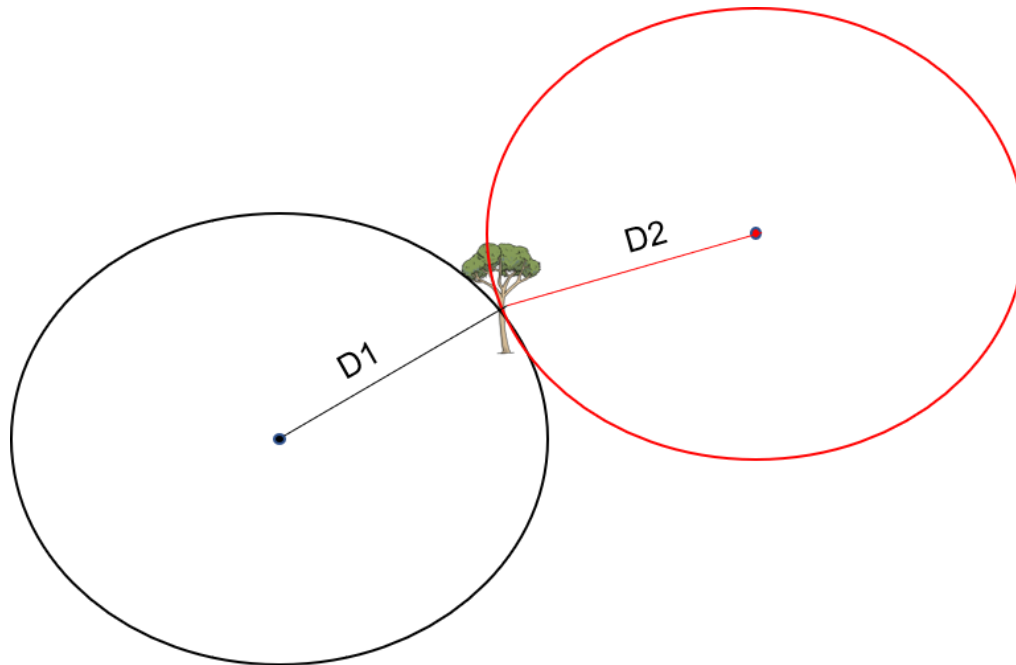
Em campo, os pontos de controle foram utilizados como referência para a delimitação da área e a obtenção de coordenadas arbitrárias de cada árvore. Para isso, foi mensurado com o uso trenas, as distâncias (metros) do tronco em relação a dois pontos de controle distintos, ou seja, duas medições por árvore (Figura 8). No processamento, com esses dados, foi possível gerar circunferências a partir dos pontos de controle que ao se tangenciarem estabeleceram a coordenada da árvore (Figura 9). As coordenadas arbitrárias foram utilizadas como avaliar as janelas de seleção Imf, na nuvem de pontos 3D, procedimento crucial para identificação correta das árvores na projeção tridimensional (Figura 9).

Figura 7 - A) Receptor GNSS T300; B) Sistema de aquisição da dados



Fonte: Elaborado pelo autor (2023)

Figura 8 – Distâncias dos pontos de controle mensuradas em campo, obtidas para georreferenciar as árvores



Fonte: Elaborado pelo autor (2023)

Figura 9 – Disposição espacial das coordenadas geográficas das árvores georreferenciadas, projetadas em ortofoto da área de estudo.



Fonte: Elaborado pelo autor (2023)

O espaço amostral possui 3300 m², isso representa 25% da área total. As árvores da borda do interior foram desconsideradas do processo de modelagem em função do “clip” realizado na nuvem de pontos, depreciando suas informações.

3.5 Modelagem da biomassa seca com casca de fuste

A modelagem da biomassa acima do solo foi realizada em função das métricas LiDAR, os resultados possibilitaram um ranqueamento de modelos de aprendizado de máquinas. Para isso foi utilizado a biblioteca Lazy Predict Regressor em Python, que ajusta modelos de aprendizado de máquinas e gera parâmetros estatísticos descritos no item 3.6, com isso é possível ranqueá-los com base em sua performance (Silva, 2023).

3.6 Estatística para verificação do ajuste dos modelos e variáveis

Os parâmetros estatísticos descritos no Tabela 4 foram utilizados para avaliar o grau de ajuste do processo de modelagem. Devido ao reduzido número de amostras (319 árvores) os modelos foram apenas treinados.

Tabela 4 - Parâmetros estatísticos utilizado para avaliar a performance dos modelos de aprendizado de máquinas

Parâmetro	Expressão matemática
Coeficiente de determinação ajustado R^2	$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$
Raiz do erro médio quadrático RMSE	$RMSE(Mg/ha) = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{n}}$

Fonte: Adaptado de Schuh (2019) e Schneider *et al.* (2009)

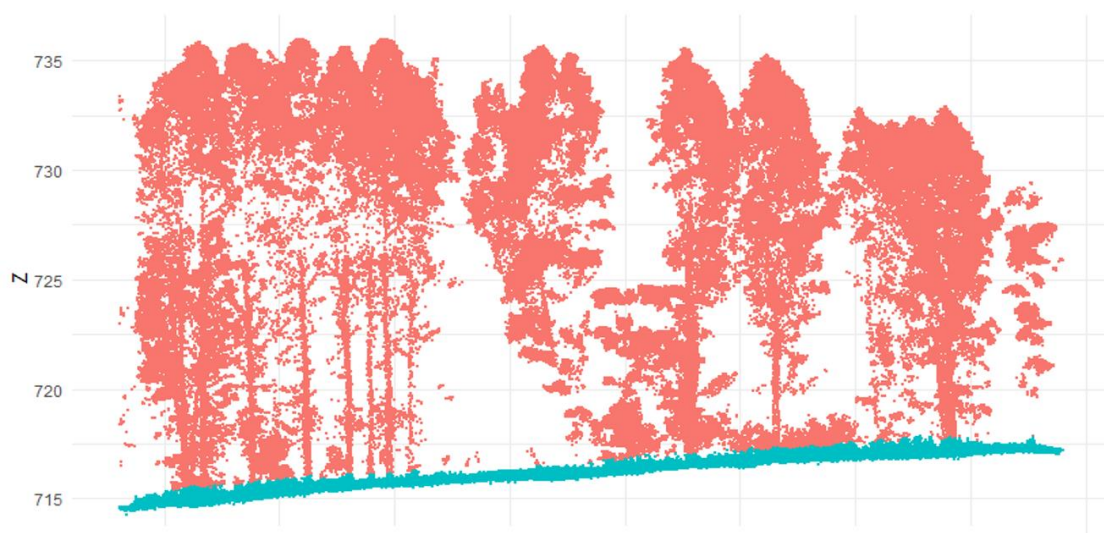
Em que: n = é o número de unidades amostrais, y_i = é o valor observado para a biomassa na unidade amostral i , $\bar{y}_i$ = é a biomassa média para a unidade amostral i , e $\hat{y}_i$ = é o valor predito para a unidade amostral i .

4 RESULTADOS E DISCUSSÃO

4.1 Produtos obtidos pelo processamento da nuvem de pontos

No perfil de elevação (Figura 10) a nuvem de pontos ainda não passou pelo processo de normalização. O resultado disso é uma escala em altitude, que varia de 715 a 735 m. Nota-se também duas diferentes colorações, que representam a classificação dos pontos, fica evidente divisão entre pontos que representam o solo e pontos de vegetação.

Figura 10 - Perfil de elevação entre pontos máximos e mínimos



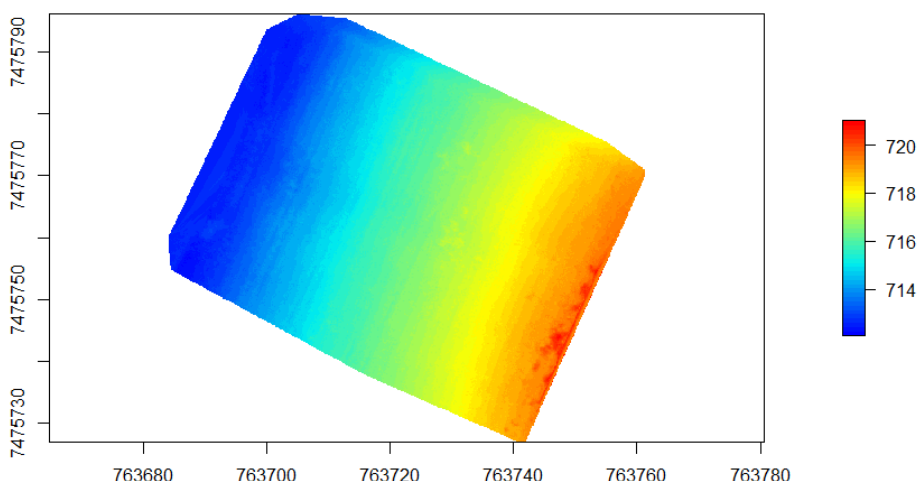
Fonte: Elaborado pelo autor (2023)

Os pontos que representam o solo foram classificados pelo algoritmo CSF, sendo possível gerar o Modelo Digital de Terro (MDT), com uma elevação de 714 a 720 m (Figura 11). O CSF baseia-se em técnicas de agrupamentos por meio de uma avaliação de regiões densas e agrupadas na nuvem de pontos que são identificadas pela criação de clusters, sendo que o input de número de clusters e a criação de valores referências de densidade de pontos são fundamentais para a classificação

dos pontos de solo (Zhang *et al.*, 2016, Yu *et al.*, 2017). Para a situação em questão foi utilizado as configurações padrão do algoritmo CSF.

Outro algoritmo de classificação de pontos de solo, o “Progressive Morphological Filtering”, não se baseia no adensamento de pontos, mas considera todos os pontos como não solo, e avalia os pontos de solo por meio de operações morfológicas que infere a respeito da distribuição espacial dos pontos. Se por um lado o CSF utiliza a clusterização dos pontos o PMF avalia a distribuição espacial dos pontos para morfológicamente classifica-los (Zhang *et al.*, 2003; Vosselman *et al.*, 2010, Jakubowski *et al.*, 2013).

Figura 11 - Modelo digital de solo gerado a partir da classificação dos pontos

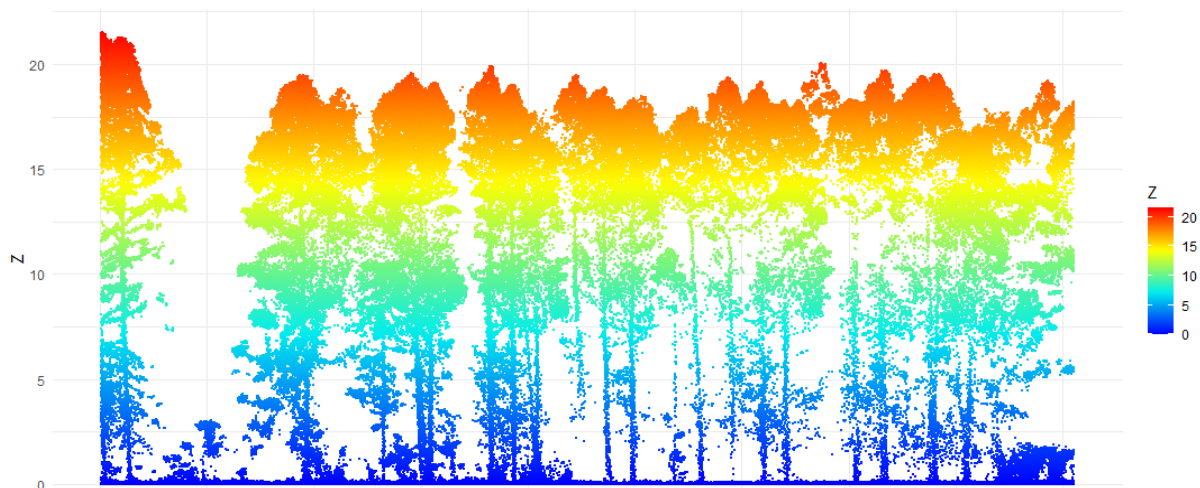


Fonte: Elaborado pelo autor (2023)

Após a normalização, é possível inferir a respeito da altura da floresta em estudo, na Figura 12 é possível notar que as árvores possuem aproximadamente 20 metros de altura. Devido a alta densidade de pontos é visível um maior detalhamento da nuvem, possibilitando a visualização do perfil de alguns fustes e também a presença de sub-bosque. A normalização da nuvem de pontos foi realizada pela interpolação dos pontos de solo por um modelo de triangulação cuja a conexão de triângulos representa a superfície, dando uma referência (o solo) para os demais pontos atribuídos como vegetação (Zhou, 2007). Existem outros métodos de

interpolação para a normalização da nuvem de pontos, como por exemplo: krigagem universal e a abordagem do vizinho mais próximo (Christakos, 2012; Goovaerts, 1997).

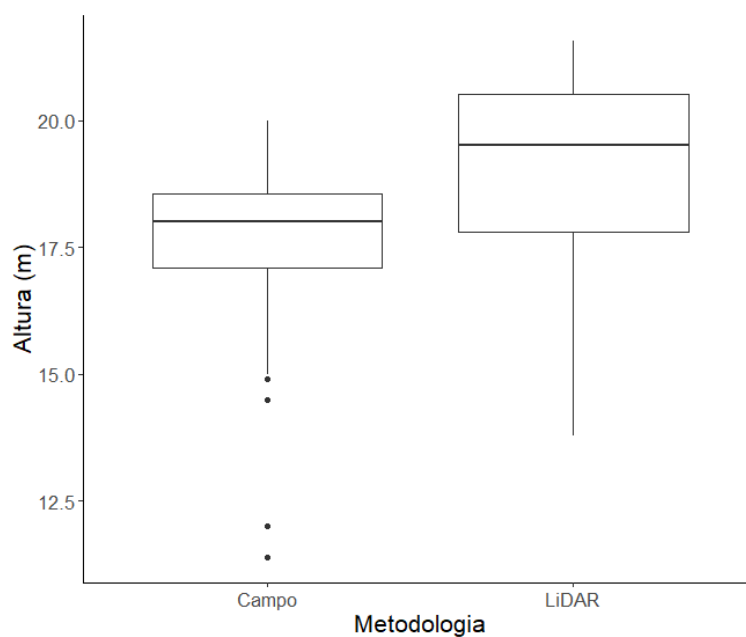
Figura 12 - Perfil da nuvem de pontos, primeiro produto obtido da nuvem normalizada, para avaliação dos dados



Fonte: Elaborado pelo autor (2023)

Para as alturas amostradas no espaço amostral (Figura 13), em campo obteve-se uma altura média de 17,6 metros ($cv = 8\%$) enquanto que com a individualização das árvores a altura média foi de 19,0 metros ($cv = 10\%$), conforme ilustrado na Figura 11. As variações entre as metodologias podem chegar a 1 metro segundo Liao *et al.* 2022 ao avaliar uma floresta de *Eucalyptus* spp.

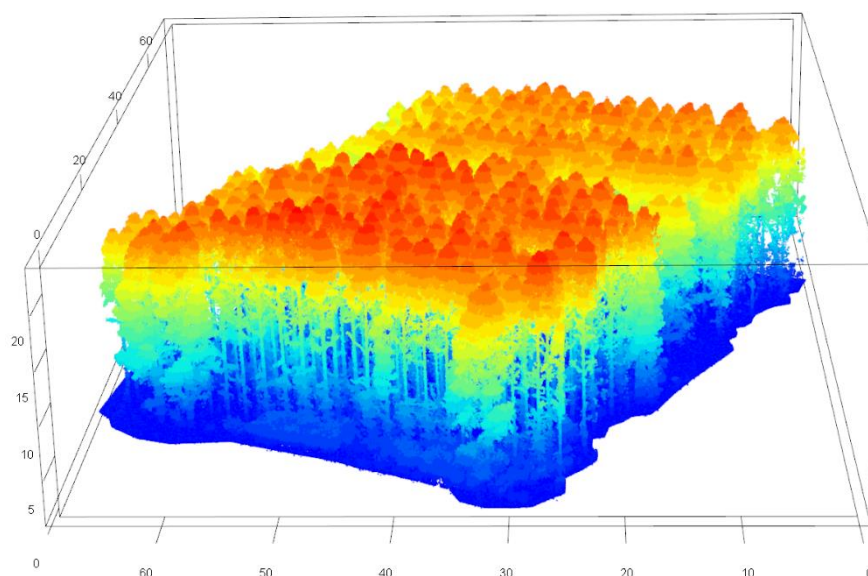
Figura 13 –Alturas obtidas em campo e as alturas obtidas dos pontos de máxima de cada árvore resultado da individualização



Fonte: Elaborado pelo autor (2023)

A projeção tridimensional (Figura 14) foi colorizada com base na altura dos pontos em relação ao solo, ou seja, cores mais quentes indicam pontos mais distantes do solo.

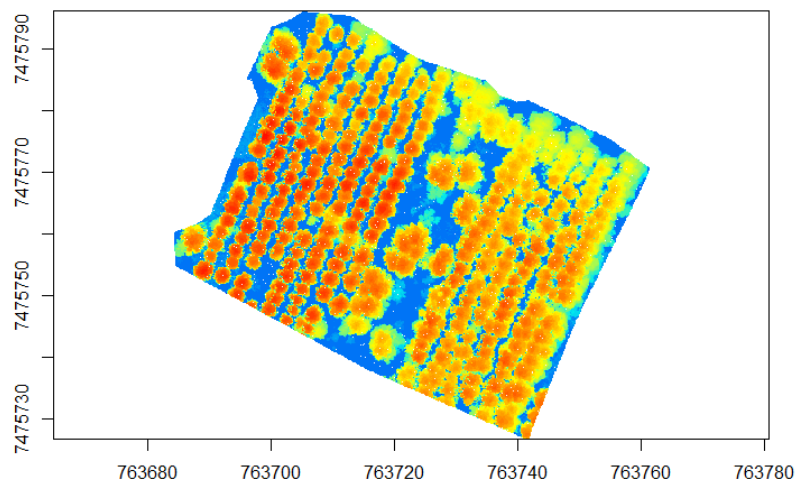
Figura 14 – Nuvem de pontos com elevado detalhamento da área de estudo



Fonte: Elaborado pelo autor (2023)

O modelo digital de Copas (Figura 15) é um dos produtos raster do processamento da nuvem e foi utilizado como input no processo de segmentação.

Figura 15 - Modelo digital de copas colorizado com base na altura das árvores tendo como referência os pontos classificados como solo



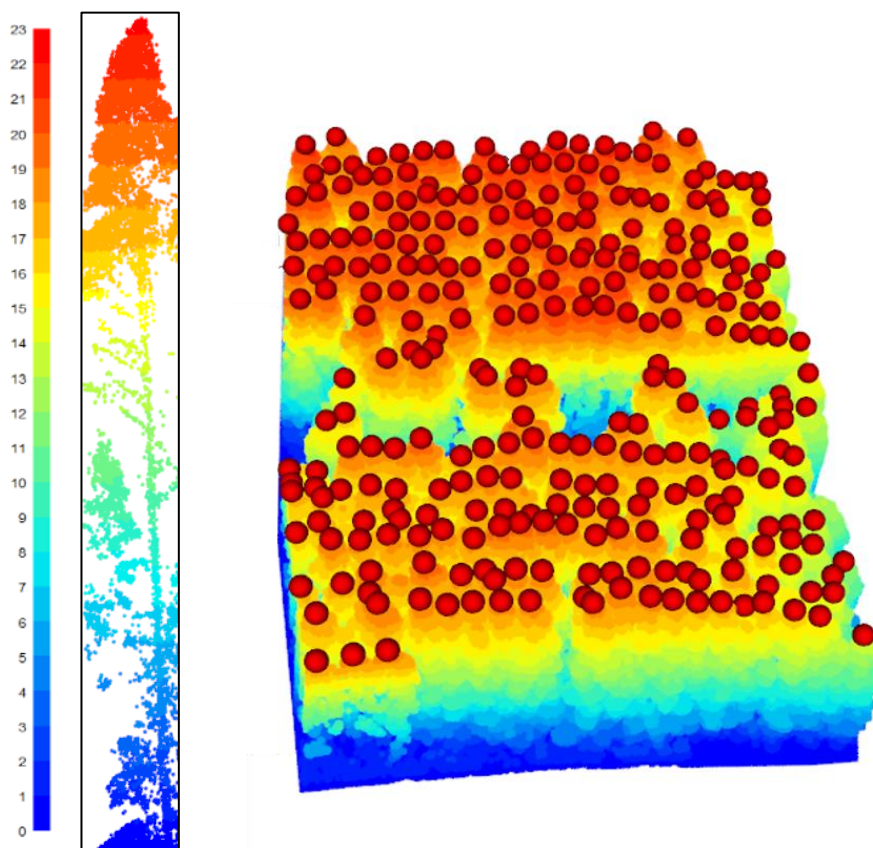
Fonte: Elaborado pelo autor (2023)

Com base na nuvem normalizada, o algoritmo Imf com janelas de 2 metros obteve um Erro Relativo (ER) de 11.9% e um Erro absoluto (EA) de 38 árvores enquanto que com janelas de 3 metros o ER foi de 32.9 % e o EA de 105 árvores, a Figura 16 ilustra o resultado obtido para o Imf com melhor performance. A Figura 17 demonstra o processo de validação com base nas árvores georreferenciadas. A densidade de pontos de uma nuvem está dentre os fatores que podem afetar a identificação dos pontos de máxima, tratando-se de UAV-LS uma elevada amostragem de pontos por m² espera ser obtido, algo que pode ser afetado em função dos parâmetros de voo na aquisição dos dados (Corte *et al.*, 2020).

Autores ao avaliarem janelas de seleção com maior range em outras condições para detecção de pontos de máxima em árvores, em cenários com maior espaçamento entre árvores, com janelas de 7 x 7 metros, obtiveram uma acurácia de 80% na identificação de árvores individuais (Mohan *et al.*, 2021). Em florestas de eucalipto, a detecção de árvores para janelas de 3 x 3 metros obteve uma performance que varia de 59,3 % a 96,7% enquanto que para janelas de 5 x 5 metros de 49,2% a

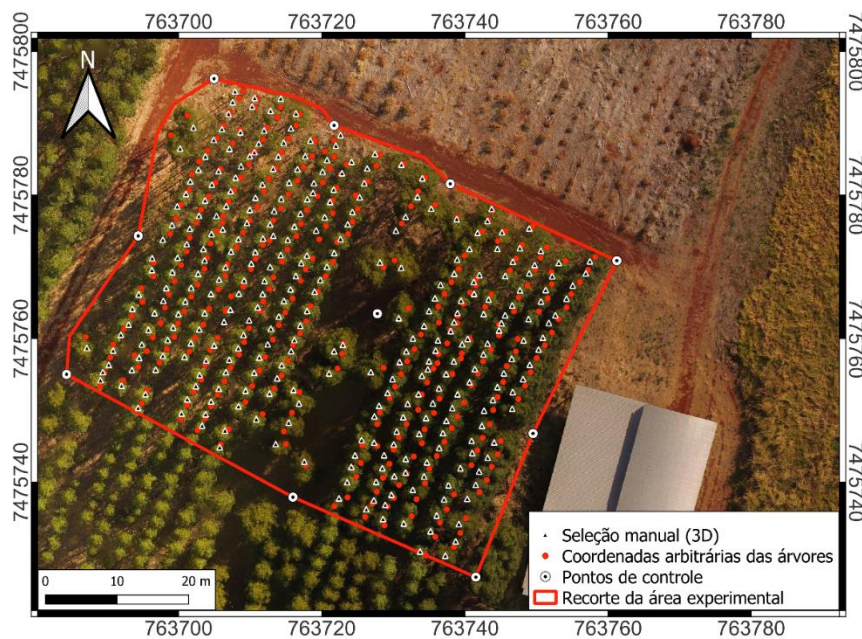
95,9% de identificação, com maiores erros de omissão para janelas mais espaçadas, chegando a 45% (Hernández *et al.*, 2018).

Figura 16 - Resultados da detecção dos pontos de máxima com janelas de 2m para a área de estudo



Fonte: Elaborado pelo autor (2023)

Figura 17 - Validação das árvores reconhecidas pelo “Imf” 2 metros em função da realidade de campo



Fonte: Elaborado pelo autor (2023)

Das 319 árvores segmentadas por Silva (2016) (Figura 18), foram extraídas 36 métricas LiDAR de altura. A seleção desse algoritmo classificador, baseia-se no fato de possuir menor tempo de processamento e uma quantidade menor de inputs (Silva *et al.* 2016).

Figura 19 – Árvores segmentadas extração das métricas LiDAR

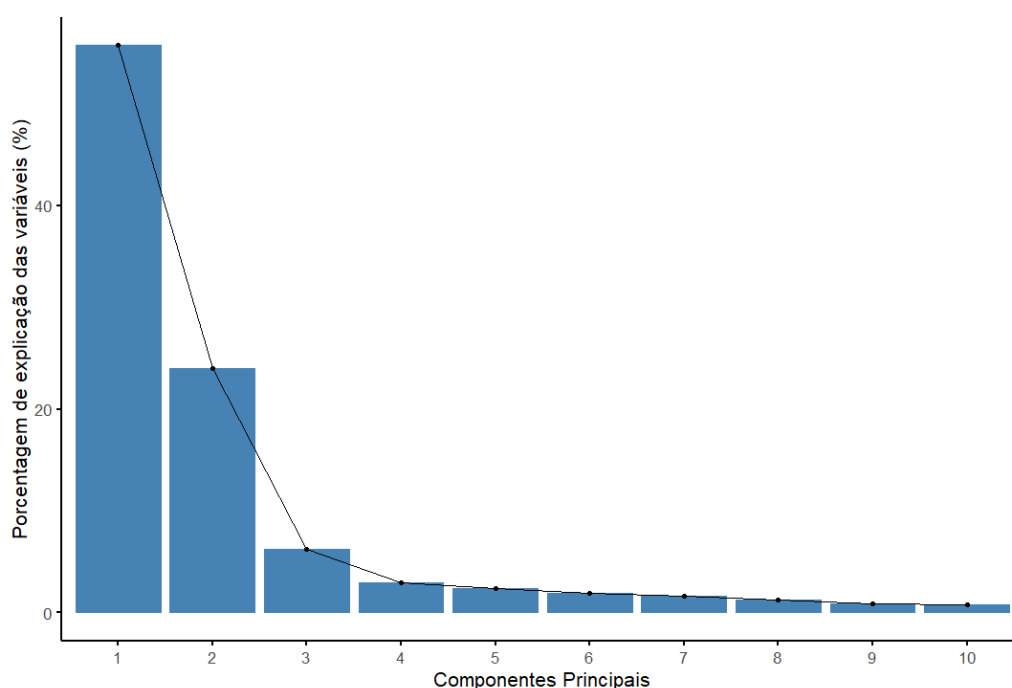


Fonte: Elaborado pelo autor (2023)

4.2 Modelagem preditiva por aprendizado de máquinas

Avaliar o comportamento entre as variáveis é um processo crucial na modelagem de dados (MURPHY, 2012), com isso as métricas LiDAR foram submetidas a análise dos componentes principais, em um cenário em que as componentes principais 1 e 2 juntas explicam 79,7 % das variáveis (Figura 19). De maneira geral, busca-se que as componentes expliquem pelo menos 80% da variação total, onde quanto maior o valor percentualmente melhor o ajuste matemático dos dados (Jhonson, Wichern, 2007).

Figura 19 – Explicação das variáveis com base nas componentes principais geradas no processo da análise de componentes principais

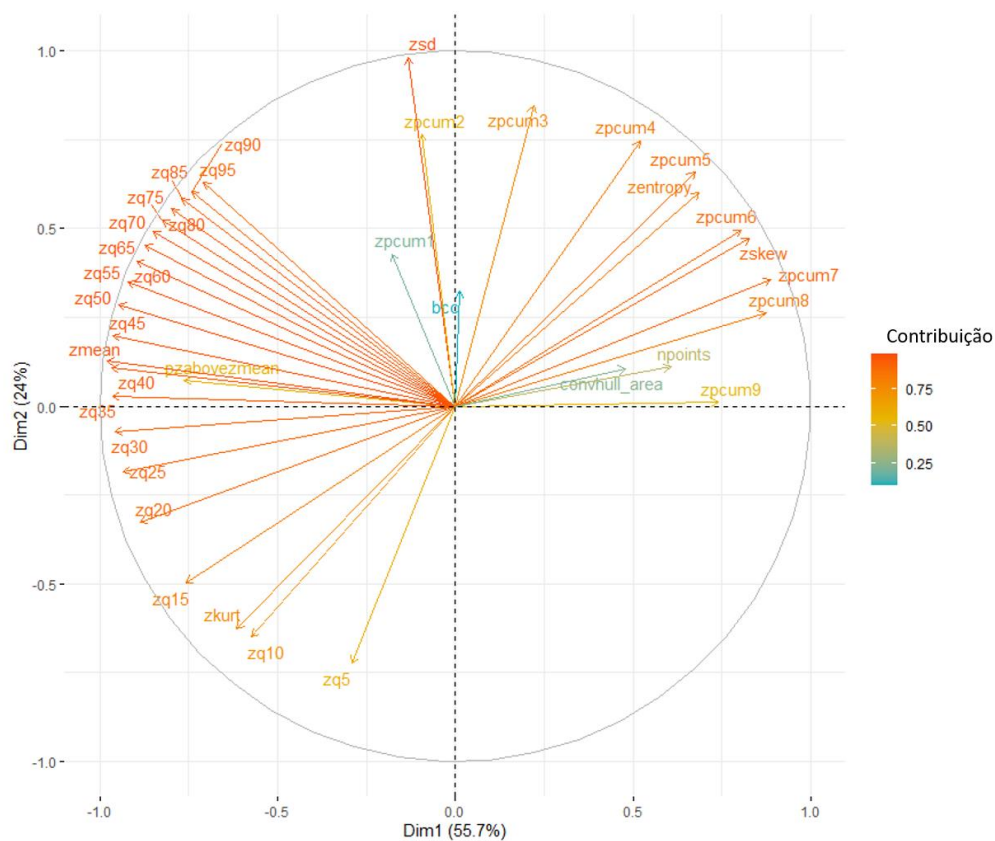


Fonte: Elaborado pelo autor (2023)

A importância de cada variável dentro da componente principal é visualizada no mapa fatorial (Figura 20). Ao explorar a correlação entre as variáveis é possível notar que a b_{cc} possui vetor com maior dimensão para a CP2, porém com uma baixa contribuição (0,25), vale destacar que, é a componente com menor percentagem de explicação do conjunto de dados. Grande parte das métricas LiDAR possuem uma correlação negativa com PC1. Destaque para as métricas $zpcum1$, $zpcum2$, zsd e

zcum3 que fatorialmente expressaram o mesmo comportamento que b_{cc} com diferentes níveis de contribuição.

Figura 20 - Mapa fatorial representa a distribuição das métricas LiDAR nas componentes principais CP1 e CP2



Fonte: Elaborado pelo autor (2023)

Sugere-se esse tipo de análise para seleção das métricas com mesmo comportamento fatorial que b_{cc} . Conforme realizado por Schuh (2019), que ao avaliar as métricas LiDAR-ALS as selecionou com base na componente com maior interação com a biomassa florestal. Para treinar os modelos todas as variáveis foram incluídas no processo de treinamento dos modelos, visto que, a variável de interesse expressou pouca interação com a componentes principais. Prospecta-se a avaliação da correlação existente entre as variáveis em um conjunto com maior número de instâncias, em um cenário em que seja possível validar os modelos em futuros trabalhos.

Dentre os modelos avaliados e sumarizados na Tabela 5, os modelos que obtiveram maior destaque em performance foram os baseados em arquitetura de árvores de decisão. Porém uma hipótese seria o super ajuste desses modelos em função do baixo número de instâncias. Outros modelos como XGBRegressor, HistGradientBoostingRegressor, GradientBoostingRegressor, LGBMRegressor também expressaram elevado R^2 aj. e menores valores de erro.

Tabela 5 - Ranqueamento dos modelos treinados com a biblioteca Lazy Predict Regressor

Model	R^2_{aj}	RMSE
DecisionTreeRegressor	1	0
ExtraTreeRegressor	1	0
ExtraTreesRegressor	1	0
XGBRegressor	1	0,02
HistGradientBoostingRegressor	0,92	10,71
GradientBoostingRegressor	0,92	11,06
LGBMRegressor	0,92	11,24
RandomForestRegressor	0,88	13,48
BaggingRegressor	0,83	16,15
AdaBoostRegressor	0,57	25,43
KNeighborsRegressor	0,4	29,97
TransformedTargetRegressor	0,31	32,05
LinearRegression	0,31	32,05
PoissonRegressor	0,31	32,15
Ridge	0,31	32,23
RidgeCV	0,29	32,66
HuberRegressor	0,28	32,74
LassoLarsIC	0,28	32,86
SGDRegressor	0,28	32,95
BayesianRidge	0,27	33,00
Lasso	0,27	33,07

LassoCV	0,27	33,09
LassoLarsCV	0,27	33,13
ElasticNetCV	0,26	33,22
LarsCV	0,26	33,32
ElasticNet	0,25	33,51
OrthogonalMatchingPursuitCV	0,25	33,56
GammaRegressor	0,24	33,75
TweedieRegressor	0,24	33,86
OrthogonalMatchingPursuit	0,22	34,13
LinearSVR	0,12	36,24
SVR	0,05	37,8
NuSVR	0,02	38,31

Fonte: Elaborado pelo autor (2023)

O Random Forest (RF) é de destaque dentre os modelos avaliados, além de obter bons valores de performance com um R^2_{aj} de 0,88 e RMSE de 13,48, diversos trabalhos de mensuração florestal modelam biomassa com informações LiDAR por RF (Corte *et al.*, 2020; Silva *et al.*, 2020). Ao avaliar a modelagem a biomassa com RF com base em dados LiDAR aerotransportado, foram obtidos R^2_{aj} que variam de 0,6 a 0,9, com RMSE entre 5 e 15%, para a estimativa de volume e biomassa florestal (Silva *et al.*, 2017, Corte *et al.*, 2022, Schuh, 2019).

5 CONCLUSÕES

Foi possível obter modelos com elevada performance preditiva a nível de árvore individual em uma população de treinamento, com destaque para os modelos baseados em árvores de decisão. No processo de individualização das árvores, a janela de seleção dos pontos de máxima de 2 metros foi a com melhor desempenho. Isso possibilitou a predição da biomassa seca de fuste por meio do ajuste de modelos de Machine Learning baseados nas métricas lidar, com destaque para os baseados em arquitetura de árvores de decisão.

Porém sugere-se que o processo de modelagem passou por um super ajuste, em que as hipóteses que o explicam são: o baixo número de instâncias. Prospecta-se que em trabalhos futuros, com o aprimoramento das técnicas de individualização de árvores e localização de pontos de máxima, será possível um maior número de estâncias para que seja possível a validação dos modelos preditivos da biomassa seca de fuste a nível de árvore individual.

REFERÊNCIAS

- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M. Köppen's climate classification map for Brazil. **Meteorol. Z.**, v. 22, p. 711-728, 2013.
- ASSMANN, E. **The principles of forest yield study**. Oxford: Pergaman Press, 1970. 506 p, 1970.
- AZEVEDO, G. B. *et al.* Multi-volume modeling of Eucalyptus trees using regression and artificial neural networks. **Plos One**, v. 15, n. 9, e0238703, 2020.
- BADIN, T. L. **Estimativa de biomassa utilizando dados lidar em floresta tropical**. Dissertação (Mestrado). Universidade Federal de Santa Maria. Rio Grande do Sul, p.71, 2018.
- BATISTA, J. L. F.; COUTO, H. T. Z.; MARQUESINO, M. Desempenho de modelos de relações hipsométricas: estudo em três tipos de Floresta. **Scientia Forestalis**, n. 60, p. 149-163, 2001.
- BATISTA, J. L. F.; COUTO, H. T. Z. do; SILVA FILHO, D. F. da. **Quantificação de recursos florestais: árvores, arvoredos e florestas**. São Paulo: Oficina de Textos, 2014. p. 361-366, 2014.
- BATISTA, T. S. **Redes Neurais Artificiais para quantificar o volume de madeira em espécie de Eucalipto, com ampliação teórica da base e dados**. Dissertação (Mestrado) - Universidade Federal do Mato Grosso do Sul (UFMS), 22 p, 2020.
- BELAND, M. *et al.* On promoting the use of lidar systems in forest ecosystem research. For. **Ecol. Manag.**, v. 450, n. 117484, 2019.
- BUENO, G.; COSTA, E. Comparação preditiva de modelos hipsométricos em plantio de Eucalipto: equação de regressão e redes neurais artificiais. **Enciclopédia Biosfera**, v. 17, n. 32, 2020.
- BURKHART, H. E.; TOMÉ, M. **Modeling Forest Trees and Stands**. Springer: Netherlands, 458 p. 2012.
- CAMPOS, J. C. C.; LEITE, H. G. **Mensuração Florestal: perguntas e respostas**. 5.ed. atualizada e ampliada. Viçosa: Editora UFV, 636 p. 2017.
- CHASMER, L.; HOPKINSON, C.; TREITZ, P. Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial LiDAR. Canadian Journal of **Remote Sensing**, v. 32, n. 2, p. 116- 125, 2006.
- CHRISTAKOS, G. **Geostatistical Analysis of Continuous Data: Prediction and Simulation**. SPRINGER, 2012.

CHRISTIAN, W.; LINDENBERGER, J., Experiences of 10 years Laser Scanning. **Photogrammetric Week 99**, p. 125-132, 1999.

CLUTTER, J.C.; FORSTON, J.C.; PIENNAR, L.V.; BRISTER, G.H.; BAILEY, R.L. **Timber management: a quantitative approach**. 3. ed. New York: John Willey & Sons, p. 333, 1983.

CORTE, A. P. D. *et al.* Forest inventory with high-density UAV-Lidar: Machine learning approaches for predicting individual tree attributes. **Comput. Electron. Agric.**, v. 179, p. 10-14, 2020a.

CORTE, A. P. D. *et al.* Measuring Individual Tree Diameter and Height Using GatorEye High-Density UAV-Lidar in an Integrated Crop-Livestock-Forest System. **Remote Sensing**, v. 12, p. 1-15, 2020b.

CORTE, A. P. D. *et al.* **Aplicações do LiDAR para o inventário florestal – enfoque unidade de área**. Atena Editora., v. 1, 118 p, 2022.

COSTA-FILHO, S. V. S. *et al.* Configuração de algoritmos de aprendizado de máquina na modelagem florestal: um estudo de caso na modelagem da relação hipsométrica. **Ci. Fl.**, Santa Maria, v. 29, n. 4, p. 1501-1515, 2019.

CUNHA-NETO, E. M. Using high-density UAV-Lidar for deriving tree height of Araucaria Angustifolia in an Urban Atlantic Rain Forest. **Urban Forestry and Urban Greening**, v. 63, p. 127-197, 2021.

DA CUNHA NETO, E. *et al.* Aprendizado de máquina e regressão linear na estimativa do volume de Eucalyptus na Amazônia Oriental. **BIOFIX Scientific Journal**, v. 6, n. 1, p. 08-14, 2021.

DAI, X. **Novel Methods for Estimating Above Ground Biomass**. Tese de (Doutorado). University of New Brunswick, 136 p, 2021.

DALPONTE, M.; COOMES, D. A. Tree-Centric Mapping of Forest Carbon Density from Airborne Laser Scanning and Hyperspectral Data. **Methods Ecol. Evol.**, v. 7, p. 1236-1245, 2016.

DASH, J. P.; PEARSE, G. D.; WATT, M. S. UAV Multispectral Imagery Can Complement Satellite Data for Monitoring Forest Health. **Remote Sens.**, v. 10, p. 12-16, 2018.

DEEPA, J. *et al.* Computational Drug Discovery Using QSAR Modeling. In: Recent Trends in Computational Intelligence and Its Application. CRC Press, 2023. p. 438-445.

DOMINGOS, P. A few useful things to know about machine learning. **Communications of the ACM**, v. 55, n. 10, p. 78-87, 2012.

ELLI, E. F.; SENTELHAS, P. C.; BENDER, F. D. Impacts and uncertainties of climate change projections on Eucalyptus plantations productivity across Brazil. **For. Ecol. Manage.**, v. 474, n. 118365, 2020.

EMBRAPA. Embrapa Solos: relatório anual 2022. 2023. 3 ed. **Brasília: Embrapa Solos**, 353 p, 2022.

EUFRATE JUNIOR, H. de. J. **Caracterização físico-química da biomassa produzida em sistemas florestais de curta rotação para geração de energia**. Tese (Mestrado em energia na agricultura), Faculdade de Ciências Agrônômica, UNESP, Botucatu-SP. 135 p, 135 p. 2015.

FAO - FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **Global Forest Resources Assessment Update 2004**. Terms and definitions (Final version). FAO Forestry Department. Forest Resources Assessment Programme (FRA). Working Paper 83/E, Rome, 2004.

FAO – Organização das Nações Unidas para Agricultura e Alimentação. Global Forest **Resources Assessment**. 2020. Disponível em: <<https://fra-platform.herokuapp.com/>>. Acesso em: 24. Mar. 2021.

GAO, A. K. *et al.* Machine-Learning-Based Virtual Screening and Ligand Docking Identify Potent HIV-1 Protease Inhibitors. **Artificial Intelligence Chemistry**, p. 100014, 2023.

GIONGO, M. *et al.* LiDAR: princípios e aplicações florestais. **Pesquisa Florestal Brasileira**, Colombo, v. 30, n. 63, p. 231-244, 2010.

GLEASON, C. J.; IM, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. **Remote Sensing of Environment**, 2012.v. 125, p. 80-91, 2012.

GOOVAERTS, P. **Geostatistical for Natural Resources Evaluation**. Oxford University Press, 1997.

GORGENS, E. B.; SILVA, A. G. P.; RODRIGUEZ, L. C. E. **Lidar**: aplicações florestais. Curitiba: CRV. 2014. 132 p.

HAO, Z. *et al.* Automated tree-crown and height detection in a young forest plantation using mask region-based convolutional neural network (Mask R-CNN). **ISPRS J. Photogramm. Remote Sens.**, v.178, p.112-123, 2021.

HE, J. *et al.* In-field Cotton Yield and Loan Rate Prediction using Machine Learning Regression Algorithms. **Beltwide Cotton Conferences**, San Antonio, TX, January 4-6, 2022.

HOUSMAN, I. W.; CHASTAIN, R. A.; FINCO, M. V. An Evaluation of Forest Health Insect and Disease Survey Data and Satellite-Based Remote Sensing Forest

Change Detection Methods: Case Studies in the United States. **Remote Sens.**, v. 10, n. 8, p. 1184, 2018.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção da Extração Vegetal e da Silvicultura (PEVS)**. IBGE:Rio de Janeiro, v. 35, p.1-8, 2020.

INDÚSTRIA BRASILEIRA DE ÁRVORES (IBÁ). **Relatório 2020**. 66 p. 2020.
Disponível em: <https://iba.org/datafiles/publicacoes/relatorios/relatorio-iba-2020.pdf>.
Acesso em: 10 fev. 2022

JAAKKOLA, *et al.* A low-cost multisensoral mobile mapping system and its feasibility for tree measurements, ISPRS J. Photogramm. **Remote Sens.**, v. 65, n. 6, p. 514-522, 2010.

JAKUBOWSKI, M. K. *et al.* Delineating individual trees from LiDAR data: A comparison of vector-and raster-based segmentation approaches. **Remote Sensing**, v. 5, n. 9, p. 4163-4186, 2013.

JORGE, L. B.; EUFRATE-JUNIOR, H. de J. **Dendrometria e inventário florestal com aplicações em R**. Editora Cubo, p.179 v.1, 2023.

JHONSON, R. A.; WICHERN, D. W. **Applied Multivariate Analysis**. 6. Ed. Upper Sanddle River: Pearson Education, Inc., 773 p, 2007.

KELLNER, J. R. *et al.* New opportunities for forest remote sensing through ultra-high-density drone lidar. **Surveys in Geophysics**, v. 40, n. 4, p. 959-977, 2019.

LARY, D. J. *et al.* Machine learning in geosciences and remote sensing. **Geoscience Frontiers**, v. 7, n. 1, p. 3-10, 2016.

LEE, J.; IM, J.; KIM, K.; Quackenbush, L.J. Machine Learning Approaches for Estimating Forest Stand Height Using Plot-Based Observations and Airborne LiDAR data. **Forests**, v. 9, 268, 2018.

LEITOLD, V. *eronika et al.* Airborne lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. **Carbon balance and management**, 2015 v. 10, n. 1, p. 1-12, 2015.

LIANG, L. *et al.* NDVI-derived forest area change and its driving factors in China. **Plos One**, v. 13, n. 10, e0205885, 2018.

LIAO, K. *et al.* Examining the role of uav lidar data in improving tree volume calculation accuracy. **Remote Sensing**, v. 14, n. 17, p. 4410, 2022.

LIAW, A.; WIENER, M. Classification and regression by randomForest. R news, 18-22, 2002.

LOETSCH, F.; ZÖHRER, F.; HALLER, K.E. **Forest Inventory**. München: BLV VerlagsgesellschaftmbH, 469p, 1973.

LUTHER, J. E. *et al.* Predicting wood quantity and quality attributes of balsam fir and black spruce using airborne laser scanner data, **Forestry**, v. 87, n. 2, p. 313-326, 2014.

MALTAMO, M. *et al.* Airborne laser scanning for tree diameter distribution modelling: a comparison of different modelling alternatives in a tropical single-species plantation. **Forestry**, v. 91, 121-131, 2018.

MARRS, J.; NI-MEISTER, W. Machine Learning Techniques for Tree Species Classification Using Co-Registered LiDAR and Hyperspectral Data. **Remote Sens.**, v. 11, 819, 2019.

MELO, A. M. *et al.* Monitoring the understory in eucalyptus plantations using airborne laser scanning. **Scientia Agrícola**, v. 78, n. 1, e20190134, 2021.

MOE, K. T. *et al.* Application of UAV photogrammetry with Lidar data to facilitate the estimation of tree locations and Dbh values for high-value timber species in Northern Japanese mixed-wood forests. **Remote Sensing**, v. 12, n. 17, p. 2865, 2020.

MONTAÑO, R. N. R. A. **Aplicação de técnicas de aprendizado de máquina na mensuração florestal**. Dissertação (Doutorado), Universidade Federal do Paraná (UFPR), 86p. 2016.

MÜLLER, F.; JAEGER, D.; HANEWINKEL, M. Digitization in wood supply – A review on how Industry 4.0 will change the forest value chain. **Comput. Electron. Agric.**, v. 162, p. 206-218, 2019.

MURPHY, K. P. Machine learning: a probabilistic perspective. **MIT press**, 2012.

NAKAI, E. S. **Quantificação da biomassa e estoque de carbono em diferentes coberturas vegetais por meio de sensoriamento remoto**. Tese (Doutorado), Universidade de São Paulo, 154 p, 2016.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society: Series A (General)**, v. 135, n. 3, p. 370-384, 1972.

NEUVILLE, R.; BATES, J. S.; JONARD, F. Estimating forest structure from UAV-mounted LiDAR point cloud using machine learning. **Remote sensing**, v. 13, n. 3, p. 352, 2021.

ODUM, E. P. **Ecologia**. Rio de Janeiro: Guanabara Koogan. 1986. 434 p. 1986.

PERSSON, A.; HOLMGREN, J.; SODERMAN, U. Detecting and measuring individual trees using an airborne laser scanner. **Photogrammetric Engineering and Remote Sensing**, v. 68, n. 9, p. 925-932, 2002.

PIRES, J. E.; AI, M. P. **Machine Learning Aplicado a Imagens da Observação da Terra para o Estudo da Floresta**. Dissertação de Mestrado. Universidade Nova de Lisboa. 112p. 2019.

PITKÄNEN, J. *et al.* Adaptive methods for individual tree detection on airborne laser based canopy height model. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 36, p.187–191, 2004.

POURSHAMSI, M. *et al.* Tropical forest canopy height estimation from combined polarimetric SAR and LiDAR using machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 172, p. 79-94, 2021.

PRETZSCH, H. *Forest dynamics, growth and yield*. New York: **Springer**, 2010. 664p.

PULITI, N.; BREIDENBACH, J.; ASTRUP, R. Estimation of Forest Growing Stock Volume with UAV Laser Scanning Data: Can It Be Done without Field Data? *Remote Sens.*, v. 12, 1245, 2020.

R CORE TEAM. R: **A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna, 2020. Disponível em: <<https://www.Rproject.org>>.

REX, F. E. *et al.* Estimating Above-Ground Biomass of *Araucaria angustifolia* (Bertol.) Kuntze Using LiDAR Data. *Floresta e Ambiente*, v. 26, n. 4, 2019.

RODRIGUES, P. F. *et al.* UAV-based LiDAR scanning for individual tree detection and height measurement in young forest permanent trials. *Remote Sensing*, v. 14, n. 1, p. 170, 2021.

ROUSSEL, J. R. *et al.* LidR: An R Package for Analysis of Airborne Laser Scanning (ALS) Data. *Remote Sens. Environ.*, v. 251, n. 112061, 2020.

RUBILAR, R. A. *et al.* Advances in silviculture of intensively managed plantations. *Curr. For. Rep.*, v. 4, p. 23-34, 2018.

RUDGE, M. L. M. *et al.* Modelling the Diameter Distribution of Savanna Trees with Drone-Based LiDAR. *Remote Sens.*, v. 13, 2021.

SAHAL, R. *et al.* Industry 4.0 towards Forestry 4.0: Fire Detection Use Case. *Sensors*, v. 21, 694, 2021.

SANDINO, J. *et al.* Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors*, v. 18, n. 944, 2018.

SANQUETTA, C. R. Métodos de determinação de biomassa florestal. In: SANQUETTA, C. R. *et al.* (Eds.). **As florestas e o carbono**. Curitiba: [s.n.], p. 119-140, 2002.

SCHUH, M. S. **Modelos preditivos de biomassa em Floresta Amazônica a partir de dados lidar**. Dissertação de Mestrado. Universidade Federal de Santa Maria, RS, 82p, 2019.

SCHUMACHER, F. X.; HALL, F. S. Logarithmic expression of timber-tree volume. *J. Agric. Res.*, v. 47, n. 9, p. 719–734, 1993.

SILVA C. A. *et al.* Efeito da densidade de pontos LiDAR na predição da altura em plantações de Pinus taeda L. *Scientia Forestalis*. Piracicaba. V. 45, n. 115, p. 481-492, 2017.

SILVA, A. B., SANTOS, C. D.; OLIVEIRA, E. F. Estimação da produção de biogás na codigestão anaeróbia de resíduos orgânicos da indústria sucroalcooleira utilizando algoritmos de aprendizado de máquina. **Revista de Energia Renovável**, v. 10, n. 2, p. 45-62, 2022.

SILVA, C. A. *et al.* Imputation of individual longleaf pine (*Pinus Palustris* Mill.) tree attributes from field and LiDAR data. *Can. J. Remote Sens.*, v. 42, p. 554-573, 2016.

SKOVSGAARD, J. P.; VANCLAY, J. K. Forest site productivity: a review of the evolution of dendrometric concepts for even aged stands. **Forestry**, v. 81, p. 13-31, 2008.

SOARES, C. P. B.; NETO, F. P.; SOUZA, A. L. **Dendrometria e inventário florestal**. 2. ed. Viçosa MG: Editora UFV, 272 p. 2012.

SOARES, P.; TOMÉ, M. Height–diameter equation for first rotation eucalypt plantations in Portugal. **For. Ecol. Manag.**, v. 166, p.99-109, 2002.

SOUSA, C. L.; PONZONI, F. J. Avaliação de índices de vegetação e de bandas TM/Landsat para estimativa de volume de madeira em floresta implantada de Pinus spp. **Simpósio Brasileiro de Sensoriamento Remoto**, v. 9, p. 1725-1732, 1998.

SOUZA, A. L.; SOARES, C. P. B. **Florestas nativas: estrutura, dinâmica e manejo**. Viçosa: UFV, 322 p, 2013.

TODA, M. *et al.* More asymmetric tree competition brings about more evapotranspiration and less runoff from the forest ecosystems: a simulation study. **Ecol. Model.**, v. 221, p. 2887-2898, 2010.

TOMÉ, M.; BURKHART, H. E. Distance-dependent competition measures for predicting growth of individual trees. **Forest Science**, v. 35, n. 3, p. 816–831, 1989.

TOMÉ, M.; VERWIJST, T. Modelling competition in short rotation forests. **Biomass and Bioenergy**, v. 11, p.177-187, 1996.

VASCONCELLOS, B. N. V. **Estimativa de variáveis dendrométricas individuais de Araucaria angustifolia por meio de dados Laser Scanner aéreo e Terrestre**. 2020. 183 f. Tese (Doutorado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba.

VENDRUSCOLO, D. G. S. *et al.* Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v. 33, n. 4, p. 556-569, 2015.

VOSSSELMAN, G.; MAAS, H. G. **Airborne and Terrestrial Laser Scanning**. Boca Raton: CRC Press, 2010.

WEISKITTEL, A. R. *et al.* Forest growth and yield modeling. **New York: John Wiley and Sons**, 2011. 415p, 2011.

XU, D. *et al.* LiDAR Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. **Forests**, v. 12, 550, 2021.

YADAV, S. Ground and Non-Ground Filtering for Airborne LIDAR Data. **International Journal of Advanced Remote Sensing and GIS**, v. 5, n. 1, p. 1500-1506, 2016.

YU, H. *et al.* An Enhanced Clustering-Based Filtering Algorithm for Airborne LiDAR Data. **Remote Sensing**, v. 9, n. 3, p. 283, 2017.

YU, R. *et al.* A machine learning algorithm to detect pine wilt disease using UAV-based hyperspectral imagery and LiDAR data at the tree level. **International Journal of Applied Earth Observation and Geoinformation**, v. 101, n.102363, 2021.

ZHANG, K. *et al.* A progressive morphological filter for removing nonground measurements from airborne LIDAR data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 41, 2003.

ZHANG, W. *et al.* An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. **Remote Sens.**, v. 8, no. 6, p. 501, 2016.

ZHOU, J. *et al.* Evaluation of Different algorithms for estimating the growing stock volume of *Pinus massoniana* Plantations using spectral and spatial information from a SPOT6 image. **Forests**, v. 11, n. 540, 2020.

ZOU, W. *et al.* A survey of big data analytics for smart forestry. **IEEE Access**, v. 7, 2019.