

BILL HERBERT ZIEGELMAIER NETO

**Classificação dos estágios de sucessão da Floresta Atlântica utilizando
dados multiplataforma orbitais e ALS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Florestal, do Centro de Ciências Agroveterinárias da Universidade do Estado de Santa Catarina, como requisito parcial para a obtenção do grau de Mestre em Engenharia Florestal.

Orientador: Marcos Benedito Schimalski

Coorientador: Veraldo Liesenberg

**LAGES
2022**

**Ficha catalográfica elaborada pelo programa de geração automática da
Biblioteca Setorial do CAV/UEDESC,
com os dados fornecidos pelo(a) autor(a)**

Ziegelmaier Neto, Bill Herbert

Classificação dos estágios de sucessão da Floresta Atlântica
utilizando dados multiplataforma orbitais e ALS / Bill Herbert
Ziegelmaier Neto. -- 2022.

93 p.

Orientador: Marcos Benedito Schimalski

Coorientador: Veraldo Liesenberg

Dissertação (mestrado) -- Universidade do Estado de Santa
Catarina, Centro de Ciências Agroveterinárias, Programa de
Pós-Graduação em Engenharia Florestal, Lages, 2022.

1. Floresta Atlântica. 2. Aprendizado de máquina. 3.
Transferência de aprendizado. 4. Landsat-8. 5. CBERS-4A. I.
Schimalski, Marcos Benedito. II. Liesenberg, Veraldo. III.
Universidade do Estado de Santa Catarina, Centro de Ciências
Agroveterinárias, Programa de Pós-Graduação em Engenharia
Florestal. IV. Título.

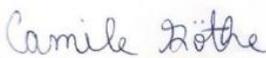
BILL HEREBERT ZIEGELMAIER NETO

**Classificação dos estágios de sucessão da Floresta Atlântica
utilizando dados multiplataforma orbitais e ALS.**

Dissertação apresentada ao curso de Mestrado em Engenharia Florestal do Centro de Ciências Agroveterinárias da Universidade do Estado de Santa Catarina, como requisito parcial para a obtenção do grau de mestre em Engenharia Florestal.

Banca examinadora (Realizado de forma presencial/remota)

Orientador: 
Prof. Dr. Marcos Benedito Schimalski
Universidade do Estado de Santa Catarina – UDESC

Membro: 
Dr^a. Camile Sothe
Universidade McMaster
Hamilton, ON, Canada

Membro: 
Prof. Dr. Leonardo Biffi
Universidade do Estado de Santa Catarina – UDESC

Lages, 28 de junho de 2022.

Dedico a minha família e ao meu amor por todo o apoio
nessa jornada.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus pela vida e saúde .

A minha família, Rolf, Izabete e Bruna por todo o apoio, e suporte necessário para chegar até aqui. Por todos os ensinamentos e conselhos sobre a vida, tudo que sou hoje é graças a vocês.

A Marina, minha companheira de vida, por toda ajuda nessa etapa e em todas as outras que já passaram e pelas que virão. Você é meu suporte, sem sua presença nada disso seria possível. A Rupi e Bebel por serem minhas companheiras, principalmente nos períodos isolados, e alegrarem todos os meus dias.

A todos meus familiares que mesmo estando longe sempre estão torcendo por mim. Aos meus amigos que estão presentes no dia-a-dia, mas também aos que estão longe, todos sempre disponíveis quando necessário, e ajudando em tudo que seja do alcance de vocês.

Ao meu orientador Marcos Schimalski, por todos os ensinamentos nessa jornada, por me proporcionar essa oportunidade, todo o crescimento pessoal e profissional que adquiri nesse período sob sua orientação foi imenso. Ao professor Veraldo meu co-orientador por toda a ajuda, sempre disponível para tirar dúvidas e compartilhando seu conhecimento. Também agradeço a todos os professores do programa de pós-graduação da Engenharia Florestal que contribuíram para o desenvolvimento desse trabalho.

A empresa Klabin por fornecer os dados e recursos para a realização dessa pesquisa.

RESUMO

Os levantamentos de informações sobre a estrutura da vegetação de forma convencional demandam alto custos financeiros e tempo de execução devido à grande heterogeneidade e quantidade de indivíduos arbóreos em florestas tropicais. Assim, faz-se necessário a utilização de tecnologias para a obtenção de informações dessas florestas, como as ferramentas de sensoriamento remoto que não necessitam de visitas a campo combinadas com métodos de *machine learning* (ML), possibilitando extrair o potencial máximo dos dados remotamente situados, para o auxílio do monitoramento da diversidade florestal. Além disso é possível empregar a transferência de aprendizagem (TA) que utiliza um modelo já treinado para a classificação de outra área, reduzindo o tempo necessário para ajustar o modelo. Contudo, para isso, é necessária criar uma abordagem que vise obter uma classificação com resultados satisfatórios utilizando dados de imagens orbitais e de *Light Detection and Ranging* (LiDAR). Esse estudo tem como objetivo avaliar a precisão da classificação dos estágios de sucessão da vegetação a partir de técnicas de ML e TA. Para isso, foram utilizados dados multiplataforma de sensoriamento remoto, sendo esses as imagens dos satélites orbitais CBERS-4A, Sentinel-2, Landsat-8 e Planet em conjunto com dados LiDAR. As áreas de estudos encontram-se nos municípios de Bocaina do Sul, Bom Retiro, Paineira, Santa Cecília e Urupema - Santa Catarina, todas pertencentes ao Bioma Mata Atlântica. Foram derivadas as informações para a criação dos *datasets* utilizando as imagens multiplataforma orbitais e os dados do Lidar. Então, foram treinados os algoritmos *Random Trees* (RT), *Support Vector Machine* (SVM) e *Maximum Likelihood Classifier* (MLC) para a classificação de uma das áreas de estudo. As demais áreas foram classificadas usando TA e classificação das imagens. A avaliação de acurácia dos resultados foi realizada gerando matrizes de confusão, com as quais foi possível computar o índice Kappa, desvio padrão e intervalo de confiança com 5% de significância. Como resultados, obtiveram-se Kappas variando de 0,49 a 0,95 para a classificação normal e 0 a 0,70 para a transferência de aprendizagem. Os dados derivados do sensor orbital Landsat-8 juntamente com dados LiDAR apresentaram os melhores resultados para o treinamento dos modelos, já para a TA os dados derivados do Sentinel-2 e LiDAR obtiveram melhor acurácia. Para treinar os modelos, o classificador SVM se apresentou melhor, porém quando aplicada a TA, nenhum dos classificador se destacou dos demais. Nosso estudo mostrou que é possível a utilização do ML e TA para a classificação dos estágios de sucessão da vegetação, auxiliando na obtenção de informações sobre a estrutura florestal de forma remota.

Palavras-chave: Floresta Atlântica, Landsat-8, CBERS-4, Planet, Sentinel-2, Aprendizado de Máquina, Transferência de aprendizagem.

ABSTRACT

The conventional way of collecting information about the vegetation structure demands high cost and execution time due to the great heterogeneity and number of individual trees in tropical forests. Thus, it is necessary to use technologies to obtain information from these forests, such as remote sensing tools that do not require field visits combined with machine learning (ML), making it possible to extract the maximum potential from remotely located data. In addition, it is possible to use transfer learning (TL) that uses a model already trained to classify another area, reducing the time needed to adjust the model. However, for this, it is necessary to create an approach that aims to obtain a classification with satisfactory results using data from orbital images and from Light Detection and Ranging (LiDAR). The study areas are in the cities of Bocaina do Sul, Bom Retiro, Paineira, Santa Cecília and Urupema - Santa Catarina. Information for creating datasets was derived using multiplatform orbital images and Lidar data. Then, the Random Trees (RT), Support Vector Machine (SVM) and Maximum Likelihood Classifier (MLC) algorithms were trained to classify one of the study areas, and for the others, TL and image classification were performed. The evaluation of the accuracy of the results was performed by generating confusion matrices, Kappa index, standard deviation and confidence interval with 5% significance. Results showed Kappa indexes ranging from 0 to 0.95 were obtained, achieving excellent ratings. The data derived from the Landsat-8 orbital sensor together with LiDAR presented the best results for the training of the models, as for the TL, the data derived from Sentinel-2 and LiDAR obtained better performance. The SVM classifier performed better for training the models, but when applied to TA, none of the classifiers stood out from the others. Our study showed that it is possible to use ML and TA to classify vegetation successional stages, helping to obtain information on vegetation structure.

Keywords: Atlantic Rainforest, Landsat-8, CBERS-4, Planet, Sentinel-2, machine learning, transfer learning.

LISTA DE ILUSTRAÇÕES

Figura 1. Perfil esquemático de sucessão florestal, com predomínio de pioneiras (A); intermediárias (B); e climácicas (C).	18
Figura 2. Obtenção de imagens por sensoriamento remoto.....	20
Figura 3. Modelo Digital de Terreno da área de estudo, gerado a partir dos dados LiDAR....	24
Figura 4. Modelo Digital de Superfície da área de estudo, gerado a partir dos dados LiDAR.	24
Figura 5. Canopy Height model para a área de estudo, gerado pela subtração do MDS pelo MDT.	25
Figura 6. Conjunto de árvores de decisão representando o funcionamento do algoritmo <i>Random Forest</i>	31
Figura 7. Representação da classificação SVM.....	32
Figura 8. Ilustração da transformação de dados não linearmente separáveis.....	33
Figura 9. Mapa de localização da área de estudo, sendo: (a) Localização dos municípios no estado de Santa Catarina; (b) Localização da área de estudo dentro dos municípios; (c) Delimitação do perímetro da área de estudo.	37
Figura 10. Fluxograma da metodologia de processamento de dados utilizada neste trabalho.	38
Figura 11. Canopy Height Model para a área de estudo.	46
Figura 12. Representação gráfica das classificações dos datasets e sensores: (a) Classificação referência; (b) Classificação da imagem CBERS-4A utilizando o classificador SVM e dataset 2*; (c) Classificação da imagem Sentinel-2 utilizando o classificador SVM e o dataset 3*; (d) Classificação da imagem Landsat-8 utilizando o classificador SVM e o dataset 4*;(e) Classificação da imagem Planet utilizando o classificador RT e o dataset 10*.....	50
Figura 13. Mapa de localização das área de estudo. (a) Mapa contendo estado de Santa Catarina e a delimitação dos municípios de Bom Retiro e Santa Cecília. (b) Delimitação das áreas de estudo I e II. (c) Localização das áreas de estudo I e II dentro do município de Bom Retiro. (d) Delimitação da área de estudo III. (e) Localização da área de estudo III dentro do município de Santa Cecília.....	60
Figura 14. Fluxograma da metodologia de processamento de dados utilizada neste trabalho.	61
Figura 15. Mapa de distribuição das parcelas do inventário florestal convencional para a área I e II.....	67
Figura 16. Mapa de distribuição das parcelas do inventário florestal convencional para a área III.	67
Figura 17. Mapa de distribuição das amostras de validação utilizados para validar a transferência de aprendizagem. (a) Área de estudo I. (b) Área de estudo II. (c) Área de estudo III.	70
Figura 18. Comparação entre a imagem testemunha (a) e imagem classificada utilizando a transferência de aprendizagem para a área 1, utilizando o classificador Máquina de vetores de suporte com a composição 3 da imagem Sentinel-2.	76
Figura 19. Comparação entre a imagem testemunha (a) e imagem classificada utilizando a transferência de aprendizagem para a área 2, utilizando o classificador Árvores aleatórias com a composição 8 da imagem Sentinel-2.	76
Figura 20. Comparação entre a imagem testemunha (a) e imagem classificada utilizando a transferência de aprendizagem para a área 3, utilizando o classificador de Verossimilhança com a composição 5 da imagem Sentinel-2.	77

LISTA DE TABELAS

Tabela 1. Características do recobrimento aéreo e do equipamento LiDAR utilizado.	40
Tabela 2. Características espectrais das bandas dos satélites CBERS-4A (WPM), Landsat-8/OLI, Sentinel-2/MSI e PlanetScope utilizadas.	40
Tabela 3. Fórmulas dos índices de vegetação utilizados no estudo.	41
Tabela 4. Características do Veículo Aéreo Não Tripulado.	42
Tabela 5. Conjuntos de dados e número de variáveis como dado de entrada para a classificação supervisionada.	43
Tabela 6. Classes encontradas dentro da área de pesquisa e suas respectivas áreas em hectares.	44
Tabela 7. Quantidade de pixels (1x1 m) para as amostras de treinamento e validação.	44
Tabela 8. Comparação da altura média das árvores mensuradas no inventário florestal convencional com a altura das árvores no CHM.	46
Tabela 9. Valores obtidos para o índice kappa, resultante da classificação supervisionada (RT, SVM e MLC) das imagens dos sensores CBERS-4A, Sentinel-2, Planet e Landsat-8. Os melhores resultados de cada dataset são apresentados em negrito.	47
Tabela 10. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem CBERS-4A, com a dataset 2* utilizando o classificador SVM.	48
Tabela 11. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Sentinel-2, com o dataset 3* utilizando o classificador SVM.	48
Tabela 12. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Landsat-8, com a dataset 4* utilizando o classificador SVM.	49
Tabela 13. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Planet, com a dataset 10* utilizando o classificador RT.	49
Tabela 14. Índice Kappa, valores mínimos e máximos para o índice e desvio padrão, para as melhores classificações das imagens CBERS-4A, Sentinel-2, Landsat-8 e Planet.	49
Tabela 15. Valor do teste z para comparar as melhores classificações de cada algoritmo de classificação, para as imagens CBERS-4A, Sentinel-2, Landsat-8 e Planet.	51
Tabela 16. Data de aquisição dos dados utilizados.	62
Tabela 17. Características do recobrimento aéreo e do equipamento LiDAR utilizado.	63
Tabela 18. Cena utilizada de cada sensor orbital para a aplicação da transferência de aprendizagem.	64
Tabela 19. Características espectrais das bandas dos satélites CBERS-4A (WPM), Landsat-8/OLI, Sentinel-2/MSI e PlanetScope utilizadas.	64
Tabela 20. Fórmulas dos índices de vegetação utilizados no estudo.	65
Tabela 21. Conjuntos de dados e número de variáveis como dado de entrada para a classificação supervisionada.	68
Tabela 22. Classes encontradas dentro da área de pesquisa e suas respectivas áreas em hectares.	69
Tabela 23. Quantidade de pixels (1x1 m) para as amostras de treinamento e validação.	70
Tabela 24. Índice Kappa para todas as imagens orbitais, composições e classificadores testados com a transferência de aprendizagem para as três áreas de estudo.	72
Tabela 25. Matriz de confusão para a transferência de aprendizagem para a área de estudo 1, utilizando o classificador Máquina de vetores de suporte com a composição 3 da imagem Sentinel-2.	73
Tabela 26. Matriz de confusão para a transferência de aprendizagem para a área de estudo 2, utilizando o classificador Árvores aleatórias com a composição 8 da imagem Sentinel-2.	74

Tabela 27. Matriz de confusão para a transferência de aprendizagem para a área de estudo 3, utilizando o classificador de Verossimilhança com a composição 5 da imagem Sentinel-2. ..	74
Tabela 28. Índice de concordância, desvio padrão, valor máximo e valor mínimo para as classificações das imagens Sentinel-2 utilizando a transferência de aprendizagem para a área 1 com o classificador Máquina de vetores de suporte com a composição 3, para a área 2 com o classificador Árvores aleatórias e a composição 8, e para a área 3 utilizando o classificador de Verossimilhança com a composição 5.....	75

LISTA DE ABREVIATURAS E SIGLAS

LiDAR – Laser Detection and Ranging

ALS – *Airborne Laser Scanner*

DAP – Diâmetro altura do peito

GNSS – *Global Navigation Satellite System*

INS - *Inertial Navigation System*

MDT – Modelo Digital de Terreno

MDS – Modelo Digital de Superfície

MDE – Modelo Digital de Elevação

CHM – *Canopy Height Model*

NDVI - Índice de Vegetação da Diferença Normalizada

EVI - Índice de Vegetação Melhorado

TGI - Índice Triangular Verde

SAVI - Índice de Vegetação Ajustado para o Solo

VARI - Índice de Vegetação Resistente à Atmosfera

CBERS - Satélite Sino-Brasileiro de Recursos Terrestres

RFE – *Recursive Features Elimination*

RT – *Random Trees*

RF – Random Forest

SVM – *Support Vector Machine*

MLC – *Maximum Likelihood Classifier*

KDD - *Knowledge Discovery in Database*

ML – Machine Learning

TA – Transferência de aprendizagem

SUMÁRIO

1	INTRODUÇÃO	15
2	HIPÓTESE	16
3	OBJETIVOS	17
3.2	OBJETIVOS ESPECÍFICOS	17
4	CAPÍTULO I - REVISÃO BIBLIOGRÁFICA	18
4.1	ESTÁGIOS DE SUCESSÃO	18
4.2	SENSORIAMENTO REMOTO	19
4.2.1	Sensoriamento Remoto no estudo da Vegetação	21
4.2.2	Light Detection and Ranging (LIDAR)	22
4.2.3	Modelo Digital de Elevação (MDE).....	23
4.2.4	Modelo Digital de Terreno (MDT)	23
4.2.5	Modelo Digital de Superfície (MDS)	24
4.2.7	Índice de Vegetação (IV) e Matriz de Co-ocorrência de Níveis de Cinza (GLCM)	26
4.3	MINERAÇÃO DE DADOS	28
4.3.1	Classificação Digital	29
4.3.2	Maximum Likelihood Classifier	30
4.3.3	Random Trees e Random Forest	30
4.3.4	Support Vector Machine (SVM)	32
	CAPÍTULO II	34
5	INTRODUÇÃO	34
6	MATERIAL E MÉTODOS	36
6.1	ÁREA DE ESTUDO	36
6.2	COLETA E PROCESSAMENTO DOS DADOS.....	37
6.2.1	LiDAR	38
6.2.2	Descrição dos dados multiespectrais de satélite	39
6.2.4	Veículo Aéreo Não Tripulado (VANT).....	42
6.2.5	Classificação de referência	42
6.2.6	Composição dos dados.....	43
6.3	CLASSIFICAÇÃO DAS IMAGENS.....	43
6.3.1	Coleta de amostras.....	43
6.3.2	Classificação Supervisionada.....	45
6.4	AVALIAÇÃO DA ACURÁCIA	45
7	RESULTADOS	45
7.1	MODELO DIGITAL DE ALTURE DE DOSSEL.....	45
7.2	CLASSIFICAÇÕES DE ESTÁGIO SUCESSIONAL DA VEGETAÇÃO	47

8	DISCUSSÃO.....	52
9	CONCLUSÕES	56
	CAPÍTULO III	57
10	INTRODUÇÃO.....	57
11	MATERIAL E MÉTODOS	59
11.1	ÁREA DE ESTUDO	59
11.2	COLETA E PROCESSAMENTO DOS DADOS.....	61
11.2.1	LIDAR	63
11.2.2	Descrição das Imagens de satélites.....	63
11.2.3	Pré-processamento imagens de satélites	64
11.2.4	Classificação de referência	66
11.2.5	Composição dos dados.....	68
11.3	CLASSIFICAÇÃO DAS IMAGENS.....	69
11.3.1	Coleta de amostras.....	69
11.3.2	Transferência de aprendizagem.....	72
11.4	ANÁLISE ESTATÍSTICA	72
12	RESULTADOS	72
13	DISCUSSÃO.....	78
14	CONCLUSÕES	80
15	CONSIDERAÇÕES FINAIS	82
	APÊNDICE	83
	REFERÊNCIAS	84

1 INTRODUÇÃO

A heterogeneidade das florestas nativas dificulta os levantamentos dendrométricos in loco, visto que elas apresentam grande diversidade biológica e densidade de indivíduos, aumentando o tempo e custo para a obtenção de informações. Para a realização do inventário florestal de forma eficaz, é necessária uma grande quantidade de recursos financeiros e humanos, além de demandar tempo em deslocamento para áreas distantes e de difícil acesso. Os custos para os levantamentos a campo costumam ser altos, com intervalos temporais grandes e irregulares (ANDRESEN *et al.*, 2007). A discriminação de diferentes tipologias florestais é difícil por meio do levantamento em campo, por isso é importante o uso de novas metodologias e técnicas de classificação que contemplem informações sobre a forma do alvo, a textura ou relações entre alvos vizinhos no mapeamento das fases de sucessões florestais (PONZONI; SHIMABUKURO, 2009).

A separação do processo de sucessão em fases distintas é uma estratégia utilizada na busca do entendimento sobre as atividades funcional das florestas (KAGEYAMA *et al.*, 1986). As modalidades de uso e normas para a supressão da vegetação são regulamentadas de forma diferenciada para cada estágio sucessional, tornando importante a classificação de forma acurada e rápida para a fiscalização e o licenciamento ambiental (SEVEGNANI *et al.*, 2012). As medições de atributos estruturais relacionados às sucessões florestais também são importantes para prever processos a longo prazo como o sequestro de carbono (SHUGART, 2000).

Por isso, surge uma demanda de investigação de novas técnicas e metodologias que sejam eficientes no mapeamento da cobertura vegetal, de modo a aperfeiçoar as técnicas tradicionais de inventário em campo (CINTRA, 2007). Uma alternativa é o uso da tecnologia LiDAR (*Light Detection and Ranging*), que consiste na emissão e recebimento de um pulso laser. Em que o tempo de retorno do pulso entre o dispositivo e os alvos é medido, permitindo a obtenção da estimativa da distância desses objetos (GIONGO *et al.*, 2010). Esses pulsos têm a capacidade de penetrar na vegetação, fornecendo informações da estrutura florestal e do relevo das áreas (JENSEN, 2009). Já o sistema LiDAR/ALS (*Airborne Laser Scanner*) é a modalidade aerotransportada, que permite a cobertura de uma grande extensão de superfície, resultando em produtos de alta resolução, podendo chegar a níveis submétricos (D'OLIVEIRA; FIGUEIREDO; PAPA, 2014).

Além de dados LiDAR, a utilização de imagens de sensores ópticos a bordo de satélites costuma ter a vantagem de ser mais rápida em relação à obtenção dos dados in loco, mas também por envolverem menores custos financeiros (JANOTH *et al.* 2007). E com a evolução tanto das características espaciais como espectrais dos sensores orbitais tem aumentado a utilização dessas ferramentas como fontes de informações para o mapeamento de uso e cobertura da terra (ADAM *et al.*, 2014).

Os algoritmos tradicionais de *machine learning* (ML) vem sendo utilizados na atualidade por apresentar bons resultados com dados de alta dimensionalidade utilizando uma quantidade menor de amostras (MOUNTRAKIS; IM; OGOLE, 2011) quando comparado ao *deep learning* (DP), além de demandar um menor custo computacional para seu processamento. Dentro do ML pode ser utilizada a transferência de aprendizagem (TA), que consiste na aplicação de modelos aprendidos anteriormente para solucionar novos problemas de forma eficiente (CHEN; ZHANG; OUYANG, 2018), e alcança bons resultados economizando tempo e evitando *overfitting* dos modelos (KUMAR; TALIB; VERMA, 2021).

Nesse sentido, diversos estudos têm mostrado que dados extraídos de LiDAR aerotransportado combinados com imagens multispectrais de satélite têm a capacidade de fornecer as informações necessárias para a classificação de estágios sucessionais de forma acurada, além de necessitar menor quantidade de mão-de-obra e menos tempo para a obtenção desses atributos, fornecendo resultados de um modo mais eficiente. Esse estudo tem como objetivo propor novas metodologias de classificação dos estágios de sucessão comparando diferentes plataformas de imagens orbitais combinadas com dados LiDAR, além de comparar também diferentes algoritmos de ML e verificar suas acurácias quando utilizados para a TA.

2 HIPÓTESE

As informações obtidas dos dados LiDAR aerotransportado e imagens multispectrais de satélite possibilitam a diferenciação dos estágios sucessionais da vegetação nativa de forma acurada.

A inclusão de dados LiDAR melhora a performance da classificação dos estágios de sucessão quando comparado ao uso isolado de dados multiespectrais.

Os modelos ML treinados em uma área de estudo podem ser empregados em outras áreas dentro da mesma fitofisionomia vegetal de forma acurada.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Testar se as informações obtidas dos dados LiDAR aerotransportado e imagens multiespectrais de diferentes satélites possibilitam a classificação dos estágios sucessionais da vegetação nativa de forma acurada, por meio de algoritmos de classificação ML e TA.

3.2 OBJETIVOS ESPECÍFICOS

- Extrair métricas relacionadas a vegetação a partir da nuvem de pontos LiDAR.
- Extrair métricas de textura e índices vegetativos das imagens de satélite.
- Determinar a acurácia de distinção dos estágios sucessionais utilizando os dados LiDAR aerotransportado e de imagens orbitais em relação ao inventário florestal convencional e fotointerpretação.
- Comparar o uso de imagens multiespectrais isoladamente e combinadas com dados LiDAR para a classificação dos estágios sucessionais.
- Testar o uso de imagens multiespectrais de diferentes satélites para a classificação dos estágios sucessionais.
- Comparar métodos de classificação paramétricos com métodos de aprendizado de máquina.
- Avaliar a acurácia da transferência da aprendizagem de um modelo de classificação para diferentes áreas.

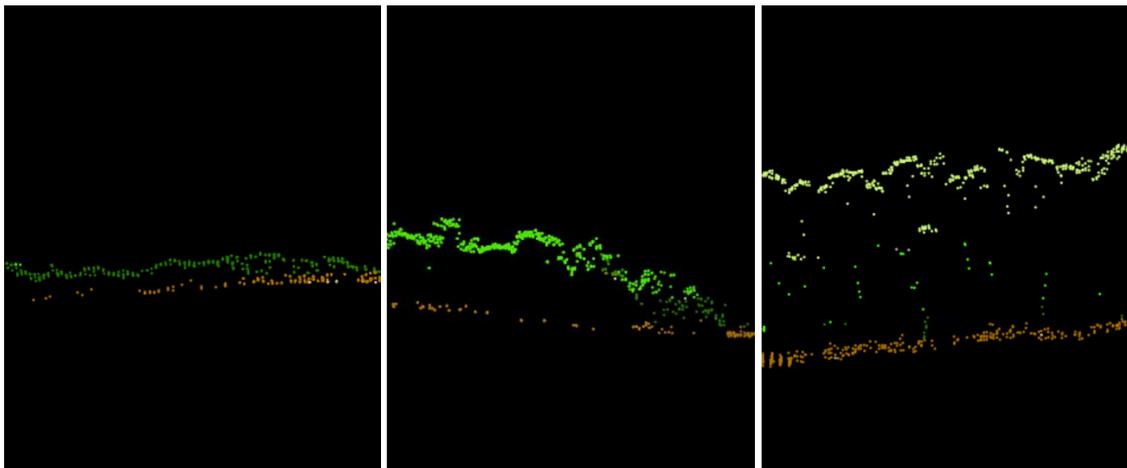
4 CAPÍTULO I - REVISÃO BIBLIOGRÁFICA

4.1 ESTÁGIOS DE SUCESSÃO

As mudanças que ocorrem nas características fisiológicas, estruturais e florísticas, antes que uma comunidade se estabilize em um local, são chamadas de sucessão vegetal (RODRIGUES, 1995). A sucessão da vegetação pode ser promovida através de eventos de perturbação no ambiente terrestre, alterando tanto a composição florística como a estrutura vegetal local (SOUSA, 1984).

Barnes *et al.* (1998) define a sucessão florestal como a mudança na composição de espécies ou substituição da biota em determinado lugar por outra de diferente natureza. A sucessão pode ser dividida em dois tipos: primária ou secundária. A primária ocorre após um evento que proporciona o crescimento de indivíduos em um lugar onde não havia vegetação. Já a secundária é quando acontece algum distúrbio na vegetação, removendo-a e proporcionando o crescimento de novos indivíduos. Então dentro das sucessões, as espécies pioneiras são as primeiras a povoarem a comunidade, possibilitando as características necessárias para a introdução das secundárias iniciais e tardias e, por fim, as espécies climácicas têm as condições para seu estabelecimento como apresentado na Figura 1.

Figura 1. Perfil esquemático de sucessão florestal obtido da nuvem de pontos LiDAR da área de estudo, com predomínio de pioneiras (A); intermediárias (B); e climácicas (C).



(a)

(b)

(c)

Fonte: Elaborado pelo autor (2022).

Diferentes modalidades das atividades florestais e políticas ambientais são influenciadas pela correta classificação dos estágios de sucessão de uma comunidade arbórea (SIMINSKI *et al.*, 2004). Na Mata Atlântica, a classificação de estágios sucessionais ganhou grande destaque, por causa da regulamentação das normas para a supressão ser diferenciada para cada estágio pela legislação. Aumentando sua importância para o licenciamento e fiscalização ambiental (SEVEGNANI *et al.*, 2012).

Segundo a Resolução CONAMA nº 04/1994 (BRASIL, 1994), para a caracterização da vegetação primária e secundária nos estágios inicial, médio e avançado de regeneração, deve ser observado um conjunto de critérios e indicadores estruturais tanto qualitativos quanto quantitativos. O estágio inicial de regeneração surge logo após o abandono do solo, em que os indivíduos possuem no máximo oito centímetros de diâmetro médio na altura do peito (DAP) e árvores com altura máxima de quatro metros. São uma vegetação com baixa diversidade de espécies e predominantemente de ervas e arbustos e duram entre seis e dez anos, dependendo da qualidade do solo. Já no estágio médio de regeneração a diversidade aumenta, as espécies de árvores pioneiras são predominantes e existe maior quantidade de arvoretas e arbustos do que herbáceas. O DAP médio é de 8 a 15 cm e altura de até 12 m. Este estágio ocorre entre seis e quinze anos depois do abandono da área. E no estágio avançado de regeneração, a fisionomia arbórea é dominante sobre as demais, formando um dossel fechado e relativamente uniforme, apresentando árvores dominantes. A altura média da vegetação é superior a 12 m, e o DAP médio varia de 15 a 25 cm.

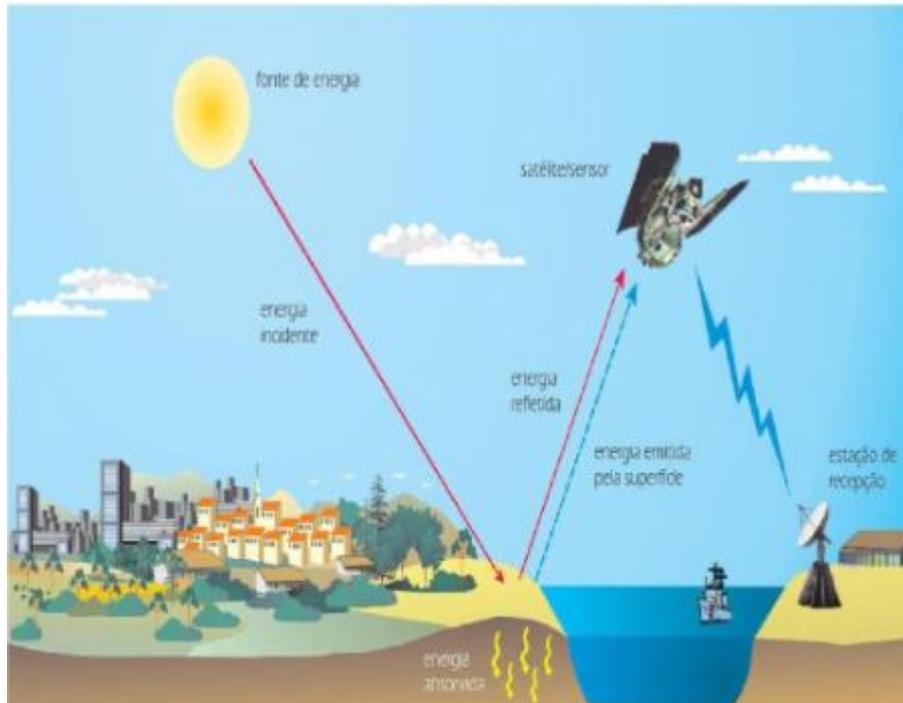
A sucessão florestal pode ser avaliada diretamente no campo via técnicas de inventário florestal tradicional. Apesar de acuradas, as classificações em campo são frequentemente limitadas a pequenas áreas (BERGENS; DRONOVA, 2007). A fim de ser útil para o manejo florestal sustentável, classificações das sucessões florestais devem cobrir maiores áreas (FALKOWSKI *et al.*, 2009).

4.2 SENSORIAMENTO REMOTO

Segundo Novo (2010) o sensoriamento remoto é definido como sendo a utilização conjunta de sensores, equipamentos para processamento de dados, equipamentos de transmissão de dados colocados a bordo de aeronaves, espaçonaves, ou outras plataformas, com o objetivo de estudar eventos, fenômenos e processos que ocorrem na superfície do planeta Terra a partir do registro e da análise das interações entre a radiação eletromagnética e as substâncias que o compõe, em

suas mais diversas manifestações. Sem ter contato com a superfície do objeto analisado, os sensores captam a energia eletromagnética, a partir da detecção e medição das mudanças impostas pelo objeto à sua redondeza (LIU, 2015).

Figura 2. Obtenção de imagens por sensoriamento remoto.



Fonte: FLORENZANO, 2007.

O sensoriamento remoto tem como principal função fornecer ferramentas capazes de facilitar a identificação e extração de informações presentes nos dados, para então realizar uma interpretação delas. As fontes de energia utilizadas pelos sensores podem ser provenientes da luz solar ou do calor emitido pela superfície do planeta, sendo estas fontes naturais de energia (FLORENZANO, 2007), ou também podem ser utilizados sensores com uma fonte artificial como o caso dos sensores *Radio Detection and Ranging* (RADAR) e LiDAR.

Um sensor tem sua qualidade especificada pela capacidade de obtenção de informações e medidas da energia eletromagnética de forma detalhada, sendo essas relacionadas com a resolução espacial, espectral e radiométrica do equipamento (GUEDES; SILVA, 2018). Com o grande avanço da tecnologia, os sistemas e sensores ópticos de satélites possuem resoluções espaciais e espectrais melhores. Satélites militares e comerciais já possuem resolução espacial na casa dos centímetros, e sensores hiperespectrais com centenas de bandas espectrais (LIU, 2015).

4.2.1 Sensoriamento Remoto no estudo da Vegetação

Nas últimas décadas, observou-se uma notável evolução das tecnologias espaciais destinadas ao monitoramento dos recursos florestais. O recente avanço de características espaciais e espectrais de sensores orbitais e o melhoramento de algoritmos de classificação têm aumentado a popularidade de dados de sensoriamento remoto como fonte para mapeamento do uso e cobertura da terra (ADAM *et al.*, 2014). O uso dessas ferramentas costuma ser vantajoso em relação aos mapeamentos em campo convencionais, devido a serem de mais rápida obtenção, e também por envolverem menores custos financeiros e humanos (JANOTH *et al.*, 2007).

Amaral *et al.* (2009) realizaram um estudo em que encontraram eficácia na utilização de imagens orbitais para a distinção do estágio sucessional inicial dos demais estágios. Berveglieri *et al.* (2018) concluíram que é possível realizar a classificação de estágios sucessionais utilizando dados multitemporais de imagens ópticas, usando apenas DSM para calcular as variações verticais do dossel florestal. Já resultados de Caughlin *et al.* (2021) mostram que imagens temporais de Landsat podem ser usadas na distinção de estágios de sucessão florestal.

As imagens provenientes de sensores ópticos possuem características únicas quanto à resolução radiométrica, espectral, espacial e temporal e têm se tornado fontes primárias para classificação da cobertura da terra e da vegetação em diversas escalas (LU *et al.*, 2014). Alguns estudos utilizaram dados de levantamento aerotransportado com LiDAR para estimativa de biomassa em florestas temperadas (LEFSKY *et al.*, 2002; NÆSSET *et al.*, 2004; ANDERSEN *et al.*, 2006). Já outros usaram esta ferramenta em florestas tropicais para levantamentos ecológicos e de avaliação de estoques de carbono (ASNER *et al.*, 2008; KENNAWAY *et al.*, 2008; ASNER *et al.*, 2010). Há também estudos que utilizaram os dados e métricas fornecidas pelo levantamento de dados LiDAR para a classificação de estágios sucessionais (FALKOWSKI *et al.*, 2009; VAN EWIJK *et al.*, 2011; KOLECKA *et al.*, 2016; GU *et al.*, 2018)

Como a sucessão florestal é um processo tridimensional, os sensores ópticos passivos são muito menos sensíveis a estrutura de copa tridimensional que os instrumentos LiDAR. Esse instrumento mede a disposição tridimensional da copa florestal, fornecendo estimativas precisas e acuradas dos atributos da estrutura florestal, isto tem o potencial para classificar a sucessão florestal através de extensas áreas. As medições LiDAR caracterizam a distribuição e densidade da vegetação dentro de diferentes parcelas do inventário florestal, tornando possível a classificação dos estágios sucessionais (FALKOWSKI *et al.*, 2009).

4.2.2 Light Detection and Ranging (LIDAR)

O princípio de funcionamento do sistema LiDAR consiste na emissão de um pulso laser com uma elevada frequência de repetição, podendo essa ser emitida de plataforma aérea, terrestre ou orbital (GIONGO *et al.*, 2010). Para cada pulso emitido, são registrados de forma precisa a localização e a orientação do ponto de lançamento e o tempo de percurso entre o aparelho emissor e o ponto atingido na superfície, o que possibilita que a distância e a localização de cada ponto sejam calculadas (WAGNER *et al.*, 2004). No momento em que cada pulso é emitido, são determinadas a posição e altitude do scanner por dados coletados por GNSS (*Global Navigation Satellite System*) e INS (*Inertial Navigation System*) (KUMAR, 2012).

Considerando que o pulso laser emitido se propaga na velocidade da luz, a posição tridimensional do ponto é determinada com base na distância e orientação entre o sensor e a superfície ou objeto detectados. Além da medição do tempo de retorno, o sensor também é capaz de medir a intensidade do sinal (BALTSAVIAS, 1999). No mapeamento de terrenos são emitidos pulsos a uma taxa aproximadamente de 10.000-100.000 pulsos/ segundo (PEREIRA; SCHIMALSKI, 2014). Após o pulso ser gerado, ele é direcionado para um espelho de varredura, que é um conjunto de lentes e espelhos óticos que emitem os pulsos laser em direção aos objetos no terreno e em seguida recebem os sinais de retorno, e os direcionam ao receptor. (GIONGO *et al.*, 2010).

O pulso laser possui a capacidade de penetrar por pequenas aberturas no dossel da floresta, fornecendo assim informações sobre a topografia do terreno, sub-bosque e dossel (JENSEN, 2009). Os múltiplos retornos dos pulsos laser permitem que diferentes camadas da estrutura vertical da paisagem sejam identificadas e mensuradas, resultando na estrutura tridimensional (REUTEBUCH *et al.*, 2005).

Desde o final da década de 1990, o uso de dados LiDAR em aplicações florestais foi voltado para a estimativa de parâmetros estruturais das florestas. Modelos digitais do terreno eram utilizados para caracterizar de forma precisa a topografia do terreno, fornecendo informações necessárias para o planejamento da infraestrutura florestal. Os dados do LiDAR têm a capacidade de fornecer dados sobre a superfície do dossel e parâmetros como altura das árvores e dimensões das copas, elementos importantes para o manejo, tanto de florestas nativas como de plantadas (TIEDE *et al.*, 2005).

Entre a geração de modelos digitais do terreno (MDT), modelos digitais de superfície (MDS) e modelos digitais de altura de copa (MDA), percebe-se a grande aplicabilidade do

LiDAR na área florestal. Tornando possível o fornecimento de informações sobre a declividade do terreno, informações de árvores isoladas, quantificação de biomassa e carbono, análise de extratos florestais, modelagem 3D de copas e com a modelagem de árvores individuais e possibilitando extrair parâmetros dessas (PEREIRA; SCHIMALSKI, 2014).

Means *et al.* (2000) acreditam que o uso do dados LiDAR requer menos trabalho para a obtenção de parâmetros florestais, já que a tecnologia permite a detecção de árvores em uma grande área florestal em um curto intervalo de tempo. O estudo e análise de árvores individuais e da homogeneidade florestal permite estimar as futuras condições de copa e dossel para a análise estrutural e florística da vegetação (SILVA *et al.*, 2017). A tecnologia LiDAR permite também realizar medidas diretas e estimativas de variáveis dendrométricas (NILSSON, 1996), assim permitindo a medição direta das alturas do dossel da floresta (GIONGO *et al.*, 2010)

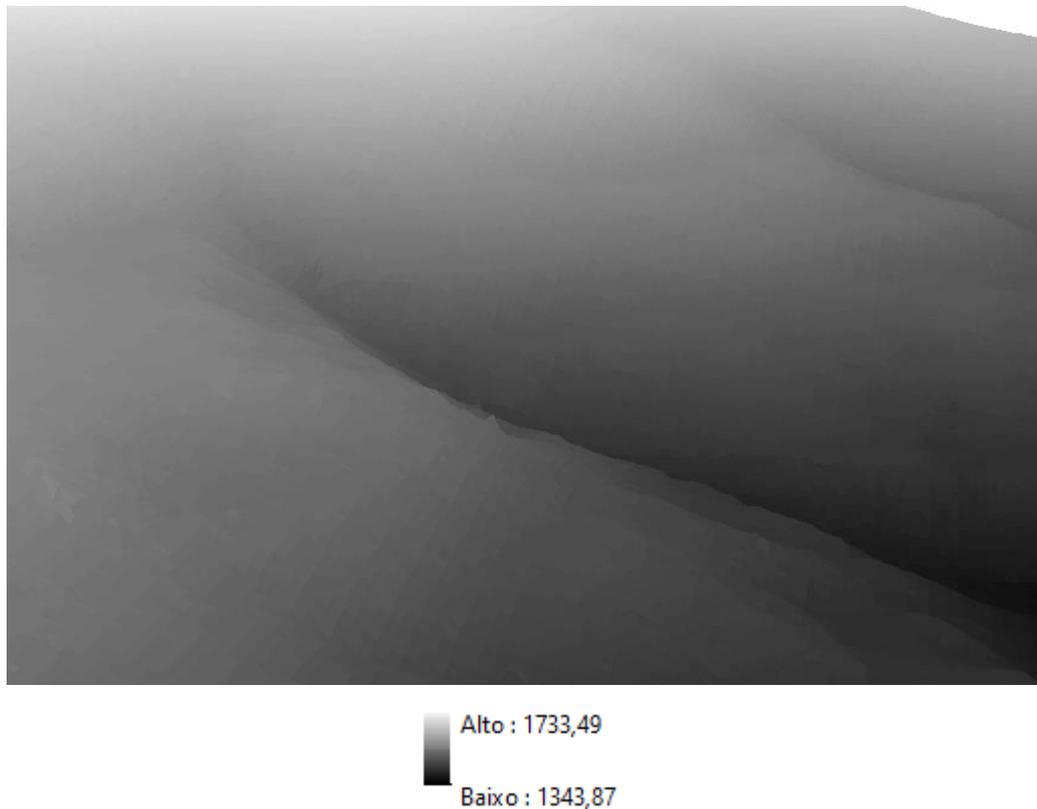
4.2.3 Modelo Digital de Elevação (MDE)

Modelos Digitais de Elevação (MDE) são a representação de maneira contínua da variação do relevo sobre espaço e podem ser subdivididos em Modelos Digitais de Superfície (MDS) e Modelos Digitais de Terreno (MDT) (LOCKS, 2017), exercendo funções importantes em diferentes aplicações como design geométrico computacional, análise de elementos finitos, sistemas de informações geográficas e robótica (FLORIANI; PUPPO, 1995).

4.2.4 Modelo Digital de Terreno (MDT)

Quando obtido a partir de dados LiDAR, esse modelo é composto apenas pelos pontos do LiDAR pertencentes ao terreno nu, sendo assim não contém informações de vegetação ou edificações (PEREIRA; SCHIMALSKI, 2014). O MDT gerado a partir do ALS pode detectar intervenções antrópicas como estradas e caminhos florestais, sendo de grande vantagem por possuir capacidade de alta acurácia detalhamento da morfologia da superfície, melhorando a definição de características do relevo (PEREIRA *et al.*, 2013). A Figura 3 apresenta um exemplo de MDT.

Figura 3. Modelo Digital de Terreno da área de estudo do capítulo II, gerado a partir dos dados LiDAR.

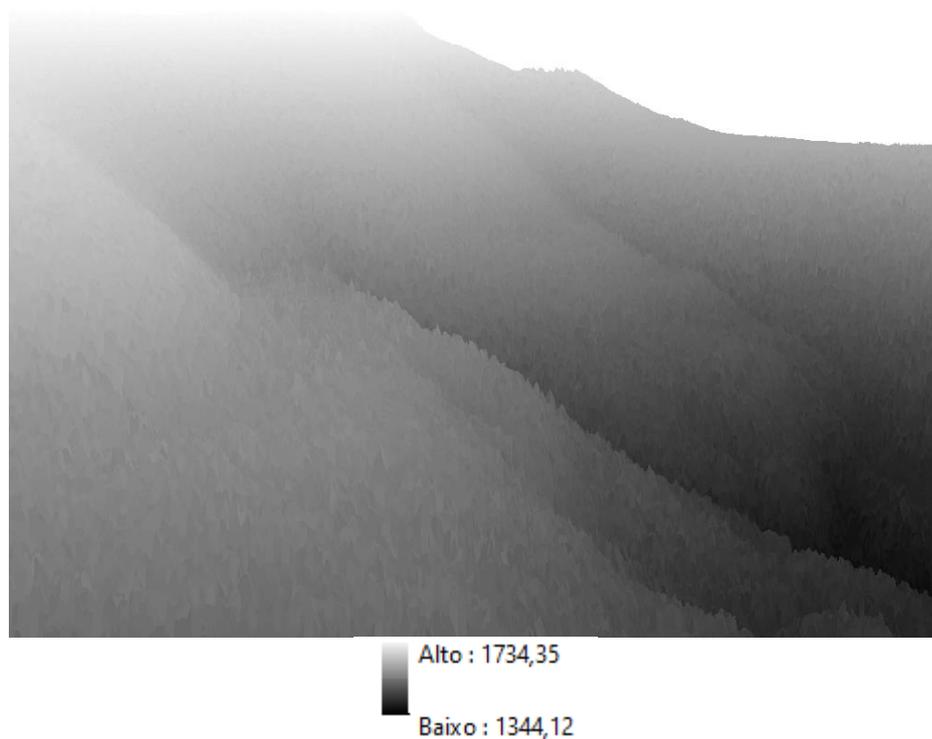


Fonte: Elaborado pelo autor (2022).

4.2.5 Modelo Digital de Superfície (MDS)

MDS contém a informação de elevação de todas as feições da paisagem, sendo a representação gráfica de feições na superfície terrestre, como árvores, construções, entre outros objetos do terreno (JENSEN, 2009; PEREIRA; SCHIMALSKI, 2014). A Figura 4 apresenta um exemplo de MDS gerado a partir da nuvem de pontos LiDAR para a área do capítulo II.

Figura 4. Modelo Digital de Superfície da área de estudo do capítulo II, gerado a partir dos dados LiDAR.

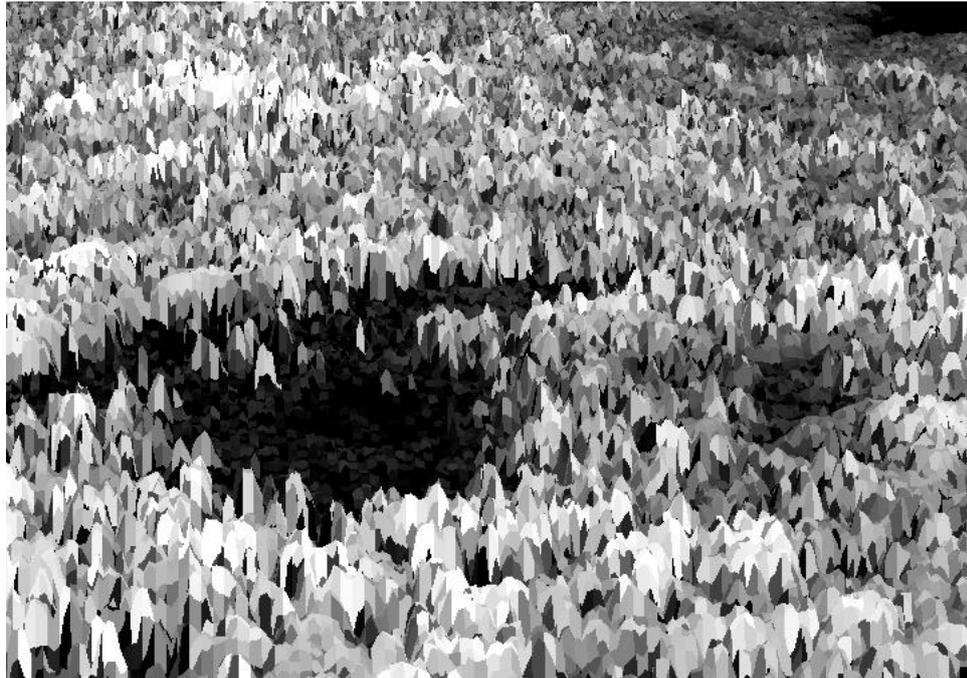


Fonte: Elaborado pelo autor (2022)

4.2.6 Modelo Digital de alturas do dossel (CHM)

Para a área florestal o CHM é um importante modelo, em que os valores das coordenadas Z do MDS são descontadas do MDT, fornecendo uma superfície que descreve a altura de cada elemento da área, como a altura de cada árvore em relação ao terreno (LOCKS, 2017). Essas informações de coordenadas podem ser extraídas dos dados LiDAR. Esse modelo (Figura 5) pode ser usado para a detecção individual de árvores, delineamento de copas de árvores, e para obter a estimativa de atributos biométricos, como biomassa e volume (SILVA *et al.*, 2017).

Figura 5. Canopy Height model para a área de estudo do capítulo II, gerado pela subtração do MDS pelo MDT.



Fonte: Elaborado pelo autor (2022).

4.2.7 Índice de Vegetação (IV) e Matriz de Co-ocorrência de Níveis de Cinza (GLCM)

A partir dos dados de satélites é possível gerar diversos dados auxiliares, como os índices de vegetação, que são dados provenientes da aritmética de bandas, ou também as informações de texturas que são calculados com a matriz de co-ocorrência de níveis de cinza, que fornecem informações sobre a diferença entre os níveis de cinza da área de estudo.

Os índices de vegetação têm como objetivo a exploração das propriedades espectrais da vegetação, com maior enfoque nas regiões do visível e infravermelho próximo. Sendo relacionados com parâmetros biofísicos da vegetação, como biomassa e área foliar, partindo do princípio de que a reflectância da vegetação nessas regiões mencionadas age de forma antagônica (PONZONI, 2009).

Os índices de vegetação são a maneira mais simples e eficiente de se realçar o sinal da vegetação, minimizando as variações na irradiância solar e os efeitos do substrato do dossel vegetal (JACKSON; HUETE, 1991). Além de ser importante para o planejamento ambiental, ressaltando comportamento espectral da vegetação e possibilitando distinguir diferentes tipos de cobertura e uso da terra (MOREIRA, 2003).

A energia do vermelho e do infravermelho próximo possui relação direta com as atividades fotossintéticas da vegetação, e com utilização de duas ou mais bandas minimizam as principais fontes de ruído, que são: variações na irradiância solar; efeitos atmosféricos; contribuições da vegetação não-fotossinteticamente ativa; contribuições do substrato, como o solo; e efeitos da composição e estrutura do dossel (FERREIRA *et al.*, 2008).

Diversos estudos vêm utilizando os índices de vegetação para o estudo do comportamento da vegetação, como o Índice de vegetação da diferença normalizada (NDVI) (CAUGHLIN *et al.*, 2021; SOTHE *et al.*, 2017), Índice de vegetação melhorado (EVI) (SILVA, 2020; GALVÃO *et al.*, 2011), Índice Triangular Verde (TGI) (PINTO, 2018), Índice de vegetação ajustado para o solo (SAVI) (PRADA, 2020) e Índice de vegetação resistente à atmosfera (VARI) (ABRAHÃO *et al.*, 2009; KUIAWSK *et al.*, 2015; Dos Santos *et al.*, 2020)

Segundo Haralick *et al.* (1973), as feições textuais são baseadas na frequência da distribuição relativa da escala de cinza. Algumas métricas tem uma relação a uma característica específica textural da imagem como homogeneidade, contraste e a presença de estruturas organizadas dentro da imagem. Outras métricas caracterizam a complexidade e natureza da transição dos tons de cinza que ocorrem na imagem.

As métricas texturais de média, dissimilaridade e contraste apresentaram um destaque para a classificação de estágios sucessionais. As imagens de Landsat-8 e RapidEye mostraram eficiência para esta classificação de estágios sucessionais utilizando métricas de textura (SOTHE *et al.*, 2017). Zhang *et al.* (2014) mostraram que a utilização de métricas texturais derivadas de imagens multiespectrais e LiDAR podem diferenciar estágios de desenvolvimento da floresta com acurácia de 89%.

Metodologias baseadas em textura são utilizadas para empregar informações espaciais em procedimentos de classificação (LU *et al.*, 2014). Segundo Kayitakire *et al.* (2006) as métricas segundo momento angular (SMA), entropia, correlação, contraste, homogeneidade e variância, são as mais utilizadas no sensoriamento remoto. Para a extração das feições de texturas, podem ser utilizadas diferentes metodologias, e para este trabalho foi empregada a matriz de concorrência dos níveis de cinza (GLCM – *Grey Level Cooccurrence Matrix*, em inglês). Em que a caracterização da textura é realizada pelas dependências e distribuições espaciais entre os tons de cinza de uma vizinhança (HARALICK, 1979). A matriz de coocorrência informa a quantidade de ocorrência de diferentes combinações de valores de intensidade dos pixels. Com o objetivo de descrever, considerando múltiplas direções, a textura

através de um conjunto de características para a ocorrência de cada nível de cinza dos pixels da imagem (ROCHA; LEITE, 2002).

As métricas de textura são calculadas com base na diferença de intensidade (níveis de cinza) entre pixels vizinhos, podendo ser escolhida a distância entre esses pixels, não necessitando ser o próximo pixel. Também se faz necessário a escolha da direção de busca, podendo essa ser de 0° , 45° , 90° e 135° (HARALICK *et al.*, 1973). A seguir são apresentados os descritores de textura utilizados neste trabalho (SOUTO JUNIOR, 2010; STEIN, 2005).

- (1) Energia: Os valores de energia estão relacionados com a uniformidade e constância da distribuição dos níveis de cinza na imagem.
- (2) Entropia: Mede a não uniformidade da textura ou o grau de desordem da distribuição dos níveis de cinza na imagem.
- (3) Correlação: Mede a dependência linear de pixels em relação a sua vizinhança
- (4) Homogeneidade: Mede o grau de variação local dos níveis de cinza entre pares vizinhos.
- (5) Contraste: Quantifica a variação local dos níveis de cinza entre pares de pixels.
- (6) Variância: Baseia-se na dispersão dos dados em torno de um valor médio, tratando especificamente das combinações do pixel de referência e do seu vizinho.

4.3 MINERAÇÃO DE DADOS

A mineração de dados é definida por Berry e Linoff (2004) como a exploração e análise de dados, por meios automáticos ou semi-automáticos, em grandes quantidades de dados, com o objetivo de descobrir regras ou padrões interessantes. Sendo parte de um processo maior de pesquisa denominado Busca por conhecimento de dados (*Knowledge Discovery in Database - KDD*) (CORTES *et al.*, 2002). As principais tarefas realizadas pela mineração de dados são: Análise de regras de associação, Análise de padrões sequenciais, Classificação, Predição, Análise de Cluster e análise de outliers (AMO, 2004).

Os bancos de dados utilizados na mineração possuem milhares de informações, impossibilitando a utilização de todas para a construção de modelos. Assim, utilizam-se amostras, que são divididas em conjunto de treinamento, teste e validação. Esta divisão tem como objetivo evitar que um modelo fique dependente de apenas um conjunto de dados, evitando o efeito conhecido por Bias (CAMILO; SILVA, 2009).

Diversas técnicas podem ser utilizadas na construção de classificadores, algumas dessas são: Árvores de decisão, Redes Neurais, Método K-Means para análise de Cluster, Método CURE para análise de Cluster, Classificação Bayesiana, Classificação baseada em regras,

Support Vector Machine (SVM), Classificação por regras de associação, Random Forest, Aprendizado tardio (CAMILO; SILVA, 2009; AMO, 2004).

Segundo Justino, Martines e Kawakubo (2017) a mineração de dados se mostrou eficiente na utilização de dados de sensoriamento remoto, em que neste caso foi aplicada para a classificação de uso e ocupação do solo. Sothe (2015) utilizou algoritmos de árvore de decisão para a separação de estágios de sucessão através de imagens do recobrimento aérea de Santa Catarina, demonstrando que essa metodologia é capaz de distinguir os estágios de sucessão com uma boa acurácia. Pinto (2018) utilizou a mineração de dados para a classificação de estágios de sucessão empregando imagens obtidas por VANT, e alcançou resultados variando de “muito bom” a “excelente”.

4.3.1 Classificação Digital

Os humanos possuem um sistema visual com grande capacidade de reconhecer padrões, entretanto uma imagem pode fornecer um volume maior de informações do que o ser humano é capaz de processar (CROSTA, 1999). O olho humano é um sensor natural que enxerga somente a luz ou energia visível. Já com o uso do sensoriamento remoto é possível obter informações de dados invisíveis, que possuem uma quantidade de níveis de cinza superiores a capacidade de detecção do olho humano (FLORENZANO, 2002).

Sabendo disso, o uso do processamento de imagens digitais fornece ferramentas para a identificação e a extração das informações contidas nas imagens, para posterior interpretação de forma facilitada. Assim, estes processos têm como fundamento medições de radiância nas faixas do espectro eletromagnético, associando os pixels aos valores numéricos que descrevem o objeto de interesse (CRÓSTA, 2002; JENSEN, 2009; LILLESAND; KIEFER; CHIPMAN, 2008). A principal função destes procedimentos é a classificação automática dos pixels com o objetivo de definir as classes de cobertura do solo (LILLESAND; KIEFER; CHIPMAN, 2008). Quanto à unidade de análise, a classificação pode ser por pixel ou por regiões. Na classificação por pixel, estes são avaliados isoladamente, e somente a informação espectral é utilizada para definir as regiões homogêneas. Fundamentam-se prioritariamente em métodos estatísticos ou determinísticos (PONZONI *et al.*, 2009). Já a classificação por regiões, que é precedida por uma etapa de segmentação, avalia os segmentos, ou seja, pixels agrupados em compartimentos homogêneos da cena.

Para realizar a classificação, são utilizados modelos matemáticos denominados algoritmos. Sendo possível fornecer ou não amostras de treinamento, que são conjuntos de pixels representativos referentes a cada uma das classes de interesse. A classificação não-supervisionada é quase totalmente automática pelo computador, necessitando de pouco ou nenhuma participação do analista (MENESES; ALMEIDA, 2002). A classificação supervisionada requer um conhecimento prévio das classes estudadas contidas na área de interesse. Portanto, o algoritmo é treinado visando distinguir as classes uma das outras, utilizando regras estatísticas pré-estabelecidas para isso (VENTURIERI, 2007).

4.3.2 Maximum Likelihood Classifier

O *Maximum Likelihood Classifier* (MLC) é um algoritmo paramétrico, associando pontos individuais da imagem a suas classes. Assumindo que essas possuem uma distribuição normal, a partir da amostragem e dos parâmetros definidos, o algoritmo computa a probabilidade estatística de um pixel desconhecido pertencer a uma ou outra classe, levando em conta a variabilidade das classes usando a matriz de covariância (LEITE; ROSA, 2012; LU *et al.*, 2014).

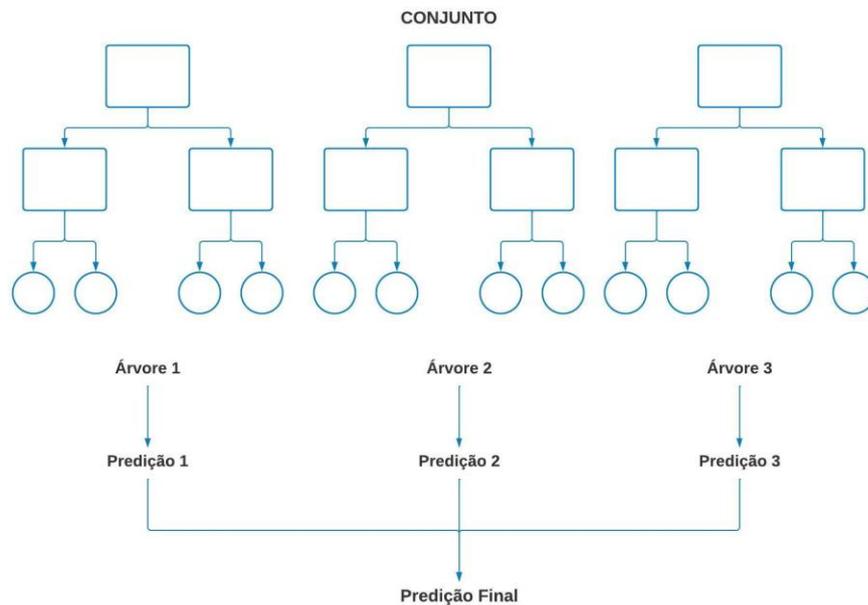
Utilizando parâmetros estatísticos, o método MLC pondera as distâncias entre as médias dos valores dos pixels das classes, e assume que todas as bandas têm distribuição normal. Para que a classificação por MLC seja precisa, é necessário que o conjunto de amostras de treinamento possuam um número elevado de pixels, permitindo uma base segura para tratamento estatístico (CRÓSTA, 2002). Nos últimos anos, diversos trabalhos vêm destacando a utilização do método MLC e sua utilização no sensoriamento remoto (CORREIA *et al.*, 2007; ALMEIDA, SANTOS e CHAVES, 2011; RODRIGUES *et al.*, 2014; MOREIRA; FARIA, 2019; RAMOS; DE MELO, 2020)

4.3.3 Random Trees e Random Forest

Random forest ou florestas aleatórias é um algoritmo supervisionado de aprendizado de máquina, que utiliza várias árvores de decisões (Figura 6) para prever um resultado. Sendo construído a partir de subconjuntos aleatórios dos dados de treinamento e seus atributos. A classificação final é decidida aplicando a regra da maioria sob os votos das classificações individuais. Para produzir previsões não correlacionadas e dissimilares, cada árvore é produzida

usando apenas amostras reduzidas (*bootstrap*) do conjunto de treinamento. Além do mais, para aumentar a diversidade entre eles o algoritmo introduz aleatoriedade na busca das melhores divisões (BREIMAN, 2001). É uma ferramenta poderosa que é usada para tomar decisões melhores, reduzir riscos e maximizar o sucesso, podendo ser usada para classificação ou regressão (SMITH, 2017).

Figura 6. Conjunto de árvores de decisão representando o funcionamento do algoritmo *Random Forest*.



Fonte: Smith, 2017.

Segundo Breiman (2001) o *Random Forest* é uma combinação de árvores preditoras, em que cada um depende dos valores dos vetores aleatórios amostrados independentemente e com a mesma distribuição para todas as árvores na floresta. Sendo uma ferramenta efetiva na predição, devido a lei dos grandes números não causando *overfitting*. O uso do tipo certo de aleatoriedade o torna um classificador e regressor preciso. De acordo com Pal (2005) o *Random Forest* é um algoritmo popular devido sua facilidade de utilização, necessitando a configuração de poucos parâmetros pelo usuário, atingindo acurácia semelhante ao SVM.

Os modelos de classificação com o *Random Forest* buscam utilizar um espaço de características que sejam representativas para a classificação de determinado tipo de classe de uso e cobertura do solo. As características são geradas a partir das bandas espectrais e temporais do produto e em seguida utilizadas com dados de entrada para o algoritmo. O resultado é um

classificador capaz de identificar uma ou várias classes com um certo nível de acurácia, de acordo com o modelo (PAIVA, 2020).

Além disso, o *Random Forest* realiza a atribuição de importância dos dados utilizados no treinamento do algoritmo. Sendo importante para maximizar o ganho de informação em cada passo e para a redução ou seleção dos atributos utilizados na construção do modelo. Essa informação pode ser obtida através da diminuição média de impureza (MDI), que é uma medida da diminuição média da impureza do nó da divisão de uma variável de todas as árvores (Han *et al.*, 2016). O MDI ($\Delta i(n)$) (eq. 2) é calculado a partir do índice GINI (i) (eq. 1)

$$i(n) = 1 - \sum p^2 \quad (1)$$

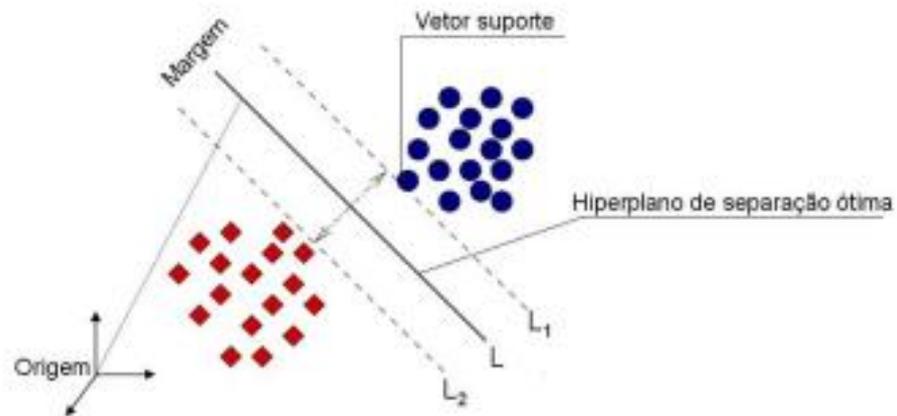
$$\Delta i(n) = i(n) - pl(nl) - pr(nr) \quad (2)$$

Random tree é uma combinação entre o simples modelo de árvores e as florestas aleatórias, sendo também um classificador supervisionado. É um algoritmo de aprendizado conjunto, que emprega a ideia de ensacamento (*bagging*) para construir um conjunto de dados aleatórios gerando uma árvore de decisão. Para a árvore padrão, cada nó é dividido usando a melhor divisão entre todas as variáveis, já no random forest, cada nó é dividido usando o melhor subconjunto de preditores escolhidos aleatoriamente naquele nó (MISHRA; RATHA, 2016).

4.3.4 Support Vector Machine (SVM)

O *Support Vector Machine* (SVM) é uma técnica embasada na Teoria de Aprendizado Estatístico, com intuito de resolver problemas de classificação de padrões (VAPNIK, 1995). É uma técnica desenvolvida para classificação binária, para superfície de decisão utiliza a construção de um hiperplano (Figura 7), buscando a separação máxima entre as classes, para padrões lineares (GONÇALVES, 2010). Além disso, SVM realiza de forma eficiente a classificação não-linear usando quatro funções, sendo elas: linear, quadrática, polinomial e função de base radial (FBR) (NASCIMENTO *et al.*, 2009).

Figura 7. Representação da classificação SVM.

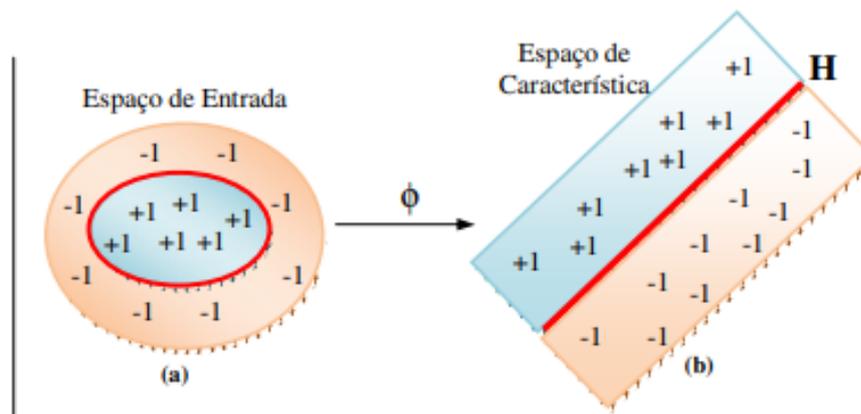


Fonte: NASCIMENTO *et al.*, 2009

Os vetores de suporte são padrões críticos que determinam o hiperplano ótimo (SMOLA *et al.*, 1999). Já a margem é obtida pela distância máxima entre o hiperplano e os vetores de suporte mais próximos a ele (GONÇALVES, 2010).

Para padrões não linearmente separáveis, os dados de entrada são transformados em um vetor de características de alta dimensionalidade, com o objetivo de realizar a separação linear (SOARES, 2008), como ilustra a Figura 8 a seguir.

Figura 8. Ilustração da transformação de dados não linearmente separáveis.



Fonte: SOARES, 2008.

O método SVM apresenta robustez em dados de grandes dimensões, possibilitando sua aplicação em dados de sensoriamento remoto, como as imagens de satélites (MONTEIRO, 2020). Segundo Nascimento (2009), o classificador SVM apresenta uma ótima separação de classes, além de apresentar uma minimização de erros pelo algoritmo.

CAPÍTULO II

Combining airborne LiDAR and spaceborne multispectral data for mapping successional forest stages in subtropical forests

Bill Herbert Ziegelmaier Neto
Marcos Benedito Schimalski
Veraldo Liesenberg
Camile Sothe

5 INTRODUÇÃO

A heterogeneidade das florestas tropicais dificulta os levantamentos dendrométricos in loco, visto que essas áreas apresentam grande diversidade biológica e densidade de indivíduos, os levantamentos de forma tradicional demandam de recursos e tempo que somam-se ainda a dificuldade de acesso a determinados locais para a obtenção de informações. Para a realização do inventário florestal de forma eficaz, é necessária uma grande quantidade de recursos financeiros e humanos, além de demandar tempo em deslocamento para áreas distantes e de difícil acesso. Os custos associados aos levantamentos a campo costumam ser altos, com intervalos temporais grandes e irregulares (ANDERSEN; REUTEBUCH.; MCGAUGHEY, 2006).

A separação do processo de sucessão em fases distintas é uma estratégia utilizada na busca do entendimento sobre a dinâmica funcional das florestas (KAGEYAMA; BRITO; BAPTISTON, 1986). Considerando a legislação brasileira, as modalidades de uso e normas para a supressão da vegetação são regulamentadas de forma diferenciada para cada estágio sucessional, tornando importante a classificação de forma precisa e rápida para a fiscalização e o licenciamento ambiental (SEVEGNANI *et al*, 2012). As medições de atributos estruturais relacionados às sucessões florestais também são importantes para prever processos a longo prazo como o sequestro de carbono (SHUGART, 2000). A classificação de diferentes estágios sucessionais em larga escala e de forma contínua é inviável por meio do levantamento em campo. Por isso, surge uma demanda de investigação de novas técnicas e metodologias que sejam eficientes no mapeamento da cobertura vegetal, de modo a aperfeiçoar as técnicas tradicionais de inventário em campo (CINTRA; OLIVEIRA; REGO, 2007). Uma alternativa é o uso da tecnologia LiDAR, que consiste na emissão e recepção de um pulso laser.

Além de dados LiDAR, nas últimas décadas também tem se observado grandes avanços em sensores multiespectrais. Com o aperfeiçoamento das resoluções espaciais e espectrais dos

sensores orbitais, tem aumentado a utilização dessas ferramentas como fontes de informações para o mapeamento de uso e cobertura da terra (ADAM *et al*, 2014).

Quanto a utilização de dados de sensoriamento remoto, alguns estudos utilizaram dados de levantamento aerotransportado com LiDAR para estimativa de biomassa em florestas temperadas (LEFSKY *et al*, 2002; NÆSSET, 2004). Já outros usaram esta ferramenta em florestas tropicais para levantamentos ecológicos e de avaliação de estoques de carbono [ASNER *et al*, 2008; KENNAWAY *et al*, 2008; ASNER *et al*, 2010]. Há também estudos que utilizaram os dados e métricas fornecidas pelo levantamento de dados LiDAR para a classificação de estágios sucessionais (FALKOWSKI *et al*, 2009; VAN EWIJK; TREITZ; SCOTT, 2011; KOLECKA *et al*, 2015; GU; CAO, 2018). Outros autores utilizaram dados de sensores óticos para a classificação do uso e cobertura da terra e da vegetação (CAUGHLIN *et al*, 2021; SOTHE *et al*, 2017 (a); SZOSTAK *et al*, 2021; MIRANDA *et al*, 2021; SOTHE *et al*, 2017 (b); SZOSTAK; HAWRYŁO; PIELA, 2018).

A disponibilidade de dados orbitais combinadas as técnicas de classificação automática destas imagens em larga escala, reduz inconsistências associadas à interpretação humana (FINLAYSON; VAN DER VALK, 1995; GUO *et al*, 2017). Algoritmos tradicionais de estatística e ML são muito utilizados pela capacidade de lidar com dados de alta dimensionalidade com um número limitado de amostras (MOUNTRAKIS; IM; OGOLE, 2017) e menor custo computacional comparado com algoritmos avançados de Deep learning. Algoritmos não-paramétricos de machine learning (ML) são utilizados para modelagem e classificação de áreas, aceitando uma variedade de dados como preditores de entrada e não fazem suposições esses dados (SEMOLINI, 2002; MAXWELL; WARNER; FANG, 2018), sendo o Random Tree (RT) e Support Vector Machine (SVM) algoritmos dessa modalidade. Já os algoritmos paramétricos assumem que a função de regressão pode ser parametrizada com um número finito de parâmetros (IZBICKI; SANTOS, 2018). E os modelos baseados em estatística e probabilidades assumem que uma distribuição gaussiana descreve os dados de entrada, permitindo derivações estatísticas descrevendo essas distribuições (LEITE; ROSA, 2012; LIU, 2015; MUSIAL; BOJANOWSKI, 2022), como o Maximum Likelihood Classifier (MLC).

Estudos têm sido realizados mostrando o potencial do uso de dados provenientes de dados de sensoriamento remoto para a classificação de estágios de sucessão (LI, G. *et al*, 2011; LU, D. *et al*, 2014; PIAZZA, A. G. *et al*, 2016; BERVEGLIERI *et al*, 2018; SOTHE, C. *et al*, 2017ab; PINTO, M. F. 2018; BISPO *et al*, 2019). Porém, nenhum dos estudos mencionados

comparou dados multiespectrais de quatro plataformas orbitais CBERS-4, Planet, Sentinel-2 e Landsat-8 combinados com informações derivadas de LiDAR, para mapeamento de estágios sucessionais. Assim, este trabalho tem por objetivo propor uma metodologia para a classificação do uso e cobertura do solo e de estágios de sucessão florestal em uma área de floresta subtropical no sul do Brasil, combinando dados de sensoriamento remoto, como o LiDAR/ALS e dados multiespectrais provenientes de diferentes sensores. Dos dados LiDAR, foi extraído o canopy height model (CHM) que foi validado com dados de campo e utilizado como atributo em algumas das abordagens de classificação. Dos quatro sensores multiespectrais, foram extraídos métricas de textura e índices de vegetação. Além de serem testados conjuntos de dados compostos por diferentes atributos e sensores, também foram comparados diferentes métodos de classificação, um método de classificação paramétrico tradicional, MLC, e dois métodos de aprendizado de máquina, RT e SVM.

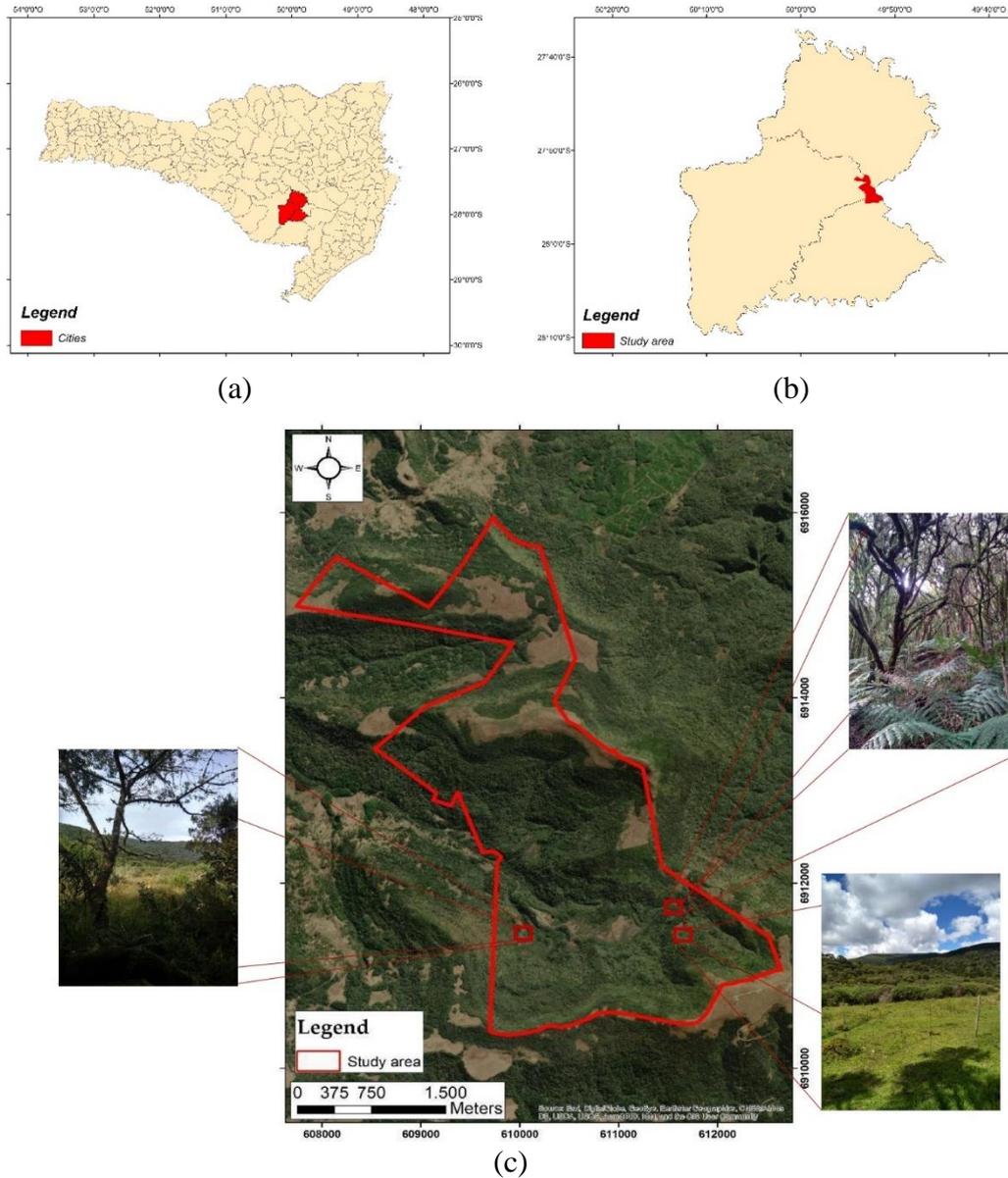
A avaliação da acurácia das classificações foi feita através de dados de inventário florestal convencional e fotointerpretação a partir das imagens da plataforma orbital CBERS-4A e de imagens VANT.

6 MATERIAL E MÉTODOS

6.1 ÁREA DE ESTUDO

A área de estudo pertence a uma empresa florestal do setor catarinense, e está localizada nos municípios de Bocaina do Sul, Painel e Urupema, no estado de Santa Catarina. Essa região possui um clima Cfb (temperado mesotérmico úmido e verão ameno) (KÖPPEN, 1948). A vegetação predominante na região é classificada como Floresta Ombrófila Mista Alto Montana (IBGE, 1992), com altitude média de 1525 metros acima do nível do mar e pertencente ao bioma Mata Atlântica. Segundo os dados de inventário florestal fornecido pela empresa as espécies predominantes da área são *Araucaria angustifolia*, *Dicksonia sellowiana*, *Drimys angustifolia*, *Drimys brasiliensis*, *Eugenia pyriformis*, *Ilex microdonta*, *Ilex paraguariensis*, *Myrcia palustris* e *Siphoneugena reitzii*. Na Figura 9 é apresentado o mapa da região de estudo, contendo os limites da propriedade.

Figura 9. Mapa de localização da área de estudo, sendo: (a) Localização dos municípios no estado de Santa Catarina; (b) Localização da área de estudo dentro dos municípios; (c) Delimitação do perímetro da área de estudo.

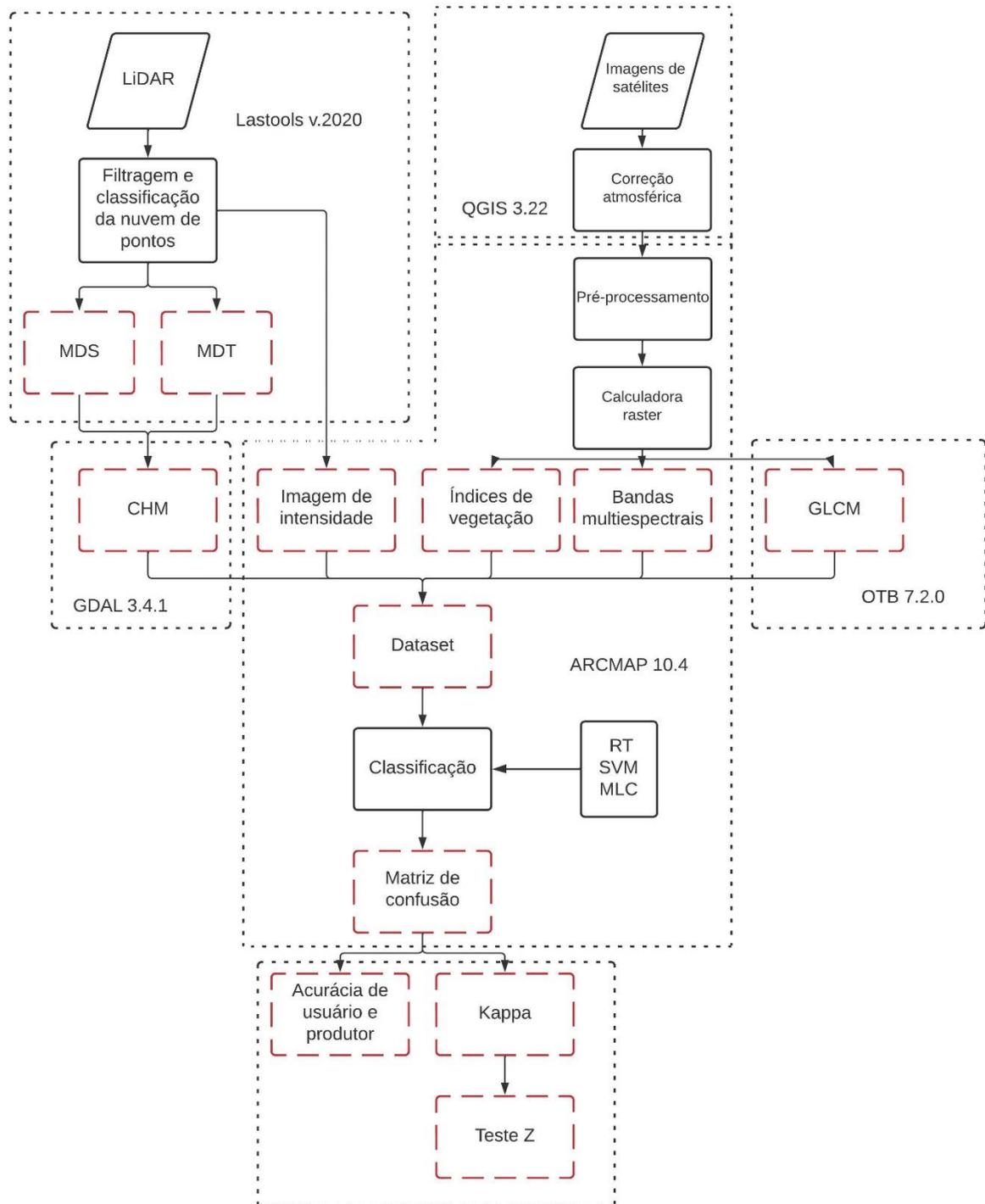


Fonte: Elaborado pelo autor (2022).

6.2 COLETA E PROCESSAMENTO DOS DADOS

Para o desenvolvimento do trabalho foram utilizadas três fontes de dados: uma nuvem de pontos LiDAR, imagens multiespectrais orbitais de quatro sensores, CBERS-4, Landsat-8, Sentinel-2 e Planet, e imagens obtidas por um UAV. A seguir são detalhados as informações e o processamento da nuvem de pontos e é apresentado o fluxograma (Figura 10) de processamento de dados, desenvolvido neste trabalho.

Figura 10. Fluxograma da metodologia de processamento de dados utilizada neste trabalho.



Fonte: Elaborado pelo autor (2022).

6.2.1 LiDAR

Para o levantamento dos dados LiDAR, foi utilizado o sensor e as configurações mostradas no Tabela 1. O recobrimento aéreo foi realizado pela empresa SAI – Serviços Aéreos Industriais

em outubro de 2019. Para a produção do CHM e o processo de classificação e filtragem da nuvem de pontos LiDAR, foi utilizada uma adaptação da metodologia desenvolvida por trabalhos anteriores (KUMAR, 2014; PEREIRA; SCHIMALKI, 2014) no Lastools v.2020.

6.2.2 Descrição dos dados multiespectrais de satélite

Das imagens orbitais do CBERS-4A (sensor WPM) e PlanetScope são datadas de 17 de abril de 2020, sendo utilizadas as cenas “CBERS_4A_WPM_20200312_206_146_L4” e “20201704_142122_94_1069_3B_AnalyticMS_SR”, respectivamente. Já a imagem do Landsat-8/OLI, é de 10 de março de 2020, correspondendo a cena “LC08_L1TP_221079_20200310_20200822_02_T1”, e para o Sentinel-2/MSI, de 24 de abril de 2020 na qual foi utilizada a cena “L2A_T22JFQ_A016366_20200424T132408”.

O download das imagens do CBERS-4A foi realizado pelo website do INPE (<http://www2.dgi.inpe.br/catalogo/explore>), sendo que essa plataforma possui 3 sensores: Multispectral Camera (MUX); Wide Field Imaging Camera (WFI) e Wide Scan Multispectral and Panchromatic Camera (WPM), esse último possui uma alta resolução espacial de 2 metros na banda pancromática, resolução radiométrica de 10 bits e espectral de 4 bandas. As imagens da plataforma Landsat-8/OLI foram obtidas pelo website USGS (<https://earthexplorer.usgs.gov/>), suas imagens possuem resolução espacial variando de 15-30 m e resolução radiométrica de 16 bits, contando também com um acervo histórico que possui imagens desde o início da década de 70.

Os dados Sentinel-2 foram obtidos através do website Copernicus (<https://scihub.copernicus.eu/dhus/#/home>), essa plataforma é constituída por um par de satélites imageadores e possui resolução espacial de (10-60 m), resolução radiométrica de 12 bits e alta resolução espectral com 12 bandas. As três plataformas citadas anteriormente disponibilizam seus dados de forma gratuita. Já as imagens do Planet são disponibilizadas de forma comercial, pelo website SCOON Geospatial (<https://www.scon.com.br/>), essa plataforma possui uma constelação de satélites com resolução radiométrica de 16 bit, 3 metros de resolução espacial e atualmente contam com 8 bandas de resolução espectral, além do diferencial da possibilidade de resolução temporal diária. Utilizou-se as imagens com as datas mais próximas a realização do recobrimento aéreo LiDAR, sendo essas imagens sem nuvem para área de interesse. Como existe um intervalo de tempo entre as datas de obtenção dos dados, pode haver diferenças das características da vegetação entre eles. Na Tabela 2, são apresentadas as bandas utilizadas juntamente com a resolução espectral de cada imagem.

Tabela 1. Características do recobrimento aéreo e do equipamento LiDAR utilizado.

Aparelho LiDAR	Optech ALTM Gemini
Comprimento de onda	1064 nm
Data de aquisição	08/10/2019
Altitude de voo	800 m
Velocidade média de voo	184 km/h
Ângulo de escaneamento	+/- 10°
Repetição do Laser Scanner	70 kHz
Frequência de scaneamento	70 Hz
Número de retornos	1-4
Intensidade	12 bits
Densidade média de pontos	15.38 pontos/m ²

Fonte: Elaborado pelo autor (2022).

Tabela 2. Características espectrais das bandas dos satélites CBERS-4A (WPM), Landsat-8/OLI, Sentinel-2/MSI e PlanetScope utilizadas.

CBERS-4A	LANDSAT-8	Sentinel-2	Planet
0.45-0.52µm(B)	0.45-0.51µm(B)	0.49 µm(B)	0.455-0.515µm(B)
0.52-0.59µm(G)	0.53-0.59 µm(G)	0.56 µm(G)	0.50-0.59µm(G)
0.63-0.69µm(R)	0.64-0.67 µm(R)	0.665 µm(R)	0.59-0.67µm(R)
0.77-0.89µm(NIR)	0.85-0.88 µm(NIR)	0.842 µm(NIR)	0.78-0.86µm (NIR)
0.45-0.90 µm (PAN)	1.57-1.65 µm(SWIR1)	0.705 µm(Red Edge 1)	
	2.11-2.29 µm(SWIR2)	0.740 µm(Red Edge 2)	
	0.50-0.68 µm(PAN)	0.783 µm(Red Edge 3)	
		0.865 µm(Red Edge 4)	
		1.61 µm(SWIR1)	
		2.19 µm(SWIR 2)	

Fonte: Elaborado pelo autor (2022).

*Nota: B:Azul, G:Verde, R: Vermelho, NIR: Infravermelho próximo, SWIR: Infravermelho de ondas curtas e PAN: Pancromática.

2.2.3 Pré-processamento imagens de satélites

A etapa de pré-processamento e obtenção dos dados derivados foram realizadas para os quatro sensores estudados neste trabalho. As imagens dos satélites CBERS-4A, Planet e Sentinel-2 já possuem correção atmosférica. Já para as imagens do Landsat-8 foi utilizado o complemento *semi-automatic classification plugin* (CONGEDO, 2021) no aplicativo QGIS 3.22. Utilizando a ferramenta *preprocessing*, a correção atmosférica foi realizada com base no arquivo .MTL, fornecido com as imagens.

Após a correção atmosférica, as imagens foram submetidas a um *layer stacking*, tornando possível a obtenção de imagens coloridas, que foram utilizadas para a extração de métricas de textura. Para as imagens que possuíam a banda pancromática foi realizado o cálculo do peso de

nitidez das bandas, esse cálculo é feito considerando a radiância dentro de cada banda (ARCGIS, 2022).

Então foi realizada a fusão entre as composições coloridas e a banda pancromática, com o objetivo de aumentar a resolução espacial das imagens para os dados RGB (*Red, Green, Blue* – Vermelho, Verde, Azul) Landsat-8 e CBERS-4A, sendo utilizado o valor de peso de cada umas das bandas calculado anteriormente e utilizando o método de Gram-Schmidt, que é um método baseado em um algoritmo para ortogonalização de vetores (LABEN, 2000).

Por fim, utilizando a ferramenta *raster calculator* foram gerados os seguintes índices de vegetação apresentados na Tabela 3, para todos os satélites, por meio de operações matemáticas entre as bandas.

Tabela 3. Fórmulas dos índices de vegetação utilizados no estudo.

Índices de Vegetação	Fórmulas	Referências
NDVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R})$	(ROUSE <i>et al</i> , 1973)
EVI	$2.5(\text{NIR}-\text{R})/(\text{L1}+\text{NIR}+\text{C1}*\text{R}-\text{C2}*\text{B}+1)$	(JUSTICE <i>et al</i> , 1998)
TGI	$-0.5*(190*(\text{R}-\text{G})-120*(\text{R}-\text{B}))$	(HUNT JR <i>et al</i> , 2013)
SAVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R}+\text{L2}) * (1+\text{L2})$	(HUETE, 1988)
VARI	$(\text{G}-\text{R})/(\text{G}+\text{R}-\text{B})$	(GITELSON <i>et al</i> , 2002)

Fonte: Elaborado pelo autor (2022).

*Nota : L1 = 1 ; C1 = 6; C2 = 7.5; L2 = 0.5.

A composição de imagens coloridas foi convertida em uma imagem de nível de cinza, a partir da aplicação de pesos para cada banda e normalização da imagem de saída utilizando a função grayscale (ARCGIS, 2022) no programa ArcGIS. Então, a imagem em níveis de cinza foi importada para o aplicativo OTB monteverdi para a extração das informações de textura. Utilizando a ferramenta Haralick (HARALICK; SHANMUGAM; DISNTEIN, 1973) de forma empírica e levando em consideração o tamanho dos pixels das imagens, foi definida a janela de busca (3x3, 5x5 e 7x7), a direção sudoeste (SOTHE et al, 2017 (a)) e como resultado foram extraídas as informações de energia, entropia, correlação, contraste, homogeneidade de variância. Esse processo realizado para todas as imagens de satélites estudadas. Ao final do pré-processamento todos os dados foram reamostrados para 1 m de resolução espacial.

6.2.4 Veículo Aéreo Não Tripulado (VANT)

Um dos métodos utilizados para avaliar a classificação de referência foi através de imagens de altíssima resolução espacial adquiridas pelo VANT Parrot Blue Grass. Na Tabela 4 são especificadas algumas características do drone.

Tabela 4. Características do Veículo Aéreo Não Tripulado-

Características	Parrot Blue Grass
Câmera	Parrot Sequoia e RGB 16 MP
Autonomia de voo (min)	25
Peso (g)	1850
Sensor multiespectral	Verde, vermelho, RedEdge e infravermelho
Sensores de navegação	GPS + GLONASS
Alcance máximo	2 km
Resolução Espacial	2 cm
Sistema de navegação inercial	Sim
Sensor ultra-sônico	Sim

Fonte: Elaborado pelo autor (2022).

Para o planejamento de voo de algumas faixas sobre área de estudo, foi utilizado o software Pix4Capture, programando-se a altura de voo de 120 m, sobreposição frontal e lateral de 80% e velocidade do voo de 5 m/s, e como resultado o aplicativo forneceu informações do tamanho do pixel de 11.31 cm/pixel e tempo aproximado de voo de 13 minutos, com base nos parâmetros estabelecidos.

Após a coleta das imagens, estas foram submetidas a um processo de ortorretificação utilizando o software WebODM, esse aplicativo foi escolhido por ser gratuito. Para a obtenção da nuvem de pontos fotogramétrica da área de interesse foi utilizado o parâmetro de processamento 'High Resolution' deixando suas opções como padrão. As imagens do VANT foram utilizadas para validar a fotointerpretação gerada a partir da imagem CBERS-4A e dos dados derivados do LiDAR. O voo com o VANT foi realizado em abril de 2021, uma data distante da qual foi realizado o levantamento com o LiDAR, podendo ocasionar uma diferença da altura da vegetação e da composição de alguns indivíduos de determinado local.

6.2.5 Classificação de referência

A produção da classificação de referência, ou verdade de campo, foi realizada com o auxílio das informações do inventário florestal convencional e do CHM (Apêndice 1), e da

fotointerpretação da área de interesse utilizando a imagem do satélite CBERS-4A, que apresenta a melhor resolução espacial entre as imagens utilizadas. Utilizaram-se as imagens geradas a partir do VANT para validar a fotointerpretação, já que foram realizados voos apenas em algumas partes da área de estudo, e essas não forneceram informações suficiente para gerar a classificação de referência. As alturas fornecidas pelo CHM foram as principais informações utilizadas para a delimitação das classes, juntamente com a análise visual da imagem CBERS-4A.

6.2.6 Composição dos dados

Para cada satélite utilizado (CBERS-4A, Sentinel-2, Landsat-8 e Planet) foram criados 11 conjuntos de dados (*datasets*), de forma a utilizar todas as informações disponíveis para a classificação. Na Tabela 5, apresentam-se as combinações das informações.

6.3 CLASSIFICAÇÃO DAS IMAGENS

Para realizar a classificação das imagens, foram criados polígonos de treinamento e validação, necessários para fornecer ao algoritmo de classificação informações sobre as classes de interesse.

6.3.1 Coleta de amostras

Para a criação das amostras, foram gerados polígonos para as diversas classes das áreas, utilizando a ferramenta *Region of interest* (ROI) do software ArcMAP, baseados apenas na análise visual. Então, esses dados foram divididos aleatoriamente entre amostras de treinamento ($\pm 70\%$) e validação ($\pm 30\%$). As classes amostradas foram “campo”, “água”, estágio inicial “SS1”, estágio intermediário “SS2” e estágio avançado “SS3”. Na tabela 6, são apresentadas as áreas obtidas a partir da classificação de referência de cada classe em hectares (ha) e na Tabela 7 a quantidade de pixels amostrados por classe. Já no Apêndice 1, é mostrada a distribuição das amostras na área de interesse.

Tabela 5. Conjuntos de dados e número de variáveis como dado de entrada para a classificação supervisionada.

	Dataset	Número de variáveis			
		CBERS-4A	Landsat-8	Sentinel-2	PLANET
1	Bandas do satélite	4	6	10	4

2	Bandas do satélite + CHM (LiDAR)	5	7	11	5
3	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR)	6	8	12	6
4	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + NDVI	7	9	13	7
5	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + EVI	7	9	13	7
6	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + TGI	7	9	13	7
7	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + SAVI	7	9	13	7
8	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + VARI	7	9	13	7
9	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 3x3	13	15	19	13
10	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 5x5	13	15	19	13
11	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 7x7	13	15	19	13

Fonte: Elaborado pelo autor (2022).

Tabela 6. Classes encontradas dentro da área de pesquisa e suas respectivas áreas em hectares.

Classe	Área (ha)
Campo	125.8
Água	0.3
SS1	118.1
SS2	349.6
SS3	378.9
Total	972.7

Fonte: Elaborado pelo autor (2022).

Tabela 7. Quantidade de pixels (1x1 m) para as amostras de treinamento e validação

Classe	Amostras		Total
	Treinamento	Validação	
Campo	164410	85816	250226
Água	975	788	1763
SS1	82332	39157	121489
SS2	144005	67030	211035
SS3	97036	49120	146156

Fonte: Elaborado pelo autor (2022).

6.3.2 Classificação Supervisionada

Para a classificação dos *datasets* da Tabela 5, foram empregados três algoritmos de classificação: o RT, SVM e MLC pelo software ArcMAP 10.4. Para treinar a classificação do algoritmo, primeiramente foram feitos testes empíricos para definir os parâmetros de cada classificador. Para o RT foram utilizados os parâmetros de 50 para o número máximo de árvores, 30 para a profundidade máxima de árvore e 1000 para o número máximo de amostras por classe, com os atributos de cor e média. Já para o algoritmo SVM foi utilizado o parâmetro de 500 como número máximo de amostras por classe e os atributos de cor e média, no aplicativo ArcMAP não está disponível a informação de qual kernel é utilizado pela ferramenta. Já o MLC utilizou apenas os atributos de cor e média.

Os classificadores citados anteriormente ajustaram um modelo de classificação utilizando como base as amostras de treinamento, sendo realizada para cada um dos *datasets* dos quatro satélites imageadores. A imagem classificada foi gerada a partir da classificação do respectivo *dataset* utilizando o modelo treinado na etapa anterior. A matriz de confusão (MILLER; NICELY, 1955) foi gerada a partir de pontos de avaliação criados dentro dos polígonos de validação. Esses pontos extraíram a informação de classes da imagem classificada e classificação de referência para a construção da matriz comparando essas duas informações.

6.4 AVALIAÇÃO DA ACURÁCIA

A partir da matriz de confusão, foi possível calcular a acurácia global, índice Kappa ponderado (COHEN, 1968), desvio padrão do índice Kappa e seus valores mínimo e máximo. Além disso, foi calculada a acurácia do produtor e usuário. Ainda foi verificado se havia diferença estatística entre os valores de altura média do inventário e do CHM, utilizando o teste t-Student, e entre os valores de Kappa, utilizando o teste z (MA; REDMOND, 1995), ambos os testes possuem 95% de significância.

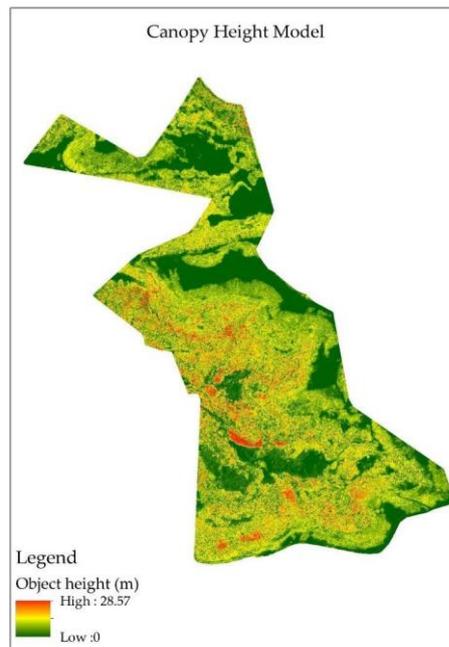
7 RESULTADOS

7.1 MODELO DIGITAL DE ALTURE DE DOSSEL

Uma das etapas desse estudo foi gerar o CHM (ou Modelo Digital de Altura de Dossel) para ser usado como atributo na classificação dos estágios sucessionais. O CHM foi gerado partir da classificação e filtragem da nuvem de pontos LiDAR (Figura 11) e subtração entre o

MDS e o MDT (PEREIRA; SCHIMALSKI, 2014). Esse produto forneceu as informações sobre a estrutura vertical da vegetação, como alternativa a obtenção de dados de campo. Na Tabela 8 é apresentada a comparação entre as alturas médias fornecidas pelo inventário florestal convencional e a altura média do CHM.

Figura 11. Modelo Digital de Altura de Dossel para a área de estudo.



Fonte: Elaborado pelo autor (2022).

As alturas médias do CHM não apresentaram diferença estatística em relação às alturas média do inventário florestal, apresentando um valor de $p = 0,41$ no teste t com significância de 95%.

Tabela 8. Comparação da altura média das árvores mensuradas no inventário florestal convencional com a altura das árvores no CHM.

Parcela	Altura média pelo inventário (m)	Altura média pelo CHM (m)
1	10.05	12.05
2	10.46	9.9
3	10.94	12.05
4	6.89	6.25
5	8.72	9.9
6	6.12	8
7	8.63	9.9

7.2 CLASSIFICAÇÕES DE ESTÁGIO SUCESSIONAL DA VEGETAÇÃO

Foi calculado o índice Kappa para cada um dos conjuntos de dados utilizando os três diferentes algoritmos (RT, SVM e MLC). Na Tabela 9, são apresentados os valores Kappa resultantes para cada dataset.

Tabela 9. Valores obtidos para o índice Kappa, resultante da classificação supervisionada (RT, SVM e MLC) das imagens dos sensores CBERS-4A, Sentinel-2, Planet e Landsat-8. Os melhores resultados de cada dataset são apresentados em negrito.

Dataset	CBERS-4A			Sentinel-2			Planet			Landsat-8		
	RT	SVM	MLC	RT	SVM	MLC	RT	SVM	MLC	RT	SVM	MLC
1	0,7 9	0,80	0,76	0,81	0,90	0,88	0,65	0,68	0,63	0,88	0,92	0,93
2	0,8 5	0,88	0,85	0,85	0,90	0,90	0,79	0,84	0,81	0,91	0,93	0,94
3	0,8 7	0,86	0,88	0,86	0,93	0,91	0,81	0,84	0,83	0,90	0,94	0,92
4	0,8 5	0,88	0,85	0,86	0,91	0,92	0,82	0,84	0,79	0,89	0,95	0,82
5	0,8 7	0,88	0,86	0,84	0,93	0,92	0,84	0,84	0,81	0,89	0,95	0,91
6	0,8 7	0,87	0,82	0,86	0,91	0,90	0,83	0,82	0,82	0,92	0,94	0,83
7	0,8 8	0,88	0,86	0,84	0,91	0,89	0,83	0,85	0,80	0,91	0,94	0,81
8	0,8 8	0,87	0,87	0,86	0,89	0,90	0,83	0,84	0,82	0,91	0,95	0,74
9	0,8 6	0,86	0,59	0,86	0,91	0,65	0,84	0,84	0,83	0,87	0,90	0,84
10	0,8 8	0,86	0,52	0,84	0,91	0,61	0,87	0,84	0,85	0,90	0,87	0,48
11	0,8 8	0,85	0,49	0,85	0,91	0,60	0,84	0,84	0,85	0,81	0,89	0,90

É possível verificar que as imagens dos satélites Sentinel-2 e Landsat-8 apresentaram resultados com maior índice Kappa, alcançando respectivamente 0,93 (Sentinel-2, classificador SVM, dataset 3 e 5) e 0,95 (Landsat-8, classificador SVM, dataset 4 e 5). Já para as imagens dos satélites CBERS-4A o melhor resultado foi de 0,88 (classificar SVM, dataset 2, 4 e 5, e classificador MLC, dataset 3), e para o Planet, o melhor resultado foi de 0,87 com o classificador RT e dataset 10.

Dentre os 3 classificadores, o SVM apresentou o maior Kappa para duas das quatro plataformas de dados utilizados no estudo, sendo essas para Sentinel-2 o dataset 3 e para o Landsat-8 com o dataset 4. Para o conjunto de dados do CBERS-4A, diversas classificações obtiveram um índice Kappa de 0,88, valor máximo para esse sensor, porém o dataset 2 com o

algoritmo SVM foi a combinação com menor quantidade de dados adicionais, que alcançou esse valor para o índice de concordância. A classificação da imagem Planet foi a única que, utilizando as informações de textura, obteve um valor de Kappa superior aos outros conjuntos de dados, sendo que o dataset 10 alcançou um Kappa de 0,87, utilizando o classificador RT. Todas as melhores classificações são consideradas excelentes (LANDIS; KOCH, 1977). Os piores resultados foram alcançados na utilização do classificador MLC em conjuntos com os datasets que utilizam as métricas de texturas, para os sensores CBERS-4A e Landsat-8, sendo esses classificados como boa (LANDIS; KOCH, 1977).

A seguir são apresentadas as matrizes de confusão e os valores de acurácia de usuário e produtor para a melhor classificação dos sensores CBERS-4A (Tabela 10), Sentinel-2 (Tabela 11), Landsat-8 (Tabela 12) e Planet (Tabela 13). Optou-se pela utilização de datasets contendo menos atributos, quando o resultado do índice Kappa foi igual.

Tabela 10. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem CBERS-4A, com a dataset 2* utilizando o classificador SVM.

Classe	Campo	Água	SS1	SS2	SS3	Total	AU
Campo	355	0	0	0	0	355	1
Água	0	9	0	0	1	10	0,90
SS1	0	0	162	6	11	179	0,91
SS2	0	0	8	253	51	312	0,81
SS3	0	0	1	9	140	150	0,93
Total	355	9	171	268	203		
AP	1	1	0,95	0,94	0,69		

Tabela 11. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Sentinel-2, com o dataset 3* utilizando o classificador SVM.

Classe	Campo	Água	SS1	SS2	SS3	Total	AU
Campo	355	0	0	0	0	355	1
Água	0	10	0	0	0	10	1
SS1	0	0	162	20	4	186	0,87
SS2	0	0	2	249	22	273	0,91
SS3	0	0	1	6	175	182	0,96
Total	355	10	165	275	201		
AP	1	1	0,96	0,93	0,92		

Tabela 12. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Landsat-8, com a dataset 4* utilizando o classificador SVM.

Classe	Campo	Água	SS1	SS2	SS3	Total	AU
Campo	355	0	0	0	0	355	1
Água	0	10	0	0	0	10	1
SS1	0	0	162	1	2	165	0.98
SS2	0	0	7	267	17	291	0.92
SS3	0	0	2	5	178	185	0.96
Total	355	10	171	273	197		
AP	1	1	0.90	0.97	0.94		

Tabela 13. Matriz de confusão, acurácia do usuário e produtor para a classificação de estágio de sucessão da imagem Planet, com a dataset 10* utilizando o classificador RT.

Classe	Campo	Água	SS1	SS2	SS3	Total	AU
Campo	355	0	1	0	0	356	0.99
Água	0	6	0	0	4	10	0.60
SS1	0	0	151	4	26	181	0.83
SS2	0	0	6	237	21	264	0.90
SS3	0	0	4	31	159	194	0.82
Total	355	6	162	272	210		
AP	1	1	0.86	0.83	0.87		

***Nota:** O conjunto de dados das composições pode ser verificado na Tabela 9.

O dataset derivados do CBERS-4A e CHM apresentou acurácia de usuário variando de 0,92 a 1 e acurácia do produtor variando de 0,90 a 1. O dataset empregando Sentinel-2, e derivados do LiDAR, obteve valor mínimo e máximo de 0,87 a 1 e 0,92 a 1, para a acurácia de usuário e produtor, respectivamente. A acurácia de usuário e produtor variou de 0,92 a 1 e 0,90 a 1 para o dataset Landsat-8 em conjunto com LiDAR e NDVI. E a classificação da imagem Planet apresentou a maior variação dentre as demais com acurácia do produtor variando de 0,83 até 1 e acurácia de usuário variando de 0,60 a 0,99.

Na Tabela 14, apresentam-se o intervalo de confiança para as melhores classificações de cada sensor orbital e LiDAR estudado, com uma significância de 95%.

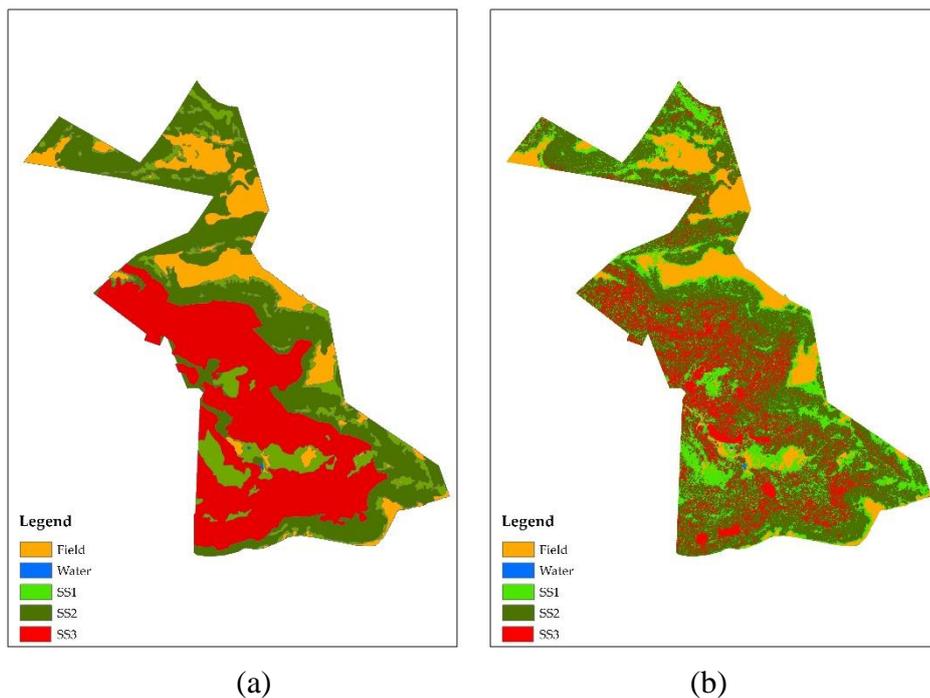
Tabela 14. Índice Kappa, valores mínimos e máximos para o índice e desvio padrão, para as melhores classificações das imagens CBERS-4A , Sentinel-2, Landsat-8 e Planet .

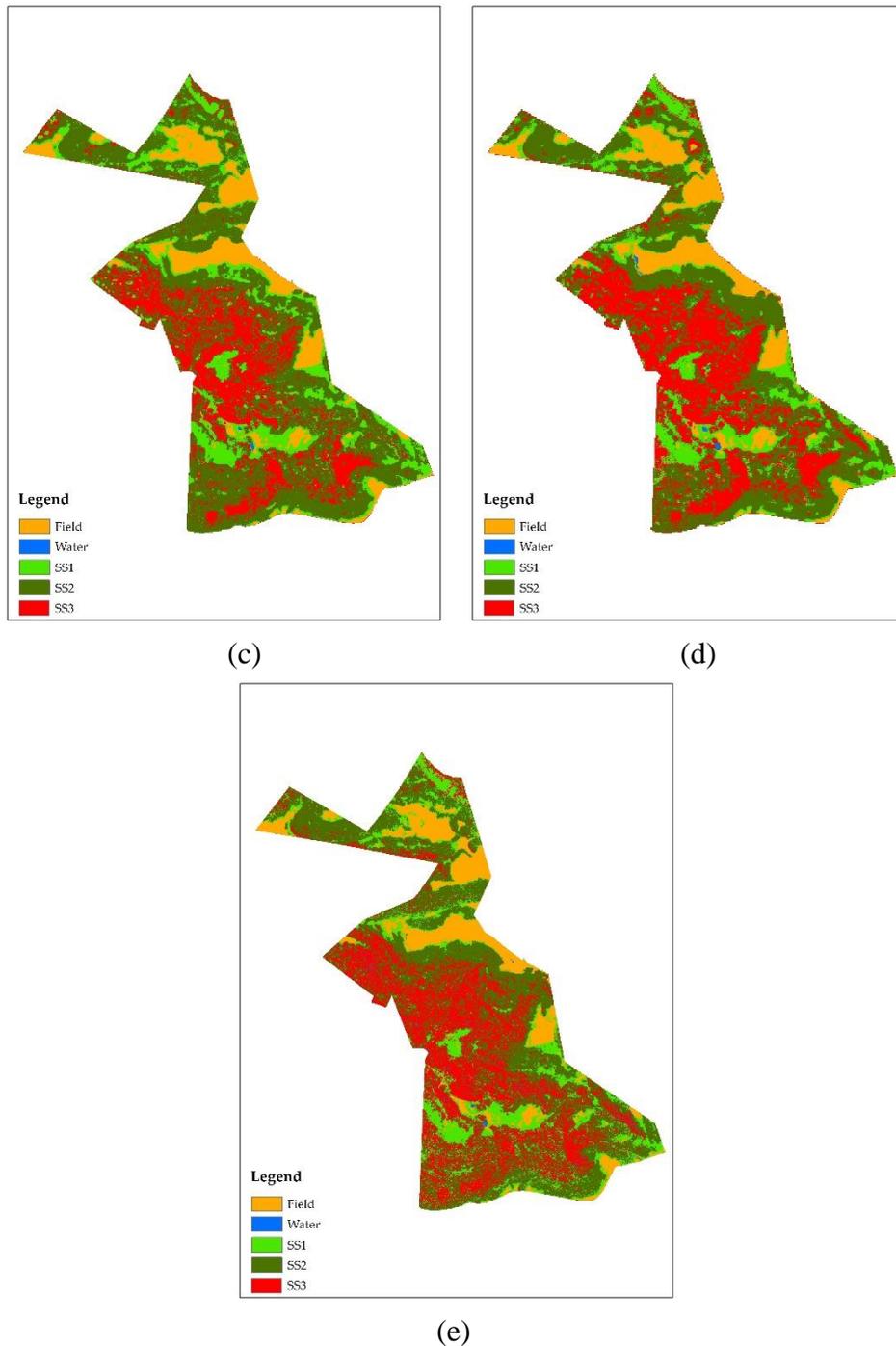
Sensor	Classificador	Dataset	Kappa ponderado	Desvio padrão	Mínimo	Máximo
CBERS-4A	SVM	2*	0.88	0.0119	0.8603	0.9068
Sentinel-2	SVM	3*	0.93	0.0097	0.9066	0.9446
Landsat-8	SVM	4*	0.95	0.0077	0.9387	0.9691
Planet	RT	10*	0.87	0.0125	0.8441	0.8931

***Nota:** O conjunto de dados dos datasets pode ser verificado na Tabela 9. Significância de 95%.

A classificação para o Landsat-8 utilizando o classificador SVM e o dataset 4 foi a que obteve menor variação com desvio padrão de 0,0077, e possui como valores mínimo e máximo 0,9387 e 0,9691, respectivamente. Com 0,0125 de desvio padrão, a imagem Planet com classificador RT e a dataset 10 foi a que apresentou maior variação, com o índice Kappa podendo alcançar 0,8931 como valor máximo e 0,8441 como valor mínimo. A seguir na Figura 12 temos a representação gráfica das classificações em comparação com a classificação de referência.

Figura 12. Representação gráfica das classificações dos datasets e sensores: (a) Classificação referência; (b) Classificação da imagem CBERS-4A utilizando o classificador SVM e dataset 2*; (c) Classificação da imagem Sentinel-2 utilizando o classificador SVM e o dataset 3*; (d) Classificação da imagem Landsat-8 utilizando o classificador SVM e o dataset 4*; (e) Classificação da imagem Planet utilizando o classificador RT e o dataset 10*.





Fonte: Elaborado pelo autor (2022).

***Nota:** O conjunto de dados das composições pode ser verificado na Tabela 9.

Na Tabela 15, comparou-se os melhores índices Kappa de cada classificador para cada um dos quatro satélites em conjunto com os dados LiDAR, empregando-se o teste Z com significância de 95%.

Tabela 15. Valor do teste z para comparar as melhores classificações de cada algoritmo de classificação, para as imagens CBERS-4A, Sentinel-2, Landsat-8 e Planet.

Sensor	Datasets	Z	Valor crítico
CBERS-4A	SVM2* - RT11*	0.3122	1.96
	SVM2* - MLC3*	0.2175	
Sentinel-2	SVM3* - RT8*	5.3904	0.7511
	SVM3* - MLC5*	0.7511	
Landsat-8	SVM4* - RT6*	3.8575	2.1378
	SVM4* - MLC2*	2.1378	
Planet	RT10* - SVM7*	1.6503	1.0931
	RT10* - MLC11*	1.0931	

* **Nota:** O conjunto de dados das composições pode ser verificado na Tabela 9. Significância do teste foi de 95%.

Os índices Kappa das melhores classificações do CBERS-4A de cada algoritmo de classificação não apresentaram diferença estatística, sendo que o valor de $z < 1.96$. Para a classificação da imagem Sentinel-2 com o algoritmo SVM e composição 3 quando comparada a classificação do algoritmo RT e composição 8, apresentou diferença estatística. Comparando os índices de concordância da melhor classificação da imagem Landsat-8, que foi SVM com composição 4, em relação aos dos algoritmos RT e MLC, para ambos a diferença foi estatisticamente significativa. Por fim, para as comparações entre Kappa das classificações da imagem Planet não apresentaram diferença estatística.

8 DISCUSSÃO

Nesse estudo, primeiramente produziu-se um CHM a partir de dados LiDAR para ser utilizado como atributo adicional além das imagens multiespectrais na classificação dos estágios sucessionais. As alturas médias obtidas pelo CHM não apresentaram diferença estatística pelo teste t quando comparadas às alturas médias avaliadas a campo pelo inventário florestal convencional. Este resultado é condizente com o estudo similar, que apresentou forte correlação entre o CHM e medidas a campo, apresentando coeficiente de determinação e RMSE, respectivamente, de 0,85 a 0,92 e 2,7 a 3,5 m (OH *et al.*, 2022). Nosso estudo também apresentou concordância ao estudo prévio que concluiu que o LiDAR/ALS possui uma boa performance para descrever diretamente a altura das árvores em florestas naturais (KOTIVUORI; KORHONEN; PACKALEN, 2016).

As classificações dos estágios de sucessão para os quatro sensores orbitais e LiDAR utilizados alcançaram resultados excelentes segundo a classificação do Kappa (LANDIS; KOCH, 1977). Para a imagem PlanetScope o melhor resultado foi de 0,87 para o índice de concordância, sendo esse resultado próximo a de um estudo prévio que obteve um Kappa de

0,90, também utilizando as imagens Planet em conjunto com dados LiDAR e o classificador RT para o mapeamento da cobertura da terra e avaliação da vegetação (SZOSTAK *et al*, 2021). Para o nosso conjunto de dados Planet, tanto as informações derivadas do LiDAR como as métricas de textura da imagem contribuíram para um aumento da acurácia da classificação, sendo também o classificador RT que alcançou o melhor resultado, porém esse não mostrou diferença estatística com as melhores classificações dos outros algoritmos utilizados.

Para a imagem do satélite CBERS-4A, a melhor classificação obteve um Kappa de 0,88, quando utilizado o classificador SVM com a dataset 2, sendo essa composta pelas bandas multiespectrais do sensor em conjunto com o CHM. Em estudo similar, foi realizada a classificação de áreas ambientais preservadas ou não preservadas utilizando redes neurais convolucionais e imagens do CBERS-4A para o treinamento do algoritmo, e imagens do Sentinel-2 para testes, e obteve uma acurácia de 0,87 (MIRANDA *et al*, 2021).

A classificação utilizando as imagens do satélite Landsat-8 foi a que obteve maior índice de concordância dentre os demais utilizados. O Kappa de 0,95 foi alcançado utilizando as imagens multiespectrais do sensor com a adição dos dados derivados do LiDAR (CHM e imagem de intensidade) e índices de vegetação, com o classificador SVM. Sendo esse resultado concordante com (SOTHE *et al*, 2017 (b)) que indicam que a utilização de dados adicionais em conjunto com as imagens Landsat-8 produz um Kappa classificado como razoável a excelente, para o estudo da vegetação. Os autores (SOTHE *et al*, 2017 (a)) mostram que a melhor classificação dos estágios de sucessão utilizando imagens Landsat-8 com um Kappa de 0,88, também com a utilização do índice de vegetação NDVI, porém em conjunto com o índice RS e métricas de textura, e utilizando o classificador RF.

A composição das bandas multiespectrais com os dados derivados do LiDAR (dataset 3) foi a que apresentou melhor resultado para o Sentinel-2, alcançando um Kappa de 0,93. Um estudo que classificou uma floresta ETF (Europe forest types) em três classes de vegetação (coníferas puras, folhosas puras e misto) obteve um kappa de 0,83, utilizando imagens do Sentinel-2 (PULETTI; CHIANUCCI; CASTALDI, 2018). Um estudo em uma região similar à estudada no presente trabalho utilizou imagens do Sentinel-2 para a classificação de estágios de sucessão e alcançou uma acurácia global de 0,98 (SZOSTAK, Hawryło; Piela, 2018).

O melhor desempenho dos satélites Sentinel-2 e Landsat-8 pode ser atribuído à melhor resolução radiométrica e à utilização de mais bandas espectrais, que podem auxiliar na discriminação da resposta espectral de cada alvo pelo sensoriamento remoto (JENSEN, 2009). Um estudo apresentou seus melhores resultados para a classificação de estágios de sucessão

com a utilização de imagens Landsat-8 do que com imagens RapidEye. O Rapideye possui características distintas do Landsat-8, como resolução espacial de 5 m, espectral de 4 bandas e radiométrica de 12 bits e os resultados foram atribuídos a melhor resolução radiométrica do Landsat-8, e a maior variabilidade espectral no interior da mesma classe pelo Rapideye (SOTHE *et al*, 2017(a))

Para a classe campo, todas as imagens obtiveram uma boa classificação com acurácia do produtor de 1 e acurácia do usuário variando de 0,99 a 1, indicando que a classe teve baixa confusão com as demais classes, e que o modelo possui uma boa separabilidade dessa classe das demais. Para a classe água, a acurácia do produtor também foi de 1, e a acurácia do usuário variou de 0,6 (Planet) a 1 (Sentinel-2, CBERS-4A e Landsat-8), indicando que apenas para a imagem Planet teve maior confusão entre essa classe e a SS3. Já para as classes de vegetação, a classe SS2 teve maior confusão com a classe SS3, em relação a acurácia do usuário variando de 0,81 a 0,92. Já a classe SS3 foi a que teve maior confusão quanto a acurácia do produtor, com uma variação de 0,92 a 0,69, sendo essa a proporção de pontos atribuídos corretamente à classe pelo classificador, tendo maior confusão com a classe SS2. Estudos anteriores (PINTO, 2018; SOTHE, 2015) também encontraram esse problema de diferenciação entre os estágios de sucessão florestal.

Quanto aos classificadores, o SVM foi o que apresentou melhores resultados em uma visão geral dos índices Kappa, sendo também o melhor classificador para as imagens Sentinel-2 e Landsat-8, concordando com os estudos de (SWETANISHA; PANDA; BEHERA, 2022; XIE; NICULESCU, 2021; JAMALI, 2021), que mostraram que esse classificador possui melhores resultados para imagens com menor resolução espacial. Já o classificador RF possui uma melhor performance com imagens de resolução maior (XIE; NICULESCU, 2021), resultado também alcançado em nosso trabalho, sendo que o classificador RT, semelhante ao RF, obteve um maior índice de concordância para a imagem Planet, porém esse não diferiu estatisticamente dos outros classificadores utilizados. O algoritmo MLC não resultou na melhor classificação para nenhuma das imagens utilizadas, porém seus resultados não possuem diferença estatística para as melhores classificações das imagens CBERS-4A, Sentinel-2 e Planet, com Kappa variando de 0,85 a 0,94 para as melhores classificações, semelhante ao estudo de (BISPO *et al*, 2019; SILVA, 2020), que apresentaram uma acurácia geral para a classificação de estágios de sucessão utilizando dados derivados do TanDEM-X e VANT de respectivamente 0,87 e 0,89 para, com esse classificador.

Observou-se que a adição de informações com as bandas multiespectrais contribuiu para a melhoria da acurácia da classificação dos estágios de sucessão. Para a maioria das classificações a adição dos dados derivados do LiDAR melhoraram a assertividade, já que esses dados fornecem informações sobre a estrutura vertical da vegetação (SOTHE *et al*, 2018). Esse resultado concorda com estudos anteriores, em que a acurácia de classificação aumenta quando os dados de LiDAR são adicionados ao modelo (DALPONTE; BRUZZONE; GIANELLE, 2012; CHO *et al*, 2012). Quanto às métricas de textura, todas apresentaram resultados semelhantes quando adicionadas a composição de classificação, sendo que as informações de textura são atributos importantes para a diferenciação das classes de vegetação (SOTHE *et al*, 2017(b)), porém em alguns casos essas informações não contribuíram com a melhoria do índice de concordância, esse problema ocorreu principalmente com o classificador MLC.

Os índices de vegetação utilizados mostraram que também podem melhorar as performances das classificações, neste trabalho esses índices se destacaram quando utilizados com o classificador SVM para a imagem Landsat-8, obtendo os melhores resultados dentre as composições utilizadas. Outros estudos (SOTHE *et al*, 2017(a); TASSETI; MALINVERNI; HAHN, 2010; SANNIGRAHI *et al*, 2022) apresentaram um incremento nos resultados da classificação com a utilização de índices de vegetação.

Em relação aos resultados gerais, nosso trabalho alcançou resultados excelentes com o maior Kappa de 0,95. Esse resultado é concordante com outros estudos para a classificação de estágios de sucessão, que obteve Kappa de 0,908 com a utilização dos sensores Landsat-8 e RapidEye (SOTHE *et al*, 2017 (b)). O estudo de (FALKOWSKI *et al*, 2009) realizou a classificação apenas dos estágios de sucessão da vegetação diferenciando em seis classes, utilizando um sensor LiDAR e obteve um Kappa de 0,95. Já a classificação utilizando apenas sensores óticos a bordo de VANT obteve Kappa de 0,946 (SILVA, 2020).

A principal limitação deste estudo foi utilizar apenas a classificação orientada a pixel, pois isso pode reduzir a acurácia para os satélites de melhor resolução espacial, limitando a capacidade de detecção de detalhes em cenas heterogêneas (PINHO; FEITOSA; KUX, 2005). Alguns trabalhos sugerem a classificação orientada a regiões para imagem com alta resolução espacial. Estudos futuros são recomendados utilizando metodologias de segmentação de imagens para realizar a classificação orientada a regiões, e utilizando outros dados derivados do LiDAR, como métricas de percentis.

9 CONCLUSÕES

A metodologia utilizada nesse estudo avaliou três diferentes algoritmos de classificação supervisionada para a diferenciação dos estágios de sucessão e uso e ocupação do solo, em que foram utilizadas diferentes bandas espectrais, índices de vegetação e textura de sensores ópticos com várias especificações em termos de resolução espacial, radiométrica e espectral, além da adição de dados derivados de LiDAR. Sendo que para esse objetivo a metodologia com maior acurácia foi utilizado o algoritmo SVM com os dados do Landsat-8 em conjunto com os dados do LiDAR.

A acurácia da classificação para todos os sensores quando utilizados de forma individual foram satisfatórias, porém com a adição dos dados do LiDAR e os de índices de vegetação e textura, a precisão de classificação foi aumentada. Quanto aos classificadores utilizados, todos alcançaram classificações excelentes, e dentre os três utilizados o SVM foi o que apresentou melhor eficácia.

Esse estudo mostrou que a altura das árvores do dossel florestal pode ser estimada de forma precisa a partir dos dados LiDAR. O CHM produzido não apresentou diferença estatística com alturas das árvores fornecidas pelo inventário florestal convencional. Além disso, ele foi um importante atributo na classificação dos estágios sucessionais da vegetação.

Todos os quatro satélites utilizados (CBERS-4A, Sentinel-2, Landsat 8 e Planet) têm o potencial para alcançar resultados excelentes para a classificação de estágios de sucessão da vegetação na área de estudo. As informações geradas a partir da nuvem de pontos LiDAR mostraram contribuir para o aumento da performance da classificação. Enquanto as informações derivadas das imagens orbitais mostraram ser eficazes com composições e algoritmos de classificação específicos.

Por fim, todos os dados quanto à metodologia apresentada neste estudo são promissores quando empregados na classificação do estágio de sucessão da floresta ombrófila mista, podendo contribuir para o entendimento das características da floresta nativa. Essa metodologia mostra também um grande potencial para uso no licenciamento e fiscalização dessas áreas.

CAPÍTULO III

Transferência de aprendizagem de modelos de machine learning utilizando dados de satélites multiespectrais e ALS para classificação de estágios de sucessão de FOM e FOD.

Bill Herbert Ziegelmaier Neto
Marcos Benedito Schimalski
Veraldo Liesenberg
Camile Sothe

10 INTRODUÇÃO

A Mata Atlântica é considerada um dos hotspots mundial, por possuir uma elevada riqueza biológica, sendo uma das prioridades a conservação de sua biodiversidade, porém atualmente conta com apenas 8% de sua extensão original no Brasil. Para garantir a proteção da sua biodiversidade é necessário o conhecimento sobre o funcionamento de seu ecossistema para que seja realizado o seu manejo (PINTO; BEDE; FONSECA, 2006). Porém, a mensuração da estrutura florestal de forma convencional possui um elevado custo, além da dificuldade de acesso em determinados locais e dependendo de tempo e planejamento para sua execução, adiciona-se ainda a questão de não recorrência ao longo do tempo dessas mensurações (ANDERSEN, H. E.; REUTEBUCH, S. E.; MCGAUGHEY, 2006).

Um dos fatos geradores dessa situação é a heterogeneidade das florestas naturais, dificultando os levantamentos dendrométricos in loco, visto que essas áreas apresentam grande diversidade biológica e densidade de indivíduos. Neste contexto, a diferenciação dos estágios de sucessão florestal é de difícil realização por meio do levantamento em campo, por isso é importante a adoção de metodologias e técnicas empregando dados remotamente situados que possibilitem a classificação e a obtenção de informações sobre a forma do alvo, a textura ou relações entre alvos vizinhos no mapeamento das fases de sucessões florestais (PONZONI, F. J.; SHIMABUKURO, 2009).

Para melhor entendimento sobre o funcionamento das florestas naturais, elas são separadas em diferentes processos de sucessão (KAGEYAMA, P. Y.; BRITO, M. A.; BAPTISTON, [s.d.]). Segundo a legislação brasileira, as normas e modalidades para o manejo da vegetação são diferentes para cada estágio sucessional, tornando sua classificação imprescindível para a fiscalização e licenciamento ambiental (SEVEGNANI, [s.d.]). Além disso, processos de sequestro de carbono também podem ser preditos através de atributos estruturais relacionados aos estágios de sucessão (SHUGART, 2000).

Dentre as diversas fontes de dados do sensoriamento remoto, o sistema LiDAR/ALS (Airborne Laser Scanner) consiste na emissão e recepção de um pulso laser, cujo sensor está embarcado em uma plataforma aerotransportada. Esse sistema é composto pelo sensor laser, um receptor de sistema global de navegação por satélite (GNSS) e um sistema de unidade de medida inercial (IMU), permitindo extrair com precisão os pontos correspondentes a superfícies e objetos no terreno (GIONGO et al., 2010). Essa tecnologia permite a cobertura de uma grande extensão de superfície terrestre, resultando em produtos de alta resolução espacial, podendo chegar a níveis submétricos (D'OLIVEIRA; FIGUEIREDO; PAPA, 2014). Além de fornecer informações da estrutura florestal e do relevo das áreas (JENSEN, 2009), constitui-se uma alternativa à coleta de dados de inventário florestal de forma tradicional.

A utilização de imagens orbitais é outra alternativa pois necessita de menos recursos e permite fornecer informações precisas em menor tempo quando comparados as técnicas de mensuração in loco (JANOTH, [s.d.]). Entre as plataformas orbitais capazes de realizar imageamento óptico terrestre cita-se o satélite Sino-Brasileiro de Recursos Terrestres (CBERS)-4A, que possui 3 sensores: Multispectral Camera (MUX); Wide Field Imaging Camera (WFI) e Wide Scan Multispectral and Panchromatic Camera (WPM). Esse sensor WPM possui uma alta resolução espacial de 2 metros na banda pancromática, resolução radiométrica de 10 bits e espectral de 4 bandas. O par de satélites Sentinel-2 possui resolução espacial de (10-60 m) e resolução radiométrica de 12 bits e alta resolução espectral com 12 bandas. O Landsat-8/OLI possui resolução espacial variando de 15-30 m e resolução radiométrica de 16 bits, é uma plataforma que tem um acervo de imagens desde o início da década de 70. Todas essas três plataformas disponibilizam as imagens gratuitamente. Uma das plataformas comerciais existentes é o Planet/Dove, que possui resolução radiométrica de 16 bits e 3 metros de resolução espacial, e atualmente conta com 8 bandas de resolução espectral, e a possibilidade de resolução temporal diária.

Para processar esses dados oriundos de plataformas orbitais faz-se necessário o uso de interpretação humana que é suscetível a diversas inconsistências quando realiza a classificação de imagens em grandes áreas, devido às limitações visuais humanas. A combinação de técnicas de processamento digital de imagens (FINLAYSON; VAN DER VALK, 1995; GUO et al., 2017), mais especificamente, machine learning em imagens orbitais, tornam-se uma alternativa aos métodos clássicos de mensuração florestal em campo. Em relação a essa metodologia, o machine learning (ML) apresenta bons resultados com dados de alta dimensionalidade utilizando uma quantidade menor de amostras (MOUNTRAKIS; IM; OGOLE, 2011) comparado ao deep learning (DP).

Quanto às metodologias de ML, podemos dividir em paramétricas e não-paramétricas. Os algoritmos paramétricos assumem que uma distribuição gaussiana descreve os dados de entrada, permitindo derivações estatísticas descrevendo essas distribuições (LEITE; ROSA, 2012; MUSIAL; BOJANOWSKI, 2022), como exemplo o algoritmo Maximum Likelihood (MLC). Já os algoritmos não-paramétricos são

utilizados para modelagem e classificação de áreas, aceitando uma variedade de dados como preditores de entrada e não fazem suposições esses dados (MAXWELL; WARNER; FANG, 2018; SEMOLINI, 2002), sendo o Random Tree (RT) e Support Vector Machine (SVM) algoritmos dessa modalidade. Dentro do ML existe uma técnica inovadora que é a transferência de aprendizagem (TA), possuindo a vantagem de economizar tempo de processamento, e podendo alcançar desempenhos superiores e evitando problemas de overfitting dos modelos de ML (KUMAR; TALIB; VERMA, 2021). Então a TL consiste na aplicação de modelos aprendidos anteriormente para solucionar um problema similar em uma nova área (CHEN; ZHANG; OUYANG, 2018). Alguns estudos já mostram as possibilidades de aplicação do TL com diferentes problemas do sensoriamento remoto como estimativa de variáveis florestais, detecção de objetos, análises de mudanças e classificação da vegetação (ASTOLA et al., 2021; CHEN; ZHANG; OUYANG, 2018; LIU et al., 2021; QIAN et al., 2020; XIE; NICULESCU, 2021; ZHU et al., 2019).

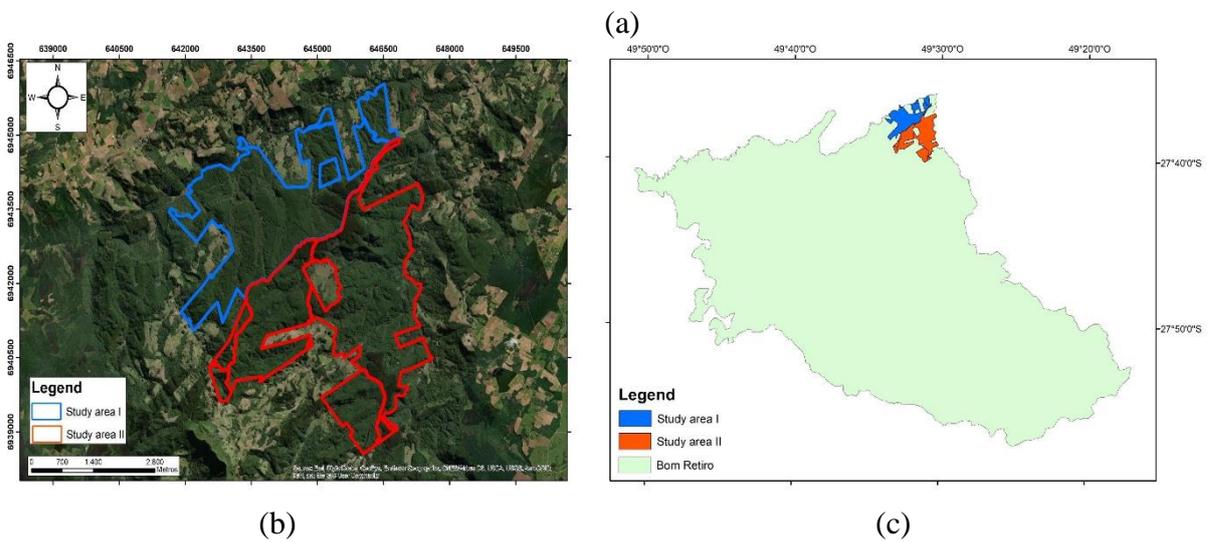
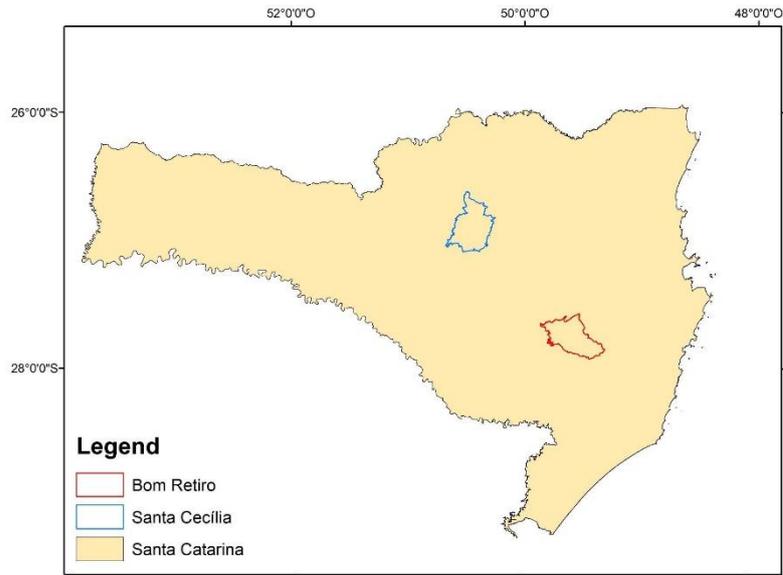
Estudos têm sido realizados mostrando o potencial da TA para a classificação de imagens orbitais (ALEM; KUMAR, 2021). Sendo o objetivo desse trabalho, a avaliação da aplicação dessas metodologias para a classificação do uso e cobertura da terra e da classificação dos estágios de sucessão em fragmentos florestais da Mata Atlântica. Para isso foram geradas informações como índices de vegetação e textura a partir de dados de e imagens obtidas de plataformas orbitais, e informações sobre a estrutura vertical da vegetação produzido com as informações do LiDAR/ALS. Então, foi determinada a acurácia de distinção dos diferentes estágios sucessionais classificados com a TA comparando uma classificação de referência gerada com dados obtidos do inventário florestal convencional e fotointerpretação. Por fim, foi avaliada a eficácia da TA de métodos de classificação paramétricos com métodos de ML.

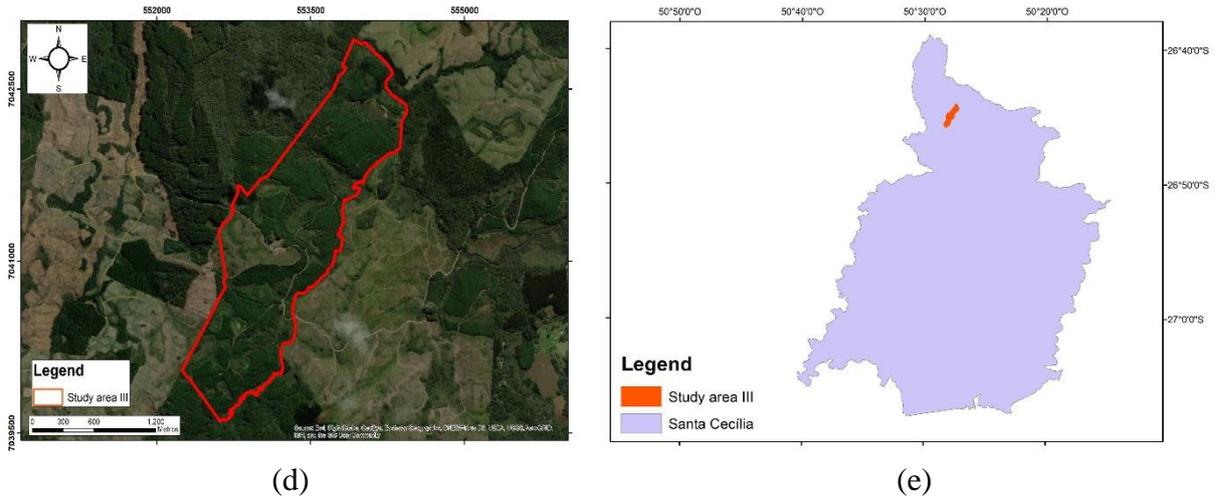
11 MATERIAL E MÉTODOS

11.1 ÁREA DE ESTUDO

O estudo de transferência de aprendizagem foi empregado em três áreas, que pertencem a uma empresa florestal. As áreas (I e II) localizam-se município de Bom Retiro e fazem parte do bioma Floresta Atlântica com a tipologia Floresta Ombrófila Densa Montana (FOD), com altitude média de 640 metros (m). Já a terceira área (III) localiza-se município de Santa Cecília e faz parte também do bioma Floresta Atlântica, porém com uma tipologia Floresta Ombrófila Mista Alto Montana (FOM), com elevação média de 1151 m (IBGE, 1992)[IBGE]. Segundo os dados do inventário florestal fornecido pela empresa, nas áreas I e II as espécies com maior ocorrência são *Pinus spp.*, *Alsephila setosa*, *Ocotea puberula* e *Mimosa scabrella*. E para a área III são as espécies *Mimosa scabrella* Benth., *Dicksonia sellowiana* Hook, *Baccharis uncinella* DC e *Myrsine coriacea* (Sw.) R.Br. ex Roem. & Schult. Na figura 1 são apresentadas as áreas de estudo e sua localização.

Figura 13. Mapa de localização das áreas de estudo. (a) Mapa contendo estado de Santa Catarina e a delimitação dos municípios de Bom Retiro e Santa Cecília. (b) Delimitação das áreas de estudo I e II. (c) Localização das áreas de estudo I e II dentro do município de Bom Retiro. (d) Delimitação da área de estudo III. (e) Localização da área de estudo III dentro do município de Santa Cecília



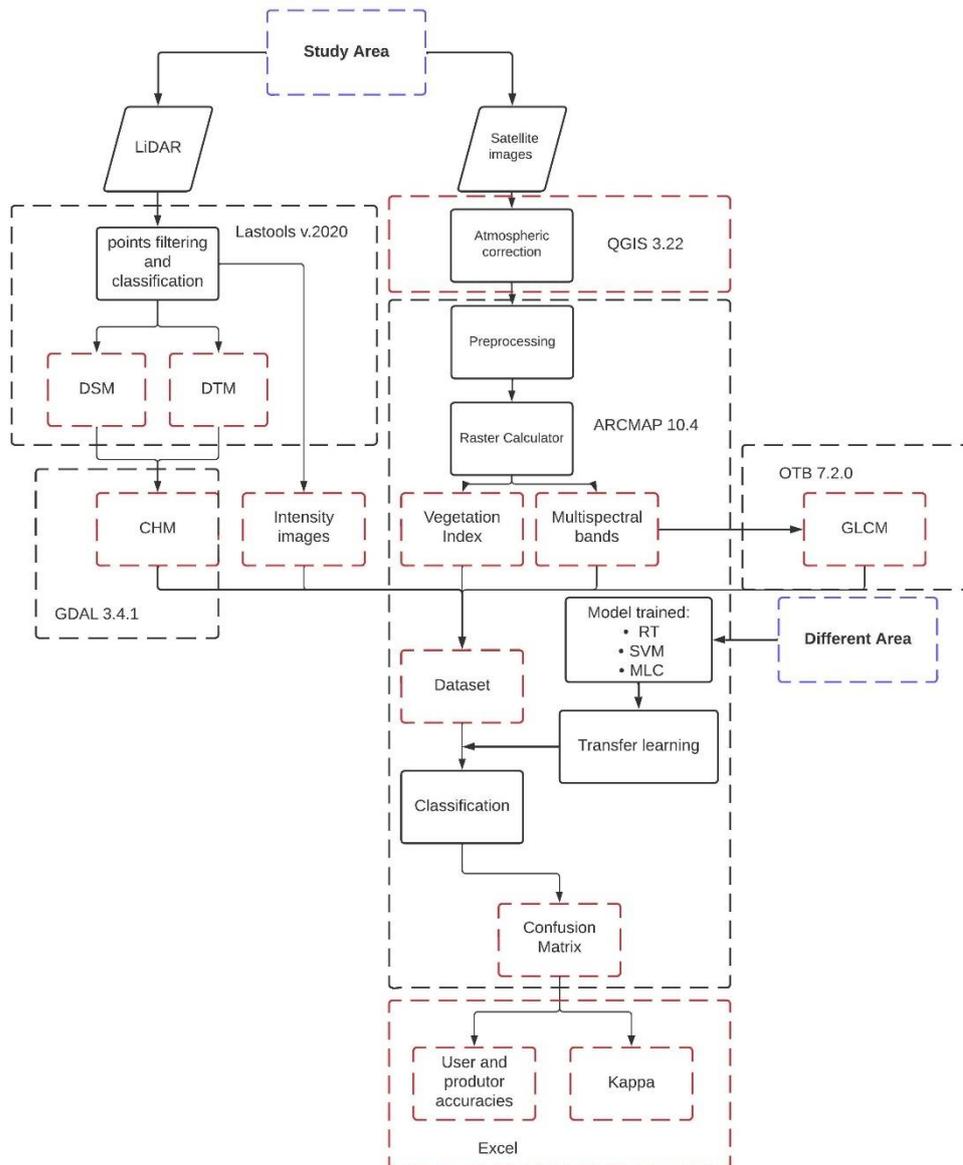


A metodologia utilizada neste estudo foi adaptada do Capítulo II, estudo que realizou o treinamento dos algoritmos, que foram utilizados para a TA.

11.2 COLETA E PROCESSAMENTO DOS DADOS

Neste estudo foram utilizadas duas fontes de dados: imagens orbitais e uma nuvem de pontos LiDAR. O fluxograma na Figura 14 detalha as etapas de processamento empregado neste trabalho.

Figura 14. Fluxograma da metodologia de processamento de dados utilizada neste trabalho.



Na tabela 16 são apresentadas as datas para as coletas dos dados remotamente situados utilizados neste estudo.

Tabela 16. Data de aquisição dos dados utilizados.

	LiDAR	Sentinel 2 (L2A)	CBERS 4A (WPM)	Planet	Landsat 8 (L1C)
Áreas I e II	16/09/2020	22/08/2020	19/09/2020	19/09/2020	11/09/2020
Área III	24/11/2019	15/03/2020	12/03/2020	12/03/2020	10/03/2020

Fonte: Elaborado pelo autor (2022).

Fonte: Elaborado pelo autor (2022).

11.2.1 LiDAR

O levantamento aéreo com aeronave tripulada dos dados LiDAR utilizou o sensor e as configurações apresentadas na Tabela 17, sendo realizado pela empresa SAI – Serviços Aéreos Industriais.

Tabela 17. Características do recobrimento aéreo e do equipamento LiDAR utilizado.

Aparelho LiDAR	Optech ALTM Gemini
Comprimento de onda	1064 nm
Altitude de voo	800 m
Velocidade média de voo	184 km/h
Ângulo de escaneamento	+/- 10°
Repetição do Laser Scanner	70 kHz
Frequência de escaneamento	70 Hz
Número de retornos	1-4
Intensidade	12 bits
Densidade média de pontos - Area I	7.680
Densidade média de pontos - Area II	7.791
Densidade média de pontos - Area III	5.989

Fonte: Elaborado pelo autor (2022).

O processo de classificação e filtragem da nuvem de pontos LiDAR para obter o Canopy Height Model (CHM) foi adaptado da metodologia utilizada nos estudos de (PEREIRA; SCHIMALSKI, 2014; VINOD KUMAR, 2014) no aplicativo Lastools v.2020 via prompt de comando.

11.2.2 Descrição das Imagens de satélites

O download das imagens do CBERS-4A foi realizado pelo website do INPE; as do Landsat-8/OLI, pelo website USGS; do Sentinel-2, pelo website Copernicus e, Planet images were obtained using the Planet Explorer platform and access to them was obtained through enrollment in the Planet's Education and Research (E&R) Program. Como critérios para a

seleção das imagens considerou-se: a cobertura de nuvens e a proximidade com a data do recobrimento aéreo LiDAR. Com exceção dos dados Planet, os outros dados orbitais foram obtidos de forma gratuita nas plataformas citadas. Na tabela 18, são apresentados os identificadores de cada cena utilizados no estudo.

Tabela 18. Cena utilizada de cada sensor orbital para a aplicação da TA.

	Área I e II	Área III
Sentinel 2 (L2A)	L2A_T22JFQ_A018082_20200822T132348	L2A_T22JER_A015794_20200315T132231
CBERS 4A (WPM)	CBERS_4A_WPM_20200919_205_148_L4	CBERS_4A_WPM_20200312_206_147_L4
Planet	20200919_133739_20_1060 e 20200919_133737_20_1060	20200312_123533_104a e 20200312_123532_104a
Landsat 8 (L1C)	LC08_L2SP_220079_20200911_20200919_02_T1	LC08_L2SP_221079_20200310_20200822_02_T1

Na Tabela 19, apresentam-se as características espectrais para cada sensor/plataforma espacial.

Tabela 19. Características espectrais das bandas dos satélites CBERS-4A (WPM), Landsat-8/OLI, Sentinel-2/MSI e PlanetScope utilizadas.

CBERS-4A	LANDSAT-8	Sentinel-2	Planet
0.45-0.52 μm (B)	0.45-0.51 μm (B)	0.49 μm (B)	0.455-0.515 μm (B)
0.52-0.59 μm (G)	0.53-0.59 μm (G)	0.56 μm (G)	0.50-0.59 μm (G)
0.63-0.69 μm (R)	0.64-0.67 μm (R)	0.665 μm (R)	0.59-0.67 μm (R)
0.77-0.89 μm (NIR)	0.85-0.88 μm (NIR)	0.842 μm (NIR)	0.78-0.86 μm (NIR)
0.45-0.90 μm (PAN)	1.57-1.65 μm (SWIR1)	0.705 μm (Red Edge 1)	
	2.11-2.29 μm (SWIR2)	0.740 μm (Red Edge 2)	
	0.50-0.68 μm (PAN)	0.783 μm (Red Edge 3)	
		0.865 μm (Red Edge 4)	
		1.61 μm (SWIR1)	
		2.19 μm (SWIR 2)	

Fonte: Elaborado pelo autor (2022).

*Nota: B:Azul, G:Verde, R: Vermelho, NIR: Infravermelho próximo, SWIR: Infravermelho de ondas curtas e PAN: Pancromática.

11.2.3 Pré-processamento imagens de satélites

A correção atmosférica para a imagem do sensor OLI/Landsat-8 empregou-se o complemento semi-automatic classification plugin (CONGEDO, 2021) no aplicativo QGIS 3.22. Com a ferramenta preprocessing, a correção atmosférica foi realizada com base no arquivo

.MTL, fornecido com as imagens. As imagens dos satélites CBERS-4A, Planet e Sentinel-2 foram fornecidas com a correção atmosférica (ESA, [s.d.]; INPE, [s.d.]; SCCON, [s.d.]; USGS, [s.d.]), e ainda todas as imagens possuem correção do relevo e são georreferenciadas. Após a correção atmosférica, as imagens foram submetidas ao *stacking* de bandas, tornando possível a obtenção de imagens coloridas. Então, foi realizado o cálculo de peso de nitidez das bandas, considerando a radiância dentro de cada banda (ARCGIS, [s.d.]), para as imagens que possuem banda pancromática. Para as imagens que possuíam a banda pancromática realizou-se a fusão entre as composições coloridas (RGB) e a banda pancromática com o objetivo de aumentar a resolução espacial das imagens. O valor de peso de cada uma das bandas calculado anteriormente foi utilizado em conjunto com o método de Gram-Schmidt, que é um método baseado em um algoritmo para ortogonalização de vetores (LABEN; BROWER, 2000).

Também foram gerados os seguintes índices de vegetação por meio da operação matemática de bandas para todas as plataformas orbitais estudadas (Tabela 20), sendo que esses índices foram utilizados por serem índices já utilizados para o estudo do comportamento da vegetação.

Tabela 20. Fórmulas dos índices de vegetação utilizados no estudo.

Índice de vegetação	Fórmula	Referências
NDVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R})$	(ROUSE <i>et al</i> , 1973)
EVI	$2.5(\text{NIR}-\text{R})/(\text{L1}+\text{NIR}+\text{C1}*\text{R}-\text{C2}*\text{B}+1)$	(JUSTICE <i>et al</i> , 1998)
TGI	$-0.5*(190*(\text{R}-\text{G})-120*(\text{R}-\text{B}))$	(HUNT JR <i>et al</i> , 2013)
SAVI	$(\text{NIR}-\text{R})/(\text{NIR}+\text{R}+\text{L2}) * (1+\text{L2})$	(HUETE, 1988)
VARI	$(\text{G}-\text{R})/(\text{G}+\text{R}-\text{B})$	(GITELSON <i>et al</i> , 2002)

Fonte: Elaborado pelo autor (2022).

*Nota : L1 = 1 ; C1 = 6; C2 = 7.5; L2 = 0.5.

Utilizando a função *grayscale*, as imagens coloridas foram convertidas em uma imagem em níveis de cinza, essa ferramenta aplica pesos para cada banda e realiza a normalização da imagem de saída (ARCGIS, [s.d.]), assim podendo extrair as métricas de textura com a ferramenta de Haralick (HARALICK; SHANMUGAM; DISNTEIN, 1973). Nesta etapa foram utilizados os seguintes parâmetros: janela de busca (5x5 e 7x7), a direção sudoeste (ZIEGELMAIER NETO, 2022), e como resultado temos as informações de energia que estão relacionados com a uniformidade e constância da distribuição dos níveis de cinza na imagem, a entropia que mede a não uniformidade da textura ou o grau de desordem da distribuição dos níveis de cinza na imagem, a correlação que mede a dependência linear de pixels em relação a sua vizinhança, homogeneidade que analisa o grau de variação local dos níveis de cinza entre

pares vizinhos, o contrasta que quantifica a variação local dos níveis de cinza entre pares de pixels, e a variância que se baseia-se na dispersão dos dados em torno de um valor médio, tratando especificamente das combinações do pixel de referência e do seu vizinho. Todas essas métricas de textura foram extraídas para cada uma das plataformas orbitais (SOUTO JUNIOR, 2010).

Ao fim da criação dos dados adicionais, todos foram reamostrados para a resolução espacial de 1m. Como a nuvem de pontos fornecida pela empresa estava recortada apenas para a área de vegetação nativa, todos os outros dados gerados foram recortados para os mesmos limites respectivos a suas áreas de estudo.

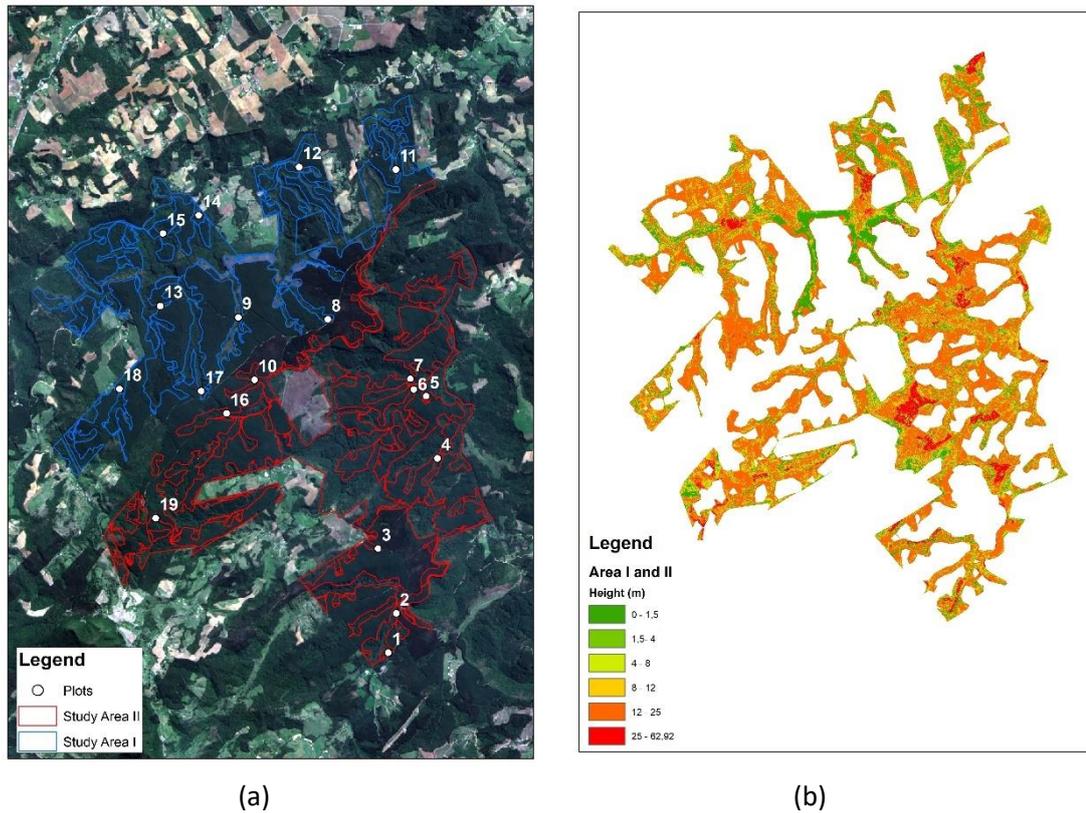
11.2.4 Classificação de referência

A classificação de referência, ou verdade de campo, foi gerada a partir da interpretação visual e dos dados do inventário convencional. Para delimitar as classes temáticas referentes a cobertura da terra e dos estágios de sucessão florestal empregou-se a composição colorida (RGB) obtidas do sensor WPM/CBERS-4A combinados com o CHM. As alturas fornecidas pelo CHM foram as principais informações utilizadas para a delimitação dos estágios de sucessão florestal que é um dos critérios definidos pela legislação brasileira para delimitar a sucessão florestal.

Segundo a Resolução CONAMA nº 04/1994 para a caracterização da vegetação primária e secundária nos estágios inicial, médio e avançado de regeneração, deve ser observado um conjunto de critérios e indicadores estruturais tanto qualitativos quanto quantitativos. O estágio inicial de regeneração surge logo após o abandono do solo, em que os indivíduos possuem no máximo oito centímetros de diâmetro médio na altura do peito (DAP) e árvores com altura máxima de quatro metros. São uma vegetação com baixa diversidade de espécies e predominantemente de ervas e arbustos e duram entre seis e dez anos, dependendo da qualidade do solo. Já no estágio médio de regeneração a diversidade aumenta, as espécies de árvores pioneiras são predominantes e existe maior quantidade de arvoretas e arbustos do que herbáceas. O DAP médio é de 8 a 15 cm e altura de até 12 m. Este estágio ocorre entre seis e quinze anos depois do abandono da área. E no estágio avançado de regeneração, a fisionomia arbórea é dominante sobre as demais, formando um dossel fechado e relativamente uniforme, apresentando árvores dominantes. A altura média da vegetação é superior a 12 m, e o DAP médio varia de 15 a 25 cm. A seguir são apresentadas as figuras contendo o limite da área

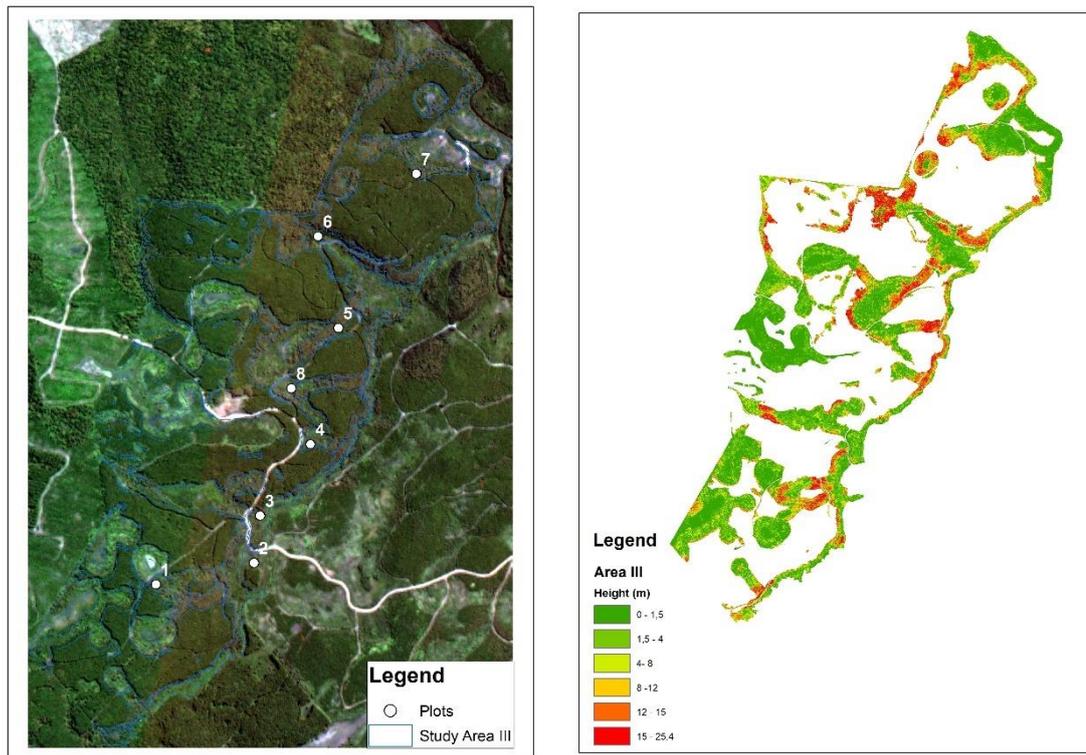
recortado com a distribuição das parcelas de inventário e a representação do CHM (Figura 15 e 16).

Figura 15. (a) Mapa de distribuição das parcelas do inventário florestal convencional para as áreas de estudo I e II com imagem RGB e delimitação do recorte de cada área. (b) CHM para as áreas de estudo I e II.



Fonte: Elaborado pelo autor (2022).

Figura 16. (a) Mapa de distribuição das parcelas do inventário florestal convencional para a área de estudo III com imagem RGB e delimitação do recorte da área. (b) CHM para a área de estudo III.



Fonte: Elaborado pelo autor (2022).

11.2.5 Composição dos dados

A criação dos datasets foi baseada utilizando a metodologia de (Capítulo II). Na Tabela 21 são apresentadas as combinações e a quantidade de variáveis para cada dataset.

Tabela 21. Combinação de dados e número de variáveis como dado de entrada para a classificação supervisionada.

	Dataset	Número de variáveis			
		CBERS-4A	Landsat-8	Sentinel-2	PLANET
1	Bandas do satélite	4	6	10	4
2	Bandas do satélite + CHM (LiDAR)	5	7	11	5
3	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR)	6	8	12	6
4	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + NDVI	7	9	13	7
5	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + EVI	7	9	13	7
6	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + TGI	7	9	13	7
7	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + SAVI	7	9	13	7

8	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + VARI	7	9	13	7
9	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 3x3	13	15	19	13
10	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 5x5	13	15	19	13
11	Bandas do satélite + CHM (LiDAR) + Intensidade (LiDAR) + GLCM 7x7	13	15	19	13

Fonte: Elaborado pelo autor (2022).

11.3 CLASSIFICAÇÃO DOS DATASETS

Para realizar a classificação das imagens, foram utilizados modelo treinados em uma área distinta as estudadas e foram criados polígonos de validação, necessários para encontrar padrões significativos nos dados das imagens que possam ser extraídos para validar a classificação.

11.3.1 Coleta de amostras

Utilizando a ferramenta Region of interest (ROI) do software ArcMAP® 10.4, foram gerados polígonos para as diversas classes temáticas previamente identificadas no processo de fotointerpretação das áreas, baseados apenas na análise visual e de dados de inventário florestal tradicional para o estágio de sucessão florestal. As classes temáticas identificadas foram: “campo”, estágio inicial “SS1”, estágio intermediário “SS2” e estágio avançado “SS3”. Na tabela 22, são apresentadas as áreas obtidas a partir da classificação de referência de cada classe em hectares (ha) e na Tabela 23 a quantidade de pixels amostrados por classe. Já na Figura 17, é mostrada a distribuição das amostras na área de interesse.

Tabela 22. Representação das classes temáticas identificadas na área de estudo e suas respectivas áreas em hectares.

Área	I	II	III	Amostra
Campo	26,9	8,9	8,3	
SS1	47,3	20,9	73,9	

SS2	139,2	46,2	22,1	
SS3	182,5	489,1	22,0	

Fonte: Elaborado pelo autor (2022).

Tabela 23. Quantidade de pixels (1x1 m) para as amostras de validação para cada área de estudo.

Área	I	II	III
Classe			
Campo	5143	1799	5143
SS1	4652	6085	4652
SS2	11867	1698	11867
SS3	15435	1700	15435

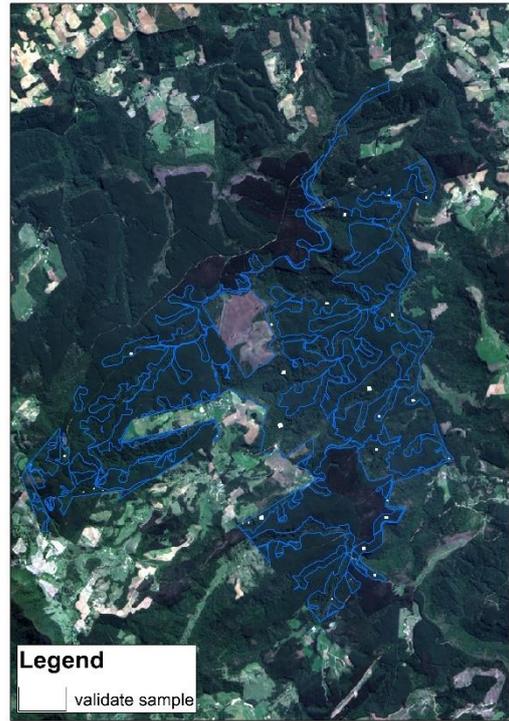
Fonte: Elaborado pelo autor (2022).

Figura 17. Mapa de distribuição das amostras para validar a transferência de aprendizagem.

(a) Área de estudo I. (b) Área de estudo II. (c) Área de estudo III.



(a)



(b)



(c)

Fonte: Elaborado pelo autor (2022).

11.3.2 Transferência de aprendizagem

Para a transferência de aprendizagem e classificação das áreas de interesse, foram utilizados os modelos treinados no estudo anterior (Capítulo II), em que foram utilizados os seguintes classificadores Random Tree (RT), Support Vector Machine (SVM) e Maximum likelihood classifier (MLC). A partir do estudo citado anteriormente, foram selecionados apenas os datasets que apresentaram melhores resultados quanto ao índice kappa para cada um dos classificadores (RT, SVM e MLC) das quatro plataformas orbitais (CBERS-4A, Sentinel-2, Landsat-8 e Planet), resultando em 12 classificações.

Dentro dos polígonos de validação foram criados 1000 pontos de avaliação de acurácia, sendo esses criados de forma aleatória e ponderado pelo tamanho amostrado de cada classe temática. A matriz de confusão (MILLER; NICELY, 1955) foi gerada a partir de pontos de avaliação de acurácia, esses pontos extraíram a informação sobre as classes da imagem classificada e da classificação de referência para a construção da matriz, comparando essas duas informações.

11.4 ANÁLISE ESTATÍSTICA

A partir da matriz de confusão, foi possível calcular a acurácia global, índice Kappa ponderado (COHEN, 1968), desvio padrão do índice Kappa e seus valores mínimo e máximo. Além disso, foi calculada a acurácia do produtor e usuário (AP e AU).

12 RESULTADOS

A Tabela 24 contém os valores do índice de concordância obtidos pela transferência de aprendizagem de todas as combinações estudadas neste trabalho.

Tabela 24. Índice Kappa para todas as imagens orbitais, composições e classificadores testados com a TA para as três áreas de estudo.

Satélite	Classificador/dataset	Kappa		
		Área I	Área II	Área III
CBERS 4A	MLC 3	0,33	0,25	0,14
CBERS 4A	RF 7	0,37	0,37	0,20
CBERS 4A	SVM 4	0,05	0,00	0,24
SENTINEL 2	MLC 5	0,38	0,40	0,70
SENTINEL 2	RF 8	0,45	0,40	0,59
SENTINEL 2	SVM 3	0,48	0,33	0,57
LANDSAT 8	MLC 2	0,08	0,03	0,64

LANDSAT 8	RF 6	0,30	0,07	0,21
LANDSAT 8	SVM 4	0,00	0,00	0,00
PLANET	MLC 11	0,26	0,33	0,17
PLANET	RF 10	0,20	0,29	0,28
PLANET	SVM 7	0,00	0,00	0,47

Fonte: Elaborado pelo autor (2022).

É possível observar que para as três áreas os datasets derivados a partir do sensor MSI (Multi-spectral Instrument) a bordo da plataforma Sentinel-2 obtiveram os melhores resultados para a transferência de aprendizagem, sendo a dataset 2* com o classificador SVM o mais eficiente para a área I com um kappa de 0,48. Para a área II o classificador que mostrou melhor resultado foi o de RT em conjunto com a dataset 7*, com um índice de concordância de 0,40, sendo esse resultado igual ao do classificador MLC com a dataset 4. Essa última combinação de dados e classificador também se mostrou eficaz na área III, atingindo um kappa de 0,70. Quanto às imagens dos satélites CBERS-4A, Landsat-8 e Planet obtiveram resultados melhores que os alcançados pelo Sentinel-2 utilizando a transferência de aprendizagem. A seguir serão apresentadas as matrizes de confusão (Tabela 25, 26 e 27), juntamente com as acurácias de usuário e produtor para as melhores classificações a partir da transferência de aprendizagem de cada área de estudo.

Tabela 25. Matriz de confusão para a transferência de aprendizagem para a área de estudo 1, utilizando o classificador SVM com o dataset 2 da imagem Sentinel-2.

Classe	Campo	Água	SS1	SS2	SS3	Total	AU
Campo	118	0	15	0	6	139	84,89
Água	0	0	0	0	0	0	0
SS1	46	0	46	14	19	125	36,8
SS2	13	3	80	78	146	320	24,38
SS3	0	0	1	13	402	416	96,63
Total	177	3	142	105	573	1000	
AP	66,67	0	32,39	74,29	70,16		

Fonte: Elaborado pelo autor (2022).

Tabela 26. Matriz de confusão para a transferência de aprendizagem para a área de estudo 2, utilizando o classificador RT com a dataset 7 da imagem Sentinel-2.

	Campo	SS1	SS2	SS3	Total	AU
Campo	35	0	0	0	35	100
SS1	4	21	8	14	47	44,68
SS2	4	72	24	34	134	17,91
SS3	0	30	82	672	784	85,71
Total	43	123	114	720	1000	
PA	81,4	17,07	21,05	93,33		

Fonte: Elaborado pelo autor (2022).

Tabela 27. Matriz de confusão para a transferência de aprendizagem para a área de estudo 3, utilizando o classificador MLC com a dataset 4 da imagem Sentinel-2.

	Campo	SS1	SS2	SS3	Total	AU
Campo	148	9	0	0	157	94,27
SS1	0	542	0	4	546	99,27
SS2	0	104	0	37	141	0
SS3	0	20	0	134	154	87,01
Total	148	675	0	175	998	
PA	100	80,3	0	76,57		

Fonte: Elaborado pelo autor (2022).

Para a área I podemos ver que a tanto para a AU a classe de vegetação com menor valor foi a SS2 com 24,38% e para AP o menor índice foi obtido para a classe SS1, com valor 32,39% para ambas. Já a classe SS3 foi a que possui melhor classificação em relação à verdade de campo com AU de 96,63%. A classe SS2 teve um maior acerto em relação ao classificador com AP de 74,29%. Na área II a classe campo teve a melhor AP de 100%, já a classe SS2 obteve a pior classificação de apenas 17,91%. Quanto a AU a classe SS3 alcançou valor de 93,33% de acurácia, e a classe SS1 teve apenas 17,07% de sucesso. E para a área III a classe SS2 teve pior AP e AU de 0% com nenhum ponto atribuído a essa classe, para a AP a classe com melhor

classificação foi a classe SS1 com 99,27% e para a AU foi a classe campo com 100% de acurácia na classificação. Sendo essas informações importantes para a estimativa das classes de vegetação, tendo papel fundamental para o entendimento do funcionamento da vegetação, auxiliando no monitoramento de áreas remanescentes da Floresta Atlântica Brasileira.

A seguir é apresentada a tabela 28 contendo o kappa, desvio padrão, valor máximo e mínimo para as melhores classificações de cada área de estudo utilizando a transferência de aprendizagem, para a área I com o classificador SVM com o dataset 2, para a área II com o classificador RT e o dataset 7, e para a área III utilizando o classificador MLC com dataset 4.

Tabela 28. Índice de concordância, desvio padrão, valor máximo e valor mínimo para as classificações das imagens Sentinel-2.

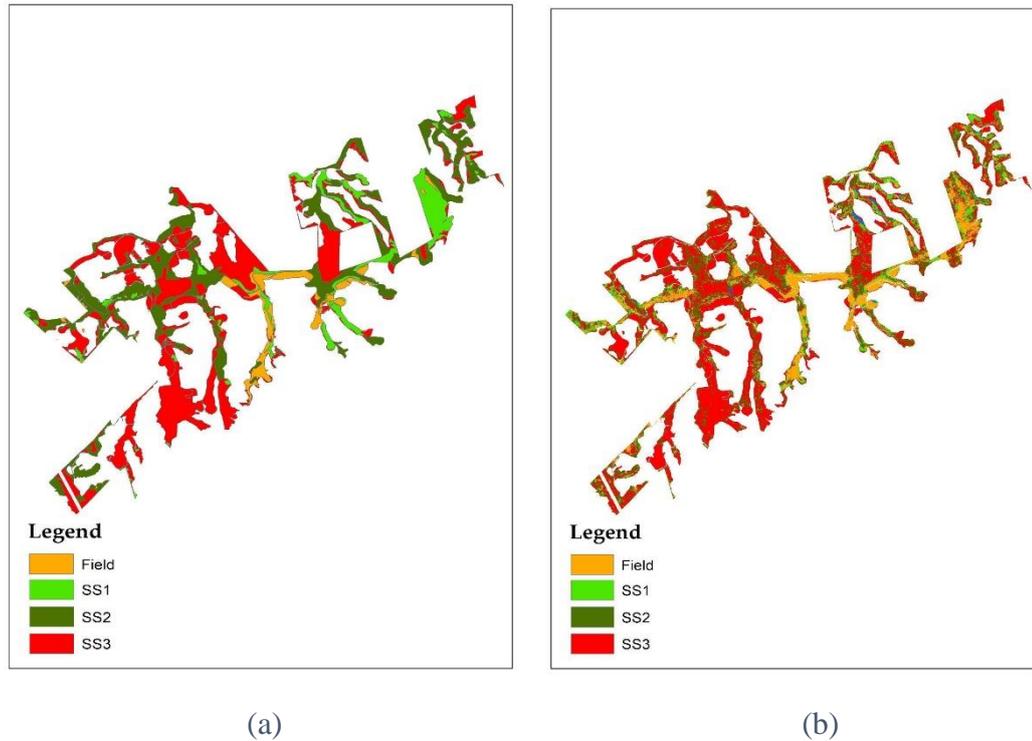
Area	Satélite	Classificador	Dataset*	Kappa Ponderado	Desvio Padrão	Mínimo	Máximo
1	Sentinel-2	SVM	2	0,48	0,02	0,44	0,52
2	Sentinel-2	RT	7	0,40	0,03	0,35	0,45
3	Sentinel-2	MLC	4	0,70	0,02	0,66	0,74

Fonte: Elaborado pelo autor (2022).

Dentre as melhores classificações para cada uma das áreas de estudo a que apresentou melhor kappa foi a classificação da área 3 com o índice de concordância alcançando um valor de 0,70 podendo variar de 0,74 a 0,66 com erro padrão de 0,02, considerando todas as classes. A área I apresentou a segunda melhor classificação baseada na transferência de aprendizagem com um kappa de 0,48, esse podendo variar de 0,52 a 0,44 com erro padrão de 0,02. E a área II foi que teve a pior classificação utilizando o algoritmo treinado em outra área, com um índice de concordância de 0,40 podendo variar entre 0,35 a 0,45 com erro padrão de 0,02. Todas as estatísticas possuem 5% de significância.

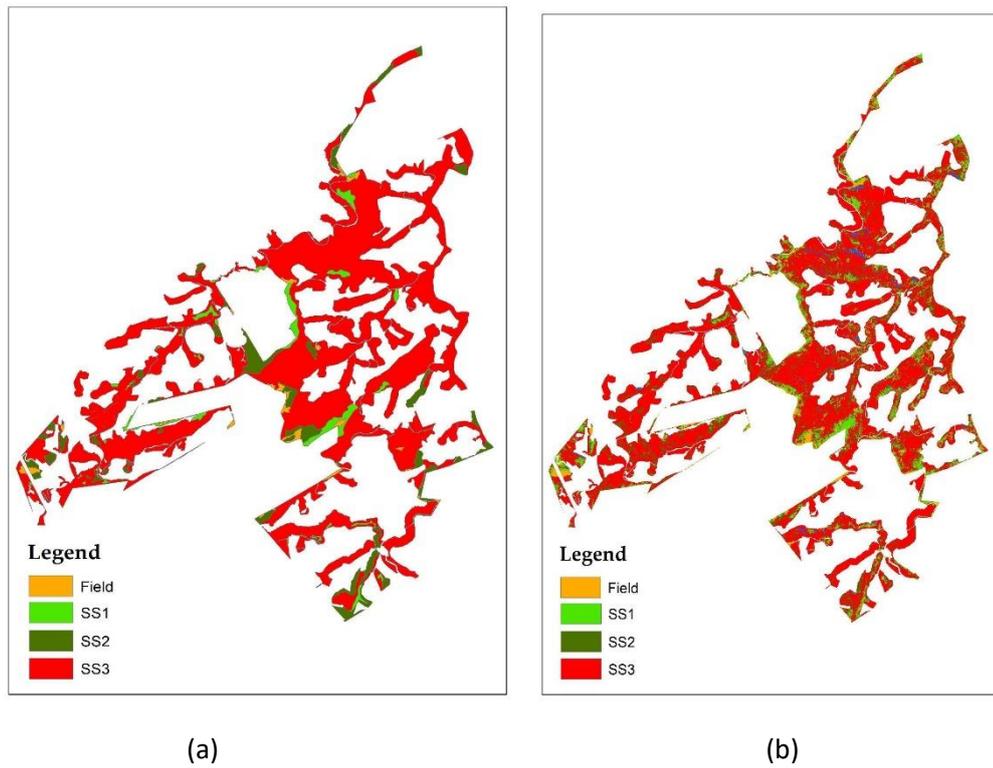
Nas figuras 18, 19 e 20 temos a comparação visual entre a imagem testemunha e a imagem classificada pela transferência de aprendizagem para a área I, II e III, respectivamente.

Figura 18. Comparação entre a classificação de referência (a) e classificação automatizada utilizando a transferência de aprendizagem para a área I, utilizando o classificador SVM com o dataset 2 da imagem Sentinel-2.



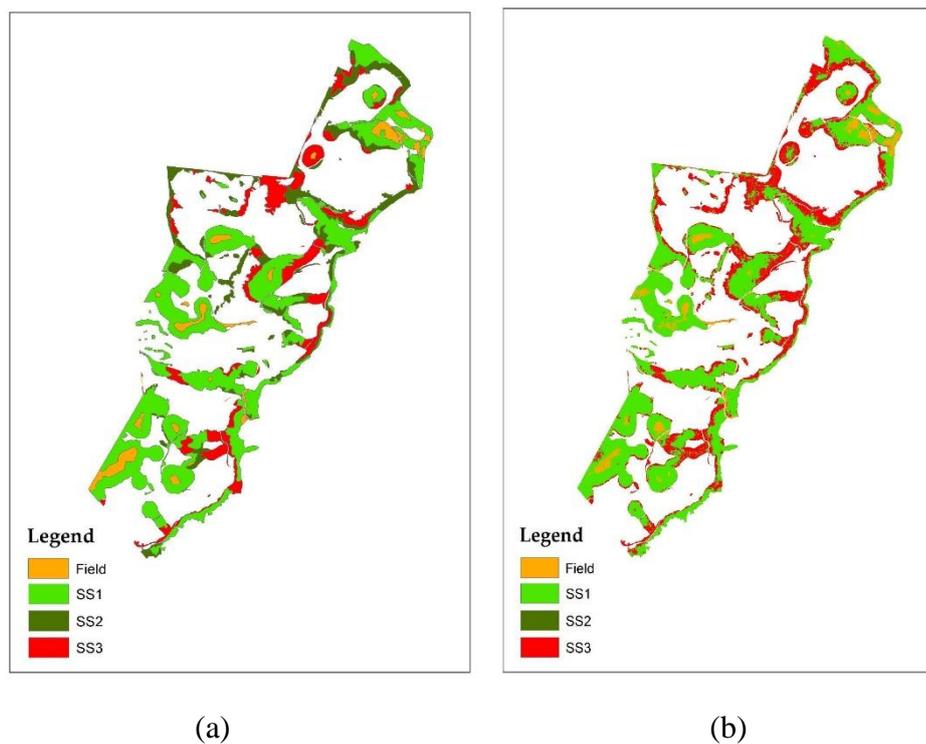
Fonte: Elaborado pelo autor (2022).

Figura 19. Comparação entre a classificação de referência (a) e classificação automatizada utilizando a transferência de aprendizagem para a área II, utilizando o classificador RT com o dataset 7 da imagem Sentinel-2.



Fonte: Elaborado pelo autor (2022).

Figura 20. Comparação entre a classificação de referência (a) e classificação automatizada utilizando a transferência de aprendizagem para a área III, utilizando o classificador MLC com o dataset 4 da imagem Sentinel-2.



Fonte: Elaborado pelo autor (2022).

12 DISCUSSÃO

Em relação às imagens de satélites utilizadas para a transferência de aprendizagem, os melhores resultados foram obtidos com as imagens orbitais provenientes do satélite Sentinel-2. Variando apenas o classificador e o dataset utilizado para cada uma delas, esse melhor desempenho pode ser devido a maior quantidade de bandas espectrais resultando em mais informações sobre cada classe. O autor (JENSEN, 2009) citam que uma maior quantidade de informação espectral e radiométrica podem auxiliar na avaliação de alvos pelo Sensoriamento Remoto. Isso pode ter possibilitado um ajuste mais robusto do algoritmo de classificação.

Em relação às três áreas de estudo, a área III apresentou um melhor desempenho na transferência de aprendizagem, com kappa de 0.70. Esse resultado pode ser devido a área estar inserida na tipologia florestal FOM, mesma em que o algoritmo foi treinado. Já as áreas I e II, se encontram na FOD, sendo que a estrutura da floresta é diferente alterando os valores espectrais de cada classe, dificultando o reconhecimento pelo algoritmo. Sendo comprovado no estudo (DLUGOSZ et al., 2005) que realizou a diferenciação de tipologias florestais utilizando segmentação da imagem Ikonos. Outro fator que pode ter influenciado na acurácia das classificações é a diferença entre as datas das imagens, o estudo (MOREIRA; ADAMI; RUDORFF, 2004) mostrou que existe diferença na resposta espectral para cultura de café entre diferentes datas.

Outra possível causa que pode ter contribuído para os resultados das classificações foi que a nuvem de pontos LiDAR utilizadas para essas áreas estava recortada, em determinados locais havia apenas uma pequena faixa de 30-40 metros de largura, dificultando a determinação das classes nessas regiões, levando em consideração as diferentes resoluções espaciais utilizadas no estudo.

A combinação de imagens que apresentou melhor resultado para a área I foi o dataset 2, utilizando o classificador SVM. Dentre os melhores resultados para as três áreas, essa foi a que utilizou a menor quantidade de dados de entrada. Para a área II o classificador que apresentou o melhor resultado foi o RT utilizando o dataset 7. E para a área III foi o dataset 4. Em todos os casos as informações do Lidar contribuíram para a classificação, e no caso das áreas II e III a utilização de índices de vegetação também foi importante. Porém, ao contrário do estudo prévio (Capítulo II) em que os modelos foram treinados, as composições utilizando o NDVI não apresentaram resultados satisfatórios para a transferência de aprendizagem.

Considerando-se os resultados obtidos para o índice kappa e de acordo com a classificação recomendada por [LANDIS E KOCH] a transferência de aprendizagem da classificação dos estágios de sucessão da vegetação apresentou resultado muito bom para a área III que apresenta o mesmo tipo de vegetação que a área em que o classificador foi treinado. Já para as outras áreas que possuem o tipologia de vegetação de FOD apresentaram a classificação do índice kappa boa (LANDIS; KOCH, 1977), mostrando que a metodologia pode ser utilizada em diferentes regiões e tipologias florestais, porém quando utilizada em uma mesma região, os resultados da classificação apresentaram uma melhor acurácia. O estudo (ALEM; KUMAR, 2021) alcançou resultados excelentes para a transferência de aprendizagem para a classificação do uso e ocupação do solo, porém esse utilizando uma rede neural convolucional (CNN), obtendo índice kappa variando de 0,92 a 0,99.

A área de estudo I teve uma grande acurácia para a classe SS3 sendo que classe alcançou 96,63% de acurácia do usuário, apenas 3% dos pontos e 0,24% foram classificados na classe SS2 e 1 na classe SS1. Para a acurácia do produtor, a classe que apresentou o melhor resultado foi a SS2, com 74,29% de classificações corretas, tendo pontos atribuídos erroneamente tanto para a classe SS1 e SS3. Também houve 3 pontos atribuídos a água, porém não há área dessa classe para esse local de estudo, sendo essa a única classificação entre as melhores para cada área que apresentou esse problema.

A área de estudo II teve uma classificação de 100% para a classe campo e 85,71% para a classe SS3, sendo essas duas as classes que possuíam melhor AU. Já a classe SS2 teve uma baixa AU resultou em apenas 17,91% de acurácia. Para a AP a classe SS3 foi a que obteve melhor classificação com 93,33%, sendo que 4,7% dos pontos pertencem a classe SS2 e 1,9% a classe SS1.

Quanto a área de estudo III a classe campo teve maior AP com 100% e a classe SS2 foi a menor com 0%, pois nenhum ponto foi atribuído a essa classe pelo classificador. Para a AU a classe SS1 alcançou 99,27% com erro de apenas 0,73% dos pontos que foram atribuídos a classe SS3. Já a classe SS2 obteve 0% com nenhum ponto classificado corretamente, confundindo principalmente com a classe SS1, sendo que 73,76% dos pontos foram atribuídos erroneamente a classe de sucessão inicial, de um total de 141 pontos que realmente pertencem ao estágio de sucessão médio. Esses resultados mostram que as acurácias obtidas para a determinação de algumas classes da vegetação utilizando a TA, alcançam resultados similares a outros estudos realizados na mesma região (PINTO, 2018; SOTHE et al., 2018, 2017; SOTHE, 2015). Esses estudos citados anteriormente aplicaram métodos tradicionais de

classificação supervisionada e também encontraram maior dificuldade para a diferenciação entre os estágios sucessionais da vegetação, utilizando diferentes sensores orbitais e Lidar. O trabalho de (BISPO et al., 2019) utilizando dados derivados de Radar mostrou que para a maioria das classificações as maiores confusões estão entre as classes de vegetação.

Uma das principais limitações desse estudo foi a utilização dos dados LiDAR recortados, acarretando a utilização das imagens óticas com as mesmas delimitações, ocasionando a falta de informações para certas regiões. Outra limitação, foi a utilização da classificação orientada a pixel, sendo essa metodologia não recomendada para a classificação de imagens de alta resolução espacial, limitando a capacidade de detecção de detalhes em cenas heterogêneas, ocasionando a redução da acurácia (PINHO; FEITOSA; KUX, 2005). Também não foi possível a utilização das imagens com mesma data em que o classificador foi treinado, podendo ter variações da resposta espectral dos alvos, confundindo o algoritmo. Recomenda-se para estudos futuros, a utilização de modelos de segmentação para realizar a classificação orientada a regiões, além de se explorar a utilização da transferência de aprendizagem em regiões de mesma tipologia florestal, utilizando imagens da mesma data, para tentar obter melhores resultados e reduzir as diferenças entre as respostas espectrais dos alvos.

14 CONCLUSÕES

Esse estudo mostrou que é possível utilizar a transferência de aprendizagem de modelos de ML para a classificação do uso e ocupação do solo e dos de estágios de sucessão florestal, com o kappa das classificações variando de 0 a 0,7, dependendo da área de estudo e do conjunto de dados e classificador.

Os melhores resultados da TA foram apresentados para a área que possui a mesma tipologia florestal em que os dados foram treinados, que é a FOM. Já para as outras duas áreas de estudos que possuem uma tipologia florestal FOD, os resultados não alcançaram a mesma acurácia, sendo as melhores classificações consideradas como boas, segundo a classificação do índice kappa de (Landis e Koch). Em comparação às plataformas orbitais utilizadas, a que apresentou melhor desempenho para todas as áreas de estudo foi o Sentinel-2. Quanto aos classificadores utilizados, nenhum se destacou, sendo que para cada área um classificador diferente alcançou o melhor resultado.

Os datasets que foram utilizados apresentaram a combinação de dados derivados das plataformas orbitais e do sensor LiDAR, mostrando que essas informações são importantes para

descrever a estrutura da floresta e assim possibilitar a classificação do estágio de sucessão, e com a TA é possível diminuir ainda mais os custos e tempo relacionados a classificação dessas áreas.

15 CONSIDERAÇÕES FINAIS

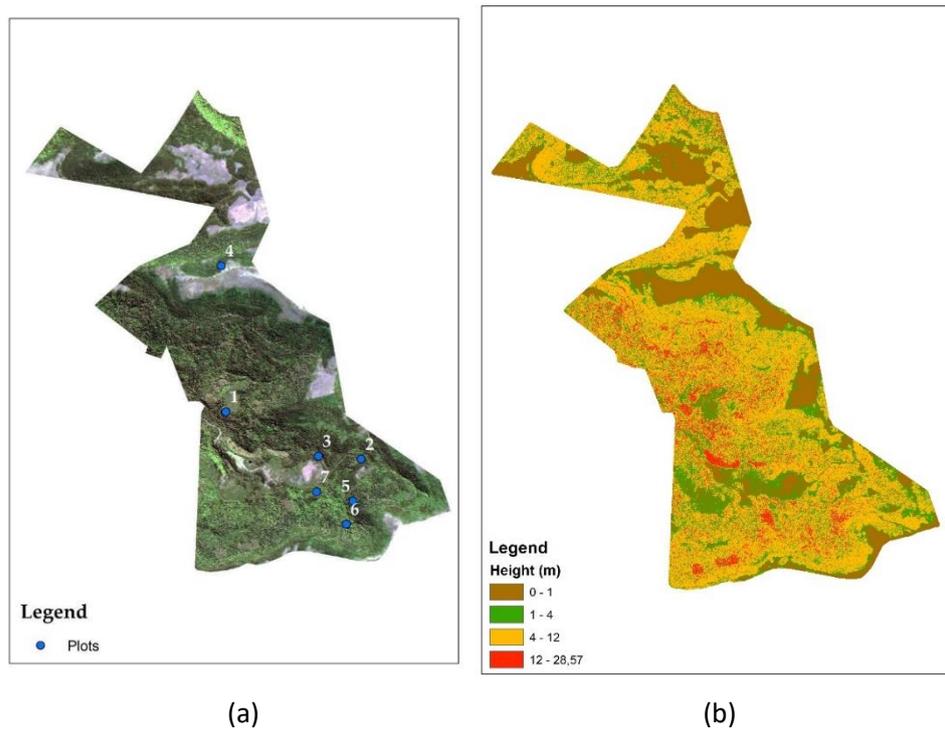
Esse estudo mostrou que a classificação dos estágios de sucessão da vegetação pode ser realizada a partir de dados de sensoriamento remoto e métodos de machine learning obtendo excelentes resultados. Todas as plataformas de imagens orbitais se provaram fontes úteis de dados para essa finalidade.

O primeiro estudo, em que foi realizado o treinamento dos modelos de classificação, apresentou resultados mais acurados utilizando dados derivados do satélite Landsat-8 em conjunto com os dados LiDAR. Já no segundo estudo, para a transferência de aprendizagem, o melhor resultado foi obtido utilizando dos dados derivados do LiDAR e Sentinel-2, mostrando que para as duas áreas a adição de informações LiDAR melhoram a performance de classificação.

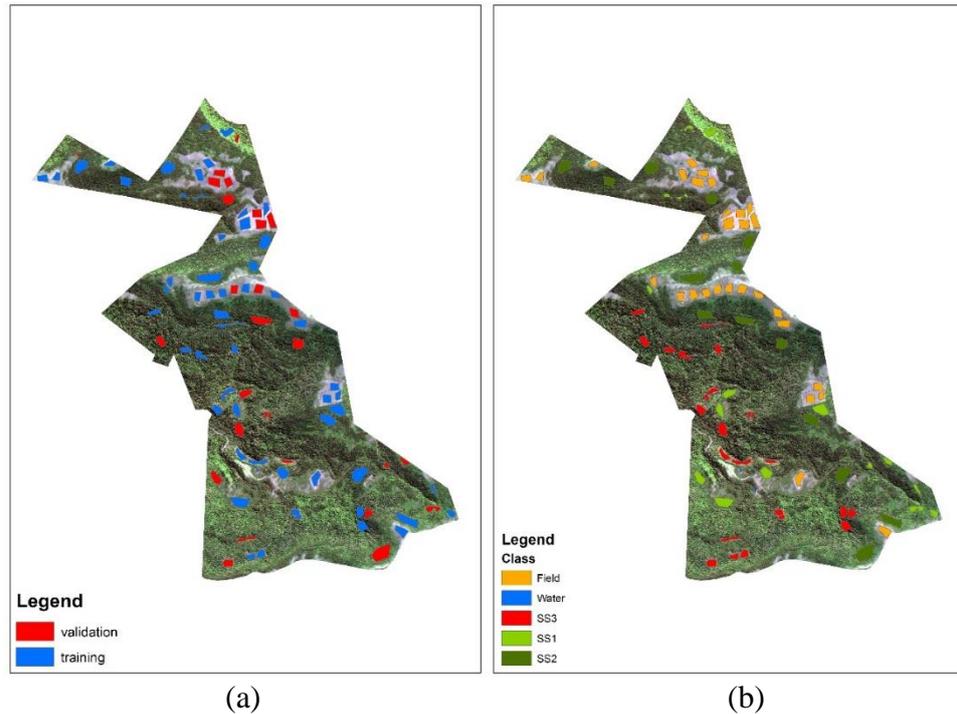
Esses resultados são importantes pois mostram que existem alternativas a coleta de dados in loco, diminuindo o tempo e custos necessários para a obtenção dessas informações, tanto pelo treinamento quanto para a transferência dos algoritmos, auxiliando gestores a entender melhor o funcionamento das florestas. Isso também pode contribuir para o licenciamento e fiscalização de forma mais efetiva dessas áreas.

APÊNDICE

Apêndice 1. (a) Mapa de distribuição das parcelas do inventário florestal convencional. (b) CHM com alturas reclassificadas utilizado para a fotointerpretação



Apêndice 2. (a) Mapa de distribuição das amostras de treinamento e validação utilizados para realizar a classificação supervisionada. (b) Mapa de distribuição das amostras separados por cada classe



REFERÊNCIAS

ABOUD NETA, S. R. *et al.* Comparação entre classificação contextual e classificação por regiões para mapeamento de uso e cobertura da terra na região da Floresta Nacional de Tapajós-PA (FLONA): utilizando imagens polarimétricas em banda L. **Simpósio Brasileiro de Sensoriamento Remoto**, v. 14, p. 7749-7756, 2009.

ADAM, E.; MUTANGA, O.; ODINDI, J.; ABDEL-RAHMAN, E.M. Land-Use/Cover Classification in a Heterogeneous Coastal Landscape Using RapidEye Imagery: Evaluating the Performance of Random Forest and Support Vector Machines Classifiers. *Int. J. Remote Sens.* **2014**, 35, 3440–3458, doi:10.1080/01431161.2014.903435.

ALEM, A.; KUMAR, S. Transfer Learning Models for Land Cover and Land Use Classification in Remote Sensing Image. *Appl. Artif. Intell.* **2021**, 00, 1–19, doi:10.1080/08839514.2021.2014192.

ALMEIDA, A. S.; SANTOS, R. L.; CHAVES, J. M. Mapeamento de uso e ocupação do solo no Município de Jeremoabo-BA: uso do Algoritmo Máxima Verossimilhança (Maxver). **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO–SBSR**, v. 15, 2011.

ANDERSEN, H. E.; REUTEBUCH, S. E.; MCGAUGHEY, R. J. Active remote sensing. In: *Computer applications in sustainable forest management*, Shao, G.; Reynolds, K.; Shao, G., The Netherlands: Springer, 2006; pp. 43-66.

ARCGIS. Available online: <https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/grayscale-function.htm>

ARCGIS. Available online: <https://desktop.arcgis.com/en/arcmap/latest/tools/data-management-toolbox/compute-pansharpen-weights.htm>

ASNER, G.P.; KNAPP, D.E.; KENNEDY-BOWDOIN, T.; JONES, M.O.; MARTIN, R.E.; BOARDMAN, J.; HUGHES, R.F. Invasive Species Detection in Hawaiian Rainforests Using Airborne Imaging Spectroscopy and LiDAR. *Remote Sens. Environ.* **2008**, *112*, 1942–1955, doi:10.1016/j.rse.2007.11.016.

ASNER, G.P.; POWELL, G.V.N.; MASCARO, J.; KNAPP, D.E.; CLARK, J.K.; JACOBSON, J.; KENNEDY-BOWDOIN, T.; BALAJI, A.; PAEZ-ACOSTA, G.; VICTORIA, E.; et al. High-Resolution Forest Carbon Stocks and Emissions in the Amazon. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16738–16742, doi:10.1073/pnas.1004875107.

ASTOLA, H.; SEITSONEN, L.; HALME, E.; MOLINIER, M.; LÖNNQVIST, A. Deep Neural Networks with Transfer Learning for Forest Variable Estimation Using Sentinel-2 Imagery in Boreal Forest. *Remote Sens.* **2021**, *13*, doi:10.3390/rs13122392.

BALTSAVIAS, E. P. Airborne laser scanning: existing systems and firms and other resources. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2, p. 164-198, 1999.

BARNES, B. V. *et al.* **Forest Ecology** 4 ed. ,1998. 774p.

BERGEN, K. M.; DRONOVA, I. Observing succession on aspen-dominated landscapes using a remote sensing-ecosystem approach. **Landscape Ecology**, v. 22, n. 9, p. 1395–1410, 2007.

BERVEGLIERI, A.; IMAI, N.N.; TOMMASELLI, A.M.G.; CASAGRANDE, B.; HONKAVAARA, E. Successional Stages and Their Evolution in Tropical Forests Using Multi-Temporal Photogrammetric Surface Models and Superpixels. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 548–558, doi:10.1016/j.isprsjprs.2018.11.002.

BISPO, P.D.C.; PARDINI, M.; PAPATHANASSIOU, K.P.; KUGLER, F.; BALZTER, H.; RAINS, D.; DOS SANTOS, J.R.; RIZAEV, I.G.; TANSEY, K.; DOS SANTOS, M.N.; *et al.* Mapping Forest Successional Stages in the Brazilian Amazon Using Forest Heights Derived from TanDEM-X SAR Interferometry. *Remote Sens. Environ.* **2019**, *232*, 111194, doi:10.1016/j.rse.2019.05.013.

BRASIL. Conselho Nacional do Meio Ambiente. 1993. Resolução CONAMA nº 10/93, de 1 de outubro de 1993. Publicada no Diário Oficial da União em 3 de nov. 1993, n. 209.

BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5-32, 2001.

CAUGHLIN, T.T.; BARBER, C.; ASNER, G.P.; GLENN, N.F.; BOHLMAN, S.A.; WILSON, C.H. Monitoring Tropical Forest Succession at Landscape Scales despite Uncertainty in Landsat Time Series. *Ecol. Appl.* **2021**, *31*, 1–18, doi:10.1002/eap.2208.

CHEN, Z.; ZHANG, T.; OUYANG, C. End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images. *Remote Sens.* **2018**, *10*, 1–15, doi:10.3390/rs10010139.

CHO, M.A.; MATHIEU, R.; ASNER, G.P.; NAIDOO, L.; VAN AARDT, J.; RAMOELO, A.; DEBBA, P.; WESSELS, K.; MAIN, R.; SMIT, I.P.J.; *et al.* Mapping Tree Species Composition in South African Savannas Using an Integrated Airborne Spectral and LiDAR System. *Remote Sens. Environ.* **2012**, *125*, 214–226, doi:10.1016/j.rse.2012.07.010.

CINTRA, D.P.; OLIVEIRA, R.R.; REGO, L.F.G. Classificação Dos Estágios Sucessionais Florestais Através de Imagens Ikonos No Parque Estadual Da Pedra Branca. *XIII Simpósio Bras. Sensoriamento Remoto* **2007**, 1627–1629.

COHEN, J. Weighted Kappa: Nominal Scale Agreement Provision for Scaled Disagreement or Partial Credit. *Psychol. Bull.* **1968**, 70, 213–220, doi:10.1037/h0026256

CONGEDO, L. Semi-Automatic Classification Plugin Documentation Release 7.9.5.1. *User Man.* **2021**, 1–225.

CONSELHO NACIONAL DO MEIO AMBIENTE BRASIL. Conselho Nacional do Meio. Resolução CONAMA nº 10/93, de 1 de outubro de 1993. p. 209.

CORREIA, V. R. M. *et al.* Uma aplicação do sensoriamento remoto para a investigação de endemias urbanas. *Cadernos de Saúde Pública*, v. 23, n. 5, p. 1015-1028, 2007.

CÔRTEZ, S. C.; PORCARO, R. M.; LIFSCHITZ, S. **Mineração de dados-funcionalidades, técnicas e abordagens.** PUC, 2002.

CRÓSTA, A. P. **Processamento digital de imagens de sensoriamento remoto.** 154p. **Unicamp**, Campinas, Brasil, 2002.

D'OLIVEIRA, M.V.N.; FIGUEIREDO, E.O.; PAPA, D. DE A. *Uso Do Lidar Como Ferramenta Para o Manejo de Precisão Em Florestas Tropicais*; 2014; ISBN 9788570354327.

DALPONTE, M.; BRUZZONE, L.; GIANELLE, D. Tree Species Classification in the Southern Alps Based on the Fusion of Very High Geometrical Resolution Multispectral/Hyperspectral Images and LiDAR Data. *Remote Sens. Environ.* **2012**, 123, 258–270, doi:10.1016/j.rse.2012.03.013.

DE MORAES NOVO, Evlyn ML. **Sensoriamento Remoto: princípios e aplicações.** Editora Blucher, 2010.

DEMIRBAŞ ÇAĞLAYAN, S.; LELOGLU, U.M.; GINZLER, C.; PSOMAS, A.; ZEYDANLI, U.S.; BILGIN, C.C.; WASER, L.T. Species Level Classification of Mediterranean Sparse Forests-Maquis Formations Using Sentinel-2 Imagery. *Geocarto Int.* **2020**, 6049, 1–20, doi:10.1080/10106049.2020.1783581.

DLUGOSZ, F.L.; ROSOT, N.C.; ROSOT, M.A.D.; OLIVEIRA, Y.M.M. de Uso Da Segmentação Por Crescimento de Regiões Em Imagem Ikonos Na Discriminação de Tipologias Da Floresta Ombrófila Mista. *An. XII Simpósio Bras. Sensoriamento Remoto*, **2005**, 1493–1500.

ESA. **Sentinel-2 data product.** [s.d.]. Disponível em: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products>.

FALKOWSKI, M.J.; EVANS, J.S.; MARTINUZZI, S.; GESSLER, P.E.; HUDAK, A.T. Characterizing Forest Succession with Lidar Data: An Evaluation for the Inland Northwest, USA. *Remote Sens. Environ.* **2009**, 113, 946–956, doi:10.1016/j.rse.2009.01.003.

FINLAYSON, C.M.; VAN DER VALK, A.G. Wetland Classification and Inventory: A Summary. *Vegetatio* **1995**, *118*, 185–192, doi:10.1007/BF00045199.

FLORENZANO, T. G. **Iniciação em sensoriamento remoto**. Oficina de Textos, 2007.

FLORIANI, L.; PUPPO, E. Hierarchical Triangulation for Multiresolution Surface Description. *ACM Transactions on Graphics (TOG)*, v. 14, n. 4, p. 363–411, 1995.

GIONGO, M.; KOEHLER, H.S.; MACHADO, S. DO A.; KIRCHNER, F.F.; MARCHETTI, M. LiDAR: Princípios e Aplicações Florestais. *Pesqui. Florest. Bras.* **2010**, *30*, 231–244, doi:10.4336/2010.pfb.30.63.231.

GITELSON, A.A.; KAUFMAN, Y.J.; STARK, R.; RUNDQUIST, D. Novel Algorithms for Remote Estimation of Vegetation Fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87, doi:10.1016/S0034-4257(01)00289-9.

GONÇALVES, A. R. **Máquina de vetores suporte**. v. 21, 2010.

GU, Z.; CAO, S.; SANCHEZ-AZOFEIFA, G.A. Using LiDAR Waveform Metrics to Describe and Identify Successional Stages of Tropical Dry Forests. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 482–492, doi:10.1016/j.jag.2018.07.010.

GUEDES, J.; DA SILVA, S. M. P. SENSORIAMENTO REMOTO NO ESTUDO DA VEGETAÇÃO: PRINCÍPIOS FÍSICOS, SENSORES E MÉTODOS. *Acta Geográfica*, v. 12, n. 29, p. 127-144, 2018.

GUO, M.; LI, J.; SHENG, C.; XU, J.; WU, L. A Review of Wetland Remote Sensing. *Sensors (Switzerland)* **2017**, *17*, 1–36, doi:10.3390/s17040777.

HAN, H.; GUO, X.; YU, H. Variable selection using mean decrease accuracy and mean decrease gini based on random forest. In: **2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)**. IEEE, 2016. p. 219-224.

HARALICK, R.M.; SHANMUGAM, K.; DISNTEIN, I. Textural Features for Image Classification 1973, 610–621.

HUETE, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309.

HUNT JR., E.R.; DORAISWAMY, P.C.; MCMURTREY, J.E.; DAUGHTRY, C.S.T.; PERRY, E.M. A Visible Band Index for Remote Sensing Leaf Chlorophyll Content at the Canopy Scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112.

IBGE. Manual técnico da vegetação brasileira. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística, **1992**. 92 p.

IZBICKI, R.; SANTOS, T.M. dos *Machine Learning Sob a Ótica Estatística*; Ufscar/Insper, 2018;

JAMALI, A. Land Use Land Cover Mapping Using Advanced Machine Learning Classifiers. *Ekol. Bratislava* **2021**, *40*, 286–300, doi:10.2478/eko-2021-0031.

JANOTH, J. *et al.* Procedimentos Baseados em Segmentação para a Análise de Mudanças e Classificação Florestais com Dados de Satélite de Alta Resolução. In: BLASCHKE, T.; KUX, H. Sensoriamento Remoto e SIG Avançados. 2nd ed.; São Paulo: Oficina de textos, Brasil, 2007; cap. 9. 303 p.

JENSEN, J.R. Sensoriamento Remoto Do Ambiente: Uma Perspectiva Em Recursos Terrestres (Tradução Da Segunda Edição). *Inf. Syst.* 2009, 598.

JUSTICE, C.O.; VERMOTE, E.; TOWNSHEND, J.R.G.; DEFRIES, R.; ROY, D.P.; HALL, D.K.; SALOMONSON, V. V.; PRIVETTE, J.L.; RIGGS, G.; STRAHLER, A.; *et al.* The Moderate Resolution Imaging Spectroradiometer (MODIS): Land Remote Sensing for Global Change Research. *IEEE Trans. Geosci. Remote Sens.* **1998**, 36, 1228–1249, doi:10.1109/36.701075.

KAGEYAMA, P. Y.; BRITO, M. A.; BAPTISTON, I. C. Estudo do mecanismo de reprodução de espécies da mata natural. In: Kageyama, P. Y. Estudo para implantação de matas ciliares de proteção na bacia hidrográfica do Passa Cinco, Piracicaba, SP. Piracicaba: DAAE/USP/FEALQ, Brasil, 1986; pp. 236.

KENNAWAY, T.A.; HELMER, E.H.; LEFSKY, M.A.; BRANDEIS, T.A.; SHERRILL, K.R. Mapping Land Cover and Estimating Forest Structure Using Satellite Imagery and Coarse Resolution Lidar in the Virgin Islands. *J. Appl. Remote Sens.* **2008**, 2, 023551, doi:10.1117/1.3063939

KOLECKA, N.; KOZAK, J.; KAIM, D.; DOBOSZ, M.; GINZLER, C.; PSOMAS, A. Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. *Remote Sens.* **2015**, 7, 8300–8322, doi:10.3390/rs70708300.

KÖPPEN, W. Climatologia, con un estudio de los climas de la tierra. Fondo de Cultura Economica, Mexico. **1948**, 496 p.

KOTIVUORI, E.; KORHONEN, L.; PACKALEN, P. Nationwide Airborne Laser Scanning Based Models for Volume, Biomass and Dominant Height in Finland. *Silva Fenn.* **2016**, 50, 1–28, doi:10.14214/sf.1567

KUMAR, S.; TALIB, M.; VERMA, P. Covid Detection from X-RAY and CT Scans Using Transfer Learning – A Study. *Proc. Int. Conf. Artif. Intell. Smart Syst.* **2021**, 85–92.

KUMAR, V. Forestry Inventory Parameters and Carbon Mapping from Airborne Lidar. 2012. 89 p. **Tese (Mestrado em Geociências e observações da terra)**- University of Twente, Enschede, 2012.

KUMAR, V. Forestry Inventory Parameters and Carbon Mapping from Airborne Lidar, Master Degree, University of Twente, 2014.

LABEN, C.A.; BROWER, B. V. PROCESS FOR ENHANCING THE SPATIAL RESOLUTION OF MULTISPECTRAL IMAGERY USING PAN-SHARPENING. U.S. Pat. 2000.

LANDIS, J.R.; KOCH, G.G. An Application of Hierarchical Kappa-Type Statistics in the Assessment of Majority Agreement among Multiple Observers Author. *Biometrics.* **1977**, 33, 363–374.

LEFSKY, M.A.; COHEN, W.B.; HARDING, D.J.; PARKER, G.G.; ACKER, S.A.; GOWER, S.T. Lidar Remote Sensing of Above-Ground Biomass in Three Biomes. *Glob. Ecol. Biogeogr.* **2002**, *11*, 393–399, doi:10.1046/j.1466-822x.2002.00303.x.

LEITE, E.F.; ROSA, R. Análise Do Uso, Ocupação E Cobertura Da Terra Na Bacia Hidrográfica Do Rio Formiga, Tocantins. *Rev. Eletrônica Geogr.* **2012**, *4*, 90–106.

LIU, M.; FU, B.; FAN, D.; ZUO, P.; XIE, S.; HE, H.; LIU, L.; HUANG, L.; GAO, E.; ZHAO, M. Study on Transfer Learning Ability for Classifying Marsh Vegetation with Multi-Sensor Images Using DeepLabV3+ and HRNet Deep Learning Algorithms. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102531, doi:10.1016/j.jag.2021.102531.

LIU, W. T. H. Aplicações de sensoriamento remoto. 2nd ed.; Oficina de Textos: São Paulo, Brasil, 2015; pp 1-539.

LOCKS, C. J. APLICAÇÕES DA TECNOLOGIA LiDAR NO MONITORAMENTO DA EXPLORAÇÃO MADEIREIRA EM ÁREAS DE CONCESSÃO FLORESTAL, 2017. 83 p. **Dissertação (Mestrado em Ciências Florestais)** – Universidade de Brasília – Brasília , 2017.

LU, D. *et al.* A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon. *GIScience & Remote Sensing*, v. 51, n. 6, p. 695-709, 2014.

Ma, Z.; Redmond, R.L. Tau Coefficients for Accuracy Assessment of Classification of Remote Sensing Data. *Photogramm. Eng. Remote Sens.* **1995**, *61*, 435–439, doi:10.1037/h0043374.

MAXWELL, A.E.; WARNER, T.A.; FANG, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817, doi:10.1080/01431161.2018.1433343.

MEANS, J. E. *et al.* Predicting forest stand characteristics with airborne scanning lidar. **Photogrammetric Engineering and Remote Sensing**, v. 66, n. 11, p. 1367-1371, 2000.

MENESES, P. R.; ALMEIDA, T. de. Introdução ao processamento de imagens de sensoriamento remoto. **Universidade de Brasília, Brasília**, 2012.

MILLER, G.A.; NICELY, P.E. An Analysis of Perceptual Confusions Among Some English Consonants HE Over-All Effects of Noise And. *J. Acoust. Soc. Am.* **1955**, *27*, 338–352.

MIRANDA, M.S.; DE SANTIAGO, V.A.; KÖRTING, T.S.; LEONARDI, R.; DE FREITAS, M.L. Deep Convolutional Neural Network for Classifying Satellite Images with Heterogeneous Spatial Resolutions. In: International Conference on Computational Science and Its Applications – ICCSA 2019; Springer, Russia, 2021; Volume 12955. pp. 519-530. <https://doi.org/10.1007/978-3-030-24311-1>.

MISHRA, A. K., RATHA B. K. Study of random tree and random forest data mining algorithms for microarray data analysis. *International Journal on Advanced Electrical and Computer Engineering* v. 3, n. 4, p. 5-7, 2016.

MONTEIRO, N. B. C. *et al.* Aprendizagem de máquina na estimativa de clorofila-a via sensoriamento remoto. 2020. 91 p. TCC (Graduação em Sistemas de informações) – Universidade Federal de Santa Catarina. 2020.

MOREIRA, M. P. *et al.* Uso de sensoriamento remoto para avaliar a dinâmica de sucessão secundária na Amazônia Central. 2003. 103 p. Dissertação (Mestrado em Ciências Agrárias) – Universidade Federal do Amazonas. 2003.

MOREIRA, M.A.; ADAMI, M.; RUDORFF, B.F.T. Análise Espectral e Temporal Da Cultura Do Café Em Imagens Landsat. *Pesqui. Agropecuária Bras.* **2004**, 39, 223–231, doi:10.1590/s0100-204x2004000300004.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support Vector Machines in Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2011**, 66, 247–259, doi:10.1016/j.isprsjprs.2010.11.001.

MUSIAL, J.P.; BOJANOWSKI, J.S. Comparison of the Novel Probabilistic Self-Optimizing Vectorized Earth Observation Retrieval Classifier with Common Machine Learning Algorithms. *Remote Sens.* **2022**, 14, doi:10.3390/rs14020378.

NÆSSET, E. Practical Large-Scale Forest Stand Inventory Using a Small-Footprint Airborne Scanning Laser. *Scand. J. For. Res.* **2004**, 19, 164–179, doi:10.1080/02827580310019257.

NASCIMENTO, R. F. F. *et al.* O algoritmo Support Vector Machine (SVM): avaliação da separação ótima de classes em imagens CCD-CBERS-2. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 14., 2009, Natal, PA. Anais...São José dos Campos: INPE, p. 2079-2086, 2009.

NILSSON, M. Estimation of Tree Heights and Stand Volume Using an Airborne LiDAR System. **Remote Sensing of Environment**, v. 56, p. 1-7, 1996.

OH, S.; JUNG, J.; SHAO, G.; SHAO, G.; GALLION, J.; FEI, S. High-Resolution Canopy Height Model Generation and Validation Using USGS 3DEP LiDAR Data in Indiana, USA. *Remote Sens.* **2022**, 14, 935, doi:https://doi.org/10.3390/rs14040935.

PAIVA, R. *et al.* Análise de metacaracterísticas para classificação de uso e cobertura do solo utilizando random forest. In: **Anais do XI Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais**. SBC, 2020. p. 71-80.

PAL, M. Random forest classifier for remote sensing classification. **International journal of remote sensing**, v. 26, n. 1, p. 217-222, 2005.

PEREIRA, A.; GOMES, A.; PEREZ-ALBERTI, A. Vantagens da utilização de dados ALS (LiDAR) na caracterização morfológica das vertentes costeiras da Galiza Espanha: aplicação ao setor de Mougás. **Revista de Geografia e Ordenamento do Território**, Porto, v. 1, n. 4, p. 223-248, 2013.

PEREIRA, J.P.; SCHIMALSKI, M.B. *LiDAR Aplicado a Florestas Naturais*; Novas Edições Acadêmicas, 2014;

PINHO, C.; FEITOSA, F.; KUX, H. Classificação Automática de Cobertura Do Solo Urbano Em Imagem IKONOS: Comparação Entre a Abordagem Pixel-a-Pixel e Orientada a Objetos. *An. XII Simpósio Bras. Sensoriamento Remoto, Goiânia* **2005**, 4217–4224.

PINTO, F. M. CLASSIFICAÇÃO DO ESTÁGIO SUCESSIONAL DA VEGETAÇÃO EM ÁREAS DE FLORESTA OMBRÓFILA MISTA (FOM) COM EMPREGO DE IMAGENS DIGITAIS OBTIDAS POR VANT (VEÍCULO AÉREO NÃO TRIPULADO). Dissertação (Mestrado em Engenharia Florestal), Universidade do Estado de Santa Catarina, 2018.

PONZONI, F. J.; SHIMABUKURO, Y. E. Sensoriamento Remoto no estudo da vegetação, 2nd ed.; São José dos Campos: Parêntese, Brasil, 2009.

PULETTI, N.; CHIANUCCI, F.; CASTALDI, C. Use of Sentinel-2 for Forest Classification in Mediterranean Environments. *Ann. Silv. Res.* **2018**, *42*, 32–38, doi:10.12899/ASR-1463.

QIAN, Y.; ZHOU, W.; YU, W.; HAN, L.; LI, W.; ZHAO, W. Integrating Backdating and Transfer Learning in an Object-Based Framework for High Resolution Image Classification and Change Analysis. *Remote Sens.* **2020**, *12*, 1–19, doi:10.3390/rs12244094.

RAMOS, R. R. D.; DE MELO JÚNIOR, J. C. F. Mapping of the current land use in part of the Irrigated Perimeter Nilo Coelho, Petrolina-PE, Brazil. *Comunicata Scientiae*, v. 10, n. 1, p. 89-97, 2019.

REUTEBUCH, S. E.; ANDERSEN, H. E.; MCGAUGHEY, R. J. Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory. *Journal of Forestry*, v. 103, n. 6, p. 286-292, 2005.

RODRIGUES, M. T. *et al.* Sensoriamento remoto e geoprocessamento aplicado ao uso da terra para avaliação entre classificadores a partir do índice Kappa. *Revista Científica Eletrônica de Engenharia Florestal*, v. 23, n. 1, p. 60-70, 2014.

RODRIGUES, R. R. A sucessão florestal. In: MORELLATO, P. C., LEITÃO FILHO, H. F. **Ecologia e preservação de uma floresta tropical urbana: Reserva de Santa Genebra**. Campinas: UNICAMP, p. 30-36, 1995.

ROUSE JR., J.W.; HAAS, R.H.; SCHELL, J.A.; DEERING, D.W. MONITORING VEGETATION SYSTEMS IN THE GREAT PLAINS WITH ERTS. *Remote Sensingcenter* **1973**, 309–317, doi:10.1021/jf60203a024.

SANNIGRAHI, S.; BASU, B.; BASU, A.S.; PILLA, F. Development of Automated Marine Floating Plastic Detection System Using Sentinel-2 Imagery and Machine Learning Models. *Mar. Pollut. Bull.* **2022**, *178*, 113527, doi:10.1016/j.marpolbul.2022.113527.

SEMOLINI, R. Support Vector Machines, Inferência Transdutiva e o Problema de Classificação, 2002.

SEVEGNANI, L. *et al.* Estádios sucessionais da Floresta Ombrófila Mista em Santa Catarina. In: VIBRANS, A. C. *et al.* Inventário Florístico Florestal de Santa Catarina. Blumenau: Edifurb, Brasil, 2012; Volume 3, cap. 9, pp. 255-271

SHUGART, H.H. Importance of Structure in the Longer-Term Dynamics of Landscapes. *J. Geophys. Res. Atmos.* **2000**, *105*, 20065–20075, doi:10.1029/2000JD900096.

SILVA, G. O. EXTRAÇÃO DE VARIÁVEIS ECOLÓGICAS DA FLORESTA OMBRÓFILA MISTA EMPREGANDO DADOS OBTIDOS POR VANT. Dissertação (Mestrado em Engenharia Florestal), Universidade do Estado de Santa Catarina, 2020.

SILVA, I. L. G. *et al.* Detecção de árvores individuais em área florestal mista de coníferas por meio de dados LiDAR aerotransportando. **Advances in Forestry Science**, v. 4, n. 2, p. 107–112, 2017.

SIMINSKI, A.; FANTINI, A.C. Classificação da Mata Atlântica do litoral catarinense em estádios sucessionais: ajustando a lei ao ecossistema. **Floresta e Ambiente**, Seropédica, v.11, n.2, p20- 25, 2004.

SMITH, C; KONING, M. **DECISION TREES: A Visual Introduction for beginners**. Blue windmill media. 2017.

SMOLA, A. J. *et al.* **Introduction to large margin classifiers**. 2000.

SOARES, H. B. Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte. 2008. 180 f. **Tese (Doutorado em Automação e Sistemas; Engenharia de Computação; Telecomunicações)** - Universidade Federal do Rio Grande do Norte, Natal, 2008.

Sothe, C. CLASSIFICAÇÃO DO ESTÁDIO SUCESSIONAL DA VEGETAÇÃO EM ÁREAS DE FLORESTA OMBRÓFILA MISTA EMPREGANDO ANÁLISE BASEADA EM OBJETOS E ORTOIMAGENS. Dissertação (Mestrado em Engenharia Florestal), Universidade do Estado de Santa Catarina, 2015.

SOTHE, C.; DE ALMEIDA, C.M.; LIESENBERG, V.; SCHIMALSKI, M.B. Evaluating Sentinel-2 and Landsat-8 Data to Map Successional Forest Stages in a Subtropical Forest in Southern Brazil. *Remote Sens.* **2017**, 9, doi:10.3390/rs9080838

SOTHE, C.; LIESENBERG, V.; DE ALMEIDA, C.M.; SCHIMALSKI, M.B. Abordagens Para Classificação Do Estádio Sucessional Da Vegetação Do Parque Nacional de São Joaquim Empregando Imagens Landsat-8 e Rapideye. *Bol. Ciencias Geod.* **2017**, 23, 389–404, doi:10.1590/S1982-21702017000300026.

SOTHE, C.; DE ALMEIDA, C.M.; SCHIMALSKI, M.B.; LIESENBERG, V. Integration of WorldView-2 and LiDAR Data to Map a Subtropical Forest Area: Comparison of Machine Learning Algorithms. *Int. Geosci. Remote Sens. Symp.* **2018**, 2018-July, 6207–6210, doi:10.1109/IGARSS.2018.8517941

SOUSA, Wayne P. The role of disturbance in natural communities. **Annual review of ecology and systematics**, v. 15, n. 1, p. 353-391, 1984.

SWETANISHA, S.; PANDA, A.R.; BEHERA, D.K. Land Use/Land Cover Classification Using Machine Learning Models. *Int. J. Electr. Comput. Eng.* **2022**, 12, 2040–2046, doi:10.11591/ijece.v12i2.pp2040-2046.

SZOSTAK, M.; HAWRYŁO, P.; PIELA, D. Using of Sentinel-2 Images for Automation of the Forest Succession Detection. *Eur. J. Remote Sens.* **2018**, 51, 142–149, doi:10.1080/22797254.2017.1412272.

SZOSTAK, M.; LIKUS-CIEŚLIK, J.; PIETRZYKOWSKI, M. Planetscope Imageries and LiDAR Point Clouds Processing for Automation Land Cover Mapping and Vegetation Assessment of a Reclaimed Sulfur Mine. *Remote Sens.* **2021**, 13, doi:10.3390/rs13142717.

TASSETTI, A.N.; MALINVERNI, E.S.; HAHN, M. Texture Analysis to Improve Supervised Classification in IKONOS Imagery. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **2010**, *38*, 245–250.

TIEDE, D.; HOCHLEITNER, G.; BLASCHKE, T. A full GIS-based workflow for tree identification and tree crown delineation using laser scanning. In: **ISPRS Workshop CMRT**. 6 p. 2005.

VAN EWIIK, K.Y.; TREITZ, P.M.; SCOTT, N.A. Characterizing Forest Succession in Central Ontario Using Lidar-Derived Indices. *Photogramm. Eng. Remote Sensing* **2011**, *77*, 261–269, doi:10.14358/PERS.77.3.261.

VAPNIK, V. N. **The nature of Statistical learning theory**. Springer-Verlag, New York, 1995.

VENTURIERI, A. *et al.* Análise da expansão da agricultura de grãos na região de Santarém e Belterra, Oeste do estado do Pará. In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO**, Florianópolis, 2007.

VENTURIERI, A. *et al.* Análise da expansão da agricultura de grãos na região de Santarém e Belterra, Oeste do estado do Pará. In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO**, Florianópolis, 2007.

WAGNER, W. *et al.* From singlepulse to full-waveform airborne laser scanners: Potential and practical challenges. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 35, p. 201–206, 2004.

XIE, G.; NICULESCU, S. Mapping and Monitoring of Land Cover/Land Use (LCLU) Changes in the Crozon Peninsula (Brittany, France) from 2007 to 2018 by Machine Learning Algorithms (Support Vector Machine, Random Forest, and Convolutional Neural Network) and by Post-Classification Comparison (PCC). *Remote Sens.* **2021**, *13*, doi:10.3390/rs13193899

ZHU, M.; XU, Y.; MA, S.; LI, S.; MA, H.; HAN, Y. Effective Airplane Detection in Remote Sensing Images Based on Multilayer Feature Fusion and Improved Nonmaximal Suppression Algorithm. *Remote Sens.* **2019**, *11*, doi:10.3390/rs11091062.