UNIVERSIDADE FEDERAL DE SANTA MARIA
CAMPUS FREDERICO WESTPHALEN
DEPARTAMENTO DE ENGENHARIA FLORESTAL
CURSO DE ENGENHARIA FLORESTAL

Geovane Sestari

APLICAÇÃO DE TÉCNICAS DE *MACHINE LEARNING* A DADOS DE SENSORIAMENTO REMOTO PARA O ESTUDO DE FLORESTAS PLANTADAS DE PINUS

Frederico Westphalen, RS
2021

Geovane Sestari

APLICAÇÃO DE TÉCNICAS DE *MACHINE LEARNING* A DADOS DE SENSORIAMENTO REMOTO PARA O ESTUDO DE FLORESTAS PLANTADAS DE PINUS

Trabalho de conclusão de curso apresentado ao curso de Engenharia Florestal, da Universidade Federal de Santa Maria, *campus* de Frederico Westphalen (UFSM-FW, RS), como requisito parcial para obtenção do grau de **Engenheiro Florestal**.

Orientador: Prof.° Dr. Fábio Marcelo Breunig

Frederico Westphalen, RS
2021

Geovane Sestari

APLICAÇÃO DE TÉCNICAS DE *MACHINE LEARNING* A DADOS DE SENSORIAMENTO REMOTO PARA O ESTUDO DE FLORESTAS PLANTADAS DE PINUS

Trabalho de conclusão de curso apresentado ao curso de Engenharia Florestal, da Universidade Federal de Santa Maria, *campus* de Frederico Westphalen (UFSM-FW, RS), como requisito parcial para obtenção do grau de **Engenheiro Florestal**.

Aprovado em 31 de agosto de 2021:

_____

Fábio Marcelo Breunig, Dr. (UFSM)
(Presidente/Orientador)

_____

Edison Rogerio Perrando, Dr. (UFSM)

_____

Wiliam Gaida, Dr. (UFSM)

Frederico Westphalen, RS
2021

DEDICATÓRIA

A minha família.

AGRADECIMENTOS

A realização desse trabalho concretiza uma fase muito especial na vida de qualquer pessoa e, como todas as conquistas, ela é compartilhada com inúmeras pessoas que, às vezes de forma inconsciente, tiveram uma grande contribuição e significado. Desta forma gostaria de registrar o meu agradecimento a:

- Aos meus pais, Ana de Lourdes Paié Sestari e Aristides Aparecido Sestari, que deram os primeiros ensinamentos na vida;

- A minha irmã Paula Aparecida Sestari Venturi e seu esposo Hermínio Venturi Neto que sempre me apoiaram e compartilharam bons pensamentos ao longo da minha trajetória. Em especial a minha irmã que sempre me mostrou o valor fundamental da educação, além de dar os melhores exemplos que eu poderia ter;

- A minha irmã Cláudia Sestari e seu companheiro Carlos Bortolotti que também sempre me deram apoio ao longo desse trajeto, em destaque pela corona até a rodoviária de Curitiba na minha primeira viagem para cidade de Frederico Westphalen, RS;

- Aos meus sobrinhos Daniel, Daiane, Miguel e Heitor por serem parte fundamental na minha vida e sempre serem fonte de inspiração pelas atitudes verdadeiras que apenas as crianças conseguem nos proporcionar;

- Ao meu orientador Fábio Marcelo Breunig por todos os ensinamentos e conversas ao longo do período da faculdade;

- A todos os demais professores do Departamento de Engenharia Florestal da Universidade Federal de Santa Maria campus Frederico Westphalen por todo o conhecimento ao longo do período de faculdade;

- Em especial a Universidade Federal de Santa Maria e as Universidades Públicas pelo ensino gratuito e de qualidade fornecido.

Enfim a todos que não citei aqui, mas que fazem parte da minha vida e que contribuíram pelo meu desenvolvimento profissional e como ser humano.

RESUMO

APLICAÇÃO DE TÉCNICAS DE *MACHINE LEARNING* A DADOS DE SENSORIAMENTO REMOTO PARA O ESTUDO DE FLORESTAS PLANTADAS DE PINUS

AUTOR: Geovane Sestari
ORIENTADOR: Fábio Marcelo Breunig

A crescente pressão antrópica sobre os sistemas naturais tem sido bastante debatida nas últimas décadas. As florestas por apresentarem, além dos seus benefícios ambientais, diferentes aplicações para a produção de produtos utilizados pelos homens sofrem constantemente impactos negativos com o aumento de demanda por matéria prima. Estudos relacionados ao aumento de produtividade das espécies comerciais fornecem apoio para que as demandas ambientais e econômicas sejam atendidas. Dados em larga escala, confiáveis e com uma boa frequência de coleta são fundamentais para a realização desses estudos, tais como os dados de sensoriamento remoto. Associados a estes, as técnicas estatísticas de aprendizado de máquinas podem aumentar a capacidade desses dados em fornecer informações e inferências aos gestores, públicos ou privados, dos recursos naturais. O presente estudo teve como objetivo avaliar a performance de algoritmos de aprendizado de máquinas na predição da idade de plantios de *Pinus taeda*. Foi utilizada uma série multitemporal dos satélites Landsat, dados SRTM e índices de vegetação para análises do comportamento espectral dos plantios. Esses dados serviram de entrada para o modelo de B*osting Learning* XGboost, para a modelagem e predição da idade dos plantios. A avaliação espectral possibilitou identificar a época de plantio, desbastes aplicados nas áreas, assim como, estabilização do crescimento. O modelo de aprendizado de máquinas obtido teve um R^2 de 0,926, RMSE de ± 1,27 e um MAE ± 0,72, como métricas de desempenho, além de apresentar uma boa consistência quando aplicado em dados de plantio que não foram utilizados durante o treinamento.

Palavras-chave: Índices de vegetação, Inteligência Artificial. NDVI. XGboost.

ABSTRACT

PINUS PLANTED FORESTS STUDY WITH MACHINE LEARNING METHODS APPLIED ON REMOTE SENSING DATA

AUTHOR: Geovane Sestari
ADVISOR: Fábio Marcelo Breunig

The anthropic pression growth on natural systems have been discuss in recently decades. More than environment benefits, wood is used to multiple manufacturing applications. The forests suffer negative impacts with raw material demand increase. Studies about increase of commercial trees species productivity provide basis to meet economic and environmental demands. Data with large scale, reliable, and high frequency as Remote Sensing Data are essential for this purpose. Complementary way, statistical machine learning techniques can improve Remote Sensing data analysis and provide insights to natural resources managers from public and private sectors. This study applied image processing techniques on Remote Sensing data and evaluate a machine learning algorithm age prediction performance on *Pinus taeda* stands. We used a multitemporal series from Landsat data, SRTM data, and vegetation index for stands spectral pattern analysis. The same data was input to boosting learning model XGboost in age modeling. The spectral analysis enabled identify planting year, thinning, and growth slowdown. The machine learning model had a 0.926 R^2, ± 1,27 RMSE, and ± 0.72 MAE performance metrics. In addition, the model had consistency when apply on data from stands than were not in training data set.

Keywords: Vegetation Index. Artificial Intelligence. XGboost.

LISTA DE FIGURAS

Figura 1 – Mapa de localização da área de estudo. 28
Figura 2 – Variação da refletância de superfície em diferentes comprimentos de ondas sobre as áreas de efetivo plantio de Pinus taeda, entre o primeiro (a – b), segundo (b – c) e terceiro (c – d) desbastes.. 35
Figura 3 – Variação do NDVI acumulado até o oitavo ano após plantio nas fazendas Santo Antônio da Platina (A), Amparo (B), São Francisco do Chapecozinho (B) e Santo Expedito (C). 38
Figura 4 – Análise de distribuição boxplot dos dados de sensoriamento remoto..... 38
Figura 5 – Análise de distribuição de frequência dos dados de sensoriamento remoto. 40
Figura 6 – Mapa de correlação de Pearson. 41
Figura 7 – Predição do modelo final na Fazenda Santo Expedito. 43

LISTA DE TABELAS

Tabela 1 – Características do sensor *Thematic Mapper* (TM) presente no Landsat 5. 18
Tabela 2 – Características do sensor *Operational and Imager* (OLI) presente no Landsat 8. . 18
Tabela 3 – Principais fatores que afetam a Função de distribuição de Reflectância Bidirecional (FDRB) da vegetação. 22
Tabela 4 – Data das imagens selecionadas por órbita ponto. 29
Tabela 5 – Parâmetros de desempenho finais do treinamento do modelo XGBoost. 42

LISTA DE ABREVIATURAS E SIGLAS

ERTS	*Earth Resources Technology Satellite*
EVI	*Enhanced Vegetation Index*
FDRB	Função de Distribuição de Reflectância Bidirecional
IBÁ	Industria Brasileira de Árvores
Landsat	*Land Remote Sensing Satellite*
MDE	Modelo Digital de Elevação
NDVI	*Normalized Vegetation Index*
NASA	*National Aeronautics and Space Administration*
OLI	*Operational and Imager*
REM	Radiação Eletromagnética
SR	Sensoriamento Remoto
SRTM	*Shuttle Radar Terrain Mission*
TIRS	*Thermal Infrared Sensor*
TM	*Tematic Mapper*

SUMÁRIO

1 INTRODUÇÃO 12
1.2 OBJETIVOS GERAIS 14
1.3 OBJETIVOS ESPECÍFICOS 15
2 REVISÃO BIBLIOGRÁFICA 16
2.1 FLORESTAS PLANTADAS 16
2.1.1 Pinus taeda 16
2.2 SENSORIAMENTO REMOTO 17
2.2.1 Sistemas sensores da série Landsat 17
2.2.2 Sensoriamento remoto da vegetação 19
2.2.2.1 Índices de vegetação 23
2.3 APRENDIZADO DE MÁQUINA 24
3 METODOLOGIA 27
3.1 ÁREA DE ESTUDO 27
3.2 AQUISIÇÃO DOS DADOS 28
3.3 ÍNDICES DE VEGETAÇÃO 30
3.4 ANÁLISE EXPLORATÓRIA DOS DADOS 31
3.5 TREINAMENTO, VALIDAÇÃO E TESTE 31
4 RESULTADOS E DISCUSSÃO 34
4.1 DADOS ESPECTRAIS 34
4.2 NDVI ACUMULADO 36
4.3 ANÁLISE EXPLORATÓRIA DOS DADOS 36
4.4 TREINAMENTO, VALIDAÇÃO E TESTE 41
5 CONCLUSÃO 44
REFERÊNCIAS 45
APÊNDICE A – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA SANTO ANTONIO DA PLATINA EM DIFERENTES ANOS, PARANÁ, BRASIL. 52
APÊNDICE B – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA SANTO EXPEDITO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL. 53
APÊNDICE C – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA AMPARO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL. 54
APÊNDICE D – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA CHAPECOZINHO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL. 55

1 INTRODUÇÃO

O debate sobre a importância das florestas para o desenvolvimento sustentável tem sido bastante enriquecido nas últimas décadas com a crescente preocupação a respeito da degradação do meio ambiente e o aumento da ocorrência de eventos ambientais extremos, pois são essenciais para o equilíbrio dos ecossistemas, atuando como um mecanismo de sequestro de carbono (TIMOTHY et al., 2016; WAHEED et al., 2018) e preservam a qualidade de outros recursos naturais que, assim como as florestas, estão sob elevada pressão antrópica, como a água e o solo (ELLISON et al., 2017; MERTEN; MINELLA, 2013).

De outra perspectiva, a madeira é matéria prima para o beneficiamento de produtos que estão sob constante aumento de demanda (PAYN et al., 2015) e ainda, representa uma fonte renovável para geração de energia (LONG et al., 2013). Conciliar as necessidades ambientais presentes e futuras com as demandas econômicas tem sido uma tarefa difícil para muitos países na definição de suas prioridades para o planejamento a curto e logo prazos (ZHANG; STENGER; HAROU, 2015). Tanto nos países desenvolvidos, com o aumento da cobertura florestal, quanto nos países em desenvolvimento, com a redução constante da taxa de desmatamento, é perceptível a busca por esse equilíbrio (SLOAN; SAYER, 2015), contudo muitos avanços são necessários principalmente no que concerne às políticas governamentais de incentivo ao setor florestal, mitigação dos impactos ambientais do desenvolvimento econômico e das mudanças climáticas (DELACOTE; KALLIO, 2016; ROSSATO et al., 2018).

Às florestas plantadas são atribuídos vários benefícios ambientais, econômicos e sociais, eles a utilização de técnicas silviculturas intensivas que possibilitam o ganho de produtividade em menores áreas de plantios, apesar de ser essa relação bastante complexa, favorece a conservação das áreas de florestas naturais que passam a ser destinas prioritariamente para o cumprimento de suas funções ecológicas e de preservação da biodiversidade (PIRARD; SECCO; WARMAN, 2016), além de ser uma fonte de geração de riquezas e empregos que movimenta a economia de vários países. No Brasil, o reconhecimento e a posição de destaque do setor florestal são destinados aos investimentos do setor privado, principalmente da indústria de polpa e papel, uma vez que as políticas nacionais de reflorestamento e controle do desmatamento ainda não se mostraram efetivas (ZHANG; STENGER; HAROU, 2015). O país dedica a maior parte de sua produção para o mercado internacional, tendo como principal produto a polpa de celulose, que é um dos principais agregados entre os produtos florestais (ROSSATO et al., 2018).

Nesse contexto, as espécies do gênero *Eucalyptus sp.* e de *Pinus sp.* correspondem a mais de 93% das áreas de florestas plantadas do Brasil (IBÁ, 2019). Na região sul do país existe certa predominância de plantios de *Pinus sp.* que é atribuída a sua boa adaptabilidade as condições de solo e clima (FERRO et al., 2018; KÄFER et al., 2018) somadas a possibilidade de crescimento durante todo o ano (RADOGLOU; TESKEY, 1997; SAMUELSON et al., 2010; ALBAUGH et al., 2018), resultando em produtividades que se aproximam do potencial máximo da taxa de crescimento estimados na literatura de algumas espécies, como é o caso do *Pinus taeda* (75 m^3/ha/ano) (FARNUM; TIMMIS; KULP, 1983; ALBAUGH et al., 2018).

O *Pinus taeda* é uma espécie de conífera de regiões quentes cultivada extensivamente em vários lugares do mundo e que possui crescimento muito superior no Brasil em comparação aos plantios na região de origem da espécie, sul e sudeste dos Estados Unidos da América (EUA), e demais locais na qual é exótico (SAMUELSON et al., 2010; ALBAUGH et al., 2018). A tendência é de que esse cenário se mantenha ao longo dos próximos anos, haja vista a frequente sistematização sobre o processo produtivo, da espécie (RODRIGUEZ et al., 2018) e do gênero (FERRO et al., 2018), seu potencial em suprir a demanda por produtos de origem florestal e seu sistema de produção eficiente (OLIVEIRA et al., 2018). Apesar disso, segue constante a busca pela homogeneização dos plantios através do entendimento das relações entre a distribuição espacial dos indivíduos com as variações topográficas, edafoclimáticas locais e regionais (CAMPOE et al., 2016; HORST et al., 2018), assim como, do efeito da competição intraespecífica (CARDOSO et al., 2013).

Com isso, surge a necessidade da obtenção de informações detalhadas sobre a distribuição do crescimento das áreas de florestas plantadas. Dados de sensoriamento remoto vem sendo utilizados para viabilizar o acompanhamento do crescimento, e outros parâmetros, das florestas em escala local (KÄFER et al., 2018), regional (PARK et al., 2016) e global (VICENTE-SERRANO et al., 2016) tendo como principais vantagens a redução dos custos quando comparados aos métodos convencionais, maior frequência na obtenção de dados e de forma espacialmente contínua. Para tal, um dos métodos criados e difundidos para obtenção de informações intrínsecas da vegetação nessa área é por meio da utilização de índices espectrais de vegetação calculados a partir de medidas de reflectância de determinadas bandas de sistema de sensores (JENSEN, 2009). Associado a isso, técnicas de análise de dados focadas no uso da inteligência artificial estão se tornando viáveis com os avanços obtidos na ciência da computação nos últimos anos (BORGES et al., 2021).

Inteligência artificial (*Artificial Intelligence* – AI) é um campo de estudos que busca, através de sistemas computacionais, reproduzir o comportamento dos seres humanos durante o

processo de tomada de decisões e reconhecimento de padrões. As técnicas de aprendizado de máquina (*Machine Learning* – ML) constituem um subcampo da AI bem consolidado, no qual são utilizadas análises estatísticas com o intuito de extrair padrões e conhecimentos a partir de um conjunto de dados (NWANGANGA; CHAPPLE, 2020; BORGES et al., 2021).

Com base nesses conceitos, a utilização de dados de sensoriamento remoto e técnicas de aprendizado de máquina, eventualmente associados a dados de campo, podem fornecer vantagens no gerenciamento dos empreendimentos florestais através da geração de informações confiáveis. Dentre elas pode-se elencar as informações sobre a idade, crescimento, técnicas silviculturais que foram ou que podem ser aplicadas nos plantios e suas interações (OLIVEIRA et al., 2021).

A idade de uma floresta implantada, normalmente é conhecida devido ao registro da data na qual ocorreu o plantio. Contudo a utilização dessa informação de forma generalizada para análises futuras do crescimento da floresta e práticas de manejo, tais como desbaste ou colheita, pode não ser ideal, uma vez que, fatores edafoclimáticos impactam no crescimento inicial e ao longo de todo o ciclo de crescimento das florestas, fazendo com que florestas com datas de plantios próximas tenham tempos de crescimento diferentes. Por um outro aspecto, quando não há informações confiáveis da data de plantio e sobre as práticas silviculturais realizadas em um determinado local, é necessário que seja empregado um esforço para o levantamento dessas informações, seja por dados coletados no campo ou por informações de sensoriamento remoto.

Com base nisso, o presente trabalho tem como hipóteses que (1) análises especificas utilizando dados de sensoriamento remoto fornecem informações relevantes para o gerenciamento de plantios florestais, tais como identificação de padrões no crescimento dos plantios, e (2) técnicas de aprendizado de máquina utilizando dados de sensoriamento remoto podem ser utilizadas para a identificação de uma idade aproximada de florestas plantadas, assim como, diferenciar áreas com datas de plantio próximas.

1.2 OBJETIVOS GERAIS

Avaliar a performance e os resultados de algoritmos de aprendizado de máquina na predição da idade de plantios de *Pinus taeda*.

1.3 OBJETIVOS ESPECÍFICOS

- Aplicar técnicas de processamento de imagens em uma série multitemporal de dados Landsat, tais como a obtenção dos índices NDVI e EVI;
- Realizar o treinamento de um algoritmo de aprendizado de máquina utilizando os dados Landsat e os índices de vegetação, extraídos de áreas com plantios de *Pinus taeda*, tendo como alvo a idade em anos a partir da data de plantio;
- Aplicar o modelo obtido na predição de idade em áreas que não foram utilizadas na construção do modelo (validar o modelo).

2 REVISÃO BIBLIOGRÁFICA

2.1 FLORESTAS PLANTADAS

O setor florestal na Região Sul do Brasil se destaca pela concentração de plantios de *Pinus sp.*, sendo os estados do Paraná (PR) e Santa Catarina (SC) os dois principais produtores. De acordo com o relatório da Industria Brasileira de Árvores (IBÁ) em 2018 o estado de SC possuía 544 mil hectares, que correspondem a 34 % das áreas plantadas com esse gênero no país e o estado do PR 672 mil hectares (42%) (IBÁ, 2019). A relevância desses números é atribuída a adaptabilidade do gênero, que na região encontra condições edafoclimáticas semelhantes à sua região de origem no hemisfério norte (FERRO et al., 2018; KÄFER et al., 2018).

2.1.1 *Pinus taeda*

As espécies de coníferas apresentam algumas peculiaridades em relação as demais espécies que influenciam diretamente no seu crescimento e ganho de biomassa. Por não apresentarem características decíduas possuem potencial de crescimento durante o ano todo, desde que não ocorram limitações de nutrição e do meio ambiente (RADOGLOU; TESKEY, 1997; SAMUELSON et al., 2010; ALBAUGH et al., 2018). Especificamente o *Pinus taeda* apresenta a longevidade de sua folhagem por volta de 18 meses (as das coníferas de clima frio possuem até 20 anos (SCHOETTLE; FAHEY; SHOETTLE, 1994)), sendo aproximadamente 20% desse tempo dedicado ao seu desenvolvimento completo. Se por um lado existe a presença das acículas durante todo o ano, o índice de área foliar (IAF) máximo só é alcançado durante os poucos meses nos quais a folhagem do ano corrente supera a produção fotossintética liquida das acículas das brotações do ano anterior. Isso ocorre devido ao lento tempo de crescimento até atingirem o seu potencial máximo, mas que pode ser compensado pelo período de crescimento maior (RADOGLOU; TESKEY, 1997).

O *Pinus taeda* é uma espécie de conífera de regiões quentes que, no cenário atual de produtos de origem florestal, possui grande demanda para suprimento de matéria prima, principalmente, para celulose e papel, assim como, para a indústria de painéis e madeira serrada (OLIVEIRA et al., 2018). Tendo como local de origem a região sul e sudeste dos Estados Unidos da América (EUA), é a principal espécie que abastece a demanda de produtos florestais no país (JOKELA; DOUGHERTY; MARTIN, 2004; SAMUELSON; JOHNSEN; STOKES,

2004; ALBAUGH et al., 2018). Em outros locais a espécie também é cultivada de forma extensiva, com destaque para a Região Sul do Brasil (SAMUELSON et al., 2010; ZANINOVICH; FONTANA; GATTI, 2016; RODRIGUEZ et al., 2018). Apesar da elevada capacidade produtiva, sabe-se que a espécie é bastante limitada pela qualidade do sítio (Zhao et al., 2016), o que não garante o crescimento homogêneo em todas as áreas e tão pouco sua capacidade genética máxima (FARNUM; TIMMIS; KULP, 1983), independente do hemisfério do plantio.

2.2 SENSORIAMENTO REMOTO

O sensoriamento remoto é considerado uma técnica de obtenção de informações de alvos a partir de sua assinatura espectral, sendo essa terminologia definida entre o início dos anos 1960s e início dos anos 1970s, uma vez que as diferentes abordagens que descreviam o processo de obtenção de dados com câmeras/sensores para o estudo de alvos a uma distância relativa não possuíam uma abrangência completa aos avanços obtidos na época. A partir desse entendimento surgiu o termo sensoriamento remoto, que pode ser descrito como a ciência/técnica ou a arte de obter informações sobre um objeto a uma certa distância considerando suas particularidades no processo de interação com a radiação eletromagnética (REM) e, se possível, sem que ocorra interferências durante a coleta (JENSEN, 2009). Outras definições restringem sua abrangência para os fenômenos terrestres (NOVO, 2002).

2.2.1 Sistemas sensores da série Landsat

Grandes avanços na área do sensoriamento remoto ocorreram na segunda metade do século XX, novas tecnologias proporcionaram o desenvolvimento de diferentes sistemas de satélites, sensores e detectores. Nesse período, aplicado ao estudo dos recursos naturais, destaca-se o lançamento da série de satélites da *National Aeronautics and Space Administration* (NASA) denominada inicialmente como *Earth Resources Technology Satellite* (ERTS) e posteriormente como *Land Remote Sensing Satellite* (Landsat). A série Landsat compreende um conjunto de 8 satélites, lançados até hoje, que fornecem imagens desde o ano de 1972 até os dias atuais, com exceção do satélite número 6 devido a uma falha durante o lançamento.

O Landsat 5 foi lançado em 1 de março de 1984 e permaneceu em operação com fornecimento de dados até 2012, após o lançamento do 8º satélite Landsat. Abordo do satélite encontrava-se o *Tematic Mapper* (TM), que é um sensor óptico mecânico de imageamento

multiespectral por varredura capturando a energia nas faixas do visível, infravermelho próximo, infravermelho de ondas curtas e infravermelho termal do espectro eletromagnético (JENSEN, 2009). As principais características do sensor são apresentadas na Tabela 1.

Tabela 1 – Características do sensor *Thematic Mapper* (TM) presente no Landsat 5.

Bandas	Resoluções			
	Espacial (m)	Temporal	Radiométrica (bits)	Espectral (µm)
1	30	16 dias	8 (valores de 0 a 255)	0,45 – 0,52
2				0,52 – 0,60
3				0,63 – 0,69
4				0,76 – 0,90
5				1,55 – 1,75
6	120			10,4 – 12,5
7	30			2,08 – 2,35

Fonte: Adaptado de Jensen (2009).

Lançado no dia 11 de fevereiro de 2013 o oitavo satélite da série Landsat, atualmente junto ao Landsat 7, está em operação com o fornecimento de dados sem perdas de informações no Brasil, possui dois sensores a bordo: o *Operational and Imager* (OLI) e o *Thermal InfraRed Sensor* (TIRS). As principais características dos produtos do sensor OLI são apresentadas na Tabela 2.

Tabela 2 – Características do sensor Operational and Imager (OLI) presente no Landsat 8.

Bandas	Resoluções			
	Espacial (m)	Temporal	Radiométrica (bits)	Espectral (µm)
1	30	16 dias	12 (valores de 0 a 4095)	0,43 – 0,45
2				0,45 – 0,51
3				0,53 – 0,59
4				0,64 – 0,67
5				0,85 – 0,88
6				1,57 – 1,65
7				2,11 – 2,29
8 (Pancromático)	15			0,50 – 0,68
9	30			1,36 – 1,38

Fonte: Adaptado de Roy et al. (2014).

Além de possuir a banda pancromática com resolução de 15 metros, o sensor OLI apresenta uma banda curta na faixa do azul (1), para maior sensibilidade da resposta da clorofila e outros materiais suspensos em corpos da água nas zonas costeiras, assim como para trabalhos envolvendo as propriedades de partículas solidas na atmosfera, e uma banda curta na faixa do infravermelho (9), para detecção das nuvens cirrus (ROY et al., 2014).

2.2.2 Sensoriamento remoto da vegetação

O meio ambiente em escala global tem sido alvo de interesse das diferentes áreas cientificas, muito devido a incipiente preocupação dos impactos negativos das ações humanas sobre os ciclos que ocorrem de forma natural. A possibilidade da visualização completa de toda a dinâmica da superfície da terra tem um papel fundamental para o entendimento dos fluxos da água, do carbono e do nitrogênio (GITELSON, 2004). A vegetação tem uma posição de destaque no funcionamento desses ciclos, pois participa diretamente por meio da fixação do CO_2 atmosférico (TIMOTHY et al., 2016), mitiga a emissão de CO_2 a longo e curto prazo (WAHEED et al., 2018), regula o regime hídrico em escala local, regional e global, e proporciona maior infiltração de água no solo (ELLISON et al., 2017), contribui para a manutenção do nitrogênio armazenado no solo, além da participação em muitos outros processos (CARMO et al., 2012). Ainda, representa elevada área de cobertura sobre a superfície terrestre (JENSEN, 2009; PAYN et al., 2015; SLOAN; SAYER, 2015), é fundamental no equilíbrio dos ecossistemas, serve como fonte de matéria prima e energia renováveis (LONG et al., 2013) e, consequentemente, possui crescente demanda para a produção de papel e celulose, painéis, carvão vegetal, madeira serrada para móveis e construção civil (IBÁ, 2019), entre outros.

Com isso, o sensoriamento remoto é fundamental no estudo das florestas e vem fornecendo a cada dia novas tecnologias e técnicas para a administração, monitoramento, estimação da biomassa acima do solo (TIMOTHY et al., 2016; DITTMANN et al., 2017; CHAO et al., 2019; ALMEIDA et al., 2019) e produtividade (PARK et al., 2016), assim como, auxilia no manejo/gestão dos recursos florestais.

Um dos principais pontos para os avanços nessa área está na caracterização do comportamento da vegetação a nível de folha e de dossel no processo de interação com o fluxo radiante incidente, que pode ser entendido como a radiação eletromagnética de uma fonte de emissão, por exemplo o Sol, que chega diretamente, ou espalhada, até a superfície foliar.

Na superfície da folha ou no seu interior o fluxo radiante pode ser absorvido pelos pigmentos, transmitido pelos espaços em sua estrutura, atingindo as camadas do estrato inferior ou o solo, ou ainda refletido. A obtenção dos respectivos valores de absorção, transmissão e reflexão pode ser feito a partir do balanço de energia quando feita a mensuração do que ocorre com todo o fluxo radiante incidente pelas equações (Equação 1) e (Equação 2) (JENSEN, 2009).

$$\frac{\phi i\lambda}{\phi i\lambda} = \left(\frac{\phi\rho\lambda}{\phi i\lambda}\right) + \left(\frac{\phi\alpha\lambda}{\phi i\lambda}\right) + \left(\frac{\phi\tau\lambda}{\phi i\lambda}\right) \quad (1)$$

$$i\lambda = \rho\lambda + \alpha\lambda + \tau\lambda \quad (2)$$

Onde, $\phi_{i\lambda}$ = fluxo radiante incidente (W); $\phi_{\rho\lambda}$ = fluxo radiante refletido (W); $\phi_{\alpha\lambda}$ = fluxo radiante absorvido (W); ϕ_{τ} = fluxo radiante transmitido (W); i_{λ} = energia incidente; ρ_{λ} = energia refletida; α_{λ} = energia absorvida; τ_{λ} = energia transmitida.

Os sistemas de sensoriamento remoto detectam a energia que é refletida, sendo as faixas de maior interesse concentradas entre 0,35 µm e 2,6 µm do espectro eletromagnético, tendo a vegetação, assim como cada material, uma resposta específica quando da interação da radiação eletromagnética de cada comprimento de onda com os alvos. Sendo esse intervalo utilizado, pois nele estão concentradas as regiões do visível, infravermelho próximo e o infravermelho médio que fornecem informações relevantes a respeito da atividade fotossintética, espaços intercelulares e conteúdo de água nas folhas, respectivamente (NOVO, 2002)

No processo de conversão da energia luminosa para energia química, fotossíntese, realizado pelas plantas existem três fatores condicionantes: presença de dióxido de carbono (CO_2), água (H_2O) e irradiância (E_{λ}) medida em W m^{-2}. Os dois primeiros são obtidos pelo ar e pelas raízes, respectivamente, enquanto a irradiância é obtida pela luz solar. A absorção da energia luminosa no interior de uma folha verde sadia é feita por pigmentos fotossintetizantes nos comprimentos da região do visível (VIS) (0,35 µm – 0,70 µm). Esses pigmentos ficam localizados no interior dos cloroplastos, que por sua vez se encontram nas células dos mesófilos paliçádico e esponjoso (NOVO, 2002; JENSEN, 2009).

Os principais pigmentos no processo de absorção são a clorofila *a* e a clorofila *b*, que possuem picos de absorção na região do azul (*a*: 0,43 µm; *b*: 0,45 µm) e do vermelho (*a*: 0,66 µm; *b*: 0,65 µm). Ambas possuem baixa absorção na região do verde, em aproximadamente 0,54 µm. Os demais pigmentos possuem picos máximos de absorção em outras bandas da região do visível, entre eles destaca-se o pigmento β-caroteno que possuí elevada taxa também na

região do azul, especificamente no comprimento de 0,45 µm. Com isso, Jensen (2009) destaca as regiões entre 0,45 e 0,52 µm e entre 0,63 e 0,69 µm como sendo ótimas para observar a vegetação. Em geral, os pigmentos compreendem a antocianina, flavonoides, carotenos e clorofilas (GITELSON; SOLOVCHENKO, 2018; GITELSON et al., 2019).

Contudo, o fluxo radiante que incide sobre a superfície foliar possuí energia em todos os comprimentos de ondas. Dentro do intervalo de interesse para o estudo da vegetação além da região do visível destacam-se as regiões do infravermelho próximo (NIR) (0,7 µm – 1,2 µm) e infravermelho de ondas curtas (SWIR) (1,3 µm – 2,5 µm). No geral, a faixa do infravermelho compreende a região de pouca absorção pela vegetação (JENSEN, 2009).

Os comprimentos de ondas do NIR ao chegarem ao dossel e interagirem com a superfície foliar dos primeiros estratos (camadas de folhas mais altas) são em maior parte refletidos (em torno de 40 – 60 %) pelo mesófilo esponjoso. Nele ocorrem as principais trocas gasosas nos processos de fotossíntese e respiração, tanto na liberação do oxigênio (O_2) quanto na assimilação do dióxido de carbono (CO_2), logo, é uma região na estrutura interna de uma folha com bastantes espaços intercelulares permitindo que o restante da energia seja transmitido (40 – 60 %) para os estratos inferiores que refletem e transmitem na mesma proporção. Quanto a absorção, pode ocorrer em pequena quantidade (5 – 10%) no intervalo entre 0,92 µm e 0,98 µm devido a presença de vapor d'água, devendo esse intervalo ser evitado (NOVO, 2002; JENSEN, 2009).

O principal interesse na utilização do NIR para o estudo da vegetação é a relação proporcional presente entre as variáveis relacionadas a biomassa e a refletância registrada pelos sensores nos comprimentos de ondas desse intervalo. O inverso ocorre na região vermelho (RED), onde a relação é inversamente proporcional (GITELSON, 2004).

Por ser a soma da energia imediatamente refletida pelo mesófilo esponjoso somado a energia refletida pelos estratos inferiores (reflectância aditiva), quanto maior a quantidade de folhas verdes, fotossinteticamente ativas, maior serão os valores de reflectância no NIR e, devido ao aumento da absorção, menores os valores de refletância no RED. Concomitante a isso, a proximidade do dossel com o solo contribui para o aumento ou redução dos valores registrados pelos sensores, pois apresenta valores baixos de refletância em ambos comprimentos de onda, que diminuem conforme o aumento do teor de umidade presente no solo. Sendo assim, conforme a vegetação se desenvolve, aumenta a quantidade de biomassa, aumenta a taxa de fotossíntese e se distância do solo, até o fechamento do dossel, os valores no NIR e no RED de reflectância se deslocam em sentidos opostos, sendo o primeiro em direção ao pico máximo e o segundo ao mínimo.

No SWIR as principais informações a serem obtidas da vegetação são associadas ao conteúdo de água presente nas folhas e dosséis, ou seja, sua turgidez (TREITZ; HOWARTH, 1999). A água apresenta forte absorção conforme o aumento dos comprimentos de onda. No intervalo entre 1,3 µm e 2,5 µm concentram-se alguns dos picos de absorção da água atmosférica (1,45 µm e 1,94 µm) e da água presente no mesófilo esponjoso (1,6 µm e 2,2 µm). Com essas informações podem ser feitas avaliações de estresses, geração de índices de vegetação que destaquem o conteúdo de água foliar, como por exemplo o *Normalized Difference Water Index* (NDWI) (GAO, 1996), entre outras aplicações.

Além das questões relacionadas a quantidade de energia absorvida, refletida e transmitida pela vegetação, os aspectos que afetam a função de distribuição da refletância bidirecional (FDRB) devem ser considerados em estudos na área do sensoriamento remoto. A FDRB se refere a relação entre a quantidade de energia refletida no ângulo de exitância e a quantidade do fluxo radiante fornecido em um ângulo de incidência (NICODEMUS, 1970; SCHAEPMAN-STRUB et al., 2006)

O primeiro fator a ser considerado refere-se a característica da superfície refletora. Quando a energia atinge um material como a vegetação a partir de um ângulo incidente, a maior porção que é refletida possui um ângulo de reflexão semelhante ao de incidência, as demais porções são refletidas para direções imprevisíveis devido a sua rugosidade. Isso caracteriza a vegetação como sendo uma superfície refletora difusa quase-perfeita, uma vez que um refletor difuso perfeito é aquele que reflete em todas as direções de maneira constante (superfície Lambertiana). Com isso, os principais fatores que influenciam na FDRB podem ser divididos em quatro partes: de iluminação, do sensor, da vegetação e do solo (Tabela 3).

Tabela 3 – Principais fatores que afetam a Função de distribuição de Reflectância Bidirecional (FDBR) da vegetação.

(continua)

Iluminação	Geometria	- Ângulo de incidência da fonte de – radiação - Azimute
	Características espectrais	- Comprimentos de onda (λ)
Sensor	Geometria	- Ângulo de visada - Azimute da direção de visada
	Sensibilidade espectral	- Comprimento de onda (λ)
	Campo de visada instantânea (IFOV)	

Tabela 3 – Principais fatores que afetam a Função de distribuição de Reflectância Bidirecional (FDBR) da vegetação.

(continuação)

Vegetação	Dossel	- Tipo (plantas ou árvores) - Fechamento (%) - Orientação (sistemática ou aleatória)
	Copas das árvores	- Tipo de copas (circular ou cônica) - Diâmetro
	Troncos ou caules	- Densidade - Diâmetro à altura do peito (DAP)
	Folha	- índice da área foliar (IAF) - Distribuição de ângulo foliar (planófilas ou erectófilas)
Solo	Textura Cor Conteúdo de umidade	

Fonte: Adaptado de Jensen (2009).

2.2.2.1 Índices de vegetação

No desenvolvimento de metodologias para modelagem de parâmetros biofísicos e de informações especificas da vegetação, uma das principais abordagens é por meio da obtenção de informações a partir da razão da refletância em diferentes regiões do espectro, que geram produtos denominados de índices de vegetação e são caracterizados por serem medidas radiométricas adimensionais (JENSEN, 2009; GALVÃO et al., 2018).

Os principais comprimentos de ondas utilizados no desenvolvimento dos índices de vegetação concentram-se nas regiões do RED e do NIR devido a relação distinta entre a reflectância com as variáveis associadas a biomassa de cada uma, caracterizando assim a intitulada "*Red Edge*" ou borda vermelha (TREITZ; HOWARTH, 1999). De forma auxiliar, informações referentes a outras regiões do espectro são utilizadas com o intuído de corrigir a ação de influências externas inerentes ao processo de aquisição de dados via sensores remotos. Um exemplo está no cálculo do *Enhanced Vegetation Index* (EVI) que utiliza a refletância na faixa do azul para a correção do efeito de partículas sólidas que causam o espalhamento da radiação na atmosfera (HUETE et al., 2002).

Dentre os índices de vegetação existentes, o *Normalized Difference Vegetation Index* (NDVI) é o mais comumente utilizado, sendo calculado a partir de medidas de reflectância das regiões do RED e do NIR possibilitando a avaliação e o monitoramento das mudanças biofísicas da vegetação. No geral, o NDVI apresenta boa estabilidade para a realização de comparações sazonais e interanuais nas mudanças de crescimento e atividade fotossintética, além de reduzir o efeito de ruídos presente nas diferentes imagens devido a normalização dos valores. Contudo, sua principal desvantagem é a não linearidade da relação do índice com os parâmetros biofísicos, fazendo com que os valores saturem principalmente em condições de moderada a alta biomassa (HUETE et al., 2002; GITELSON, 2004; KÄFER et al., 2018).

Da mesma forma, o EVI também é um dos principais índices de vegetação utilizado no estudo de parâmetros biofísicos da vegetação. Além das regiões do RED e do NIR, nele são utilizados os valores da região do azul (BLUE) e parâmetros que auxiliam na correção da interferência dos efeitos atmosféricos de aerossóis e do solo no sinal detectado pelos sensores. O EVI foi desenvolvido para melhorar a detecção de mudanças em condições de moderada a alta biomassa, assim como, em situações de não cobertura do solo pelo dossel (HUETE et al., 2002).

2.3 APRENDIZADO DE MÁQUINA

A inteligência artificial (*Artificial Intelligence* – AI) surgiu como um campo de estudo na década de 1950, com o objetivo de reproduzir o comportamento humano na tomada de decisões e reconhecimento de padrões, por meio da utilização de sistemas computacionais. Dentro desse campo de estudos, há diferentes segmentos que foram desenvolvidos, e viabilizados conforme os recursos computacionais disponíveis em cada momento, desde o seu surgimento. Dois dos principais e, atualmente, de maior destaque são as técnicas de aprendizado de máquina (*Machine Learning* – ML) e aprendizado profundo de máquinas (*Deep Learning*) (BORGES et al., 2021).

O aprendizado de máquinas é um campo mais amplo da AI no qual algoritmos específicos para cada tipo de situação (Ex.: classificação, regressão ou agrupamento), são treinados com base em dados históricos que expliquem o fenômeno de interesse e nos quais se saiba o resultado da variável de interesse ou não. Quando existe um resultado associado com os dados, normalmente são utilizados os algoritmos de aprendizado supervisionado. Caso contrário, são utilizados os modelos de aprendizado não supervisionado. Da mesma forma, as técnicas de aprendizado profundo de máquinas, constituem um subgrupo da ML no qual,

normalmente, se é necessário um volume maior de dados, uma vez que os problemas a serem solucionados são mais complexos computacionalmente, tais como, visão computacional, reconhecimento de voz, entre outros (NWANGANGA; CHAPPLE, 2020; BORGES et al., 2021).

A aplicação das diferentes técnicas de IA tem sido disseminada nos últimos anos com o advento de ferramentas *open source* que, com o aumento de popularidade no meio acadêmico e empresarial, formam uma extensa rede de colaboradores. As principais ferramentas voltadas para a ciência de dados e IA, e que possuem comunidades consistentes e em crescimento nos últimos anos, são as linguagens de programação R (R Core Team, 2021) e Python (Python, 2021). Além delas, linguagens bem conhecidas como MATLAB, Java e C/C++, assim como, comunidades mais novas, como a da linguagem Julia, também possuem destaque no desenvolvimento e implementação de modelos de IA (GAO et al., 2020).

Considerando o ambiente da linguagem R, a maior parte dos algoritmos estão implementados em pacotes de funções que são testadas pela comunidade que utiliza e contribui para a manutenção da ferramenta. Dentre eles estão os modelos supervisionados de regressão linear e regressão logística (R Core Team, 2021) e os modelos supervisionados de classificação *k-Nearest Neighbors* (VENABLES; RIPLEY, 2003), *Naïve Bayes* (MEYER et al., 2021) e o modelo de classificação e regressão *Decision Trees* (THERNEAU; ATKINSON, 2019).

Além dos modelos simples, na ferramenta também estão implementadas técnicas que utilizam vários algoritmos para a resolução de um determinado problema, denominados de *ensemble learnings*. Como exemplo, o modelo de *bagging learning Random Forest* (LIAW; WIENER, 2002) e o modelo de *boosting learning Gradient Boosting* (CHEN et al., 2020). Sendo o modelo de *Gradient Boosting*, implementado no pacote *XGBoost*, bastante popular em competições de *machine learning* por obter boa performance em uma gama diversificada de problemas, de regressão ou de classificação. A estrutura básica desse modelo, assim como o modelo *Randon Forest*, são árvores de decisões. Contudo, as árvores de decisão são sucessivamente dispostas de uma forma que a posterior impulsione o ajuste obtido pela anterior (*Boosting Learning*) (CHEN; GUESTRIN, 2016).

Além dos pacotes específicos para implementação de determinados algoritmos, no R, assim como nas demais linguagens, existem pacotes que dão o suporte para a avaliação e melhoria da performance dos modelos de IA. O principal deles é o pacote *caret*, que possibilita a utilização de diferentes modelos, técnicas de reamostragem (*Cross-Validation*, tais como *k-Fold Cross-Validation*, *Leave-One-Out Cross-Validation* e *Random Cross-Validation*, e *Bootstrap Sampling*) e entre outras facilidades (KUHN, 2021). De forma complementar, o

pacote *raster* e suas dependências possibilitam o manuseio de dados georreferenciados, assim como, a predição com os modelos de IA nesses dados (HIJMANS, 2021).

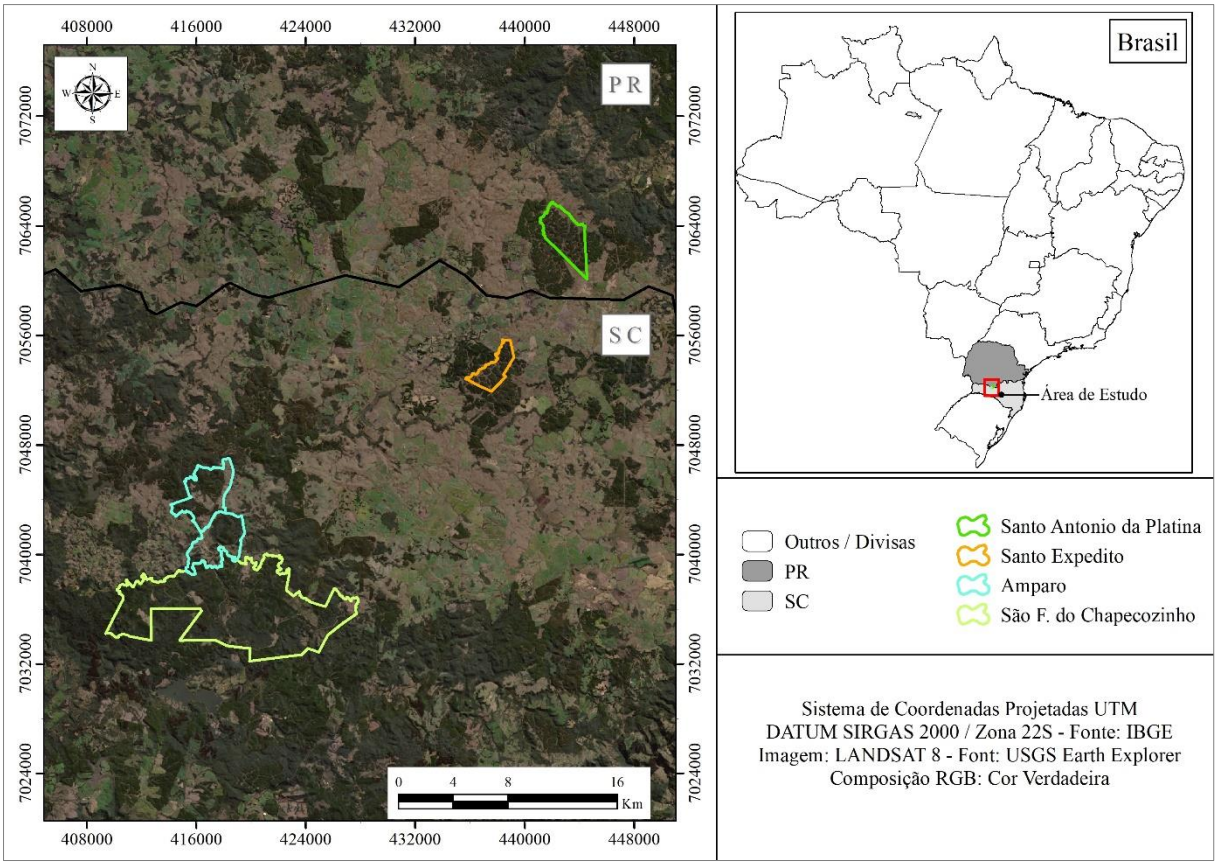
3 METODOLOGIA

3.1 ÁREA DE ESTUDO

A área de estudo corresponde a um conjunto de florestas plantadas com *Pinus taeda* no ano de 2002 com área equivalente a 1.885 ha de efetivo plantio em um espaçamento de 2 x 2,5 metros da empresa Adami S/A localizadas nos estados do Paraná e de Santa Catarina. Os povoamentos situam-se em diferentes fazendas, conforme detalhado na Figura 1. Por serem povoamentos dispersos encontram-se em duas classificações climáticas de acordo com a classificação de Köppen-Geiger (KOTTEK et al., 2006): [1] clima subtropical tipo Cfa com temperatura média inferior a 18° no mês mais frio e no mês mais quente temperatura média maior que 22°C; [2] clima temperado tipo Cfb com temperatura média inferior a 18° no mês mais frio e no mês mais quente temperatura média menor que 22°C (WREGE et al., 2012). Em relação a precipitação, a região apresenta variação anual entre 1.600 mm e 1.900 mm (ALVARES et al., 2013). Os solos predominantes são classificados como Cambissolos e Neossolos Litólicos (DORTZBACH et al., 2016), sendo ambos caracterizados por serem poucos desenvolvidos no processo de pedogênese apresentando o início da formação do horizonte B e a sua ausência, respectivamente (SANTOS et al., 2018).

Os plantios possuem como finalidade a fabricação de embalagens, peças para acabamentos residenciais e painéis, papel, pasta químico-mecânica e portas. Como forma de manejo, são realizadas intervenções silviculturais de desrama e desbaste em um ciclo de corte de aproximadamente 18 anos. No total são realizados desbastes entre o oitavo e nono ano, aos doze e aos quinze anos após o plantio, sendo o primeiro por meio de um sistema sistemático na quinta linha e seletivo nas demais para remoção de 50 % das árvores e os segundo e terceiro aplicados em um sistema de desbaste seletivo (PERRANDO et al., 2020).

Figura 1 – Mapa de localização da área de estudo.



Fonte: elaborado pelo Autor (2021).

3.2 AQUISIÇÃO DOS DADOS

Para a execução do estudo foram utilizados diferentes conjuntos de dados, dentre eles estão: arquivos com o delineamento das áreas de efetivo plantio, modelo digital de elevação (MDE) e a série multitemporal de produtos Landsat. Em relação a delimitação das áreas de efetivo plantio, talhões de *Pinus taeda*, estes dados foram disponibilizados pela empresa Adami S.A., sendo eles pertencentes a base de dados georreferenciados da empresa. Os MDTs fornecem informações importantes para a análise do comportamento temporal das áreas de plantio de acordo com as variações de altitude, declividade e ângulo de iluminação solar das áreas. Para tal, foram utilizados os dados da *Shuttle Radar Terrain Mission* (SRTM) obtidos a partir da plataforma USGS para toda a área de com resolução de 30 metros, compatível com os dados da série Landsat.

Os dados correspondentes a série multitemporal Landsat dos satélites número 5 e 8, foram adquiridos na plataforma *Earth Explorer* da USGS de froma gratuita entre os anos de

2002 e 2019. Esses produtos são disponibilizados nos níveis de processamento 1 e 2, onde o segundo corresponde aos produtos com correção atmosférica, de iluminação e geométrica, sendo assim mais consistente para análises temporais, além disso, os produtos nível 2 já estão convertidos para refletância de superfície que é uma parte importante do preparo dos dados para geração dos índices de vegetação, uma vez que, esses índices consideram a REM que está sendo diretamente refletida pela superfície de cobertura da terra.

Durante a aquisição, as datas das imagens foram utilizadas como critério de escolha. Dessa forma, foram selecionadas imagens dentro do menor intervalo de datas, entre os anos, possível. Devido ao posicionamento dos plantios foram utilizadas imagens das órbitas ponto 222/78 e 222/79, tendo como critério a mesma data de coleta entre as órbitas ponto do mesmo ano. Em resumo, as imagens foram coletadas entre as datas de 22/maio e 09/setembro (Tabela 4). Para o ano de 2012, devido a não disponibilidade de dados, foi utilizada uma interpolação linear para a para preenchimento dos valores durante a análise das medidas de reflectância extraídas.

Tabela 4 – Data das imagens selecionadas por órbita ponto.

Ano	Satélite	N° de dias no ano (órbita-ponto)		Dia / mês
		222-78	222-79	
2002	Landsat 5	219	219	07/agosto
2003		222	222	10/agosto
2004		161	161	09/junho
2005		211	211	30/julho
2006		166	166	15/junho
2007		201	201	20/julho
2008		252	252	08/setembro
2009		142	142	22/maio
2010		241	241	29/agosto
2011		228	228	16/agosto
2012	---------- Sem dados disponíveis ----------			
2013	Landsat 8	217	217	05/ agosto
2014		252	252	09/setembro
2015		159	159	08/junho
2016		226	226	13/agosto
2017		196	196	15/ julho
2018		199	199	18/ julho
2019		186	186	05/julho

Fonte: elaborado pelo Autor (2021).

3.3 ÍNDICES DE VEGETAÇÃO

Os índices de vegetação calculados com as imagens Landsat foram o *Normalized Difference Vegetation Index* (NDVI) (Equação 3) (ROUSE et al., 1974) e o *Enhanced Vegetation Index* (EVI) (Equação 4) (HUETE et al., 2002; JENSEN, 2009).

$$\mathrm{NDVI} = \frac{\rho\mathrm{NIR} - \rho\mathrm{RED}}{\rho\mathrm{NIR} + \rho\mathrm{RED}} \tag{3}$$

$$\mathrm{EVI} = \mathrm{G} * \frac{\rho\mathrm{NIR} - \rho\mathrm{RED}}{\rho\mathrm{NIR} + \mathrm{C1} * \rho\mathrm{RED} - \mathrm{C2} * \rho\mathrm{Blue} + \mathrm{L}} \tag{4}$$

Onde, ρNIR = Refletância de superfície no intervalo de comprimento de ondas do infravermelho de ondas curtas; ρRED = Refletância de superfície no intervalo de comprimento de ondas do vermelho; ρBlue = Refletância de superfície no intervalo de comprimento de ondas do azul; G = fator de ganho; C_1 e C_2 = coeficientes de ajuste da banda azul e vermelha para correção do espalhamento atmosférico por partículas sólidas; L: fator de ajuste para os solos.

De acordo com a literatura os valores de L, C_1, C_2 e G são, respectivamente: 1,0; 6,0; 7,5; e 2,5 (HUETE et al., 2002; JENSEN, 2009). As escalas para ambos os índices variam entre -1 e 1.

De forma complementar, foi aplicada uma técnica de identificação de regiões com padrões de NDVI acumulados distintos ao longo do tempo. Para tal, a série histórica do 3° (2004) até o 8° ano (2009) (primeiro desbaste) foi somada e normalizada, com a divisão pelo seu valor máximo. O início no 3° ano foi escolhido para evitar a influência da resposta do solo e de vegetações de gramíneas. A partir dessa análise, algumas inferências, devido à natureza do índice, podem ser obtidas, tais como áreas em estágios mais ou menos avançados de desenvolvimento, e ser fonte de apoio na tomada de decisão antes da aplicação de desbastes ou cortes rasos nas áreas. Um ponto importante a ser considerado refere-se à saturação dos índices. Em um trabalho com uma variedade distinta de pinus, Käfer et al. (2018) identificaram o início da saturação a partir de 10 ano para o NDVI e EVI. Isso indica o potencial desses índices retratarem o efetivo desenvolvimento dos plantios em análise.

3.4 ANÁLISE EXPLORATÓRIA DOS DADOS

Foram extraídas as informações das imagens Landsat, SRTM e índices de vegetação dentro das delimitações de efetivo plantio com uma borda de 30 metros em direção ao centro dos talhões, de forma a evitar a mistura espectral com os alvos do entorno. Uma vez com os dados tabulados, foram aplicadas técnicas de análise exploratória dos dados (EDA) com o intuito de identificar previamente *outliers*, medidas estatísticas e relacionamentos que auxiliem na construção do modelo de aprendizado de máquina. Os dados da fazenda Santo Expedito foram mantidos separados a partir desse momento para serem utilizados apenas no teste do modelo final a ser obtido.

A primeira técnica aplicada foi a análise por meio de gráficos *boxplot* para avaliação da distribuição dos dados por meio de medidas como o valor mínimo, máximo, mediana, média, intervalo entre quartis (IQR) e os 1° e 3° quartis. Por meio dessa análise é possível identificar valores extremos que, em alguns casos, são ruídos que podem prejudicar o procedimento de modelagem. Como forma de evitar esses dados, foram desconsiderados os valores inferiores ou superiores a 1,5 IQR somado ao 1° e ao 3° quartis, respectivamente.

Na sequência, foram analisadas a distribuição de frequência dos dados por meio de histogramas e a relação entre as variáveis por meio de um mapa de correlação de Pearson. Os histogramas auxiliam na identificação e validação da forma na qual os dados estão distribuídos. Alguns modelos matemáticos possuem como pressuposto que os dados se aproximem a normalidade, enquanto outros apenas performam melhor nessa situação. O mapa de correlação, por sua vez auxiliam na identificação de relações nos dados e pode ser utilizado como forma de seleção das variáveis como maior poder de predição.

3.5 TREINAMENTO, VALIDAÇÃO E TESTE

Após a avaliação e filtragem dos dados, os dados foram divididos em 75% das observações para o treinamento e 25% para o teste. Essa divisão é comumente realizada para que seja evitado o *overfitting*, quando o treinamento resulta em um modelo com uma alta capacidade preditiva no conjunto de dados utilizados no treinamento, porém não consegue bons resultados em situações reais (NWANGANGA; CHAPPLE, 2020).

Para que, durante o processo de treinamento do modelo de aprendizado de máquina, fosse possível avaliar o ajuste do modelo com a base de treinamento, de forma complementar foi utilizada a técnica de reamostragem *Cross-Validation*, mais especificamente a *k-Fold Cross-*

Validation. Nessa técnica o treinamento do modelo é realizado várias vezes (k), mantendo-se sempre uma subdivisão dos dados de treinamento de fora (*fold*) para que as métricas de performance possam ser computadas (validação do modelo). Ao término do treinamento, uma média das métricas pode ser obtida e analisada. Nesse caso, o parâmetro k=10 foi definido, com base nas capacidades computacionais disponíveis e no número de observações disponíveis (NWANGANGA; CHAPPLE, 2020).

Uma vez a base de dados pronta e essas especificações definidas, o modelo de aprendizado de máquinas *Extreme Gradient Boosting* (*XGBoost*) foi treinado, tendo com variável alvo (*target*) a idade em anos após a data de plantio dos talhões e como variáveis preditoras (*features*) as imagens Landsat, os índices de vegetação NDVI e EVI e o modelo digital de terreno SRTM.

Em relação aos dados Landsat, foram utilizadas as bandas de 1 a 5 e a banda 7 do Landsat 5 e as correspondentes bandas de 2 a 7 do Landsat 8, de acordo com as regiões do espectro eletromagnético. Para padronização da nomenclatura, foram utilizadas siglas B1, B2, B3, B4, B5 e B6 como referência para as respectivas bandas de cada sensor.

Durante o procedimento de treinamento, validação e teste do modelo, foram utilizados como parâmetros de avaliação da performance a Raiz do Erro Quadrático Médio (RMSE) (Equação 5), o Erro Médio Absoluto (MAE) (Equação 6) e o Coeficiente de Determinação (R^2) (Equação 7).

$$RMSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \tag{5}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |x_i - \hat{x}_i| \tag{6}$$

$$R^2 = 1 - \frac{RSS}{TSS} \tag{7}$$

$$RSS = \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \tag{8}$$

$$TSS = \sum_{i=1}^{N} (x_i - \bar{x}_i)^2 \tag{9}$$

Onde: N = número de observações; i = variável i; x_i = valor observado para a variável i; $\hat{x}_i$ = valor estimado para a variável i; $\bar{x}_i$ = média da variável i; RSS = soma dos resíduos ao quadrado; e TSS = soma total dos quadrados.

Após o treinamento, validação e teste, o modelo final aceito foi utilizado em um exercício prático de predição da idade do pinus, em anos, diretamente nos dados da Fazenda Santo Expedito.

4 RESULTADOS E DISCUSSÃO

4.1 DADOS ESPECTRAIS

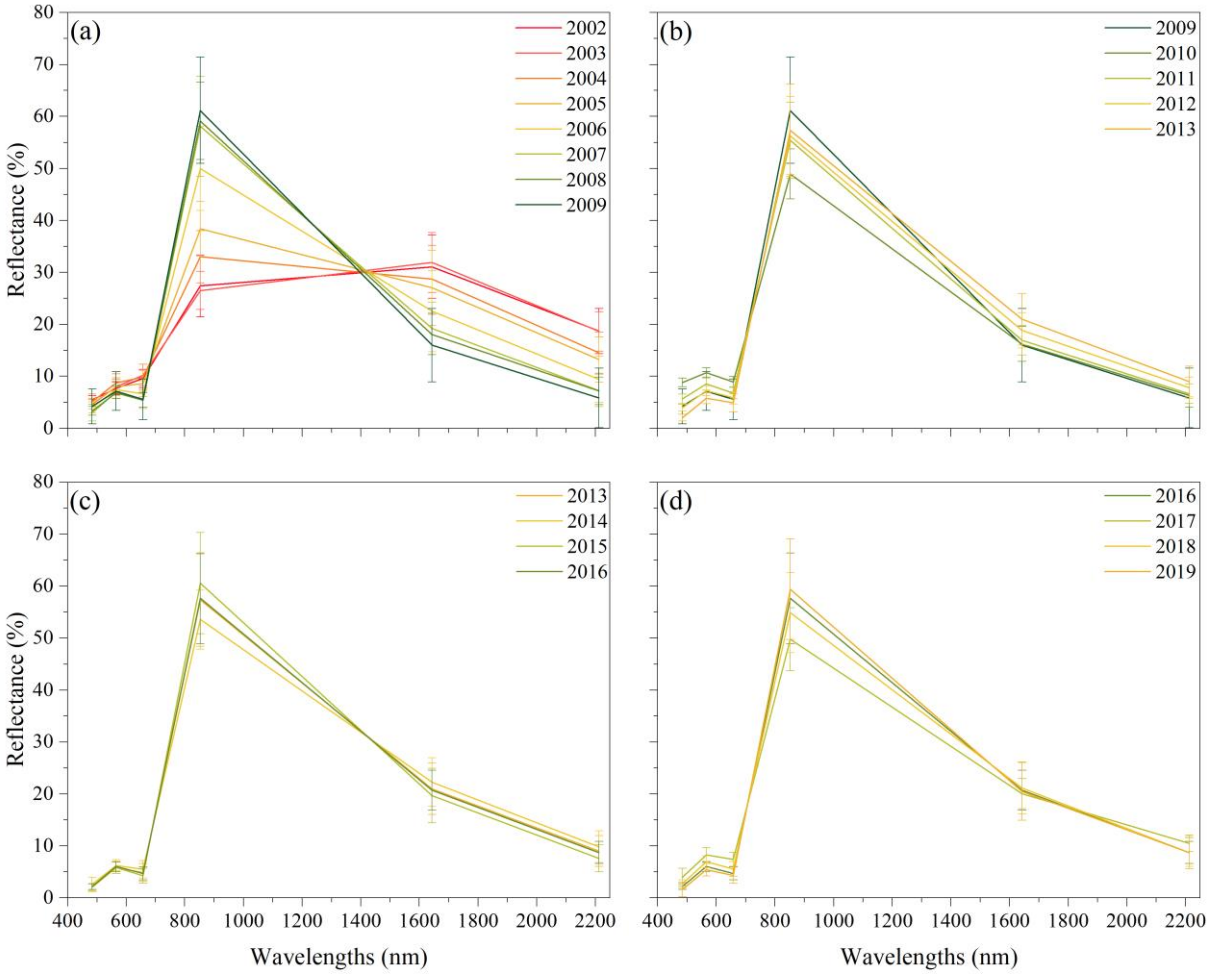
As curvas referentes a resposta espectral nos diferentes comprimentos de ondas geradas, apresentaram maiores valores percentuais no comprimento de onda do NIR, exceto no ano de implantação, aumentando gradativamente conforme o crescimento dos plantios e, consequentemente, maior densidade de copa e de dossel. Esse comportamento é reflexo do aumento da atividade fotossintética, onde boa parte da REM no intervalo do NIR é refletida pelas primeiras camadas do dossel e posteriormente através da reflectância aditiva. Na faixa do visível, os pigmentos passam a promover uma maior absorção da REM com o desenvolvimento do pinus e fechamento do dossel (JENSEN, 2009).

A Figura 2a retrata da melhor forma o comportamento do crescimento dos plantios por apresentar o comportamento espectral no ano de implantação (2002), no qual a maior parte da REM é refletida diretamente pelo solo apresentando valores baixos em todos os comprimentos de onda, típico de solo preparado para o plantio com pequena mistura espectral com a vegetação rasteira e não fotossinteticamente ativa (NPV). Conforme o crescimento e o vigor das plantas avança no tempo, a resposta espectral detectada se molda gradativamente ao padrão de reflectância da vegetação até o fechamento completo do dossel, nesse caso por volta de 2007 ou idade de 5 anos (Apêndices 1 ao 4) e atingindo os picos nos intervalos de comprimentos de ondas do vermelho e do NIR (JENSEN, 2009; KÄFER, 2018).

Por se tratar de áreas manejadas, conforme ocorrem as intervenções silviculturais os padrões de reflectância podem se modificar de acordo com o tipo de intervenção e sua intensidade. As Figuras 2b, 2c e 2d representam as variações causadas pelo manejo dos plantios, mais especificamente pela realização de desbastes aos 8 (2009), 12 (2013) e 15 (2016) anos, uma vez que os plantios sofrem três desbastes antes do corte raso por volta dos 18 anos. Como pode ser observado, entre o ano no qual são programados os desbastes no sistema de manejo aplicado sobre os plantios e o ano posterior, ocorre concomitantemente uma redução na reflectância na faixa do NIR e um aumento na faixa do vermelho. Essas alterações são explicadas em boa parte pelo aumento da radiação eletromagnética que atravessa o dossel e chega ao solo, devido a redução de IAF e consequentemente, da absorção por pigmentos. Logo, o primeiro desbaste (Figura 2b) é o de melhor visualização por ser realizado com maior intensidade reduzindo o número de árvores em 50 % do total.

Posterior a entrada do desbaste, o padrão espectral tende a representar a retomada do crescimento, aumento do vigor da vegetação e da atividade fotossintética com a diminuição da reflectância no intervalo do comprimento de onda do vermelho e, simultaneamente, com o aumento no intervalo do infravermelho próximo. Esse padrão se repete em todos os anos com desbastes programados.

Figura 2 – Variação da refletância de superfície em diferentes comprimentos de ondas sobre as áreas de efetivo plantio de *Pinus taeda*, entre o primeiro (a – b), segundo (b – c) e terceiro (c – d) desbastes.



Fonte: elaborado pelo Autor (2021).

4.2 NDVI ACUMULADO

A segunda análise aplicada foi com base na série de NDVI acumulado, normalizado pelo valor máximo do pixel, levando a um intervalo de variação entre 0 e 1. Na Fazenda Santo Antônio da Platina (Figura 3a) há a dominância de valores entre 0,78 e 0,84 com manchas, principalmente nas bordas dos talhões entre 0,72 e 0,78. Na Fazenda Santo Expedito (Figura 3c) houve uma predominância entre 0,72 e 0,78, enquanto nas Fazendas São Francisco do Chapecozinho e Amparo (Figura 3b) a predominância ocorreu na faixa entre 0,84 e 1.

Desde que validado com informações de campo, inventário florestal e de acordo com as necessidades de manejo dos produtos desejados, essa distinção das áreas pela série histórica do índice de vegetação pode fornecer ponderações relevantes nos tempos de manejo da área. Tais como, a possibilidade de uma entrada tardia do desbaste nas fazendas das Figuras 3a e 3c, uma antecipação nas fazendas da Figura 3b ou a existência de algum limitante que impactou ou esteja impactando o pleno vigor das áreas de forma homogênea.

4.3 ANÁLISE EXPLORATÓRIA DOS DADOS

Na primeira etapa da EDA, após o tratamento dos *outliers*, foram realizadas as demais análises exploratórias. Destaca-se novamente que nesse procedimento, assim como em toda a EDA, não foram utilizados os dados da Fazenda Santo Expedito.

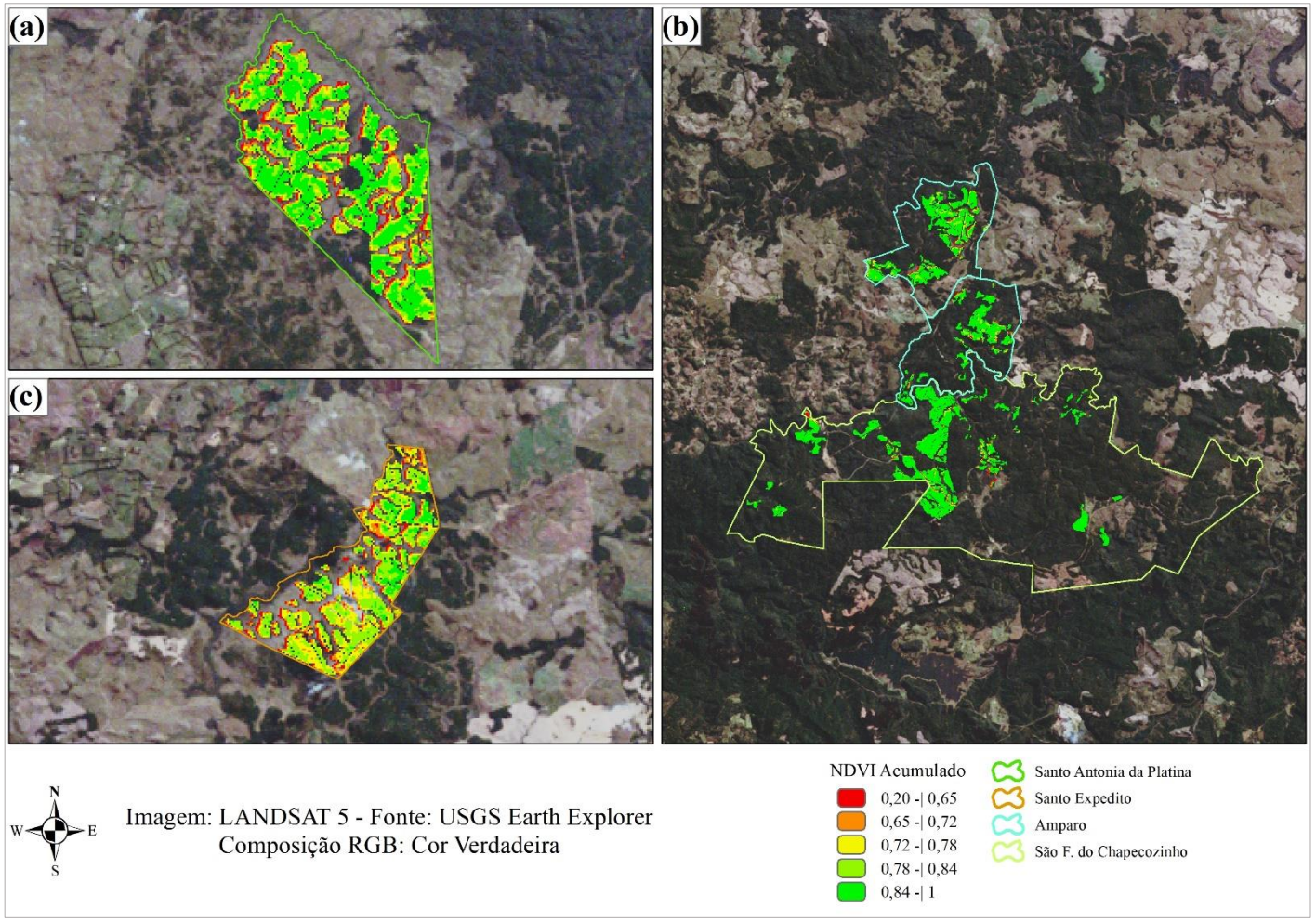
A distribuição dos dados foi feita com base nas diferentes variáveis obtidas por meio das medidas de reflectância, índices de vegetação e pela idade dos talhões (Figura 4). Mesmo com a aplicação da técnica de remoção de *outliers*, é possível observar a permanência de observações nas regiões extremas dos *boxplots*. Esses valores foram mantidos na análise, pois podem representar aspectos importantes da variabilidade dos plantios, assim como, podem agregar uma maior capacidade de generalização do modelo de aprendizado supervisionado em situações reais (teste do modelo). Nos dados SRTM foi aplicada a técnica de remoção de outliers, contudo não foram apresentados em uma distribuição temporal pois o mesmo MDE foi utilizado em todos os anos. Além disso, é interessante destacar a distinção entre as médias das séries temporais dos índices de vegetação. É possível notar que o NDVI apresentou uma maior sensibilidade as mudanças ocorridas nos plantios, tendo quedas acentuadas nas idades de desbastes nos plantios.

As distribuições de frequências do índice de vegetação EVI e da banda número 4 das imagens Landsat, após a remoção dos outliers, apresentaram curvas próximas a normalidade,

considerando que ambos apresentam alta correlação apesar da normalização sugerida pelo índice (GALVÃO et al., 2011). As Bandas 1, 2, 3, 5 e 6 apresentam curvas assimétricas à direita, enquanto os dados NDVI apresentaram curva assimétrica à esquerda. Os dados de elevação SRTM, ficaram concentrados em dois grupos, curva bimodal, com valores médio na faixa de 1100 e 1300 metros (Figura 5).

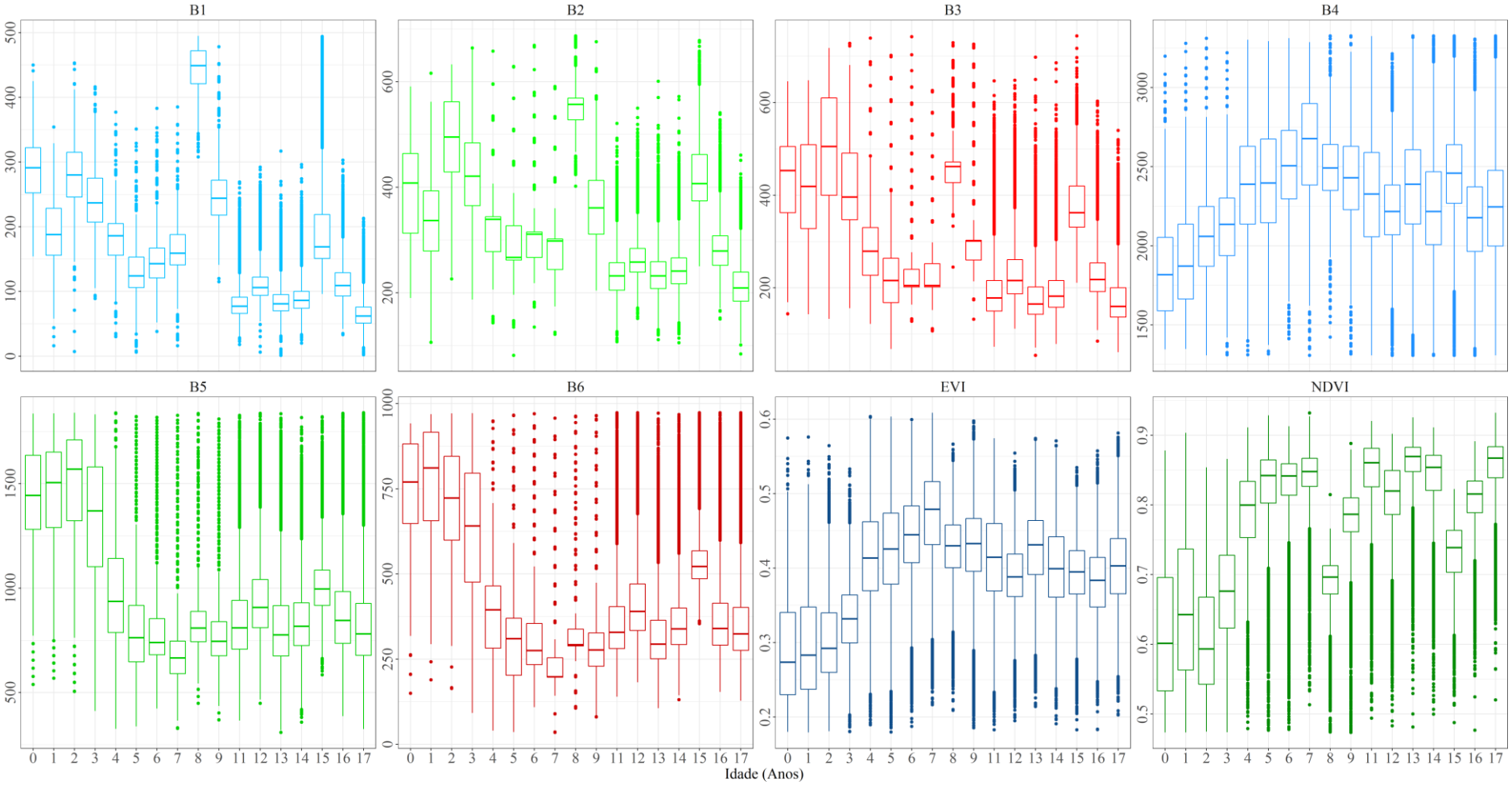
Uma vez que o modelo utilizado no treinamento não possui como pressuposto a normalidade dos dados de entrada, as situações com distribuições assimétricas e bimodal não sofreram nenhum tipo de procedimento de transformação, limpeza adicional ou criação de classes de distribuição categóricas.

Figura 3 – Variação do NDVI acumulado até o oitavo ano após plantio nas fazendas Santo Antônio da Platina (a), Amparo (b), São Francisco do Chapecozinho (b) e Santo Expedito (c).



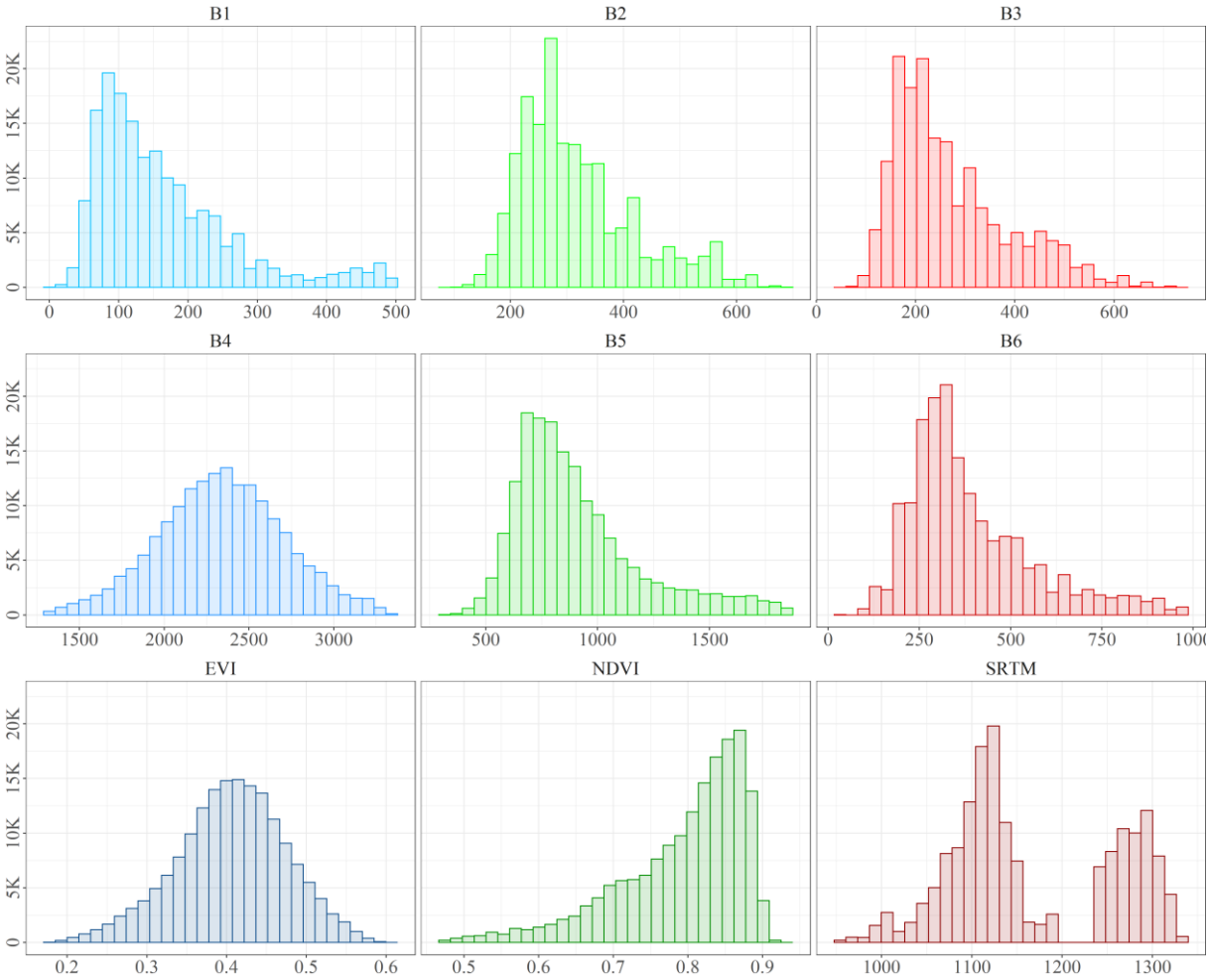
Fonte: elaborado pelo Autor (2021).

Figura 4 – Análise de distribuição *boxplot* dos dados de sensoriamento remoto.



Fonte: elaborado pelo Autor (2021).

Figura 5 – Análise de distribuição de frequência dos dados de sensoriamento remoto.



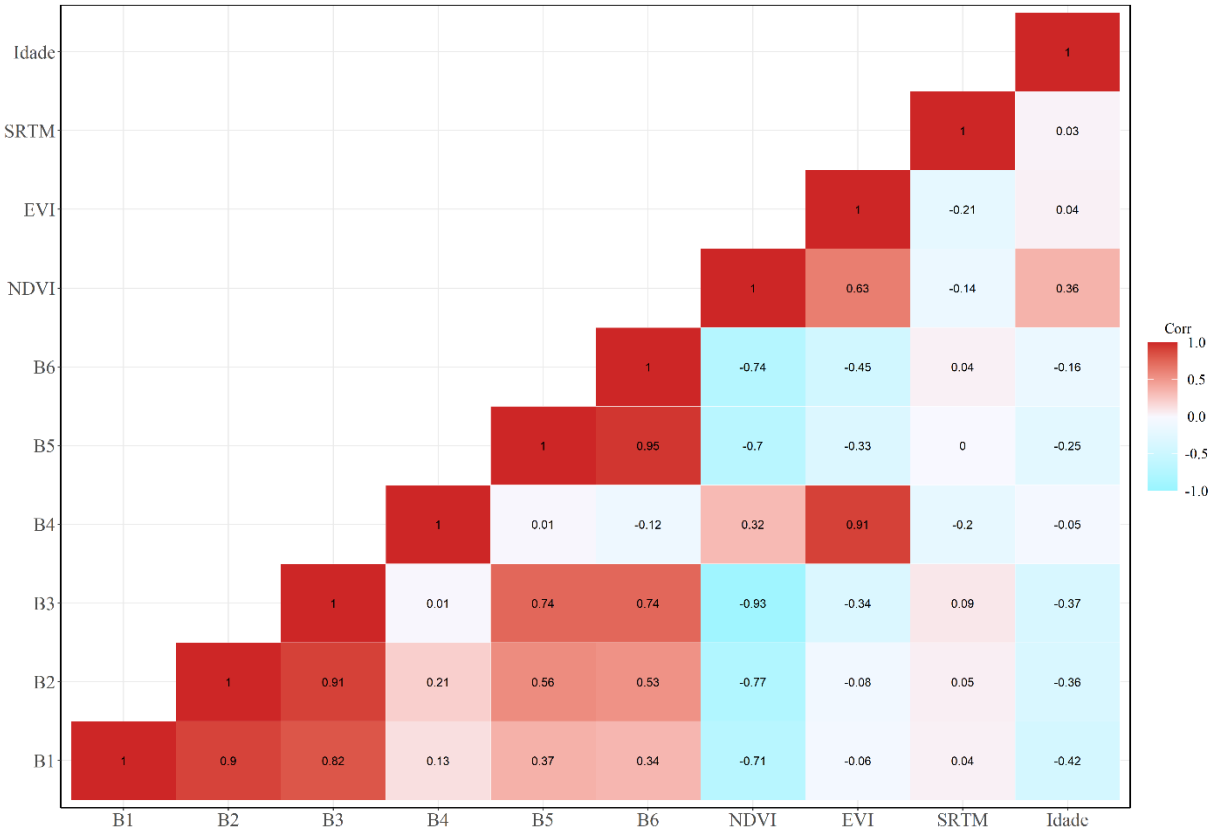
Fonte: elaborado pelo Autor (2021).

A última análise exploratória realizada, foi a construção do mapa de correlação de *Pearson*. Nesse caso, a variável de interesse foi utilizada para que fosse possível identificar ao seu relacionamento com as demais variáveis. De maneira geral, as variáveis apresentaram correlações fracas (<= 0,3) ou médias (<= 0,5) com a variável de interesse. É possível destacar as bandas 1, 2 e 3 por apresentarem correlações médias negativas, enquanto o NDVI teve uma correlação média positiva. Dentro do conjunto de dados disponíveis, essas variáveis apresentaram maior potencial preditivo.

Além, disso foi verificado a existências de correlações fortes entre as variáveis de entrada do modelo, como por exemplo entre o NDVI e a Banda 3 e do EVI com a Banda 4 dos sensores Landsat. Em alguns modelos de regressão, esse relacionamento, denominado de multicolinariedade, precisa ser tratado para reduzir a redundância dos dados utilizados no

treinamento do modelo. No entanto, para o modelo em questão não existe essa necessidade (NWANGANGA; CHAPPLE, 2020).

Figura 6 – Mapa de correlação de Pearson.



Fonte: elaborado pelo Autor (2021).

4.4 TREINAMENTO, VALIDAÇÃO E TESTE

Após o procedimento de treinamento e validação, o modelo foi testado no conjunto de dados de testes afim de se obter as medidas de performance finais. Os parâmetros obtidos estão apresentados na Tabela 5, onde foi possível observar que houve um bom ajuste do modelo pelo valor do R^2. Da mesma forma, os resultados obtidos em relação do RMSE e ao MAE ficaram entre 8 meses e 1 ano e 3 meses aproximadamente. Considerando a idade média dos dados de treinamento por volta de 10 anos, os erros percentuais em relação à média foram de 13% e 7%, respectivamente. Considerando apenas essas métricas, o modelo final obtido apresentou uma boa capacidade em solucionar o problema em questão. Cabe colocar em um projeto de implantação, o plantio de um talhão pode durar meses. Ainda assim, apenas o ano é tomado

como referência. Talhões grandes tendem a apresentar mais problemas dessa natureza. Assim, o RMSE pode ser parcialmente explicado pela variação intrínseca dos plantios de pinus.

Tabela 5 – Parâmetros de desempenho finais do treinamento do modelo *XGBoost*.

Método	R^2	RMSE (anos)	MAE (anos)
boosting learning	0,926	± 1,27	± 0,72

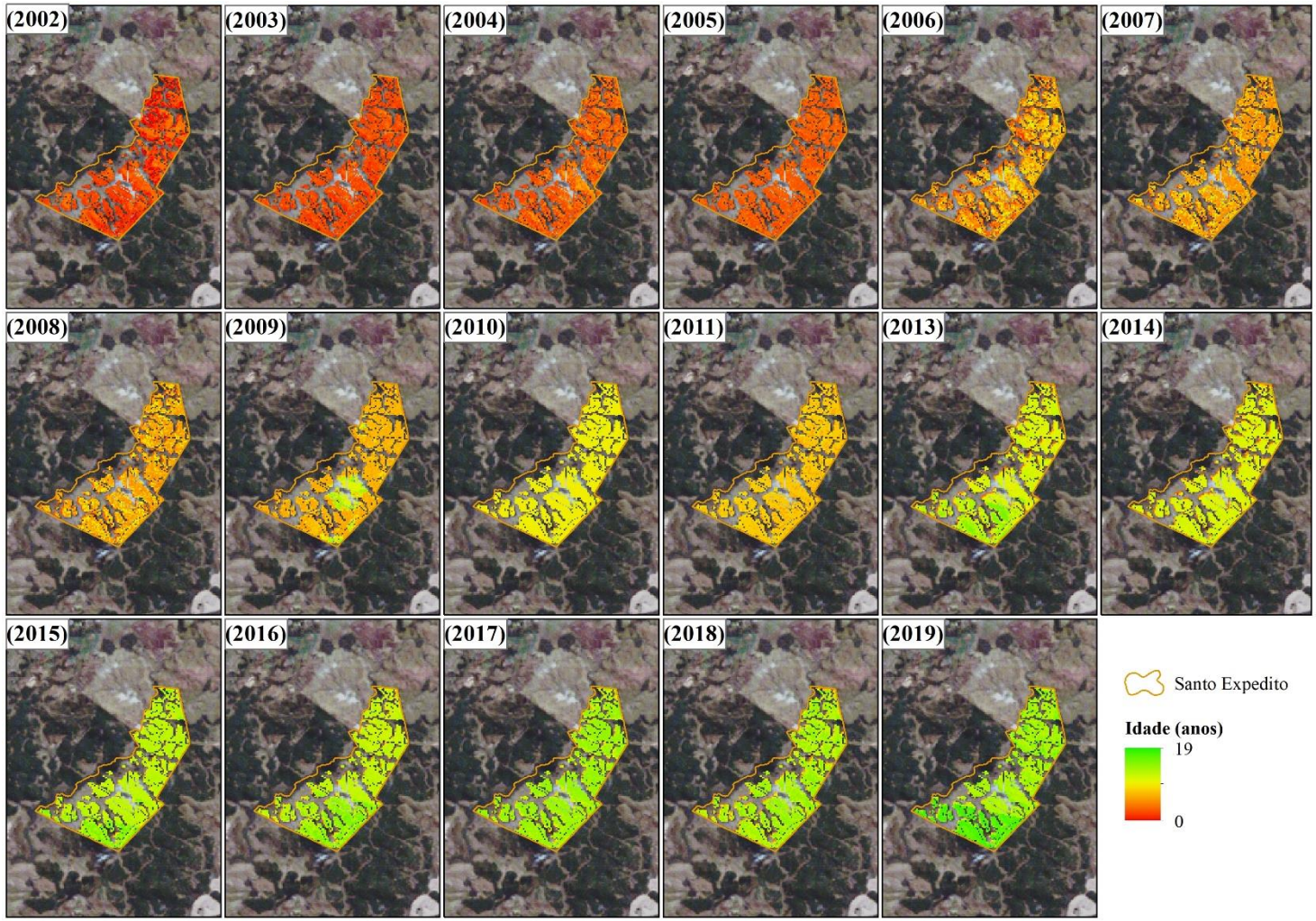
Fonte: elaborado pelo Autor (2021).

Como última etapa de avaliação do modelo obtido, foi construído um cenário de aplicação em uma situação real, procedimento conhecido como *deployment*. Para tal, foi realizada a predição da idade do pinus com uso dos dados de sensoriamento remoto da Fazenda Santo Expedito a nível de pixel, sem a necessidade de extração e tabulação, com a seguinte sequência de imagens: B1, B2, B3, B4, B5, B6, NDVI, EVI e SRTM. Como é possível observar na Figura 7, as predições obtidas foram consistentes em relação a idade dos talhões.

Oliveira et al. (2021) em um estudo no qual foram utilizados dados de sensoriamento remoto e técnicas de ML associados para distinção e identificação de espécies de eucalipto com base nos seus padrões de crescimento, obtiveram bons resultados, com média de acurácia entre todos os modelos que foram testados de 0,66. Nesse estudo específico, o problema em questão envolvia uma classificação ao invés de uma regressão, contudo os resultados obtidos evidenciam uma boa capacidade preditiva dos modelos de ML em problemas que envolvam dados espectrais e a vegetação.

Os autores ainda destacaram que o modelo que obteve maior acurácia foi o *Randon Forest*. Como dito anteriormente, o modelo apresentado nesse estudo apresenta a mesma estrutura básica, árvores de decisão. No geral, ambos normalmente apresentam desempenhos semelhante, desta forma, algoritmos que possuem estrutura semelhante são um bom ponto inicial para estudos semelhantes (OLIVEIRA et al., 2021).

Figura 7 – Predição do modelo final na Fazenda Santo Expedito.



Fonte: elaborado pelo Autor (2021)

5 CONCLUSÃO

As análises de caracterização das medidas de reflectância e índices de vegetação aplicadas diretamente sobre os dados de sensoriamento proporcionaram uma boa caracterização do comportamento dos povoamentos de *Pinus taeda* ao longo do seu ciclo de rotação, sendo possível identificar o início do crescimento (plantio), o aumento do vigor vegetativo, aplicação dos desbastes e estabilização crescimento. No geral as bandas B1, B2 e B3 das imagens Landsat e o índice NDVI, apresentaram maior potencial para predição da idade em anos de *Pinus taeda*. O modelo de ML utilizado para a inferência da idade dos plantios apresentou um comportamento consistente desde o treinamento até a predição em dados novos, sendo obtidos um R^2 de 0,926, RMSE de ± 1,27 e um MAE ± 0,72.

REFERÊNCIAS

ALBAUGH, T. J.; FOX, T. R.; MAIER, C. A.; CAMPOE, O. C.; RUBILAR, R. A.; COOK, R. L.; RAUMOND, J. E.; ALVARES, C. A.; STAPE, J. L. A common garden experiment examining light use efficiency and heat sum to explain growth differences in native and exotic *Pinus taeda*. **Forest Ecology And Management**, [s.l.], v. 425, p.35-44, out. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2018.05.033.

ALMEIDA, D. R. A.; STARK, S. C.; CHAZDON, R.; NELSON, B. W.; CESAR, R. G.; MELI, P.; GORGENS, E. B.; DUARTE, M. M.; VALBUENA, R.; MORENO, V. S.; MENDES, A. F.; AMAZONAS, N.; GONÇALVES, N. B.; SILVA, C. A.; SCHIETTI, J.; BRANCALION, P. H. S. The effectiveness of lidar remote sensing for monitoring forest cover attributes and landscape restoration. **Forest Ecology And Management**, [s.l.], v. 438, p.34-43, abr. 2019. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2019.02.002.

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEZ, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, [s.l.], v. 22, n. 6, p.711-728, 1 dez. 2013. Schweizerbart. http://dx.doi.org/10.1127/0941-2948/2013/0507.

BORGES, A. F. S.; LAURINDO, F. J. B.; SPÍNOLA, M. M.; GONÇALVES, C. A. M. The strategic use of artificial intelligence in the digital era: systematic literature review and future research directions. **International Journal Of Information Management**, [S.L.], v. 57, n. 1, p. 1-16, abr. 2021. Elsevier BV. http://dx.doi.org/10.1016/j.ijinfomgt.2020.102225.

CAMPOE, O. C.; MUNHOZ, J. S. B.; ALVAREZ, C. A.; CARNEIRO, R. L.; MATTOS, E. M.; FEREZ, A. P. C.; STAPE, J. L. Meteorological seasonality affecting individual tree growth in forest plantations in Brazil. **Forest Ecology And Management**, [s.l.], v. 380, p.149-160, nov. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2016.08.048.

CARDOSO, D. J.; LACERDA, A. E. B.; ROSOT, M. A. D.; GARRASTAZÚ, M. C.; LIMA, R. T. Influence of spacing regimes on the development of loblolly pine (*Pinus taeda* L.) in Southern Brazil. **Forest Ecology And Management**, [s.l.], v. 310, p.761-769, dez. 2013. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2013.09.021.

CARMO, J. B.; NETO, E. R. S.; DUARTE-NETO, P. J.; OMETTO, J. P. H. B.; MARTINELLI, L. A. Conversion of the coastal Atlantic forest to pasture: Consequences for the nitrogen cycle and soil greenhouse gas emissions. **Agriculture, Ecosystems & Environment**, [s.l.], v. 148, p.37-43, fev. 2012. Elsevier BV. http://dx.doi.org/10.1016/j.agee.2011.11.010.

CHAO, Z.; LIU, N.; ZHANG, P.; YING, T.; SONG, K. Estimation methods developing with remote sensing information for energy crop biomass: A comparative review. **Biomass And Bioenergy**, [s.l.], v. 122, p.414-425, mar. 2019. Elsevier BV. http://dx.doi.org/10.1016/j.biombioe.2019.02.002.

CHEN, T.; GUESTRIN, C. XGBoost. **Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining**, [S.L.], v. 22, n. 1, p. 785-794, 13 ago. 2016. ACM. http://dx.doi.org/10.1145/2939672.2939785.

CHEN, T.; HE, T.; BENESTY, M.; KHOTILOVICH, V.; TANG, Y.; CHO, K. C.; MITCHELL, R.; CANO, I.; ZHOU, T.; LI, M.; XIE, J.; LIN, M.; GENG, Y.; LI, Y. xgboost: Extreme Gradient Boosting. 2020. **R package version 1.5.0.1**. https://github.com/dmlc/xgboost

DELACOTE, P.; KALLIO, A. M. I. Forests and climate: new insights from forest sector modeling. **Journal Of Forest Economics**, [s.l.], v. 23, p.1-3, abr. 2016. Now Publishers. http://dx.doi.org/10.1016/j.jfe.2016.01.003.

DITTMANN, S.; THIESSEN, E.; HARTUNG, E. Applicability of different non-invasive methods for tree mass estimation: a review. **Forest Ecology And Management**, [s.l.], v. 398, p.208-215, ago. 2017. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2017.05.013.

DORTZBACH, D.; PEREIRA, M. G.; ANJOS, L. H. C.; FONTANA, A.; NETO, E. C. S. Genesis and classification of soils from subtropical mountain regions of Southern Brazil. **Revista Brasileira de Ciência do Solo**, [s.l.], v. 40, p.1-15, nov. 2016. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/18069657rbcs20150503.

ELLISON, D.; MORRIS, C, E.; LOCATELLI, B.; SHEIL, D.; COHEN, J.; MURDIYARSO, D.; GUTIERREZ.; V.; NORRDWIJK, M.; CREED, I. F.; POKORNY, J.; GAVEAU, D.; SPRACKLEN, D. V.; TOBELLA, A. B.; ILSTEDT, U.; TEULING, A. J.; GEBREHIWOT, S. G.; SANDS, D. C.; MUYS, B.; VERBIST, B.; SPRINGGAY, E.; SUGANDI, Y.; SULLIVAN, C. A. Trees, forests and water: cool insights for a hot world. **Global Environmental Change**, [s.l.], v. 43, p.51-61, mar. 2017. Elsevier BV. http://dx.doi.org/10.1016/j.gloenvcha.2017.01.002.

FARNUM, P.; TIMMIS, R.; KULP, J. L. Biotechnology of forest yield. Science, [s.l.], v. 219, n. 4585, p.694-702, 11 fev. 1983. **American Association for the Advancement of Science (AAAS)**. http://dx.doi.org/10.1126/science.219.4585.694.

FERRO, F. S.; SILVA, D. A. L.; ICIMOTO, F. H.; LAHR, F. A. R.; GONZÁLIZ-GARCÍA, S. Environmental life cycle assessment of industrial pine roundwood production in Brazilian forests. **Science Of The Total Environment**, [s.l.], v. 640-641, p.599-608, nov. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.scitotenv.2018.05.262.

GALVÃO, L. S.; SANTOS, J. R.; ROBERTS, D. A.; BREUNIG, F. M.; TOOMEY, M.; MOURA, Y. M. On intra-annual EVI variability in the dry season of tropical forest: a case study with modis and hyperspectral data. **Remote Sensing Of Environment**, [S.L.], v. 115, n. 9, p. 2350-2359, set. 2011. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2011.04.035.

GALVÃO, L. S.; EPIPHANIO, J. C. N.; BREUNIG, F. M.; FORMAGGIO, A. R. Crop Type Discrimination Using Hyperspectral Data. In: THENKABAIL, P. S.; LYON, J. G.; HUETE, A. (ed.). **Biophysical and Biochemical Characterization and Plant Species Studies**. 2. ed. S.L.: CRC Press, 2018. Cap. 6. p. 178-206.

GAO, B. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing Of Environment**, [s.l.], v. 58, n. 3, p.257-266, dez. 1996. Elsevier BV. http://dx.doi.org/10.1016/s0034-4257(96)00067-3.

GAO, K.; MEI, G.; PICCIALLI, F.; CUOMO, S.; TU, J.; HUP, Z. Julia language in machine learning: algorithms, applications, and open issues. **Computer Science Review**, [S.L.], v. 37, p. 100254-100267, ago. 2020. Elsevier BV. http://dx.doi.org/10.1016/j.cosrev.2020.100254.

GITELSON, A. A. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. **Journal Of Plant Physiology**, [s.l.], v. 161, n. 2, p.165-173, jan. 2004. Elsevier BV. http://dx.doi.org/10.1078/0176-1617-01176.

GITELSON, A.; VINÃ, A.; SOLOVCHENKO, A.; ARKEBAUER, T.; INOUE, Y. Derivation of canopy light absorption coefficient from reflectance spectra. **Remote Sensing Of Environment**, [s.l.], v. 231, p. 111276, set. 2019. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2019.111276.

GITELSON, A.; SOLOVCHENKO, A. Non-invasive Quantification of Foliar Pigments: principles and implementation. In: THENKABAIL, P. S.; LYON, J. G.; HUETE, A. (ed.). **Hyperspectral Indices and Image Classifications for Agriculture and Vegetation**. 2. ed. S.L.: CRC Press, 2018. Cap. 5. p. 1-332.

HIJMANS, R. J. raster: Geographic Data Analysis and Modeling. 2021. **R package version 3.4-13**. https://CRAN.R-project.org/package=raster

HORST, T. Z.; DALMOLIN, R. S. D.; CATEN, A. T.; MOURA-BUENO, J. M.; CANCIAN, L. C.; PEDRON, F. A.; SCHENATO, R. B. Edaphic and topographic factors and their relationship with dendrometric variation of *Pinus taeda* L. in a high altitude subtropical climate. **Revista Brasileira de Ciência do Solo**, [s.l.], v. 42, p.1-16, 6 dez. 2018. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/18069657rbcs20180023.

HUETE, A.; DIDAN, K.; MIURA, T.; RODRIGUEZ, E. P.; GAO, X.; FERREIRA, L. G. Overview of the radiometric and biophysical performance of the MODIS vegetation index. **Remote Sensing Of Environment.** New York, p. 195-213. mar. 2002.

IBÁ (Org.). **Relatório Ibá 2019.** São Paulo, SP: Indústria Brasileira de Árvores, 2019. 80 p. Disponível em: < https://iba.org/datafiles/publicacoes/relatorios/iba-relatorioanual2019.pdf>. Acesso em: 03 fev. 2020.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente:** uma perspectiva em recursos terrestres. 2. ed. São José dos Campos, SP: Parêntese, 2009. 598 p. Tradução: José Carlos Neves Epiphanio (coordenador) [et al.].

JOKELA, E. J.; DOUGHERTY, P. M.; MARTIN, T. A. Production dynamics of intensively managed loblolly pine stands in the southern United States: a synthesis of seven long-term experiments. **Forest Ecology And Management**, [s.l.], v. 192, n. 1, p. 117-130, abr. 2004. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2004.01.007.

KÄFER, P. S.; REX, F. E.; BREUNIG, F. M.; BALBINOT, R. Modeling *Pinus elliottii* growth with multitemporal Landsat data: a study case in Southern Brazil. **Boletim de Ciências Geodésicas**, [s.l.], v. 24, n. 3, p.286-299, ago. 2018. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1982-21702018000300019.

KOTTEK, M.; GRIESER, J.; BECK, C.; RUDOLF, B.; RUBEL, F. World Map of the Köppen-Geiger climate classification updated. **Meteorologische Zeitschrift**, [s.l.], v. 15, n. 3, p.259-263, 10 jul. 2006. Schweizerbart. http://dx.doi.org/10.1127/0941-2948/2006/0130.

KUHN, M. caret: Classification and Regression Training. 2021. **R package version 6.0-88.** https://CRAN.R-project.org/package=caret

LIAW, A.; WIENER, M. Classification and Regression by randomForest. 2002. **R News 2 (3)**, 18--22.

LONG, H. LI, Z; WANG, H.; JIA, J. Biomass resources and their bioenergy potential estimation: A review. **Renewable And Sustainable Energy Reviews**, [s.l.], v. 26, p.344-352, out. 2013. Elsevier BV. http://dx.doi.org/10.1016/j.rser.2013.05.035.

MERTEN, G. H.; MINELLA, J. P. G. The expansion of Brazilian agriculture: Soil erosion scenarios. **International Soil And Water Conservation Research**, [s.l.], v. 1, n. 3, p.37-48, dez. 2013. Elsevier BV. http://dx.doi.org/10.1016/s2095-6339(15)30029-0.

MEYER, D.; DIMITRIADOU, E.; HORNIK, K.; WEINGESSEL, A.; LEISCH, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2021. R package version 1.7-8. https://CRAN.R-project.org/package=e1071

NICODEMUS, F. E. Reflectance Nomenclature and Directional Reflectance and Emissivity. **Applied Optics**, [S.L.], v. 9, n. 6, p. 1474, 1 jun. 1970. The Optical Society. http://dx.doi.org/10.1364/ao.9.001474.

NWANGANGA, F.; CHAPPLE, M. **Practical Machine Learning in R**. Indianapolis: John Wiley & Sons, 2020.

OLIVEIRA, B. R.; SILVA, A. A. P.; TEODORO, L. P. R.; AZEVEDO, G. B.; AZEVEDO, G. T. O. S.; BAIO, F. H. R.; SOBRINHO, R. L.; JUNIOR, C. A. S.; TEODORO, P. E. Eucalyptus growth recognition using machine learning methods and spectral variables. **Forest Ecology And Management**, [s.l.], v. 497, n. 1, p. 1-8, out. 2021. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2021.119496.

OLIVEIRA, R. K.; HIGA, A. R.; SILVA, L. D.; SILVA, I, C.; GONÇALVES, M. P. M. Emergy-based sustainability assessment of a loblolly pine (*Pinus taeda*) production system in southern Brazil. **Ecological Indicators**, [s.l.], v. 93, p.481-489, out. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.ecolind.2018.05.027.

PARK, T. GANGULY, S.; TOMMERVIK, H.; EUSKIRCHEN, E. S.; HOGDA, K.; KARLSEN, S. R.; BROVKIN, V.; NEMANI, R. R.; MYNENI, R. B. Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. **Environmental Research Letters**, [s.l.], v. 11, n. 8, p.1-11, 27 jul. 2016. IOP Publishing. http://dx.doi.org/10.1088/1748-9326/11/8/084001.

PAYN, T.; CARNUS, J.; FREER-SMITH, P.; KIMBERLEY, M.; KOLLERT, W.; LIU, S.; ORAZION, C.; RODRIGUEZ, L.; SILVA, L. N.; WINGFIELD, M. J. Changes in planted forests and future global implications. **Forest Ecology And Management**, [s.l.], v. 352, p.57-67, set. 2015. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2015.06.021.

PERRANDO, E. R.; BREUNIG, F. M.; GALVÃO, L. S.; BOSTELMANN, S. L.; MARTARELLO, V.; STRAUBE, J.; CONTE, B.; SESTARI, G.; BURGIN, M. R. B.; MARCO, R. D. Evaluation of the Effects of Pine Management on the Water Yield and Quality in Southern Brazil. **Journal Of Sustainable Forestry**, [S.L.], v. 40, n. 3, p. 217-233, 6 abr. 2020. Informa UK Limited. http://dx.doi.org/10.1080/10549811.2020.1746916.

PIRARD, R.; SECCO, L. D.; WARMAN, R. Do timber plantations contribute to forest conservation? **Environmental Science & Policy**, [s.l.], v. 57, p.122-130, mar. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.envsci.2015.12.010.

Python. Python programming language. **Python Software Foundation**, Wilmington, Estados Unidos da América. 2021. URL https://www.python.org/.

RADOGLOU, K.; TESKEY, R. O. Changes in rates of photosynthesis and respiration during needle development of loblolly pine. **Tree Physiology**, [s.l.], v. 17, n. 7, p.485-488, 1 jul. 1997. Oxford University Press (OUP). http://dx.doi.org/10.1093/treephys/17.7.485.

R Core Team. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. 2021. URL https://www.R-project.org/.

RODRIGUEZ, D. R. O.; ANDRADE, G. C.; BELLOTE, A. F. J.; TOMAZELLO-FILHO, M. Effect of pulp and paper mill sludge on the development of 17-year-old loblolly pine (*Pinus taeda* L.) trees in Southern Brazil. **Forest Ecology And Management**, [s.l.], v. 422, p.179-189, ago. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2018.04.016.

ROSSATO, F. G. F. S.; SUSAETA, A.; ADAMS, D. C.; HIDALGO, I. G.; ARAUJO, T. D.; QUIROZ, A. Comparison of revealed comparative advantage indexes with application to trade tendencies of cellulose production from planted forests in Brazil, Canada, China, Sweden, Finland and the United States. **Forest Policy And Economics**, [s.l.], v. 97, p.59-66, dez. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.forpol.2018.09.007.

ROUSE, J. W.; HASS, R. W.; SCHELL, J. A.; DEERING, D. W.; HARLAN, J. C. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. **Progress Report RSC 1978-2.** College Station: Texas & A. M. University. 1974. 87p.

ROY, D. P.; WULDER, M. A.; LOVELAND, T. R.; WOODCOCK, C. E.; ALLEN, R. G.; ANDERSON, M. C.; HELDER, D.; IRONS, J. R.; JOHNSON, D. M.; KENEDY, R.; SCAMBOS, T. A.; SCHAAF, C. B.; SCHITT, J. R.; SHENG, Y.; VERMOTE, E. F.; VELWARD, A. S.; BINDSCHADLER, R.; COHEN, W. B.; GAO, F.; HIPPLE, J. D.; HOSTERT, P.; HUNTINGTON, J.; JUSTICEM C. O.; KILIC, A.; KOVALSKYY, V.; LEE, Z. P.; LYMBURNER, L.; MASEK, J. G.; MCCORKEL, L.; SHUAI, Y.; TREZZA, E.; VOGELMANN, J.; WYNNE, R. H.; ZHU, Z. Landsat-8: Science and product vision for terrestrial global change research. **Remote Sensing Of Environment**, [s.l.], v. 145, p.154-172, abr. 2014. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2014.02.001.

SAMUELSON, L. J.; JOHNSEN, K.; STOKES, T. Production, allocation, and stemwood growth efficiency of *Pinus taeda* L. stands in response to 6 years of intensive management. **Forest Ecology And Management**, [s.l.], v. 192, n. 1, p.59-70, abr. 2004. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2004.01.005.

SAMUELSON, L. J.; EBERHARDT, T. L.; BUTNOR, J. R.; STOKES, T. A.; JOHNSEN, K. H. Maximum growth potential in loblolly pine: results from a 47-year-old spacing study in Hawaii. **Canadian Journal Of Forest Research**, [s.l.], v. 40, n. 10, p.1914-1929, out. 2010. Canadian Science Publishing. http://dx.doi.org/10.1139/x10-133.

SANTOS, H. G.; JACOMINE, P. K. T.; ANJOS, L. H. C.; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; FILHO, J. C. A.; OLIVEIRA, J. B.; CUNHA, T. J. F. **Sistema brasileiro de classificação dos solos.** 5. ed. Brasília, DF: Embrapa, 2018. 590 p.

SCHOETTLE, A. W.; FAHEY, T. J.; SHOETTLE, A. W. Foliage and fine root longevity of pines. **Ecological Bulletins**, Copenhagen, DK, n. 43, p.136-153, 1994.

SLOAN, S.; SAYER, J. A. Forest Resources Assessment of 2015 shows positive global trends, but forest loss and degradation persist in poor tropical countries. **Forest Ecology and Management**, [s.l.], v. 352, p.134-145, set. 2015. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2015.06.013.

SCHAEPMAN-STRUB, G.; SCHAEPMAN, M. E.; PAINTER, T. H.; DANGEL, S.; MARTONCHIK, J. V. Reflectance quantities in optical remote sensing—definitions and case studies. **Remote Sensing Of Environment**, [s.l.], v. 103, n. 1, p. 27-42, jul. 2006. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2006.03.002.

THERNEAU, T.; ATKINSON, B. rpart: Recursive Partitioning and Regression Trees. 2019. **R package version 4.1-15**. https://CRAN.R-project.org/package=rpart

TIMOTHY, D.; ONISIMO, M.; CLETAH, S.; ADELABU, S.; TSITSI, B. Remote sensing of aboveground forest biomass: A review. **Tropical Ecology**, [s.l.], v. 2, n. 57, p.125-132, mar. 2016

TREITZ, P. M.; HOWARTH, P. J. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. **Progress In Physical Geography**, [s.l.], v. 23, n. 3, p.359-390, 1 set. 1999. SAGE Publications. http://dx.doi.org/10.1191/030913399673417384.

VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. 4. ed. New York: Springer, 2003. 498 p.

VICENTE-SERRANO, S. M.; CAMARERO, J. J.; OLANO, J. M.; MATÍN-HERNÁNDEZ, N.; PENÃ-GALLARDO, M.; TOMÁS-BURGUERA, M.; GAZOL, A.; AZORIN-MOLINA, C.; BHUYAN, U.; KENAWY, A. E. Diverse relationships between forest growth and the Normalized Difference Vegetation Index at a global scale. **Remote Sensing Of Environment**, [s.l.], v. 187, p.14-29, dez. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2016.10.001.

WAHEED, R.; CHANG, D.; SARWAR, S.; CHEN, W. Forest, agriculture, renewable energy, and CO2 emission. **Journal Of Cleaner Production**, [s.l.], v. 172, p.4231-4238, jan. 2018. Elsevier BV. http://dx.doi.org/10.1016/j.jclepro.2017.10.287.

WREGE, M. S.; STEINMETZ, S.; JÚNIOR, C. R.; ALMEIDA, I. R. (Ed.). **Atlas climático da região sul do Brasil:** estados do Paraná, Santa Catarina e Rio Grande do Sul. 2. ed. Brasília, DF: Embrapa, 2012. 334 p. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-

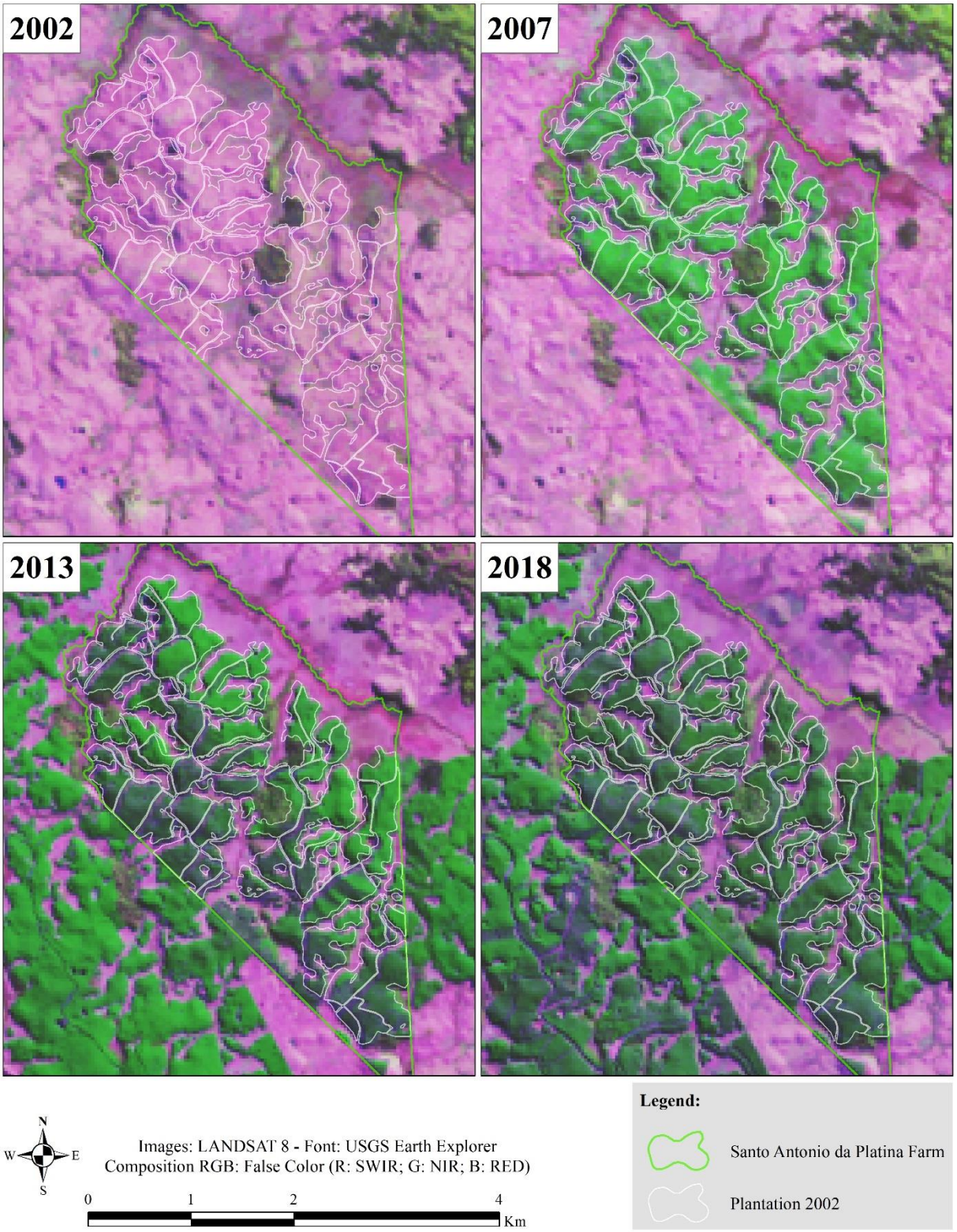
/publicacao/1045852/atlas-climatico-da-regiao-sul-do-brasil-estados-do-parana-santa-catarina-e-rio-grande-do-sul>. Acesso em: 26 mar. 2019.

ZANINOVICH, S. C.; FONTANA, J. L.; GATTI, M. G. Atlantic Forest replacement by non-native tree plantations: Comparing aboveground necromass between native forest and pine plantation ecosystems. **Forest Ecology And Management**, [s.l.], v. 363, p.39-46, mar. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2015.12.022.

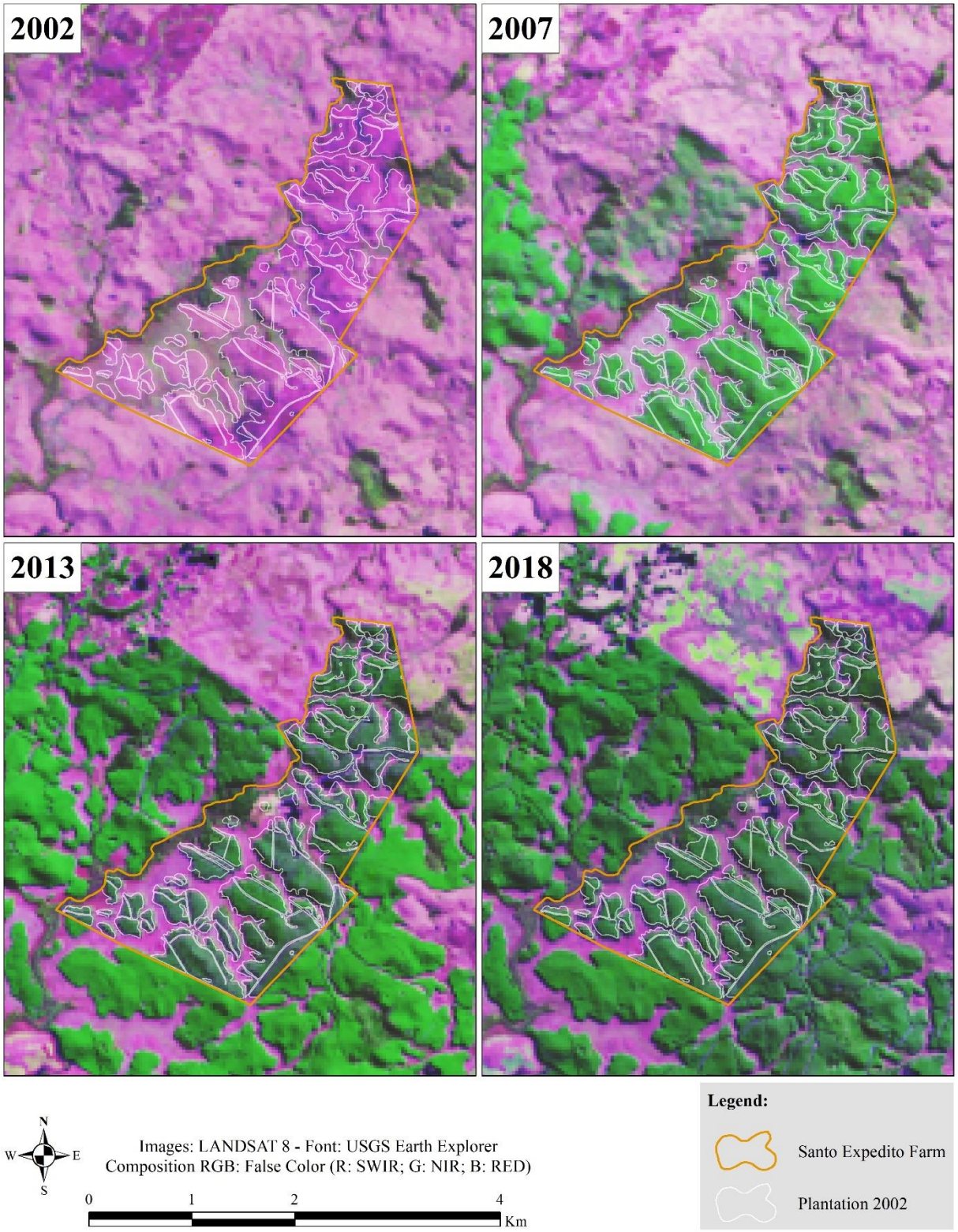
ZHANG, D.; STENGER, A.; HAROU, P. A. Policy instruments for developing planted forests: Theory and practices in China, the U.S., Brazil, and France. **Journal Of Forest Economics**, [s.l.], v. 21, n. 4, p.223-237, dez. 2015. Now Publishers. http://dx.doi.org/10.1016/j.jfe.2015.09.004.

ZHAO, D.; KANE, M.; TESKEY, R.; FOX, T. R.; ALBAUGH, T. J.; ALLEN, H. L.; RUBILAR, R. Maximum response of loblolly pine plantations to silvicultural management in the southern United States. **Forest Ecology And Management**, [s.l.], v. 375, p.105-111, set. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.foreco.2016.05.035.

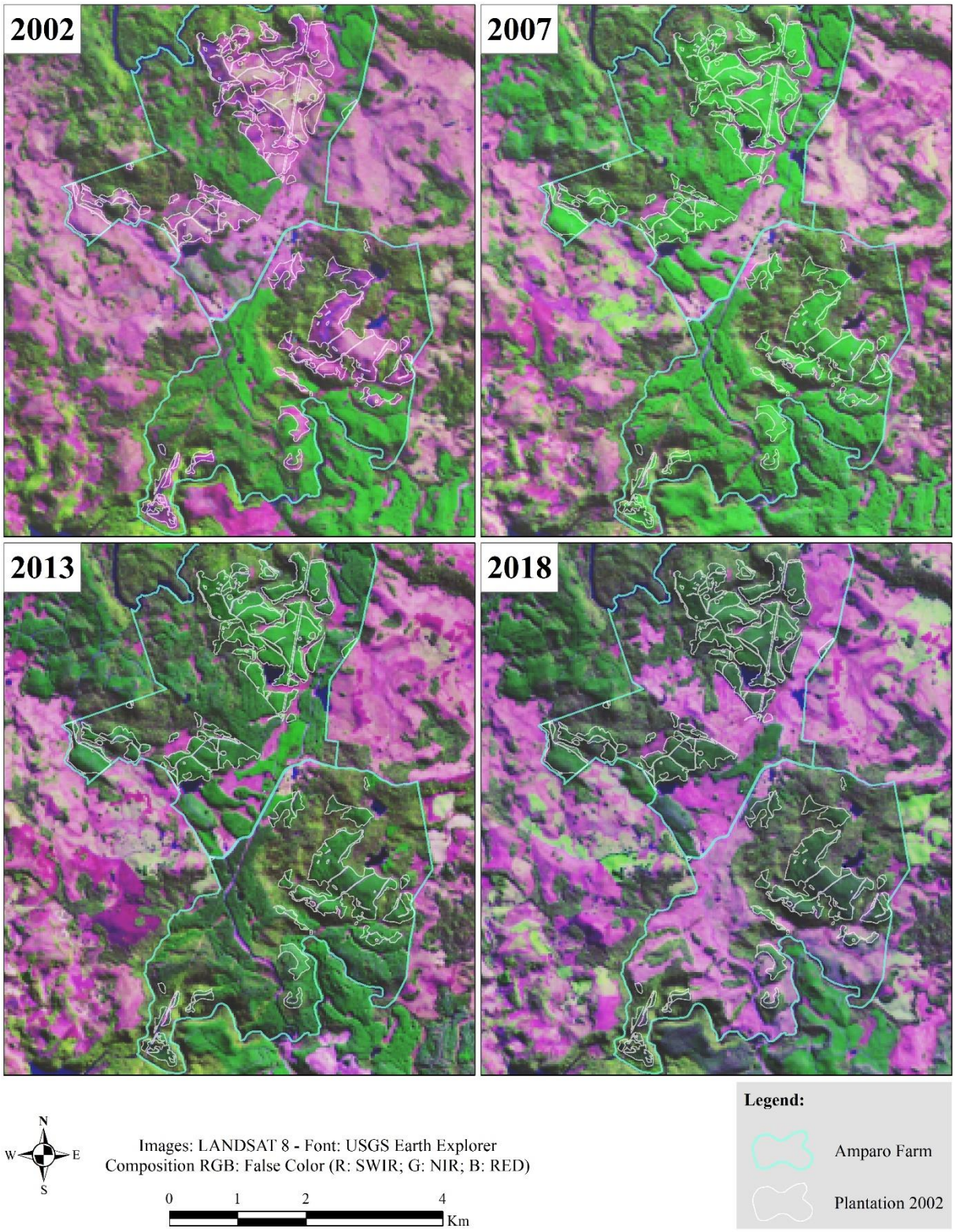
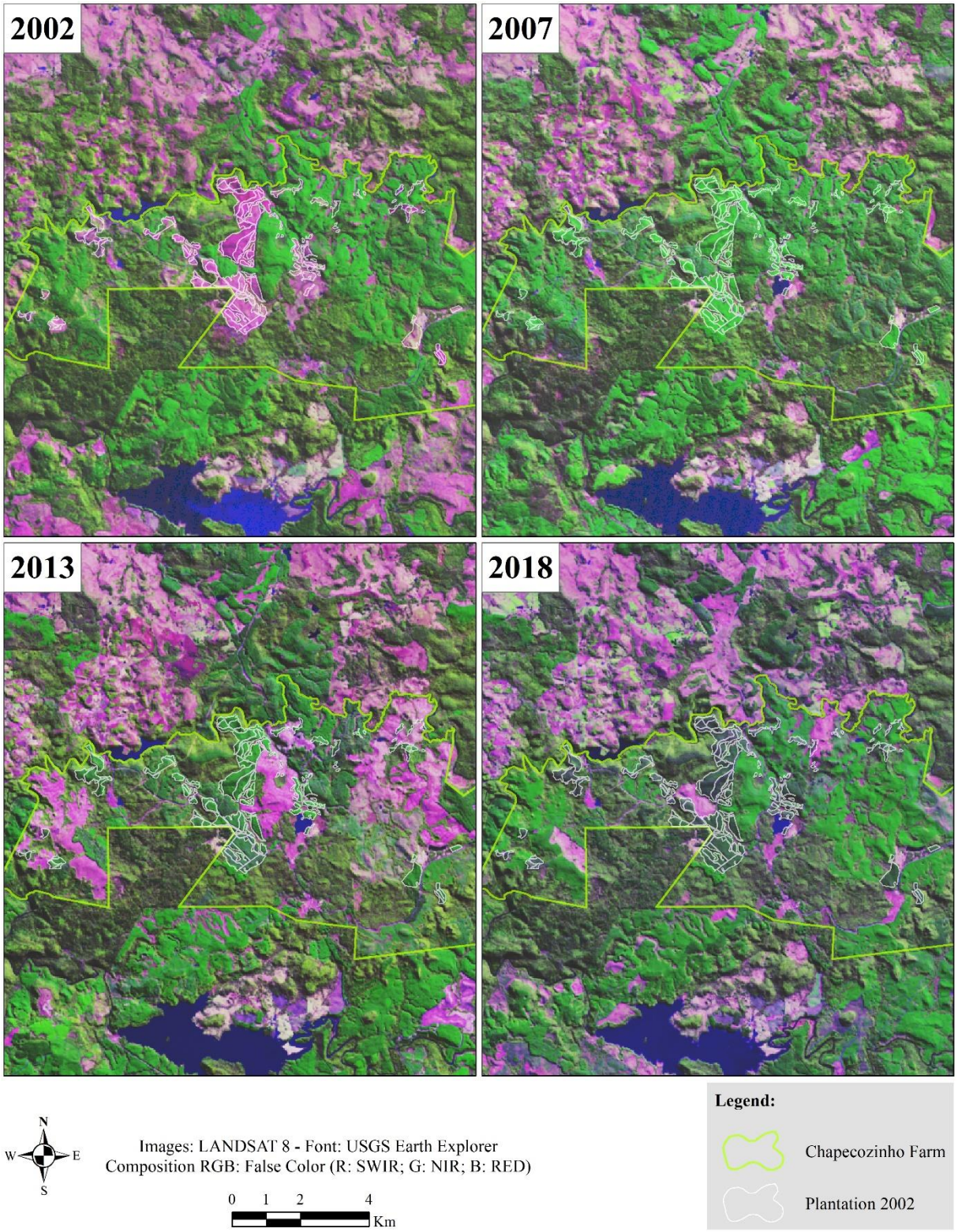
APÊNDICE A – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA SANTO ANTONIO DA PLATINA EM DIFERENTES ANOS, PARANÁ, BRASIL.



APÊNDICE B – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA SANTO EXPEDITO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL.



APÊNDICE C – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA AMPARO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL.



APÊNDICE D – ÁREAS DE EFETIVO PLANTIO IMPLANTADAS NO ANO DE 2002 DA FAZENDA CHAPECOZINHO EM DIFERENTES ANOS, SANTA CATARINA, BRASIL.