

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA URBANA

**PANORAMA AMBIENTAL DAS ÁREAS VERDES
INTRAURBANAS NOS MUNICÍPIOS DA REGIÃO
METROPOLITANA DE RIBEIRÃO PRETO ATRAVÉS DE
INTELIGÊNCIA GEOESPACIAL**

BRENO MALHEIROS DE MELO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Urbana da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutor em Engenharia Urbana.

Orientação: Prof. Dr Fábio Noel Stanganini.

São Carlos

2025

Melo, Breno Malheiros de

Panorama ambiental das áreas verdes intraurbanas nos municípios da região metropolitana de Ribeirão Preto através de inteligência geoespacial / Breno Malheiros de Melo -- 2025.
110f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador (a): Fábio Noel Stanganini

Banca Examinadora: Julio Cesar Pedrassoli, Luciano Aparecido Barbosa

Bibliografia

1. Região Metropolitana de Ribeirão Preto. 2. Sensoriamento Remoto. 3. Random Forest. I. Melo, Breno Malheiros de. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Engenharia Urbana

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Breno Malheiros de Melo, realizada em 19/09/2025.

Comissão Julgadora:

Prof. Dr. Fábio Noel Stanganini (UFSCar)

Prof. Dr. Julio Cesar Pedrassoli (WMU)

Prof. Dr. Luciano Aparecido Barbosa (UNICAMP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Engenharia Urbana.

AGRADECIMENTOS

Agradeço, em primeiro lugar, à minha família, em especial aos meus pais, Selma e Aluizio, pelo amor, apoio e carinho incondicionais ao longo de todo o meu processo de formação.

À Isabela, pelo amor, companhia e apoio nos momentos mais desafiadores desta caminhada.

Ao meu orientador, Prof. Dr. Fabio Noel Stanganini, pela confiança, orientação e apoio ao longo da pesquisa.

Aos colegas, amigos e funcionários do Programa de Pós-Graduação em Engenharia Urbana (PPGEU), pelas trocas de informações, experiências e aprendizados mútuos.

Aos amigos do MapBiomias, em especial Edimilson, Julio e Eduardo, fundamentais em minha trajetória no sensoriamento remoto.

Ao Prof. Marcel Fantin, figura disruptiva, por todos os ensinamentos, pela amizade e pelo apoio constante ao longo dos anos. Aos colegas da Urbverde, agradeço pelas ricas trocas de experiências e pelas oportunidades compartilhadas.

Aos professores Diego (UFSCar), Marcos (UFSCar) e Fernando (USP), pelos ensinamentos e conselhos.

Aos amigos da Maitá ATHIS, pelos valiosos spoilers sobre o universo da pós-graduação e pelo constante empenho em transformar em prática, nas cidades, aquilo que se pesquisa e desenvolve na universidade.

À universidade pública brasileira, por me ensinar a ser resiliente e por representar um espaço de construção coletiva do conhecimento. Em especial, agradeço à Universidade de São Paulo (USP), *alma mater*, à Universidade Federal de São Carlos (UFSCar) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), que proporcionaram as condições necessárias para a realização desta pesquisa.

RESUMO

O presente trabalho de pesquisa tem como principal objetivo a análise da expansão urbana e das transformações ocorridas nas áreas verdes urbanas na Região Metropolitana de Ribeirão Preto (RMRP), fazendo uso dos princípios da inteligência geoespacial e aproveitando as inovações tecnológicas como a computação em nuvem, o sensoriamento remoto e o aprendizado computacional (*Machine Learning*). Esse estudo é motivado pelo rápido crescimento urbano global, que pode acarretar impactos ambientais significativos, como o aumento das emissões de gases de efeito estufa e a degradação do meio ambiente. Isso realça a importância crucial de desenvolver cidades sustentáveis e resilientes, demandando novas abordagens para monitorar e compreender a dinâmica do ambiente urbano. A pesquisa se concentra na identificação precisa das áreas verdes na RMRP, na reconstrução histórica das transformações ocorridas nas áreas verdes urbanas por meio do sensoriamento remoto, na avaliação do impacto dessas áreas no processo de urbanização regional e na exploração das possibilidades de utilizar os dados obtidos para subsidiar um planejamento regional integrado. O estudo enfoca especialmente as áreas urbanas e as áreas ambientalmente vulneráveis, como as unidades de conservação e as zonas de recarga do aquífero.

Palavras-chaves: Região Metropolitana de Ribeirão Preto; Urbanização; Sensoriamento Remoto; Google Earth Engine; Random Forest; Landsat

ABSTRACT

The present research aims to analyze urban expansion and the transformations in urban green areas within the Metropolitan Region of Ribeirão Preto (RMRP). It leverages the principles of geospatial intelligence and incorporates technological innovations such as cloud computing, remote sensing, and machine learning. This study is driven by the rapid global urban growth, which can lead to significant environmental impacts, including increased greenhouse gas emissions and environmental degradation. This underscores the crucial importance of developing sustainable and resilient cities, requiring new approaches to monitor and understand urban environmental dynamics. The research focuses on the accurate identification of green areas within the RMRP, the historical reconstruction of changes in urban green spaces through remote sensing, the assessment of the impact of these areas on the regional urbanization process, and the exploration of how the resulting data can support integrated regional planning. The study places particular emphasis on urban areas and environmentally vulnerable zones, such as conservation units and aquifer recharge areas.

Keywords: Metropolitan Region of Ribeirão Preto; Urbanization; Remote Sensing; Google Earth Engine; Random Forest; Landsat

LISTA DE ILUSTRAÇÕES

Figura 1 - Região Metropolitana de Ribeirão Preto	17
Figura 2 - Inserção regional e Estrutura Metropolitana	18
Figura 3 - Crescimento (2010-2020) da área urbanizada em Ribeirão Preto	19
Figura 4 - Exposição das áreas de recarga à urbanização	20
Figura 5 - Uso e cobertura do solo em 1985 e 2017	23
Figura 6 - Exemplo de fluxo de trabalho envolvendo diversas ferramentas - Fluxograma de método de identificação de áreas incendiadas	24
Figura 7 - Linha do tempo do programa Landsat, com Landsat 1 em 1972 e Landsat Next previsto para 2030	26
Figura 8 - Bandas espectrais dos sensores Landsat	27
Figura 9 - Fluxograma de geração de amostras usando conjuntos de dados auxiliares e índices espectrais	29
Figura 10 - Decomposição da importância de uma variável ao longo de k interações dela com as demais	31
Figura 11 - Fluxograma de geração de amostras usando conjuntos de dados auxiliares e índices espectrais.	32
Figura 12 - Fluxo metodológico da pesquisa	33
Figura 13 - Fluxo de Produção	36
Figura 14 - Exemplo de amostras das classes utilizadas	37
Figura 15 - Exemplo experimental de recuperação de dados	41
Figura 16 - Exemplo experimental de recuperação de imagem com nuvem	41
Figura 17 - Exemplo experimental de recuperação de imagem com nuvem através de criação de mosaico sem nuvens	42
Figura 18 - Aplicativo desenvolvido para validação visual de amostras	43
Figura 19 - Mapa de Cobertura da Terra para Ribeirão Preto em 2024	45

Figura 20 - Transições de classe	46
Figura 20 - Série histórica de dados	47
Figura 21 - Situação dos setores censitários da RMRP	49
Figura 22 - Série temporal – Alta densidade urbana	50
Figura 23 - Série temporal – Baixa densidade urbana	51
Figura 24 - Série temporal – Núcleos urbanos	52
Figura 25 - Série temporal – Unidades de Conservação (UCs)	53
Figura 26 - Série temporal – Áreas de Vulnerabilidade do Aquífero	54

LISTA DE TABELAS

Tabela 1 - Principais índices espectrais	28
Tabela 2 - Bandas e índices espectrais utilizados	35
Tabela 3 - Acurácia do Produtor	47
Tabela 4 - Acurácia do Consumidor	48

LISTA DE SIGLAS

RMRP - Região Metropolitana de Ribeirão Preto

TST - Temperatura da Superfície Terrestre

GEE - Google Earth Engine

AWS - Amazon Web Service

GIS - Geographic Information System (Sistema de Informação Geográfica)

SIG - Sistemas de Informação Geográfica

PDI - Processamento Digital de Imagens

PDUIs - Planos de Desenvolvimento Urbano Integrado

ODIs - Objetivos de Desenvolvimento Sustentável

NDVI - Normalized Difference Vegetation Index (Índice de Vegetação por Diferença Normalizada)

NDWI - Normalized Difference Water Index (Índice de Água por Diferença Normalizada)

MNDWI - Modified Normalized Difference Water Index

UI - Urban Index

NDBI - Normalized Difference Build Up Index

SAERP - Secretaria de Água e Esgoto de Ribeirão Preto

UHI - Urban Heat Island (Ilha de Calor Urbana)

MODIS - Moderate Resolution Imaging Spectroradiometer

ASTER - Advanced Spaceborne Thermal Emission and Reflection Radiometer

AVHRR - Advanced Very High Resolution Radiometer

SEVIRI - Spinning Enhanced Visible and Infrared Imager

GEOINT - Geospatial Intelligence

HUMINT - Human Intelligence

MASINT - Measurement and Signatures Intelligence

COMINT - Communication Intelligence

ELINT - Electronic Intelligence

SIGINT - Signal Intelligence

IMINT - Imagery Intelligence

SUMÁRIO

1. INTRODUÇÃO	15
1.1 A Região Metropolitana de Ribeirão Preto.....	16
1.2 Questões da Pesquisa	20
2. OBJETIVOS	21
2.1. Objetivo geral.....	21
2.2. Objetivos específicos	21
3. FUNDAMENTAÇÃO TEÓRICA.....	21
3.1. Inteligência Geoespacial	21
3.2. Áreas Verdes Urbanas.....	25
3.3. Sensoriamento Remoto	26
3.4. Algoritmos de Aprendizagem de Máquina	30
3.5. Amostragem.....	31
4. MATERIAIS E MÉTODOS DE PESQUISA	33
4.1. Mosaico.....	34
4.2. Produção de Amostras	36
4.3. Classificador	37
4.4. Classificador	38
4.5. Filtros	39
4.6. Validação	41
5. RESULTADOS	44
5.1. Contexto das Áreas urbanas.....	44
5.2. Áreas Urbanas – IBGE.....	48
5.3. Unidade de Conservação.....	51
5.4. Áreas de vulnerabilidade do Aquífero	53

6. DISCUSSÃO	55
6.1. O Aumento das áreas verdes	55
6.2. O impacto do PDUI.....	57
7. CONSIDERAÇÕES FINAIS.....	60
8. REFERÊNCIAS	62
APÊNDICE – FLUXO DE CLASSIFICAÇÃO.....	71
ALGORÍTIMO A1 – Amostragem Automática.....	72
ALGORÍTIMO A2 – Filtro De Amostras	78
ALGORÍTIMO A3 – Classificação.....	90
ALGORÍTIMO A4 – Filtros De Classificação.....	98
ALGORÍTIMO A5 – Aplicativo de Validação Visual.....	104

1. INTRODUÇÃO

A urbanização, impulsionada inicialmente pela Revolução Industrial no século XIX, intensificou-se ao longo do século XX e, nas últimas décadas, culminou na histórica inversão demográfica na qual a população urbana ultrapassou a rural. Atualmente, pouco mais da metade da população mundial vive em áreas urbanas, e projeta-se que esse percentual alcance 68 % até 2050 (UN-Habitat, 2022). Esse rápido adensamento, muitas vezes desordenado, tem ampliado pressões ambientais, resultando no aumento das emissões de gases de efeito estufa, na escassez de recursos naturais e na degradação dos ecossistemas (Norman, 2022). Nesse contexto, a construção de cidades resilientes e sustentáveis torna-se imperativa.

As cidades resilientes não apenas estão preparadas para enfrentar os desafios impostos pelas mudanças climáticas, mas também são projetadas para se adaptar e se recuperar de eventos adversos, como desastres naturais e crises econômicas. Já as cidades sustentáveis adotam práticas que visam reduzir emissões de gases de efeito estufa, conservar recursos naturais e promover o uso eficiente de energia e água. A combinação de resiliência e sustentabilidade nas políticas urbanas é essencial para garantir um futuro habitável para as gerações futuras, mitigando os impactos ambientais da urbanização acelerada e proporcionando uma maior qualidade de vida. Cidades que integram espaços verdes e incentivam o contato diário com a natureza (cidades biofílicas) também fortalecem a coesão social e a saúde mental, criando ambientes urbanos que são mais robustos, saudáveis e conectados com o ecossistema local (Beatley & Newman, 2013).

Para que esses princípios se concretizem, é necessário que o planejamento urbano vá além de soluções pontuais e incorpore diretrizes que assegurem o uso equitativo e racional do solo urbano. Nesse contexto, o Estatuto da Cidade (LEI FEDERAL 10.251, Brasil, 2001) surge no início do século XXI como um marco jurídico que busca orientar o crescimento e o desenvolvimento das cidades brasileiras. Por meio de seus instrumentos, essa legislação promove uma ocupação mais ordenada do território, ao mesmo tempo em que visa garantir a função social da cidade e da propriedade, reafirmando o papel do espaço urbano como um bem coletivo que deve atender às necessidades de todos os cidadãos (Instituto Pólis, 2005).

Complementando esse arcabouço legal, o Estatuto da Metrópole (Lei Federal 13.089/2015) e os Planos de Desenvolvimento Urbano Integrado (PDUIs) oferecem mecanismos fundamentais para o planejamento e a governança nas Regiões Metropolitanas do país. Esses instrumentos reforçam a importância de uma abordagem integrada entre os

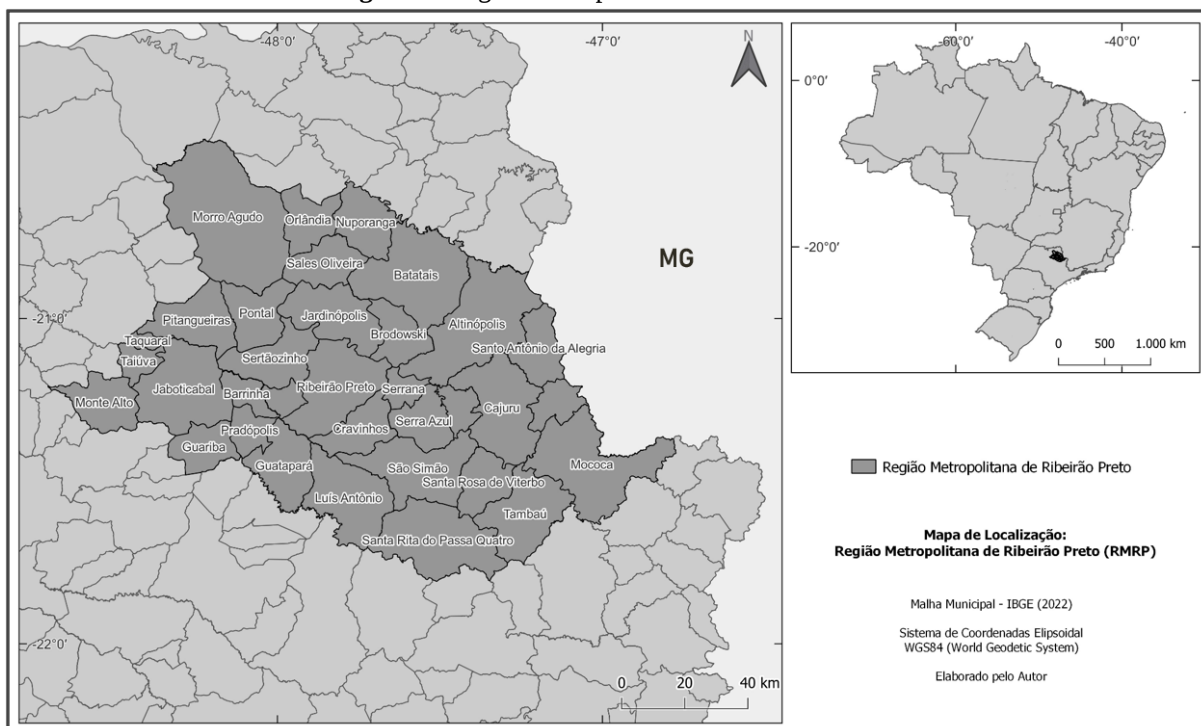
municípios, superando a fragmentação de políticas urbanas e promovendo uma gestão mais cooperativa. Os PDUIs, em particular, permitem a elaboração de estratégias territoriais que conciliam crescimento urbano, justiça social e sustentabilidade ambiental, contribuindo diretamente para a construção de cidades mais resilientes e com maior capacidade de adaptação às transformações do século XXI (Costa, 2018).

Observando essa série de instrumentos, o monitoramento da dinâmica urbana se destaca como questão estratégica central no que diz respeito às políticas públicas que buscam a construção de cidades resilientes e sustentáveis em consonância com os Objetivos de Desenvolvimento Sustentável (ODS). Assim, pesquisas que envolvem o mapeamento e a interpretação detalhada da expansão urbana e seu impacto no meio natural, com a aquisição de dados recentes e atualizados, contribuem de forma significativa para a elaboração de políticas públicas urbanas voltadas ao meio-ambiente e mitigação dos impactos gerados pelas ações antrópicas.

Em meados da década de 80, Foresti (1986) já indicava a necessidade de se buscar novos métodos, empregando tecnologias mais adequadas para detectar e investigar, em tempo quase real, a expansão urbana e as alterações ambientais, contribuindo para o aumento da eficiência da gestão pública e das instâncias regulatórias. No decorrer da evolução dos instrumentos jurídicos de controle e fiscalização das cidades, essa necessidade vem se tornando cada vez mais clara. Dentro desse contexto, para compreender o impacto dos instrumentos supracitados torna-se necessário a escolha de um local de estudo que permita uma ampla avaliação das políticas orientadas ao controle urbano e seu impacto ambiental nas áreas verdes urbanas. Assim, a presente pesquisa indica a Região Metropolitana de Ribeirão Preto (RMRP) como local de estudo.

1.1 A Região Metropolitana de Ribeirão Preto

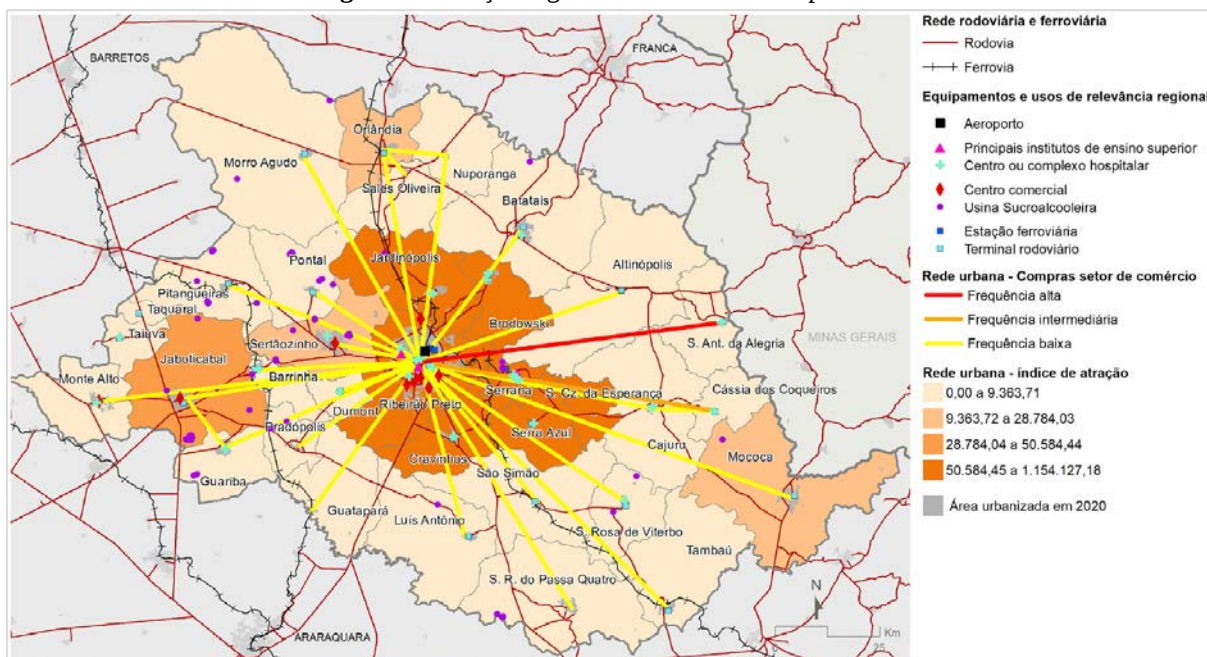
Instituída pela Lei Complementar Estadual nº. 1.290 de 06 de julho de 2016, a Região Metropolitana de Ribeirão Preto (RMRP) é uma área de notável dinamismo econômico e diversidade geográfica. Composta por 34 municípios (Figura 1), a RMRP se destaca como um polo prestador de serviços de grande importância, abrangendo uma ampla gama de setores econômicos. Entre os setores econômicos preponderantes, destacam-se o comércio, saúde, além da sólida base industrial, com ênfase nas atividades sucroalcooleiras e metalúrgicas, que desempenham um papel fundamental na economia local, gerando empregos e impulsionando o desenvolvimento tecnológico (São Paulo, 2022).

Figura 1: Região Metropolitana de Ribeirão Preto.

Fonte: Adaptado de São Paulo (2022).

Uma das características mais marcantes da RMRP é sua heterogeneidade, os municípios que compõem esta região são notáveis pela diversidade de tamanho e pelo padrão de ordenamento urbano variado que apresentam. Em uma mesma região metropolitana, encontram-se cidades de portes extremamente distintos, desde pequenas vilas a centros urbanos altamente desenvolvidos, assim como municípios em processo de conurbação (Hayashi, 2020). Essa heterogeneidade oferece a oportunidade de analisar como diferentes modelos de desenvolvimento urbano coexistem e interagem.

Conforme o Diagnóstico da Região Metropolitana de Ribeirão Preto, apresentado na elaboração do Plano de Desenvolvimento Urbano Integrado (PDUI), Ribeirão Preto constitui o principal núcleo urbano da RMRP, destacando-se como polo econômico e prestador de serviços de alcance regional e intermunicipal. A cidade polariza a dinâmica regional (Figura 2), exercendo forte influência econômica e concentrando fluxos de pessoas e mercadorias.

Figura 2: Inserção regional e Estrutura Metropolitana.

Fonte: Fipe (2021).

Essa centralidade é reforçada pela infraestrutura de transporte, incluindo o aeroporto, que conecta a região a outros centros importantes, como São Paulo, Belo Horizonte, Rio de Janeiro e Brasília (Fipe, 2001). Além disso, a RMRP é marcada pela presença de municípios como Sertãozinho, que complementam a dinâmica econômica ao atuar como polos industriais voltados para a agroindústria regional (Fipe, 2001).

Outro ponto relevante na caracterização da RMRP é a interdependência entre os meios urbano e rural, refletida tanto na economia quanto na configuração espacial. Alguns municípios são majoritariamente urbanos, enquanto outros mantêm uma forte conexão com suas atividades agrícolas, que continuam a desempenhar um papel significativo na economia regional (Hayashi, 2020). Conforme observado por Elias (2003, 2013), essa relação simbiótica é o resultado de um longo processo de modernização agrícola e expansão do capitalismo no campo. A especialização agrícola, particularmente no cultivo da cana-de-açúcar, consolidou essa interdependência, criando uma dinâmica territorial complexa.

Essa diversidade também pode ser entendida à luz da história da urbanização no interior paulista. A expansão da economia cafeeira e a modernização agrícola moldaram a rede urbana, resultando em uma configuração espacial desigual, com municípios de diferentes tamanhos e funções. A RMRP reflete essa diversidade, abrigando tanto cidades com forte presença industrial quanto municípios com economia predominantemente agrícola. Também observa-se o enfrentamento a desafios comuns a áreas metropolitanas em rápida expansão, como a gestão

de questões socioespaciais, a necessidade de um planejamento urbano mais integrado e a busca por um desenvolvimento sustentável (Elias, 2013; Firkowski, 2013). A criação da região em 2016 foi um marco importante na busca por soluções conjuntas, incentivando a cooperação intermunicipal e a formulação de políticas públicas regionais (Feldman, 2014).

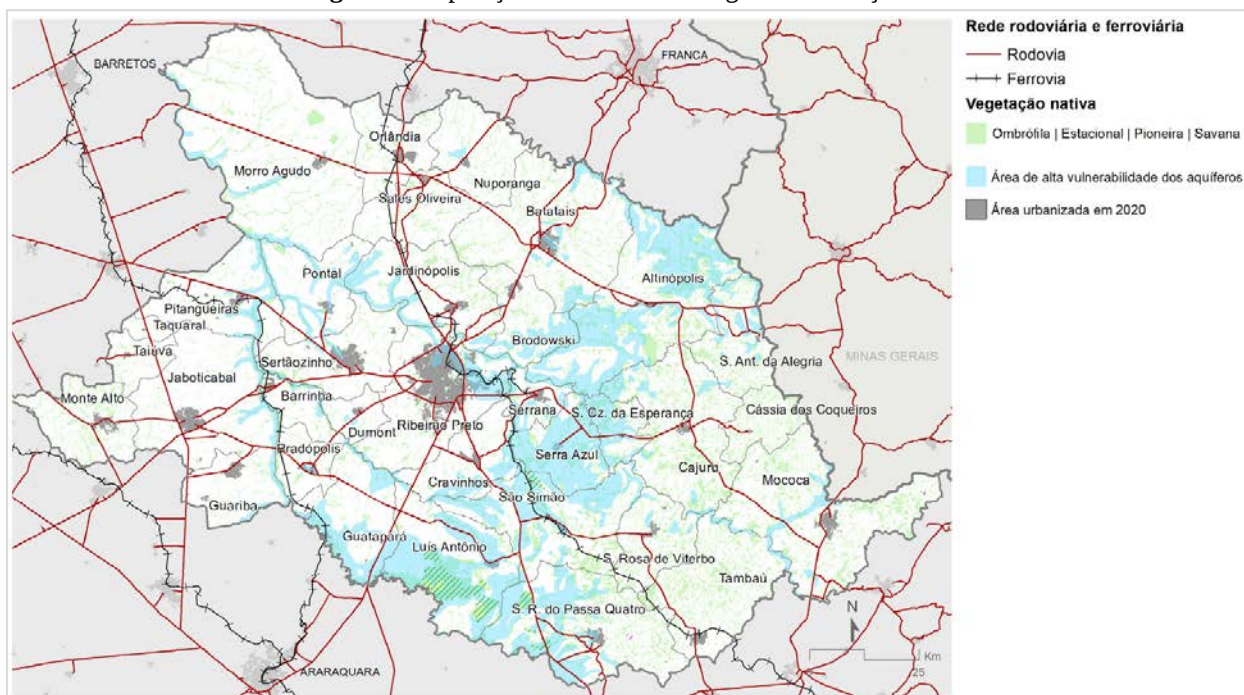
A interação entre os ambientes urbano e rural amplia os desafios da região, ao mesmo tempo em que oferece oportunidades únicas para uma gestão ambiental integrada. Essa dinâmica provoca transformações no uso do solo e na paisagem, que precisam ser monitoradas e geridas de forma eficiente. O distanciamento entre as manchas urbanas propicia o processo de expansão da área urbana em suas zonas periféricas, processo predominado pela formação de loteamentos. Para o caso do município de Ribeirão Preto (Figura 3), observa-se a concentração de condomínios fechados ao sul e formação de empreendimentos de moradia popular na região noroeste (Fipe, 2021).

Figura 3: Crescimento (2010-2020) da área urbanizada em Ribeirão Preto.



Fonte: Adaptado de Fipe (2021).

A expansão urbana desafia a resiliência dos sistemas de abastecimento da região. Ribeirão Preto, por exemplo, dependem intensamente da captação de água por meio de poços artesianos. De acordo com dados do SAERP (2020), o abastecimento da cidade é realizado integralmente por 118 poços profundos, que retiram água diretamente do Aquífero Guarani. Essa situação evidencia a vulnerabilidade das áreas de recarga do aquífero, bem como a resiliência do sistema frente à contínua expansão urbana sobre esses territórios, conforme ilustrado na figura 4.

Figura 4: Exposição das áreas de recarga à urbanização.**Fonte:** Fipec (2021).

Assim, dado o contexto de crescimento das cidades e seu impacto nas áreas naturais, o uso de ferramentas de inteligência geoespacial torna-se essencial para a gestão sustentável da expansão urbana e da preservação dos recursos naturais na região. Por meio de tecnologias avançadas de sensoriamento remoto e análise espacial, é possível monitorar as mudanças no uso e ocupação da terra, identificando áreas de alta vulnerabilidade às diversas condicionantes existentes. Essas ferramentas permitem a criação de cenários prospectivos que auxiliam na tomada de decisões, contribuindo para a elaboração de políticas públicas direcionadas à proteção ambiental e ao desenvolvimento sustentável da região.

1.2 Questões da Pesquisa

Diante do contexto apresentado, a presente pesquisa pretende responder questões como:

- I. Como identificar áreas verdes em Regiões Metropolitanas, tendo como exemplo o caso de RMRP?
- II. Como reconstruir a história das mudanças das áreas verdes urbanas usando sensoriamento remoto?
- III. Como os dados produzidos podem auxiliar no planejamento regional integrado?

2. OBJETIVOS

2.1. Objetivo geral

O presente projeto tem por objetivo desenvolver uma metodologia automatizada e replicável para analisar a expansão urbana e as alterações nas áreas vegetadas no contexto espacial da Região Metropolitana de Ribeirão Preto (RMRP) a partir do suporte de ferramentas baseadas em sensoriamento remoto, algoritmos de aprendizado de máquina baseados no conceito de inteligência artificial, e computação em nuvem, entre os anos de 1985 e 2024 de forma a prover uma interpretação pormenorizada desse período no que diz respeito ao crescimento das cidades e de suas áreas verdes.

2.2. Objetivos específicos

- I. Mapear a evolução anual das áreas urbanas a partir de 1985 com a utilização de inteligência artificial e técnicas de sensoriamento remoto;
- II. Quantificar alterações de cobertura em unidades de conservação vizinhas às áreas urbanas e em áreas de vulnerabilidade do aquífero;

3. FUNDAMENTAÇÃO TEÓRICA

3.1. Inteligência Geoespacial

A inteligência geoespacial, também conhecida como inteligência geográfica ou geointeligência (GEOINT), desempenha um papel fundamental em diversas áreas de tomada de decisão, oferecendo uma abordagem valiosa para entender e solucionar problemas com componentes geográficos. Ela se baseia na exploração e análise de informações geoespaciais, incluindo imagens e dados referenciados no espaço e no tempo.

De acordo com a *Australian Geospatial-Intelligence Organisation* (AGO, 2023) e com o Manual de Fundamentos – Inteligência Militar (Ministério da Defesa, 2015), a inteligência geoespacial, ou GEOINT, envolve a análise de imagens coletadas de satélites ou aeronaves para identificar recursos, descrever atividades e interpretar eventos em uma determinada localização. Além disso, a GEOINT inclui a coleta e análise de informações sobre recursos terrestres, utilizando sistemas de informações geográficas para classificar, examinar e tirar conclusões de dados que, de outra forma seriam difíceis de relacionar com fenômenos ou acontecimentos históricos de interesse. Em essência, a inteligência geoespacial combina informações precisas

de localização com atributos de recursos naturais e artificiais, fornecendo uma base sólida para uma ampla gama de análises e decisões.

Segundo Bacastow e Bellafiore (2009, *apud* Dias, 2019), inteligência geoespacial é uma disciplina que integra diversas fontes de informação geográfica, tornando-se uma peça-chave no contexto da inteligência tradicional. Ela não apenas complementa, mas também enriquece outras formas de inteligência, como a HUMINT (*Human Intelligence*), a MASINT (*Measurement and Signatures Intelligence*), a COMINT (Comunicação Intelligence), a ELINT (*Electronic Intelligence*), a SIGINT (*Signal Intelligence*) e a IMINT (*Imagery Intelligence*)¹. Além disso, a GEOINT tem a capacidade única de apresentar informações em um contexto espaço-temporal, considerando não apenas o "onde" e o "o que", mas também o "quando". Com a evolução das geotecnologias, softwares de geoprocessamento e sistemas de informações geográficas, a inteligência geoespacial tornou-se ainda mais poderosa, permitindo a análise de dados georreferenciados em escalas variadas, a criação de produtos visuais personalizados e interativos, e a identificação de mudanças ao longo do tempo, tornando-a uma ferramenta essencial para a tomada de decisões em diversas áreas, desde a defesa até a gestão ambiental.

Nesse contexto, a existência de informações precisas sobre o histórico de uso e cobertura do solo são fundamentais para o planejamento e gestão sustentável urbana e rural, assim como para a formulação de políticas públicas. A exemplo, é possível citar iniciativas de mapeamento como o Inventário Florestal Nacional (2000 a 2018) e os mapas de uso e cobertura do Instituto Brasileiro de Geografia e Estatística (IBGE) (2000 a 2018).

Essas iniciativas, em geral, valem-se do uso de pré-processamento e melhoria de imagens de satélite combinados com fotointerpretação visual, sendo esse conjunto de ferramentas proibitivo para o desenvolvimento de grandes análises temporais devido ao custo e tempo de execução (Souza *et al.* 2020).

Com a popularização da computação em nuvem na última década, novas alternativas à fotointerpretação visual se tornaram viáveis. Com isso, algoritmos baseados em conceitos de aprendizagem de máquina e inteligência artificial passaram a ser amplamente utilizados em

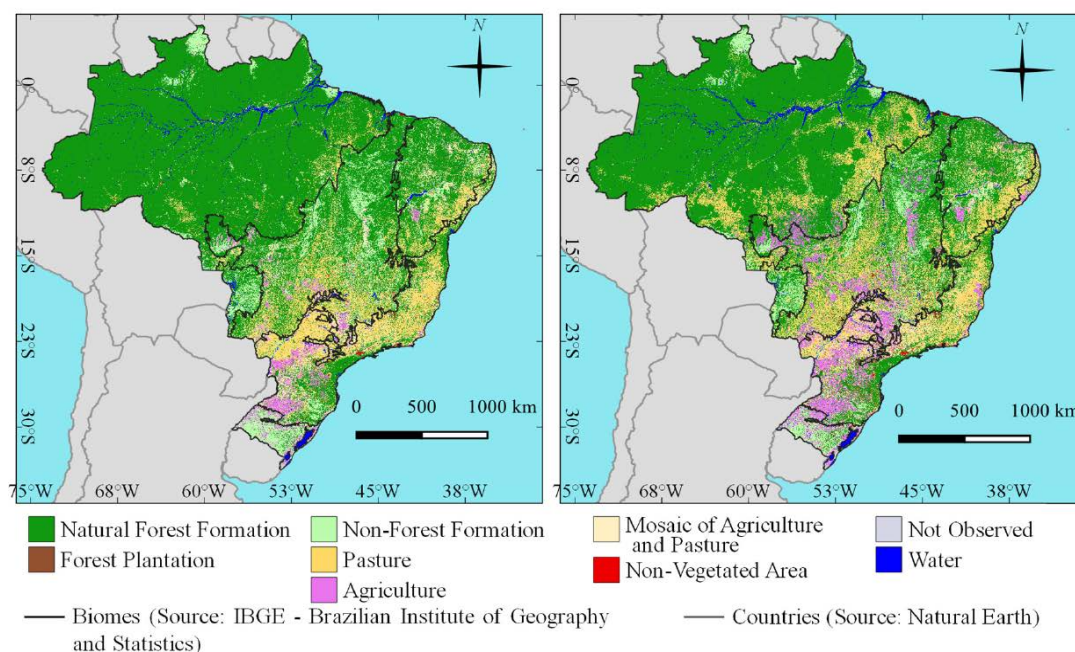
¹ HUMINT envolve a coleta de informações por fontes humanas, MASINT abrange diversas fontes, incluindo nuclear, óptica, radiofrequência, acústica, sísmica e ciências dos materiais; COMINT foca na coleta de informações de comunicações individuais, como telefonemas e mensagens online; ELINT se concentra em fontes eletrônicas não comunicativas, como telemetria de satélites e sensores de radares; SIGINT envolve a interceptação de sinais de comunicação; IMINT se dedica à coleta de informações por meio de imagens de satélites ou aeronaves. (Dias, 2019)

larga escala, inclusive para produção cartográfica de mapas de uso e cobertura. Criado em 2010, o *Google Earth Engine* (GEE) é um dos principais exemplos de plataforma utilizada para fins de geoprocessamento, seguido de outras iniciativas para suporte de *big data* como a *Microsoft Azure* (2010) e *Amazon Web Service* (AWS, lançada em 2006). (Zhao *et. al.* 2021)

De uma perspectiva de evolução do ferramental, temos o GEE como uma nova alternativa de produção de informações geográficas no campo da inteligência geoespacial. Segundo Velastegui-Montoya *et al.* (2023) as metodologias de pesquisa estão em constante mudança com o intuito de desbravar a fronteira do conhecimento, dentro desse contexto temos o GEE como uma poderosa plataforma de processamento, de uso gratuito para pesquisa e educação, que possibilita a manipulação de dados espaciais e execução de algoritmos de alta complexidade em escala global, viabilizando sua aplicação em distintos campos do sensoriamento remoto.

Como exemplo de uso dessa tecnologia, temos a iniciativa MapBiomás, que produz anualmente dados de uso e cobertura do solo (Figura 5) para todo o país, por meio do uso de classificação supervisionada (*Random Forest*) em imagens de média resolução da série Landsat (Souza *et al.* 2020).

Figura 5: Uso e cobertura do solo em 1985 e 2017

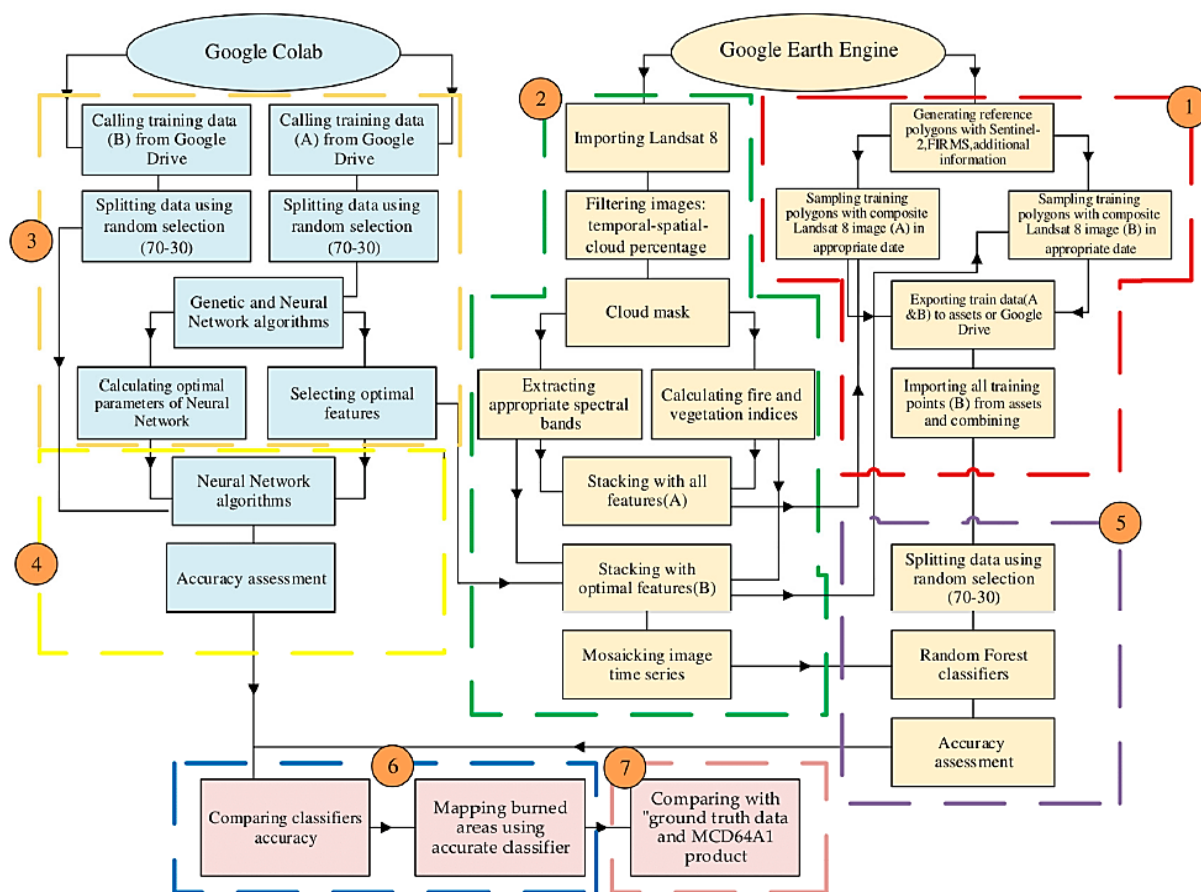


Fonte: Souza *et. al.* (2020).

Essa ferramenta possui inúmeras capacidades dentro do contexto técnico-científico, que desafiam os limites da pesquisa científica ao permitir a coleta de dados e a automatização de

análises em larga escala. Essa característica abre as portas para a integração do GEE em fluxos de trabalho complexos (Figura 6).

Figura 6: Exemplo de fluxo de trabalho envolvendo diversas ferramentas - Fluxograma de método de identificação de áreas incendiadas



Fonte: Gholamrezaie *et al.* (2022).

Fluxos que envolvem uma variedade de elementos, incluindo Sistemas de Informação Geográfica (SIG) bem estabelecidos, como o QGIS e o ARCGIS, além do uso de linguagens de programação, como Python, para atividades de *Data Science* e Processamento Digital de Imagens (PDI) envolvendo modelos de predição e aprendizagem computacional, como é o caso de técnicas de *Machine Learning* e *Deep Learning* (Amani *et. al.* 2020; Gholamrezaie *et. al.* 2022; Wang *et. al.* 2020).

Portanto, percebe-se que a inteligência geoespacial representa uma faceta essencial na tomada de decisões em várias áreas, fornecendo uma abordagem valiosa para entender e resolver problemas. O conceito de combinar informações espacialmente localizadas, associada às novas tecnologias e plataformas de processamento em nuvem, torna possível o desenvolvimento de estratégias para análises complexas, possibilitando a coleta e tratamento

de dados georreferenciados em larga escala por meio de técnicas de sensoriamento remoto. Essa capacidade vem mostrando seu potencial de revolucionar o meio técnico científico, e se posiciona como uma poderosa estratégia no âmbito de auxiliar as tomadas de decisão na gestão ambiental das áreas urbanas, em especial ao acompanhamento da dinâmica das áreas vegetadas.⁴

3.2. Áreas Verdes Urbanas

Conforme dispõe o Art. 8º, § 1º, da Resolução CONAMA Nº 369/2006, considera-se área verde de domínio público "o espaço de domínio público que desempenhe função ecológica, paisagística e recreativa, propiciando a melhoria da qualidade estética, funcional e ambiental da cidade, sendo dotado de vegetação e espaços livres de impermeabilização" (Brasil, 2006).

Como apontado por Bargos e Matias (2019), não há consenso na definição do termo Áreas verdes, esse por vezes acaba sendo utilizando como sinônimos de espaços livres e áreas de lazer. Essas distinções, acabam por muitas vezes impactando na comparação de índices de áreas verdes (IAV) aplicados por diferentes autores. Nesse sentido Cavalheiro e Del Picchia (1992) e Nucci (2008) apresentam as áreas verdes como uma subcategoria de espaços livres, destacando especialmente função social desses espaços.

Além de suas funções ecológicas, as áreas verdes urbanas desempenham papéis cruciais na mitigação de ilhas de calor, na absorção de águas pluviais, na redução da poluição do ar e na promoção da biodiversidade urbana (Loboda; Angelis, 2017). A presença dessas áreas é vital para a promoção de atividades recreativas e culturais, bem como para o bem-estar psicológico dos habitantes urbanos. A proximidade e o acesso a áreas verdes urbanas estão diretamente relacionados à redução dos níveis de estresse e ao aumento da qualidade de vida dos moradores (Beatley; Newman, 2013; Nucci, 2008; Zhong *et al.* 2020).

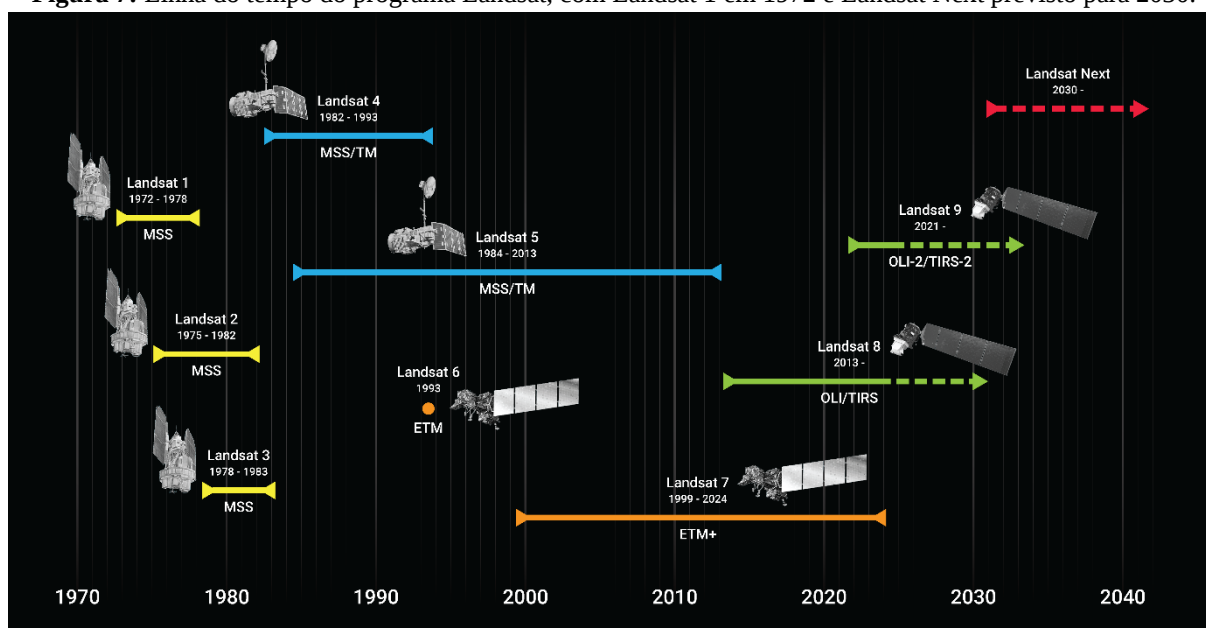
Portanto, a gestão e a expansão das áreas verdes urbanas devem ser uma prioridade nas políticas públicas de planejamento urbano, visando a sustentabilidade e a qualidade de vida nas cidades. A criação de novos parques e a preservação de áreas existentes são estratégias essenciais para garantir que os benefícios ecológicos, sociais e econômicos dessas áreas sejam plenamente aproveitados pela população.

3.3. Sensoriamento Remoto

O Sensoriamento Remoto é a principal tecnologia para a produção de mapeamentos históricos da cobertura em áreas extensas e diversas. Essa técnica permite a obtenção de informações comparáveis a partir de imagens de satélite disponíveis desde a década de 1980. Recentes avanços no tratamento de imagens de satélite, como novas técnicas de filtragem, classificação e processamento em nuvem, aumentaram significativamente a capacidade de mapear, quantificar e qualificar mudanças no uso e na cobertura do solo em escala global (Hansen *et al.*, 2013; Song *et al.*, 2018).

O programa Landsat (Figura 7) é o mais bem-sucedido dos programas espaciais voltados para a observação da Terra, utilizando diversos satélites e sensores, teve seu início em 1972.

Figura 7: Linha do tempo do programa Landsat, com Landsat 1 em 1972 e Landsat Next previsto para 2030.



Fonte: Nasa (2024).

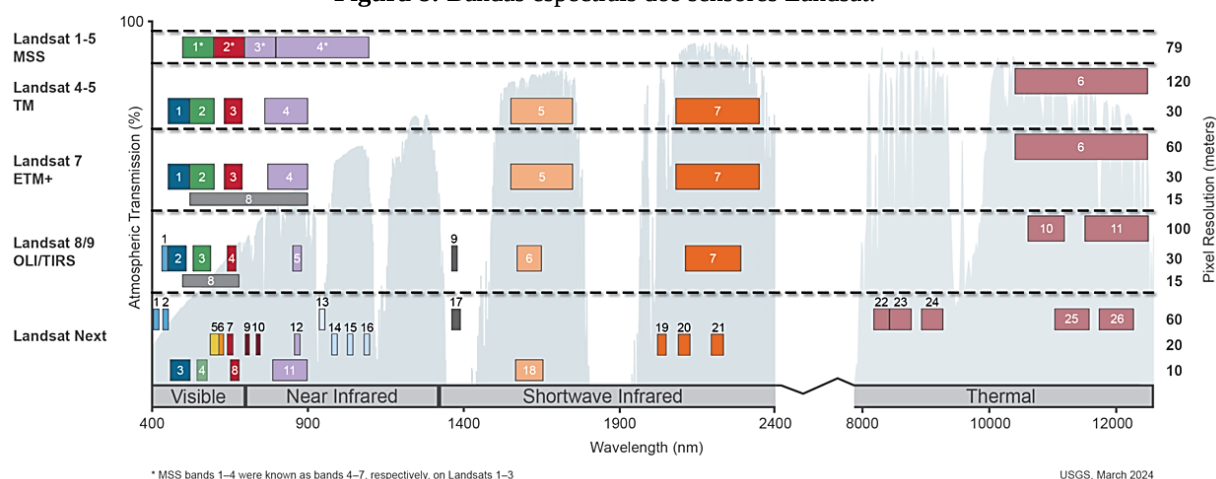
A disponibilização gratuita do acervo a partir de 2005 abriu novas possibilidades para o mapeamento de recursos naturais, o monitoramento do uso e cobertura do solo, e a análise da dinâmica da vegetação em diferentes escalas, desde o nível municipal até o global (Shimabukuro; Ponzoni, 2017).

Paralelamente à evolução dos algoritmos, o acesso a satélites de média resolução tornou-se mais fácil. Como descrito por Claverie *et al.* (2018), as possibilidades de combinar sensores, especialmente os do Landsat 8 e Sentinel-2, resultaram na criação de uma constelação virtual de satélites de refletância de superfície, ampliando os horizontes das análises

relacionadas ao uso e cobertura do solo. De forma similar, houve também um aumento na disponibilidade de sensores de alta resolução, permitindo um aprimoramento dos mapeamentos em escala mais detalhada.

As informações adquiridas por meio desses sensores orbitais podem ser organizadas em três domínios de informação, que compreendem as dimensões do espectro, espaço e tempo (Barbosa; Novo; Martins, 2019; Diniz, 2021; Shaw; Burke, 2003). Dentro do domínio espectral, temos a relação de absorção, reflexão ou emissão de energia dos objetos, onde a maioria das aplicações se dá através da quantidade de energia magnética refletida. Nesse contexto, os sensores orbitais são habilitados a distintas faixas de comprimento de onda (Figura 8), as quais auxiliam na detecção de diferentes alvos.

Figura 8: Bandas espectrais dos sensores Landsat.



Fonte: USGS (2024).

As distintas faixas de comprimento de onda permitem aos sensores capturar informações que variam conforme as características físicas e químicas dos objetos observados. Cada tipo de superfície ou material, como a vegetação, a água, o solo e estruturas urbanas, possui uma "assinatura espectral" específica, ou seja, um padrão de resposta em diferentes comprimentos de onda. Essas assinaturas espectrais tornam-se uma ferramenta essencial para identificar, mapear e monitorar os alvos, com aplicações em diversas áreas, como agricultura, gestão hídrica e monitoramento ambiental (Barbosa; Novo; Martins, 2019). Como exemplo, a vegetação saudável reflete mais energia no infravermelho próximo (NIR) em relação a visível, enquanto corpos d'água apresentam uma baixa reflectância no NIR e maior absorção no visível. Já as áreas urbanas têm uma alta reflectância no espectro visível e infravermelho médio, devido à predominância de materiais como concreto e asfalto.

A abordagem descrita possibilita a utilização de índices espectrais que, por meio da combinação de diferentes bandas, destacam alvos específicos. Esse processo constitui uma etapa inicial para a separação de diferentes alvos em uma imagem, utilizando uma metodologia rápida, não supervisionada e escalável (Javed *et al.*, 2021). Nesse contexto, há uma variedade de índices projetados para diferentes alvos e sensores; a Tabela 1 apresenta alguns dos principais.

Tabela 1: Principais índices espectrais

Nome	Equação	Referência
Normalized Difference Vegetation Index	$NDVI = (NIR-RED) / (NIR+RED)$	Huete (1988)
Enhanced Vegetation Index	$EVI = 2,5(NIR-RED) / (NIR+6RED-7,5BLUE+1)$	Huete <i>et al.</i> (1997)
Normalized Difference Water Index	$NDWI = (NIR-SWIR) / (NIR+ SWIR)$	Gao (1996)
Modified Normalized Difference Water Index	$MNDWI = (GREEN-SWIR) / (GREEN + SWIR)$	Xu (2007)
Urban Index	$UI = ((SWIR2-NIR) / (SWIR2+NIR)+1) * 100$	Kawamura; Jayamanna; Tsujiko (1997)
Normalized Difference Build-Up Index	$NDBI = (SWIR2-NIR) / (SWIR2+NIR)$	Zha <i>et al.</i> (2003)

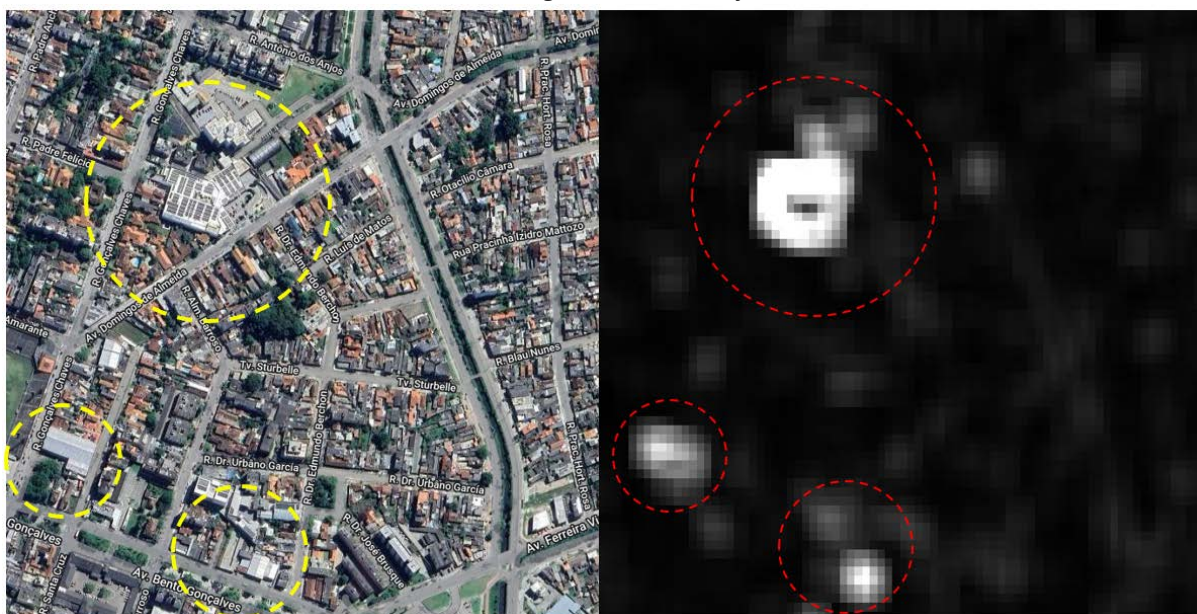
Fonte: Adaptado pelo autor com base nas referências citadas

No domínio espacial, a análise se concentra na forma, localização e distribuição dos objetos observados, utilizando informações que podem variar em diferentes resoluções espaciais, dependendo da capacidade do sensor. O nível de detalhamento da imagem, determinado pela resolução espacial, define o quão precisas serão as informações sobre o tamanho, contorno e distribuição dos elementos na superfície terrestre. Resoluções espaciais mais altas capturam detalhes finos, permitindo identificar características individuais, enquanto resoluções mais baixas fornecem uma visão mais ampla e contextualizada das áreas observadas (Barbosa; Novo; Martins, 2019; Shaw; Burke, 2003).

A textura é um elemento fundamental para a interpretação de informações visuais e um recurso valioso na classificação de imagens. A textura é descrita em termos de variação tonal e espacial dos pixels, compreendendo propriedades como granulação, regularidade e continuidade. Essas características permitem a identificação de superfícies de acordo com sua

organização estrutural e sua relação com o entorno (Figura 9). Além disso, textura e tom são conceitos inter-relacionados. O tom refere-se às diferentes intensidades de cinza em uma imagem, enquanto a textura agrega informações sobre a disposição espacial dessas variações de tom. Textura e tom não são independentes, mas coexistem em uma relação de complementaridade, em que o tom representa a intensidade básica e a textura fornece informações adicionais sobre a distribuição espacial dessas intensidades (Haralick; Shanmugam; Dinstein, 1973; He; Wang, 1990).

Figura 9: Contraste produzido por *Gray Level Co-occurrence Matrix* (GLCM) para identificação de pixels que se distinguem da vizinhança.



Fonte: Elaborado pelo autor.

O domínio temporal envolve a análise de mudanças e padrões ao longo do tempo, aproveitando a capacidade de sensores orbitais de capturar imagens repetidas de uma mesma área em diferentes momentos. Esse domínio permite monitorar dinâmicas ambientais, sazonais e antrópicas, oferecendo uma perspectiva essencial para entender processos contínuos e suas variabilidades temporais. A análise temporal apoia-se na identificação de eventos que ocorrem de forma cíclica ou pontual, facilitando a compreensão de fenômenos naturais e de atividades humanas. A variação temporal é frequentemente empregada no monitoramento de fenômenos sazonais, como o ciclo anual de crescimento da vegetação, onde índices de vegetação, como o *Normalized Difference Vegetation Index* (NDVI), são utilizados para avaliar mudanças na biomassa e no vigor das plantas ao longo das estações. Além disso, o domínio temporal é crucial na detecção de alterações ambientais, como o desmatamento, a degradação de áreas agrícolas e

a expansão urbana, pois permite registrar a sequência e a magnitude das modificações (Barbosa; Novo; Martins, 2019; Diniz, 2021).

3.4. Algoritmos de Aprendizagem de Máquina

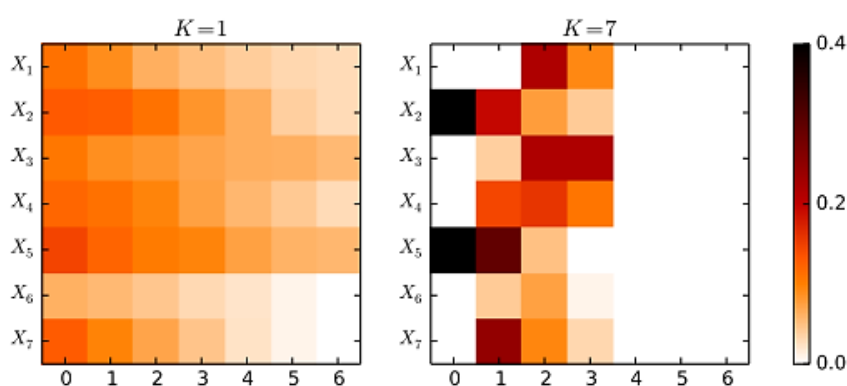
Nas últimas décadas algoritmos de aprendizagem de máquina ganharam foco em detrimento as classificações paramétricas tradicionais, em especial pela qualidade dos dados produzidos e redução de custos de implementação desses métodos (Diniz *et al.*, 2021, Gorelick *et al.*, 2017, Justus *et al.*, 2018, Maxwell *et al.*, 2018). Isso abriu novas possibilidades para o monitoramento ambiental, incluindo o mapeamento do uso e cobertura do solo em áreas urbanas, onde a velocidade e a capacidade de processamento são cruciais para análises em tempo real e em larga escala. "A implementação de algoritmos de *machine learning* intensificou-se na última década, passando a substituir os classificadores paramétricos tradicionais devido ao seu melhor desempenho em alvos com alta variabilidade de informações espectrais. No entanto, o recorrente uso de classificadores paramétricos continua a ser representativa em volume dada a sua ampla integração a softwares de processamento de imagens. Cenário que passa a se alterar com a popularização de ferramentas associados ao uso de inteligência artificial (Amani *et al.*, 2020, Maxwell *et al.*, 2018).

Semelhante ao início dos algoritmos de *machine learning*, as abordagens de *deep learning* têm mostrado resultados impressionantes e vêm ganhando popularidade devido à criação recente de pacotes que facilitam sua implementação. Com uma arquitetura inspirada no funcionamento dos neurônios, os algoritmos de redes neurais destacam-se especialmente na classificação de objetos que, além de apresentarem grande variabilidade espectral, necessitam também de informações contextuais, como geometria e rugosidade, para serem identificados. No entanto, apresentam como principal ponto negativo a infraestrutura de treinamento dos modelos, a qual tende a evoluir com o aumento da área mapeada e variabilidade das classes (Cresson, 2019; Zhu *et al.*, 2017).

No campo dos classificadores supervisionados não contextuais (*machine learning*), *Random Forest* e *Support Vector Machine* (SVM) destacam-se sobre os demais. No conjunto de artigos analisados, observou-se uma recorrência de estudos que visam comparar a eficácia de ambos os classificadores. Como indicado por Maxwell *et al.* (2018), as acurácias de ambos os classificadores são próximas, apresentando valores superiores aos demais classificadores comumente utilizados no contexto de *machine learning* para a classificação de alvos vegetais e minerais.

O algoritmo de *Random Forest* é reconhecido como uma técnica robusta em aprendizado de máquina, especialmente valiosa para tarefas de classificação e regressão devido à sua capacidade de construir modelos precisos com relativa simplicidade (Louppe, 2014). Entre os fatores que tornam o *Random Forest* atrativo, destaca-se o uso de árvores de decisão não-paramétricas, o que permite modelar relações complexas entre variáveis sem a necessidade de pressupor uma distribuição específica dos dados (Figura 10). Dessa forma, o *Random Forest* consegue captar interações intrincadas, apresentando-se como uma solução adaptável para diferentes tipos de dados e contextos.

Figura 10: Decomposição da importância de uma variável ao longo de k interações dela com as demais.



Fonte: Louppe (2014).

A técnica se destaca também pela capacidade de lidar com dados heterogêneos, incluindo variáveis categóricas e contínuas, o que amplia consideravelmente suas aplicações e reduz a necessidade de transformações complexas ou de pré-processamento específico dos dados. Durante o treinamento das árvores, o *Random Forest* realiza uma seleção automática de atributos relevantes por meio do cálculo da impureza de Gini, ganho de informação ou erro quadrático médio (MSE), diminuindo a influência de variáveis irrelevantes e o ruído nos dados, o que contribui para a obtenção de previsões mais precisas e para a interpretação do modelo. Outro aspecto vantajoso é a simplicidade de interpretação dos resultados. A técnica utiliza uma abordagem baseada em agregação, em que a decisão final para cada observação é determinada pela maioria dos votos entre as árvores, o que facilita a interpretação por profissionais de diversas áreas e permite a análise dos resultados com maior confiabilidade, especialmente quando comparado a modelos mais complexos.

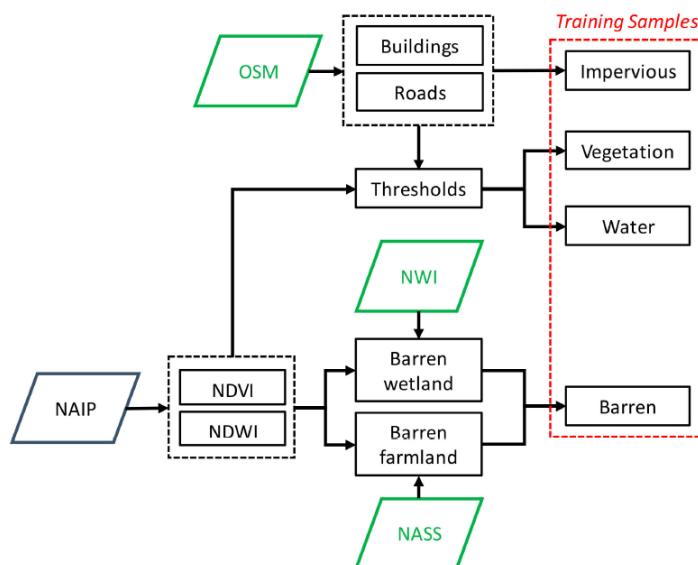
3.5. Amostragem

A amostragem é uma etapa crucial no treinamento de classificadores supervisionados, especialmente no contexto do sensoriamento remoto, onde a qualidade do conjunto de dados

influencia diretamente a eficácia do modelo. Em tarefas de classificação, a diversidade e representatividade dos dados rotulados são essenciais para que o classificador possa aprender com precisão as fronteiras entre diferentes classes de objetos ou áreas. No entanto, a rotulagem de grandes volumes de dados é um processo dispendioso e demorado, envolvendo inspeção visual ou levantamentos de campo. Em imagens de alta resolução espacial ou espectral, comuns no sensoriamento remoto, a variabilidade dentro das classes (como diferentes tipos de vegetação ou materiais de construção) aumenta a complexidade da classificação. Isso torna o processo de construção de um conjunto de treinamento ainda mais exigente, tanto em termos de esforço humano quanto de recursos financeiros (Tuia *et al.*, 2011).

Para contornar essa problemática, diversas soluções têm sido exploradas, como apresentado por Huang *et al.* (2015) e Zhang; Gorelick; Zimba (2020). Uma abordagem inicial consiste na utilização de dados de referência provenientes de bases governamentais ou colaborativas, como o Open Street Maps (Figura 11).

Figura 11: Fluxograma de geração de amostras usando conjuntos de dados auxiliares e índices espectrais.



Fonte: Zhang; Gorelick; Zimba (2020)

Essa estratégia demonstra uma contribuição significativa em ambientes urbanos, pois oferece suporte para objetos não cadastrados nessas bases. Além disso, conforme discutido por Santos Junior (2023), a seleção de amostras pode ser realizada por meio de limiares de índices espectrais ou por meio da filtragem de dados que desviam da tendência padrão. Outra possibilidade é a utilização de aprendizagem ativa apresentada por Tuia *et al.*, (2011). A aprendizagem ativa permite que o modelo identifique automaticamente os exemplos com maior incerteza, ou seja, aqueles que o classificador tem mais dificuldade em categorizar.

corretamente. Assim, o processo de rotulagem concentra-se nos dados mais relevantes para o modelo, reduzindo o volume de amostras necessárias e os custos associados à obtenção de dados rotulados.

4. MATERIAIS E MÉTODOS DE PESQUISA

Os procedimentos metodológicos adotados neste projeto estarão vinculados aos objetivos específicos anteriormente descritos. Conforme classificação apresentada por Silveira e Córdova (2009), o presente projeto apresenta natureza *aplicada* com uma abordagem predominantemente *quantitativa*, uma vez que se propõe a descrever uma realidade a partir de dados numéricos utilizando ferramentas estatísticas implementadas na área do sensoriamento remoto. Para isso o projeto se valerá das fases apresentadas na Figura 12, e melhores detalhadas em sequência.

Figura 12: Fluxo metodológico da pesquisa.



Fonte: Elaborado pelo Autor.

- I. Exploratória: se relaciona ao adensamento do repertório, onde será consultada e revisada a bibliografia referentes a técnicas de sensoriamento remoto, *Machine Learning* e computação em nuvem. Também serão analisados relatórios técnicos, artigos teses e dissertações produzidas a respeito da RMRP, além de legislações

urbanística e ambientais que permitam prover informações de interesse da dissertação a respeito do território de estudo;

- II. Experimental: se relaciona ao desdobramento dos objetivos, nessa fase seguirá a elaboração e validação de rotinas de processamento de dados, com o intuito de identificar as morfologias de interesse, e técnicas computacionais que podem ser aplicadas para esse processo em ferramentas de uso gratuito (GEE, QGIS, Google Colab). Assim como, também será realizada a organização de repositório de dados (rotinas e arquivos) em ambiente em plataforma de uso gratuito, à exemplo GitHub, para armazenar e publicar as informações e rotinas de processamento dos dados;
- III. Analítica: se relaciona a análise dos resultados, etapa de realização das análises quantitativas e qualitativas pertinentes a interpretação das informações produzidas. Nesse momento será avaliada a acurácia do dado assim como observado fenômenos que possam estar associados a mudanças ambientais ou políticas.

4.1. Mosaico

A seleção do sensor Landsat para este estudo se deve, principalmente, à sua longa série temporal de observações, que se estende desde a década de 1980. Embora sensores mais recentes, como o Sentinel-2, ofereçam melhor resolução espacial e maior frequência de revisita, sua cobertura temporal ainda é limitada a partir de 2015. Considerando que o objetivo central desta pesquisa envolve a análise de transformações territoriais ao longo de décadas, a amplitude histórica oferecida pelos sensores Landsat possui maior alinhamento ao tipo de abordagem temporal adotada.

Nesse sentido, todo processo de classificação se desenhou com base em mosaicos anuais contendo bandas espectrais nativas do satélite Landsat e índices calculáveis a partir delas (Tabela 2).

Tabela 2: Bandas e índices espectrais utilizados.

Nome em Português	Sigla	Tipo	Descrição	Fórmula
Azul	blue	Banda espectral	Região do azul do espectro visível; útil para análise de corpos d'água e solo exposto.	—
Verde	green	Banda espectral	Região do verde; sensível à vegetação saudável e usada em índices aquáticos.	—
Vermelho	red	Banda espectral	Região do vermelho; importante para contraste com vegetação.	—
Infravermelho próximo	nir	Banda espectral	Alta refletância em vegetação vigorosa; essencial para análise de biomassa.	—
Infravermelho de onda curta 1	swir1	Banda espectral	Útil para detectar umidade do solo, vegetação e áreas urbanas.	—
Infravermelho de onda curta 2	swir2	Banda espectral	Indicado para análise de queimadas e composição do solo.	—
Índice de Vegetação (NDVI)	ndvi	Índice espectral	Mede o vigor da vegetação.	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$
Índice de Água (MNDWI)	mndwi	Índice espectral	Realça corpos d'água, reduz interferência de áreas urbanas.	$(\text{GREEN} - \text{SWIR1}) / (\text{GREEN} + \text{SWIR1})$
Índice de Construção (NDBI)	ndbi	Índice espectral	Destaca áreas urbanizadas com base na refletância no SWIR.	$(\text{SWIR1} - \text{NIR}) / (\text{SWIR1} + \text{NIR})$
Índice de Solo Nu (BSI)	bsi	Índice espectral	Evidencia superfícies com solo exposto.	$((\text{SWIR1} + \text{RED}) - (\text{NIR} + \text{BLUE})) / ((\text{SWIR1} + \text{RED}) + (\text{NIR} + \text{BLUE}))$
Fração de vegetação (SMA)	veg	Componente SMA	Representa a proporção de vegetação por decomposição espectral.	Derivado por decomposição espectral linear (SMA)
Fração de substrato (SMA)	sub	Componente SMA	Representa a proporção de solo claro ou superfícies expostas.	Derivado por decomposição espectral linear (SMA)
Fração de sombra (SMA)	dark	Componente SMA	Representa sombras, água profunda ou materiais escuros.	Derivado por decomposição espectral linear (SMA)

Fonte: Elaborado pelo Autor.

Como etapa preparatória à classificação, os mosaicos Landsat utilizados neste estudo foram submetidos a um processo de normalização espectral baseado nos percentis 5 (p5) e 95 (p95) das bandas originais. Essa estratégia foi adotada com o objetivo de padronizar os valores espectrais das bandas *blue*, *green*, *red*, *nir*, *swir1* e *swir2*, harmonizando os dados provenientes de diferentes sensores (Landsat 5, 7, 8 e 9) e anos (1985–2024). A normalização visa reduzir

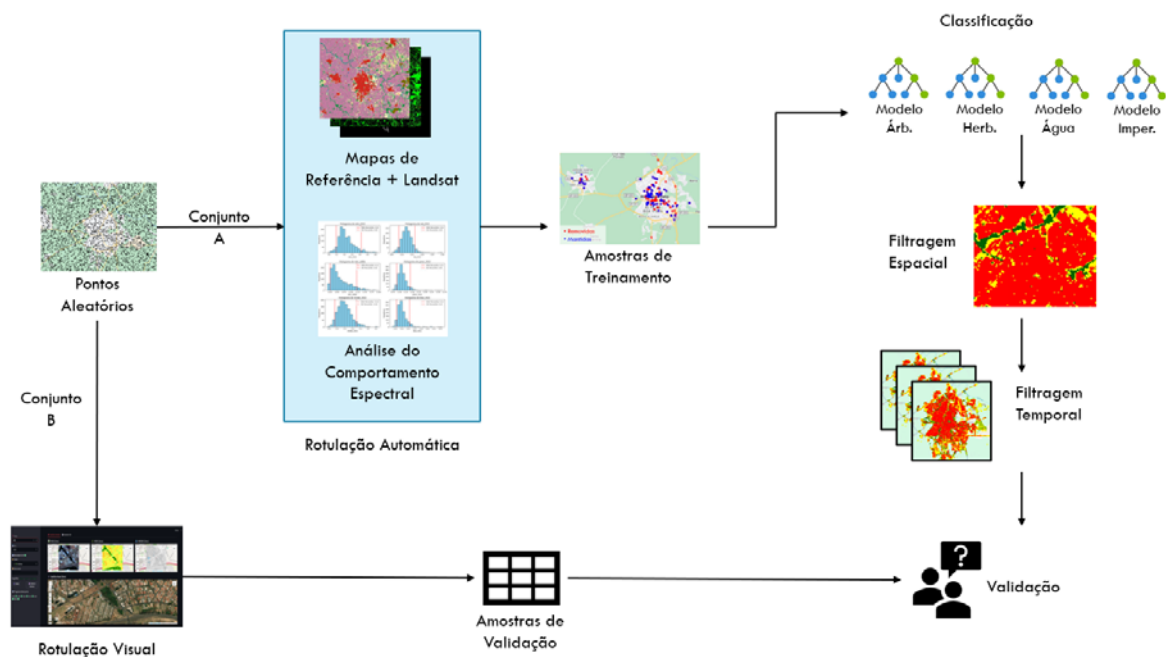
variações radiométricas não associadas a mudanças reais no uso e cobertura da terra, como aquelas causadas por diferenças instrumentais, condições atmosféricas ou sazonais, garantindo assim uma base de dados temporalmente coerente e comparável.

A escolha por essa abordagem se fundamenta na necessidade de melhorar a consistência dos dados ao longo do tempo e entre sensores, favorecendo o desempenho de classificadores como o Random Forest. Ao aplicar a normalização com base em percentis, os valores extremos de reflectância são comprimidos, minimizando o impacto de outliers, reflexos especulares e alvos altamente refletivos como áreas urbanas ou solo exposto. Essa técnica também contribui para destacar a faixa útil da reflectância, otimizando a separação entre diferentes classes temáticas (Brown *et al.* 2022).

4.2. Produção de Amostras

A aplicação do algoritmo de classificação supervisionada *Random Forest* para a categorização de imagens de satélite requer amostras que orientem o modelo sobre os padrões de resposta espectral das classes de interesse. Para tal, amostras de treinamento são geradas automaticamente (Figura 13) com base em mapas de referência, como o Mapa de Uso e Cobertura da Coleção 9 do MapBiomass e *Dynamic World*.

Figura 13: Fluxo de Produção



Fonte: Elaborado pelo Autor.

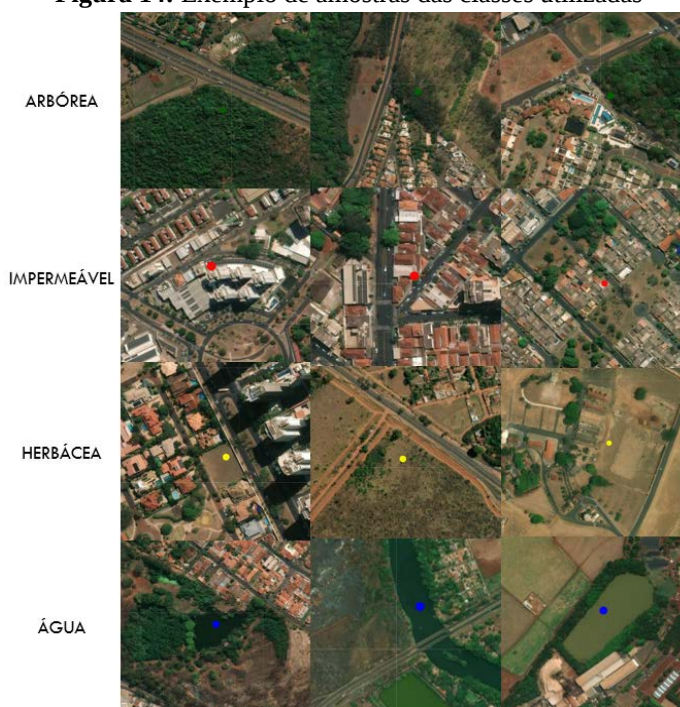
Nesse contexto, as classes de interesse são primeiramente selecionadas nos mapas de referência mencionados. Em seguida, um conjunto pontos aleatórios é gerado, sendo o conjunto A composto 30.000 pontos, incorporando tanto a classe do mapa de referência quanto as informações de resposta espectral do mosaico Landsat a ser classificado, enquanto o conjunto B fica reservado as amostras de validação.

Após a geração desse conjunto de pontos, realiza-se um processo de filtragem. Para cada classe de interesse, uma série de índices espectrais, como o Índice de Vegetação por Diferença Normalizada (NDVI) e o Índice de Água por Diferença Normalizada (NDWI), bem como as bandas espectrais originais dos sensores, são analisados. As informações resultantes são então integradas ao *dataset* dos pontos aleatórios. Com base nesse conjunto enriquecido de dados espectrais, procede-se à identificação e remoção de *outliers* para cada classe de interesse por meio da remoção do percentil 5 e 95 de cada índice. Essa etapa tem por intuito remover do conjunto amostral possíveis dados classificados erroneamente.

4.3. Classificador

Para o processo de classificação, será utilizado o algoritmo Random Forest, com o objetivo de categorizar as amostras em quatro classes principais: vegetação arbórea, vegetação herbácea, áreas impermeabilizadas e corpos hídricos (Figura 14).

Figura 14: Exemplo de amostras das classes utilizadas



Fonte: Elaborado pelo Autor com base em imagens do Google Earth.

As quatro classes consideradas no processo de classificação representam diferentes tipos de cobertura do solo. A vegetação arbórea corresponde a áreas com vegetação mais densa, incluindo não apenas florestas e matas, mas também praças, parques urbanos e ruas altamente arborizadas. A vegetação herbácea abrange campos naturais, gramíneas, pastagens e outras coberturas com plantas de pequeno porte, além de incluir áreas de uso agrícola e pecuário, como lavouras e pastagens cultivadas. As áreas impermeabilizadas referem-se a superfícies construídas ou modificadas pelo homem, como ruas, edificações e calçadas, onde a infiltração de água no solo é mínima. Por fim, os corpos hídricos abrangem rios, lagos, represas e outros ambientes onde a presença de água é dominante, de forma permanente ou temporária. Essas classes foram definidas de modo a capturar os principais padrões de uso e cobertura do solo relevantes à análise em questão.

O algoritmo será treinado com um conjunto de dados selecionado a partir de um subconjunto dos dados filtrados durante a etapa de amostragem, de forma a garantir que as amostras representem adequadamente as características de cada classe. A configuração do algoritmo Random Forest, incluindo o número de árvores de decisão e o número máximo de folhas, será determinada de forma iterativa. Em cada iteração, será avaliado o impacto dessas variáveis na precisão do modelo e no tempo de processamento, buscando-se otimizar o equilíbrio entre acurácia e eficiência computacional. Esse processo de análise iterativa permitirá selecionar uma configuração que maximize a capacidade de generalização do modelo, mantendo, ao mesmo tempo, um desempenho computacional adequado aos objetivos da classificação.

4.4. Classificador

No contexto da classificação supervisionada de imagens Landsat, a resolução espacial de 30 m impõe um desafio intrínseco: muitos pixels são mistos, representando uma mescla de coberturas (vegetação, solo exposto, áreas urbanas, etc.). Em tais casos, a assinatura espectral dominante tende a prevalecer na classificação, como quando um pixel contendo uma árvore isolada sobre asfalto exibe reflectância que lembra vegetação, embora a cobertura majoritária seja pavimento.

Essa limitação é reconhecida como uma consequência da escala espacial típica de sensores de resolução média. Com isso, o pixel classificado assume uma única categoria, frequentemente ignorando componentes minoritários, o que pode distorcer resultados em áreas urbanas heterogêneas e de transição (Arai. 2020). Independente da resolução espacial utilizada,

a questão do pixel misto se mantém existente, especialmente em paisagens complexas. Como abordado em Nghiyalwa *et al.* (2021), análises subpixel, como o uso de *Spectral Mixture Analysis* (SMA) são comuns para estimar frações de cobertura dentro de pixels, mostrando-se particularmente útil quando o sensor apresenta resolução média (30 m). SMA permite decompor espectralmente cada pixel em frações.

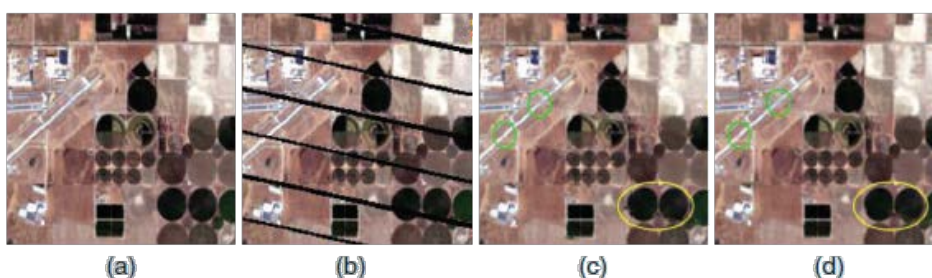
No presente estudo, a abordagem incorpora explicitamente as frações espectrais (componentes *veg*, *sub*, *dark*) no mosaico, o que permite ao classificador *Random Forest* receber informação sobre a composição interna dos pixels mistos. Ainda assim, reconhece-se que a etapa final de classificação atribui apenas uma classe por pixel, o que reflete uma redução da variabilidade interna em prol de uma representação categórica, mas acompanha uma interpretação consciente dessa limitação metodológica.

4.5. Filtros

A produção de dados espaciais está sujeita a efeitos sazonais e condições atmosféricas, que frequentemente resultam em perda de informações em determinadas áreas. Adicionalmente, a natureza orbital dos sensores, aliada à ausência de manutenção e ao tempo de vida útil limitado, pode contribuir para o surgimento de ruídos e lacunas nos dados capturados. Nesse contexto, surgem os processos de reconstrução e filtragem apresentados por SHEN *et al.* (2015), que desempenham um papel essencial na consolidação das informações em meio a essas limitações. Esses processos podem ser classificados em quatro categorias principais, conforme a fonte da informação complementar utilizada na reconstrução das lacunas:

- I. **Métodos baseados em dados espaciais:** Estes métodos utilizam as informações espaciais das áreas adjacentes para preencher lacunas nos dados (Figura 15).

Figura 15: Exemplo experimental de recuperação de dados. (a) Imagem original. (b) Imagem simulada corrompida por linhas mortas. Imagens recuperadas utilizando os seguintes métodos: (c) Interpolação por Kriging; (d) Método de difusão

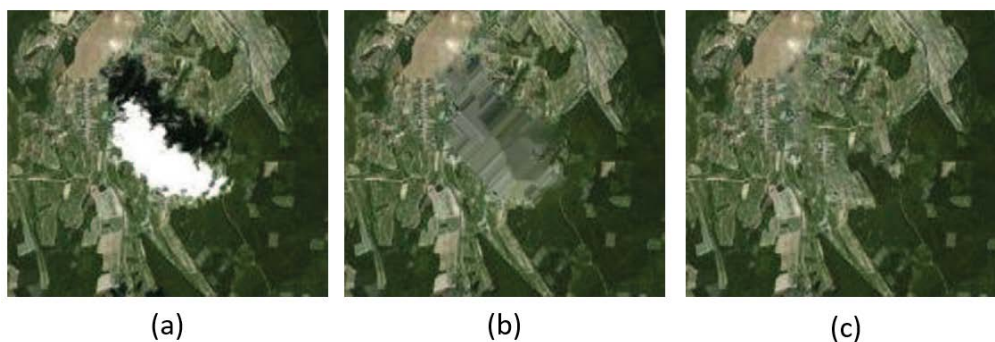


Fonte: Adaptado de Shen *et al.* (2015)

Partem do pressuposto de que os dados ausentes compartilham estrutura estatística ou geométrica com os dados adjacentes, permitindo que informações de proximidade sejam interpoladas. Técnicas comuns incluem interpolação e métodos de difusão, que buscam suavizar transições e evitar artefatos visuais nas junções de dados reconstruídos;

- II. **Métodos baseados em dados espectrais:** Aproveitam a redundância de informações nas diferentes bandas espectrais das imagens multiespectrais e hiperespectrais para preencher dados em bandas específicas (Figura 16).

Figura 16: Exemplo experimental de recuperação de imagem com nuvem. (a) Imagem original; (b) Interpolação por Kriging; (c) Algoritmo baseado em Exemplos de Criminisi

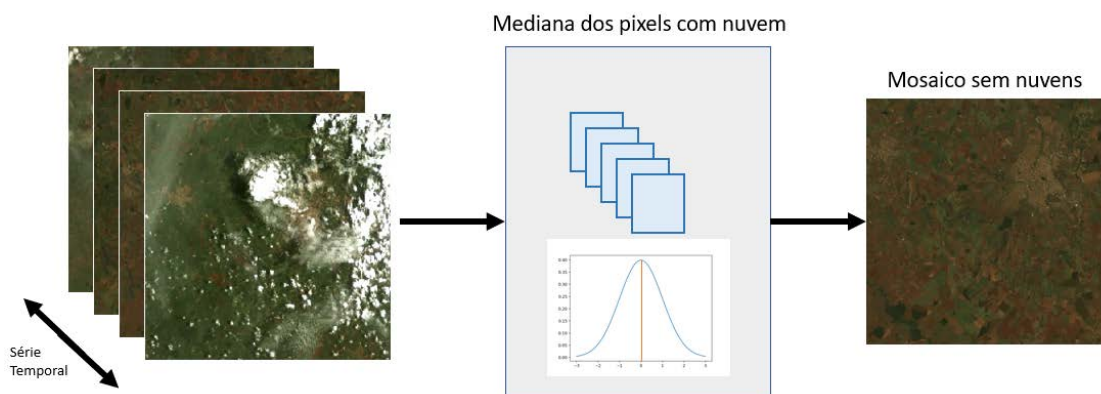


Fonte: Adaptado de Shen *et al.* (2015)

Isso é possível quando algumas bandas estão completas e apresentam correlação espectral com bandas corrompidas. Essa abordagem modela relações latentes entre as bandas para gerar uma reconstrução fidedigna das áreas danificadas, assumindo que a integridade espectral permite esta correlação direta;

- III. **Métodos baseados em dados temporais:** Estes métodos dependem de dados obtidos em múltiplos períodos, usando imagens da mesma área adquiridas em diferentes datas para compensar perdas de dados causadas por variações temporais, como cobertura de nuvens ou falhas de sensor (Figura 17).

Figura 17: Exemplo experimental de recuperação de imagem com nuvem através de criação de mosaico sem nuvens



Fonte: Elaborado pelo Autor.

A abordagem temporal pode ser limitada pela escolha do intervalo de tempo entre as aquisições, que deve ser suficiente para captar alterações sazonais sem comprometer a correlação;

- IV. **Métodos híbridos:** Integram informações espaciais, espectrais e temporais, combinando as vantagens de cada abordagem. Técnicas híbridas permitem reconstruções mais robustas e precisas ao explorar as complementaridades entre os diferentes domínios de correlação. Essas técnicas são especialmente eficazes para lidar com regiões de dados ausentes que apresentam complexidades em várias dimensões, oferecendo uma solução abrangente e adaptável para a recuperação de dados.

4.6. Validação

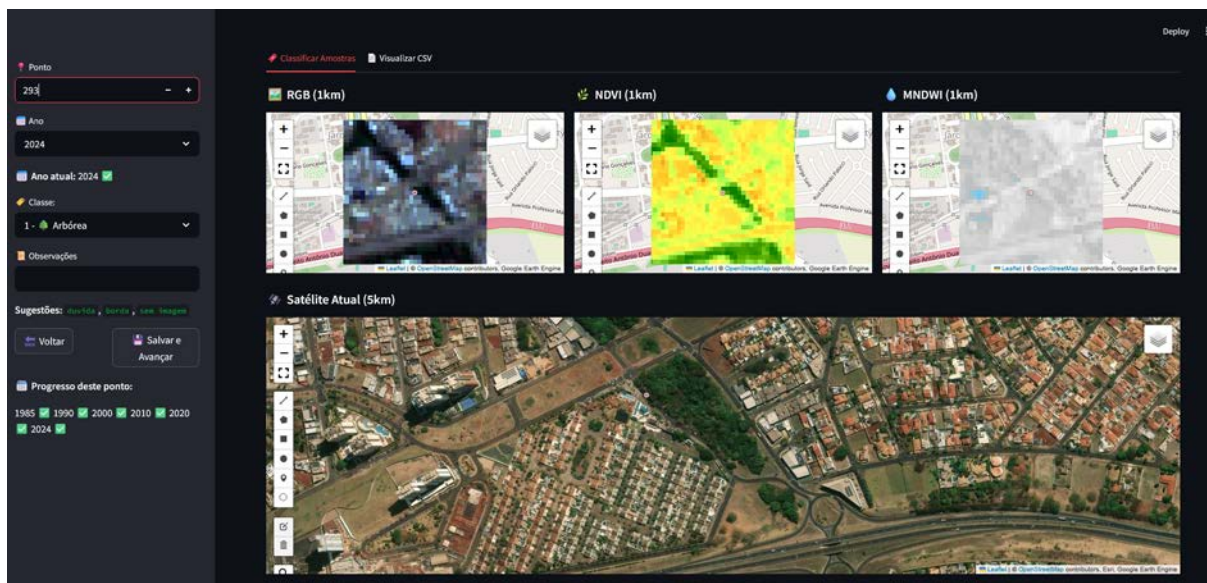
Em detrimento do baixo custo, o uso de dados provenientes do sensoriamento remoto orbital vem se tornando mais comum, assim como seu processamento por meio de algoritmos computacionais. A validação dos resultados desses processos é de extrema importância para se compreender questões relacionadas à qualidade da informação apresentada.

Para isso, a análise de acurácia de um dado tem como princípio a comparação cruzada entre o produto produzido e um dado de referência, sendo necessariamente ambos os conjuntos de dados representativos de um mesmo espaço amostral em uma perspectiva espacial assim como temporal. Esse processo de obtenção dos dados de referência em geral pode ser realizado por meio de imagens de satélite, desde de que se respeite a compatibilidade de escala com o produto a ser validado (Olofsson et al. 2014, Finegold et al. 2016).

A primeira etapa do processo de validação consiste na construção de um conjunto de amostras. Este conjunto de amostras será estratificado com base no quantitativo de áreas das classes do mapa de uso e cobertura da terra produzido pelo MapBiomias para a RMRP. Essa estratificação prévia permite uma representação aproximada adequada das diferentes classes de interesse durante a validação do processo de amostragem automática e da classificação produzida pelo algoritmo Random Forest. Com base em um *dataset* de pontos randômicos, distinto do de treinamento, um conjunto de validação é organizado, com um nível de confiança de 99% e uma margem de erro de 3%. Vale salientar que o presente conjunto de validação é integralmente removido do utilizado para o treinamento do classificador.

Para o processo de validação visual das amostras, foi desenvolvido um aplicativo interativo em Python que integra a visualização de imagens Landsat e o registro de classificações atribuídas a pontos amostrais (Figura 18). A interface foi construída com a biblioteca Streamlit, e combina ferramentas de geoprocessamento (Google Earth Engine, GeoPandas, Shapely) com bibliotecas de ciência de dados (Pandas, NumPy) e visualização geoespacial (geemap).

Figura 18: Aplicativo desenvolvido para validação visual de amostras



Fonte: Elaborado pelo Autor.

O sistema conecta-se diretamente a um catálogo de mosaicos anuais do satélite Landsat hospedado no Google Earth Engine, permitindo ao usuário inspecionar visualmente as amostras para diferentes anos. Para cada ponto e ano selecionado, são exibidas composições em cores naturais (RGB), além dos índices espectrais NDVI (Normalized Difference Vegetation Index) e MNDWI (Modified Normalized Difference Water Index), com o objetivo de aprimorar a

distinção entre vegetação, áreas impermeáveis e corpos d'água. Também é disponibilizada uma imagem de satélite em alta resolução como suporte adicional à interpretação visual.

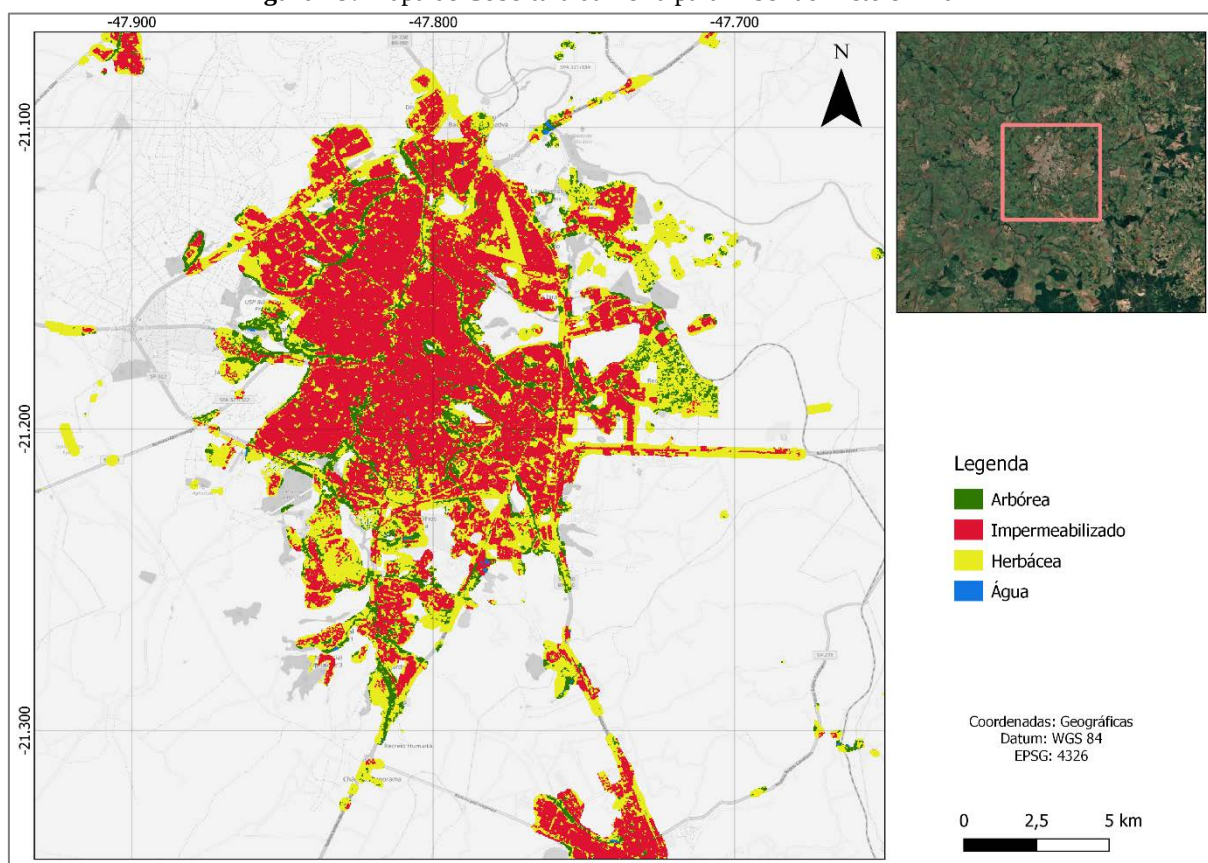
Cada ponto amostral possui um identificador único e pode ser classificado pelo usuário em uma das categorias temáticas do projeto, com a possibilidade de registrar observações adicionais sobre a amostra, como dúvidas ou limitações de imagem. As informações são armazenadas em um arquivo CSV, o qual é atualizado continuamente ao longo do processo. Para garantir a integridade dos dados, o sistema implementa uma rotina de backup automático periódico, que cria cópias versionadas do arquivo de saída, prevenindo perdas decorrentes de falhas inesperadas. A interface apresenta mecanismos para navegação entre pontos e anos, controle de progresso por amostra e uma aba complementar para visualização tabular dos dados já salvos.

5. RESULTADOS

5.1. Contexto das Áreas urbanas

Com o desenvolvimento² e aplicação do método proposto, foi possível identificar automaticamente as áreas vegetadas da RMRP por meio de uma classificação supervisionada utilizando o algoritmo Random Forest e amostras geradas de forma aleatória com o apoio de mapas de referência e filtragem. Para a construção de um mapa que abrangesse tanto a área urbana quanto seu entorno imediato, adotou-se como estratégia de filtragem espacial o índice de vias e estruturas desenvolvido por Justiniano *et al.* (2022). Esse índice funciona de maneira semelhante a um mapa de calor da densidade de vias, permitindo a delimitação da área urbana e de suas proximidades com base na concentração de ruas e avenidas. Na figura 19 apresenta-se parte do resultado obtido.

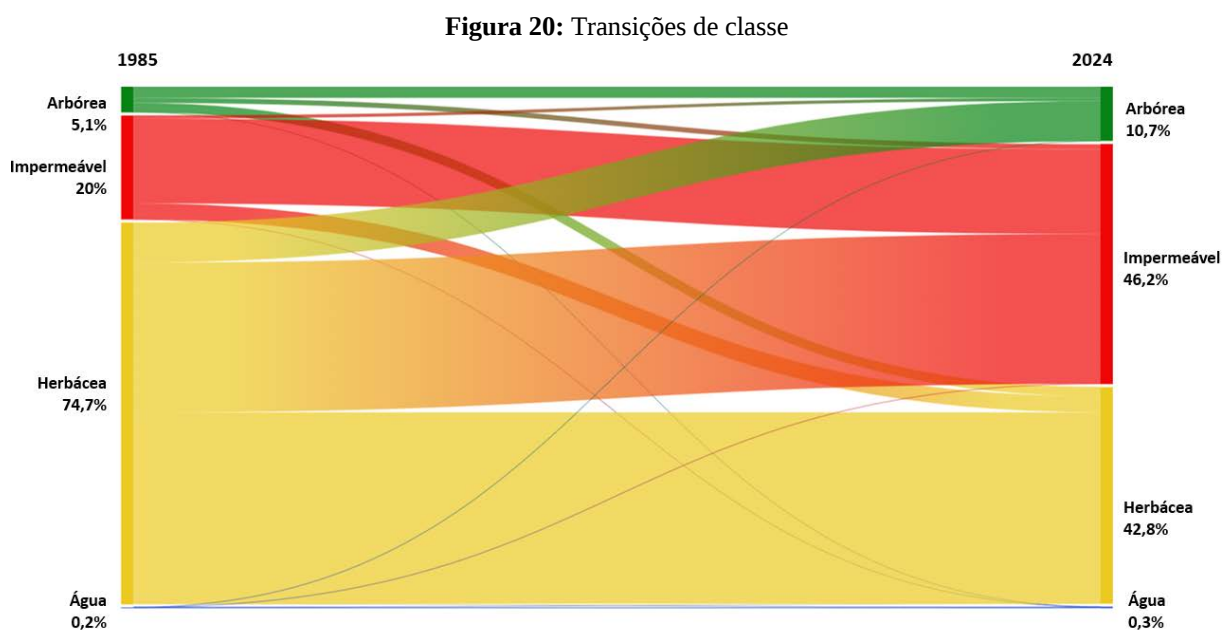
Figura 19: Mapa de Cobertura da Terra para Ribeirão Preto em 2024



Fonte: Elaborado pelo Autor

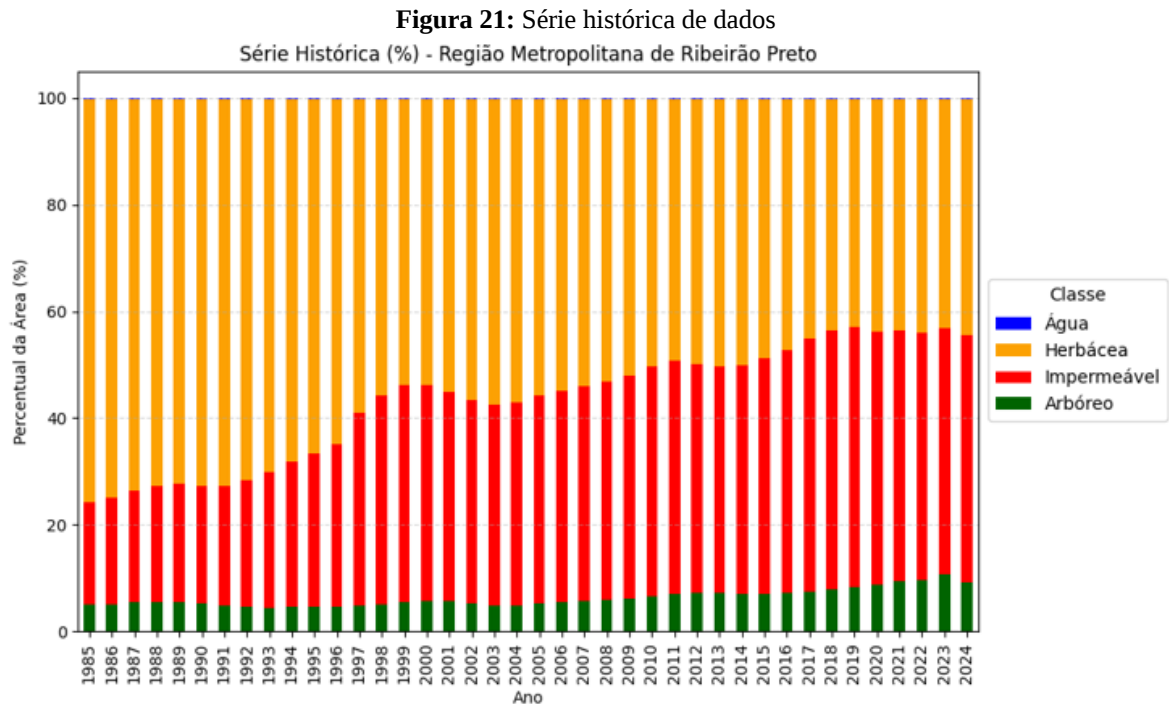
² Detalhamento dos códigos desenvolvidos encontram-se no Apêndice. Resultados da classificação encontram-se disponíveis em: <https://brenomalheiros-ufscar.projects.earthengine.app/view/classificacao-rmrp-1985-2024>

Como apresentado na Figura 20, ao observar todo o território da Região Metropolitana, constata-se o predomínio da classe de vegetação herbácea, a qual corresponde às áreas de pastagem e culturas de pequeno porte. No ano de 1985, essa classe ocupava aproximadamente 74,7% da superfície da RMRP. Apesar de uma leve redução ao longo das décadas, ela permanece relevante em termos quantitativos, representando ainda 42,8% da área mapeada em 2024. Em contraste, classe água aparece com baixa representatividade em termos de cobertura superficial quando comparada as demais classes, nas análises posteriores essa mantém valores abaixo de 1% com pequenas oscilações em valores absolutos na série temporal.



Fonte: Elaborado pelo Autor

Paralelamente, nota-se um crescimento expressivo da vegetação arbórea, que aumentou de 5,1% em 1985 para cerca de 10,7% em 2024, um acréscimo de quase 6 pontos percentuais ao longo de 40 anos. Esse aumento pode refletir processos de regeneração natural ou esforços de reflorestamento em determinadas áreas. Já as áreas impermeáveis, associadas à urbanização, também apresentaram crescimento, passando de 20% para quase 46,2% no mesmo período. A classe água manteve-se praticamente constante, oscilando levemente entre 0,18% e 0,28% de superfície total. A variação das classes nesse período pode ser observada na Figura 21.



Para a etapa de validação, foram utilizados 628 pontos amostrais aleatórios em cada um dos seis anos analisados, totalizando 3.768 amostras. Essa amostragem buscou representar adequadamente a variabilidade espacial e temporal da área de estudo. Inicialmente, avaliou-se a acurácia do produtor (Tabela 3), a qual expressa a proporção de amostras de uma classe real que foram corretamente classificadas, ou seja, indica a capacidade do classificador em evitar omissões dessa classe.

Tabela 3: Acurácia do Produtor (*Producer's Accuracy*)

Classe	1985	1990	2000	2010	2020	2024
Arbóreo	50	58,06	61,29	69,23	85,37	92,86
Impermeável	58,49	68,42	83,33	78,75	78,89	76,29
Herbácea	90,91	94,44	82,76	80,43	87,8	97,14
Água	64,71	82,35	88,24	100	100	100

Fonte: Elaborado pelo autor.

Em seguida, calculou-se a acurácia do consumidor (Tabela 4), que representa a proporção de pixels classificados em determinada classe que realmente pertencem a ela, ou seja, está relacionada à confiabilidade do dado classificado para o usuário final.

Tabela 4: Acurácia do Consumidor (*User's Accuracy*)

Classe	1985	1990	2000	2010	2020	2024
Arbóreo	83,33	94,74	76	93,1	94,59	95,12
Impermeável	81,58	88,64	86,96	85,14	92,21	98,67
Herbácea	60,61	68	69,57	59,68	62,07	58,62
Água	100	100	100	100	100	100

Fonte: Elaborado pelo autor.

A análise dos resultados de acurácia revela uma evolução consistente no desempenho do classificador Random Forest ao longo da série temporal de 1985 a 2024. Contudo, observa-se uma acurácia sensivelmente inferior nos anos iniciais, especialmente em 1985 e 1990, tanto para a acurácia do produtor quanto para a do consumidor. Essa limitação inicial pode ser atribuída a uma série de fatores técnicos e metodológicos amplamente documentados na literatura.

Em primeiro lugar, as imagens Landsat mais antigas, obtidas por sensores como o MSS e o TM, apresentam restrições importantes em termos de resolução radiométrica, espectral e espacial. O Landsat MSS operava com apenas 4 bandas espectrais e resolução de 80 metros, enquanto o TM, embora mais avançado, ainda possuía resolução de 30 metros e apenas 8 bits de quantização radiométrica (Tucker, 1980; Irons *et al.*, 1985). Essas características limitavam a capacidade de distinguir entre alvos com assinaturas espectrais semelhantes, comprometendo a separabilidade entre classes. Além disso, a relação sinal-ruído (SNR) desses sensores era consideravelmente inferior à dos sensores mais modernos, como o OLI do Landsat 8, que apresenta SNR até 10 vezes maior (Roy *et al.*, 2014). Isso implica que variações sutis na refletância, sobretudo em alvos de baixa intensidade espectral — como água turva, áreas sombreadas ou vegetação rala — eram frequentemente ofuscadas pelo ruído eletrônico e atmosférico, prejudicando a confiabilidade da classificação.

A presença de interferências atmosféricas também é uma limitação relevante nos dados antigos. As imagens Landsat das décadas de 1980 e 1990 frequentemente sofriam com a presença de névoa, aerossóis e variações atmosféricas que, mesmo após filtragem de nuvens e composição de mosaicos anuais, resultavam em distorções radiométricas. Chávez Jr. (1988) já apontava a necessidade de métodos específicos para correção atmosférica em imagens Landsat TM, sinalizando a fragilidade desses dados em sua forma bruta. Outro fator determinante para a baixa acurácia nos anos iniciais é a alta similaridade espectral entre certas classes, exacerbada pelas limitações dos sensores da época. Por exemplo, Dash *et al.* (2023) destacam que solos

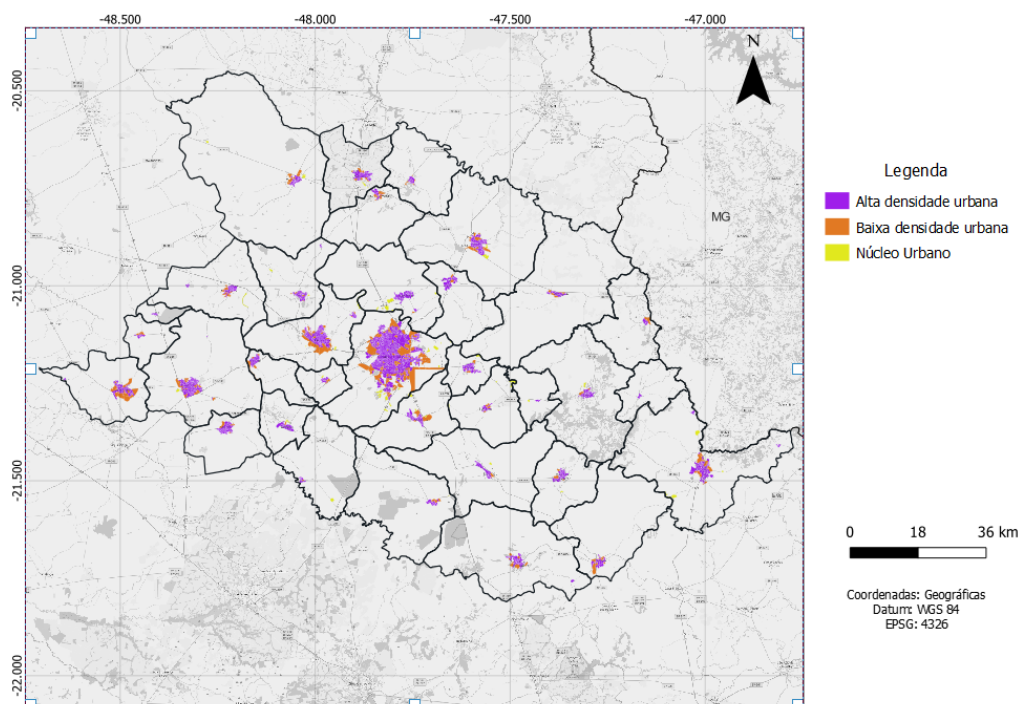
expostos em áreas agrícolas podem apresentar assinaturas espectrais muito semelhantes às de superfícies urbanas, especialmente quando analisados com bandas amplas e baixa resolução radiométrica. Situações semelhantes ocorrem na distinção entre vegetação herbácea e áreas arbóreas em regeneração ou secundárias.

A partir dos anos 2000, observa-se uma melhora significativa nas métricas de acurácia, refletindo o avanço tecnológico dos sensores, com maior fidelidade radiométrica, melhores bandas espectrais e produtos já corrigidos atmosférica e geometricamente, além da maior disponibilidade de dados de referência confiáveis para treinamento. Essa melhoria reforça a importância de considerar o contexto tecnológico e metodológico de cada época ao interpretar séries temporais de uso e cobertura da terra baseadas em classificação supervisionada com imagens Landsat.

5.2. Áreas Urbanas – IBGE

Com base na tipologia territorial proposta pelo IBGE (2022), as áreas urbanas são classificadas em três situações distintas: área urbana de alta densidade de edificações, área urbana de baixa densidade de edificações e núcleo urbano. Essa classificação (Figura 22) leva em conta tanto a morfologia urbana quanto aspectos funcionais, como densidade construtiva, infraestrutura disponível e relação espacial com a malha urbana principal.

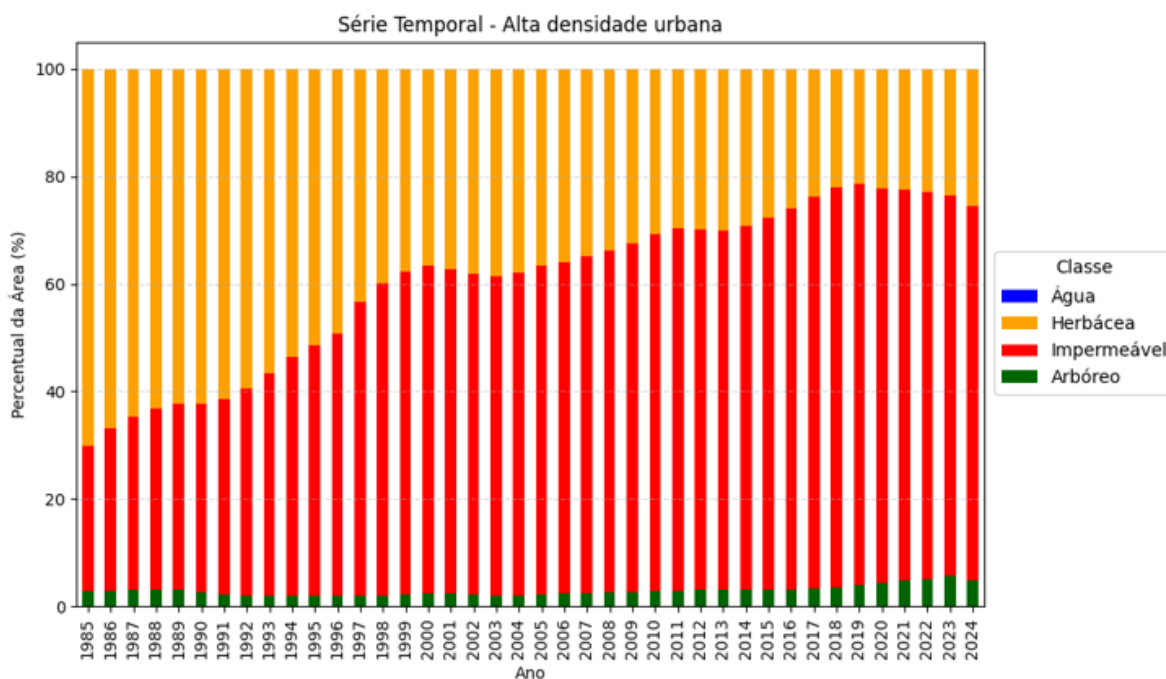
Figura 22: Situação dos setores censitários da RMRP



Fonte: Adaptado de IBGE (2022)

As áreas urbanas de alta densidade correspondem a setores consolidados das cidades, com grande concentração de edificações, presença contínua de infraestrutura urbana e elevada densidade demográfica. São, geralmente, regiões centrais e bairros formalizados, caracterizados por ocupação intensa do solo, verticalização e baixa proporção de áreas livres. A Figura 23 apresenta os resultados para cada uma das classes mapeadas nesse contexto.

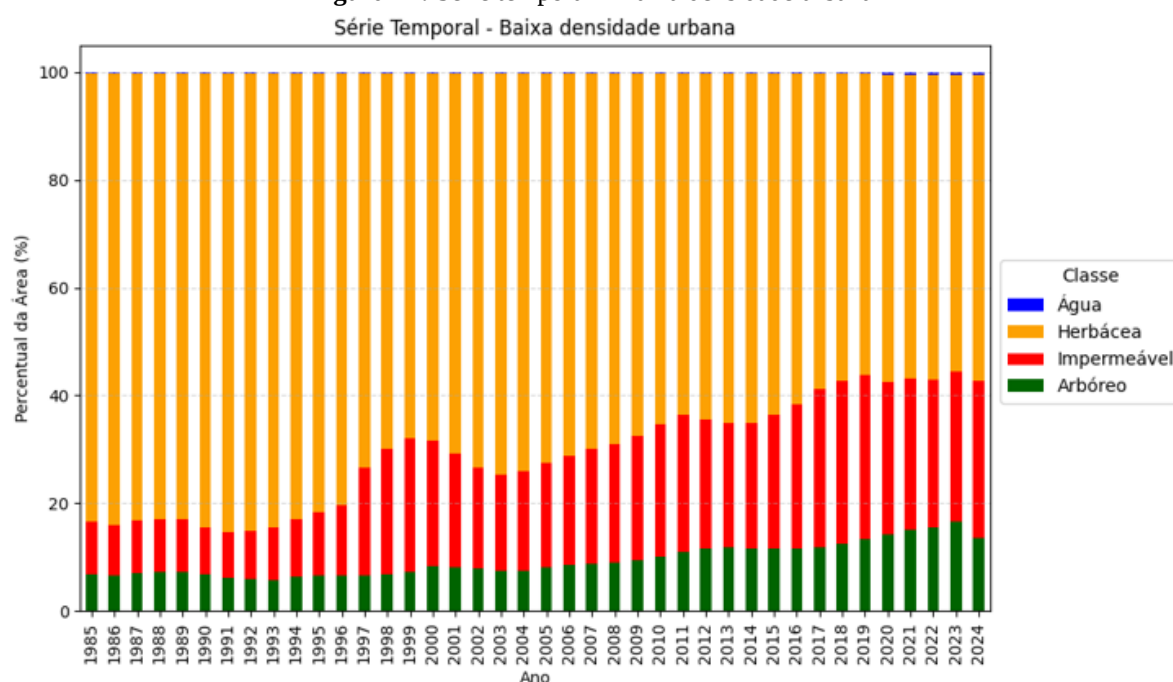
Figura 23: Série temporal – Alta densidade urbana



Fonte: Elaborado pelo autor.

Entre 1985 e 2024, essa categoria apresentou um crescimento expressivo da classe impermeável (+133,5%), refletindo a intensificação da ocupação urbana e o processo de adensamento físico. Paralelamente, observou-se uma redução significativa das áreas herbáceas (-64,9%), indicando substituição de coberturas vegetadas por edificações, vias pavimentadas e outras infraestruturas. A cobertura arbórea, por sua vez, apresentou aumento de +101,3%, o que pode estar relacionado à arborização viária e à preservação de parques urbanos, ou ainda ao aprimoramento na detecção por sensores mais modernos em anos recentes.

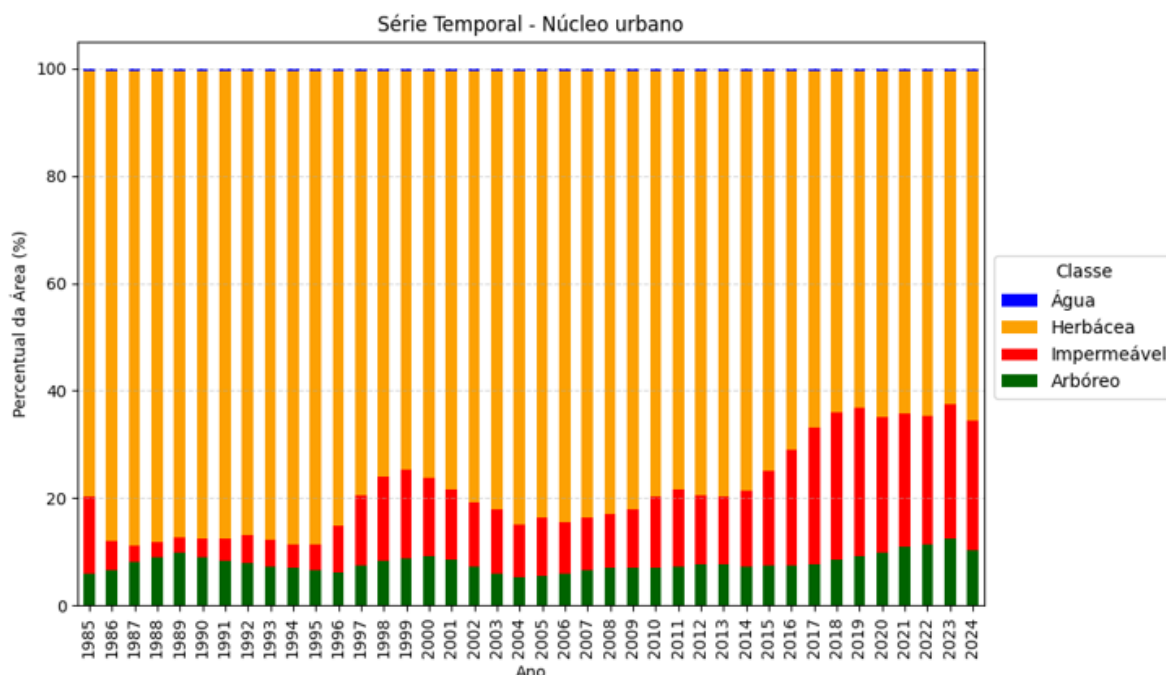
As áreas urbanas de baixa densidade englobam setores com padrão construtivo menos concentrado, geralmente localizados em zonas de expansão urbana, bairros periféricos ou áreas em processo de consolidação. Essas regiões são caracterizadas por espaçamento entre edificações, presença de terrenos vagos ou vegetação, e infraestrutura urbana em desenvolvimento. A Figura 24 apresenta os resultados para cada uma das classes mapeadas nesse contexto.

Figura 24: Série temporal – Baixa densidade urbana

Fonte: Elaborado pelo autor.

Nessas áreas, a superfície impermeável aumentou 198,4%, evidenciando a expansão horizontal da malha urbana sobre terrenos anteriormente vegetados ou não ocupados. A classe arbórea cresceu 152,0%, sugerindo a incorporação de áreas vegetadas durante o processo de urbanização ou ações de arborização em loteamentos e condomínios horizontais. Em contrapartida, a cobertura herbácea reduziu-se em 34,3%, o que demonstra uma conversão gradual de solos naturais em uso urbano.

Os núcleos urbanos são definidos como aglomerados urbanos isolados, não contíguos à malha urbana principal, mas com características típicas de urbanização. Compreendem desde condomínios fechados e núcleos de veraneio até pequenas vilas urbanizadas que se formam fora do contínuo urbano consolidado. A Figura 25 apresenta os resultados para cada uma das classes mapeadas nesse contexto.

Figura 25: Série temporal – Núcleos urbanos

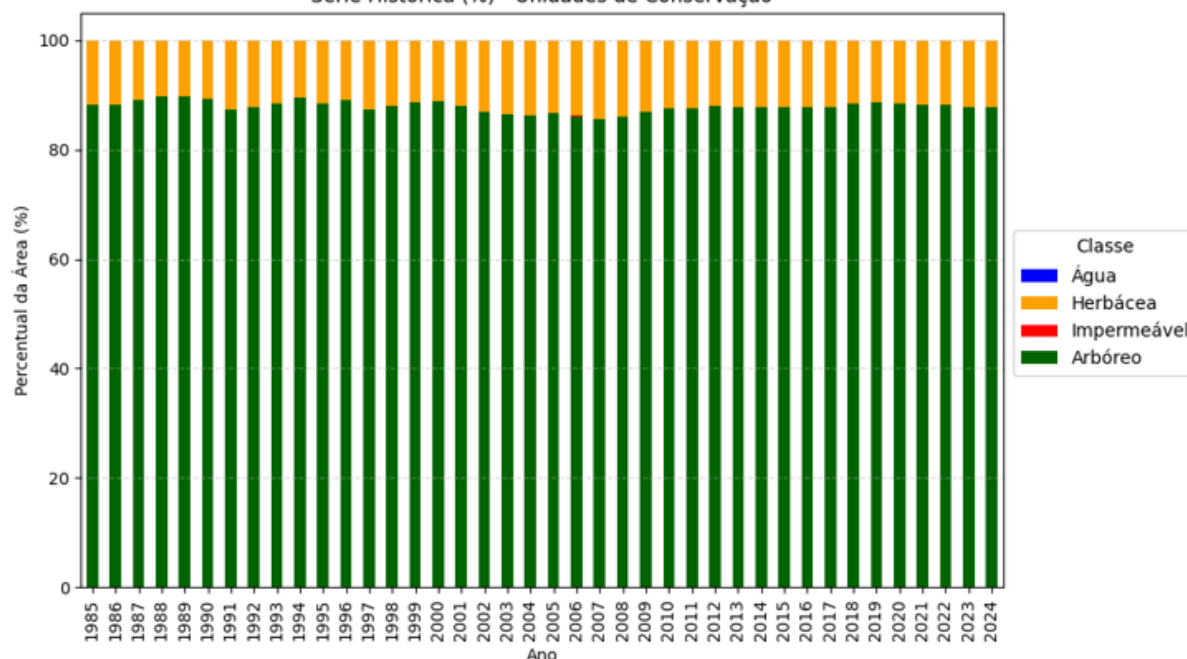
Fonte: Elaborado pelo autor.

Nesse tipo de área, observou-se o crescimento mais acentuado da classe impermeável, quase triplicando ao longo dos 40 anos analisados. Esse resultado evidencia um processo acelerado de urbanização em áreas antes periféricas ou rurais, hoje integradas ao tecido urbano. A cobertura arbórea aumentou 88,7%, o que pode indicar manutenção parcial da vegetação original ou inserção de elementos paisagísticos em novos empreendimentos. A redução da classe herbácea (-28,8%) confirma a tendência de urbanização desses espaços.

5.3. Unidade de Conservação

A análise da dinâmica do uso e cobertura do solo nas Unidades de Conservação (UCs) da Região Metropolitana de Ribeirão Preto, com base na série histórica compreendida entre os anos de 1985 e 2024, evidencia uma estabilidade relativa na dominância da vegetação arbórea, a qual se manteve como a principal forma de cobertura ao longo de todo o período (Figura 26).

Figura 26: Série temporal – Unidades de Conservação (UCs)
Série Histórica (%) - Unidades de Conservação



Fonte: Elaborado pelo autor.

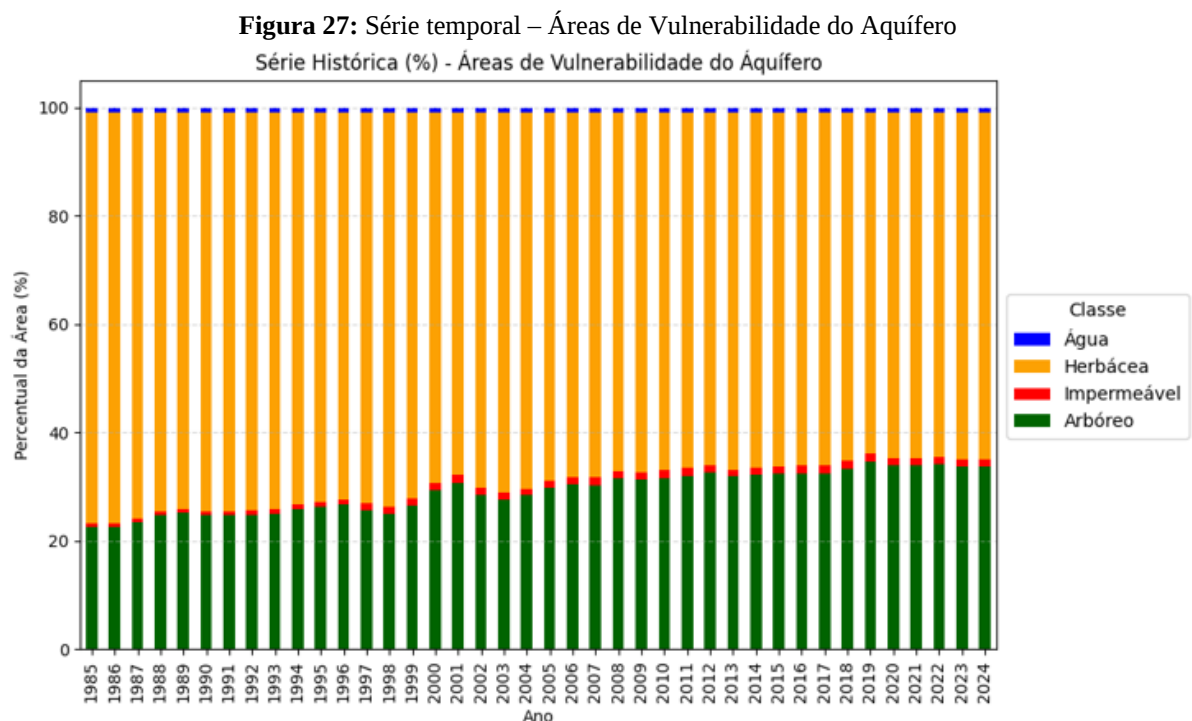
Em 1985, aproximadamente 88,2% da área das UCs era coberta por vegetação arbórea, o que corresponde a cerca de 12.456 km². Esse percentual oscilou de forma moderada ao longo das décadas, mantendo-se, de maneira geral, acima de 86%, e encerrando o período analisado com uma cobertura de 87,8% em 2024, equivalente a 12.392 km². Essa constância sugere que, diferentemente de regiões mais susceptíveis à pressão urbana ou agrícola, as unidades de conservação analisadas mantiveram sua função de proteção e conservação da vegetação nativa, mesmo diante das dinâmicas regionais de uso do solo. A pequena redução observada entre os valores máximos (cerca de 89,8% em 1989) e os valores finais pode indicar variações sazonais, ajustes de classificação ou processos localizados de degradação ou substituição de cobertura vegetal, mas não sugere uma tendência estrutural de perda florestal significativa.

Em paralelo, a vegetação herbácea apresentou comportamento oscilante, refletindo talvez variações ecológicas de curto prazo ou dinâmicas intermediárias entre regeneração e degradação. Em 1985, essa classe ocupava 1.647 km² (11,7% da área), valor que permaneceu relativamente estável ao longo dos anos, com variações entre aproximadamente 10% e 13%. Em 2024, a área herbácea atinge 1.711 km², o que representa 12,1% da superfície total das UCs. Essas flutuações indicam que parte do território pode estar sujeita a processos naturais de sucessão secundária, ou ainda à influência de distúrbios que mantêm as áreas em estágio inicial

de regeneração, como queimadas, desmatamentos localizados ou interferências antrópicas de baixa intensidade.

5.4. Áreas de vulnerabilidade do Aquífero

A análise da dinâmica de uso e cobertura do solo nas áreas de vulnerabilidade do aquífero da Região Metropolitana de Ribeirão Preto (RMRP), conforme definidas no Plano de Desenvolvimento Urbano Integrado (PDUI), revela transformações significativas no período de 1985 a 2024, especialmente no que diz respeito à vegetação arbórea e à ocupação antrópica (Figura 27).



Fonte: Elaborado pelo autor.

Essas áreas, por sua importância estratégica na recarga do aquífero Guarani e outros sistemas subterrâneos, requerem atenção especial quanto ao uso do solo, sobretudo no que diz respeito à presença de vegetação nativa e à intensidade das atividades agropecuárias e urbanas. No início da série, em 1985, a paisagem era marcada pela expressiva predominância de áreas classificadas como herbáceas, que ocupavam cerca de 75,9% do território vulnerável (aproximadamente 286.600 km²). Quando se considera que essa classe abrange, além de vegetação campestre, também pastagens e zonas de cultivo, ela se configura como um forte indicador do uso antrópico do solo. Ao longo do período, essas áreas sofreram uma redução significativa, atingindo 63,9% da superfície em 2024 (cerca de 241.700 km²). Essa diminuição

de mais de 44.000 km² sugere um processo de redução da pressão agropecuária direta, especialmente em áreas de menor aptidão agrícola ou já degradadas, que vêm sendo progressivamente abandonadas ou reconvertidas para cobertura florestal.

Em paralelo, houve um crescimento substancial da vegetação arbórea, cuja cobertura passou de 22,6% (85.568 km²) em 1985 para 33,7% (127.165 km²) em 2024. Esse aumento expressivo, de aproximadamente 41.600 km², pode ser atribuído tanto à regeneração natural em áreas anteriormente desmatadas, processo típico de sucessão secundária, quanto a ações planejadas de reflorestamento e à obrigatoriedade legal de recomposição de vegetação APPs, especialmente após a promulgação do novo Código Florestal. A presença crescente de cobertura arbórea em áreas de recarga hídrica é particularmente positiva, pois contribui para a proteção do solo, regulação do ciclo hidrológico, aumento da infiltração e melhora da qualidade da água subterrânea.

A expansão urbana, representada pelas áreas impermeáveis, também apresentou crescimento, mas de forma mais contida. Essa classe passou de 0,58% em 1985 para 1,39% em 2024, totalizando um aumento de cerca de 3.000 km². Embora esse valor seja relativamente pequeno em termos percentuais, sua significância hidrológica é elevada, já que áreas impermeáveis comprometem diretamente a capacidade de recarga do aquífero e ampliam os riscos de poluição difusa. A maior parte desse crescimento está concentrada em zonas periurbanas e de expansão urbana desordenada, particularmente nos municípios mais populosos da RMRP.

6. DISCUSSÃO

6.1. O Aumento das áreas verdes

Entre 1985 e 2024, a RMRP vivenciou um processo gradual, porém consistente, de recomposição da vegetação arbórea, marcando uma inflexão na trajetória histórica de supressão da cobertura vegetal nativa. Este fenômeno se explica, em grande medida, pela consolidação de um arcabouço jurídico-ambiental multiescalar, articulado entre os níveis federal, estadual e municipal, que ao longo de quase quatro décadas passou a regular, restringir e, sobretudo, induzir práticas de conservação e restauração florestal. Em paralelo, políticas públicas setoriais e a crescente mobilização de atores locais contribuíram para operacionalizar e dar efetividade a tais dispositivos legais. Dessa forma, o aumento das áreas arborizadas na RMRP não pode ser compreendido como um fenômeno isolado ou espontâneo, mas sim como o resultado de um conjunto articulado de normas, planos e ações que convergiram para transformar a lógica de uso e ocupação do território, tanto em zonas urbanas quanto rurais.

O ponto de partida desse processo remonta à década de 1980, quando o Brasil começou a consolidar sua legislação ambiental moderna. A promulgação da Constituição Federal de 1988 representou um marco institucional decisivo ao incorporar, em seu artigo 225, o direito de todos ao meio ambiente ecologicamente equilibrado e ao estabelecer o dever do Poder Público e da coletividade em defendê-lo e preservá-lo. Essa garantia constitucional orientou a criação de uma série de leis federais com foco na preservação da vegetação nativa, entre as quais se destacam o Código Florestal, em suas versões de 1965 e 2012, a Lei da Mata Atlântica (Lei nº 11.428/2006) e a Lei de Crimes Ambientais (Lei nº 9.605/1998). Tais normativas instituíram dispositivos fundamentais, como a obrigatoriedade de manutenção e recuperação de APPs e Reservas Legais (RLs), a proibição da supressão de vegetação em estágios avançados de regeneração e a responsabilização penal e administrativa por danos ambientais. Na prática, essas exigências passaram a condicionar a regularidade ambiental de imóveis rurais e empreendimento urbanos, forçando proprietários e gestores públicos a adotar medidas de reflorestamento e compensação.

O Novo Código Florestal (Lei nº 12.651/2012), ao introduzir o Cadastro Ambiental Rural (CAR) e os Programas de Regularização Ambiental (PRA), representou um salto qualitativo na capacidade do Estado de monitorar e induzir a recomposição florestal em larga escala. O cruzamento de dados georreferenciados permitiu identificar com precisão passivos ambientais nas propriedades e estabelecer metas específicas de recomposição florestal, especialmente em margens de rios e nascentes. Na região de Ribeirão Preto, tais dispositivos

resultaram na reativação de matas ciliares degradadas, na recuperação de fragmentos desconectados e na constituição de corredores ecológicos ao longo das principais bacias hidrográficas, como as do rio Pardo e do rio Mogi-Guaçu. Além disso, a Lei da Mata Atlântica garantiu proteção especial à vegetação secundária regenerante, presente em diversos capões florestais urbanos e rurais da RMRP, impedindo sua supressão para novos empreendimentos sem rigorosos critérios técnicos e legais, o que também inibiu a substituição de áreas arborizadas por usos urbanos intensivos.

No âmbito estadual, o Estado de São Paulo desenvolveu políticas ambientais pioneiras que atuaram como catalisadoras da recuperação florestal. A implementação do ICMS Ecológico a partir de 1993 passou a transferir maiores recursos aos municípios que mantinham áreas legalmente protegidas, como unidades de conservação, reservas legais regularizadas e fragmentos com vegetação nativa. Esse mecanismo fiscal, posteriormente expandido com a Lei nº 17.348/2021, teve impacto direto na decisão de diversos municípios da RMRP em instituir áreas verdes protegidas, como a Estação Ecológica Municipal Guarani, criada em 2018 em Ribeirão Preto, e zonas de amortecimento para remanescentes relevantes em Jardinópolis, Sertãozinho e Batatais. Adicionalmente, a Política Estadual de Mudanças Climáticas (Lei nº 13.798/2009) e o Programa Nascentes fomentaram projetos de restauração ecológica em APPs e bacias prioritárias, canalizando recursos públicos e privados para o plantio de espécies nativas em áreas estratégicas para a conectividade ecológica e proteção de recursos hídricos, muitos deles localizados na região de Ribeirão Preto.

No plano municipal, a promulgação da Lei nº 7.159/1995, em Ribeirão Preto, foi um divisor de águas ao institucionalizar um sistema normativo próprio de arborização urbana e proteção das áreas verdes. A referida norma estabeleceu critérios para poda e supressão arbórea, definiu proporções obrigatórias de reposição vegetal (chegando a exigir até 25 mudas para cada árvore suprimida em casos específicos) e reconheceu como unidades de conservação municipais as APPs, RLs e fragmentos vegetais urbanos relevantes. Essa lei foi decisiva para consolidar uma cultura jurídica e administrativa de valorização das árvores na paisagem urbana e periurbana da cidade. Em complemento, o Plano Diretor revisado em 2018 incorporou diretrizes ambientais orientadas à criação de parques lineares, proteção de fundos de vale e ampliação da cobertura arbórea nos bairros periféricos, especialmente em regiões carentes de infraestrutura verde. Tais instrumentos não apenas impediram o avanço do processo de desmatamento urbano, mas também estimularam a regeneração natural e a implantação de novos corredores vegetados.

Além das normativas, o contexto de aumento da vegetação arbórea foi favorecido por um conjunto de práticas administrativas e mobilizações sociais. Iniciativas como o Programa Município VerdeAzul (PMVA), a implantação de Planos Diretores de Arborização Urbana e a criação de unidades de conservação municipais com manejo ativo, como a própria Estação Ecológica Guarani e a Área de Proteção Ambiental do Morro do São Bento, oficializada em 2024, demonstram que a legislação encontrou ressonância institucional e apoio político para sua implementação. A participação do Ministério Público, exigindo compensações ambientais em casos de supressão legal de vegetação, e o engajamento da sociedade civil em projetos de plantio e manutenção de áreas verdes, contribuíram para consolidar a efetividade normativa e legitimar as políticas públicas voltadas à restauração ecológica. A própria recomposição da Estação Ecológica de Ribeirão Preto (Mata de Santa Tereza), parcialmente devastada por incêndios em 2014, é um exemplo emblemático da convergência entre obrigações legais e ação coletiva para a recuperação de ecossistemas urbanos estratégicos.

6.2. O impacto do PDUI

A criação da Região Metropolitana de Ribeirão Preto e a subsequente elaboração do Plano de Desenvolvimento Urbano Integrado configuraram, ao menos em nível programático, uma inflexão relevante no enfrentamento das questões socioambientais regionais. Ainda que se observe um avanço no sentido de criar instâncias técnicas e federativas de articulação entre os entes municipais, e que o processo tenha representado uma oportunidade para diagnosticar criticamente os efeitos acumulados da expansão urbana desregulada, da fragmentação ecológica e da vulnerabilidade de recursos estratégicos como o Aquífero Guarani, o alcance prático dessas formulações permanece, em grande medida, restrito ao plano das intenções. O diagnóstico técnico que embasou o PDUI delinea um cenário ambiental marcado por desequilíbrios profundos na distribuição da cobertura vegetal nativa, amplamente comprometida nos núcleos urbanos mais densamente ocupados, e pela fragilidade dos mecanismos regulatórios destinados à proteção dos ecossistemas remanescentes, os quais se apresentam não apenas espacialmente fragmentados, mas também desprovidos de conexões ecológicas funcionais.

Dentre os principais elementos estruturantes identificados, ressaltam-se a descontinuidade das áreas vegetadas, a escassez de Unidades de Conservação efetivamente instituídas e manejadas, a intensa pressão antrópica sobre zonas de recarga do Aquífero Guarani, que se sobrepõem a porções urbanizadas de diversos municípios, bem como os impactos ambientais oriundos de processos erosivos, da crescente impermeabilização dos solos

e da carência de uma lógica ecossistêmica nos instrumentos de planejamento territorial. Tal conjunto de vulnerabilidades é agravado pelo fato de que uma parcela considerável dos municípios integrantes da RMRP apresenta planos diretores desatualizados ou mesmo inexistentes, o que evidencia a ausência de uma infraestrutura normativa apta a viabilizar, de forma eficaz, os princípios de precaução e sustentabilidade territorial enunciados nos documentos estratégicos do PDUI.

Nesse panorama, a vegetação urbana é reiteradamente destacada nos relatórios e propostas como um vetor ambiental de grande relevância, não apenas por seu valor estético e simbólico, mas por seu papel funcional na provisão de serviços ecossistêmicos fundamentais à qualidade de vida urbana, como a regulação térmica, a absorção de águas pluviais, a melhoria da qualidade do ar e a promoção da saúde mental e do bem-estar coletivo. Apesar disso, os dados disponíveis indicam que a cobertura vegetal urbana da RMRP se encontra substancialmente aquém dos padrões recomendados por organismos internacionais, com distribuição profundamente desigual entre bairros e municípios, o que aprofunda as disparidades socioambientais e contribui para o agravamento das ilhas de calor e da vulnerabilidade climática das populações urbanas mais expostas.

Frente a esse quadro, o Caderno de Propostas do PDUI apresenta um conjunto coerente e tecnicamente bem fundamentado de diretrizes organizadas em quatro grandes objetivos estratégicos: garantir a segurança hídrica regional; promover a conectividade ecológica e a conservação de ecossistemas remanescentes; aprimorar a infraestrutura de saneamento urbano e rural; e ordenar o uso e a ocupação do solo com base em critérios ambientais e geotécnicos. Contudo, embora essas propostas sinalizem um desejável reposicionamento do planejamento urbano-metropolitano, sua efetividade esbarra na ausência de instrumentos legais vinculantes e de arranjos institucionais capazes de operacionalizá-las no território.

A exemplo a ideia de um “urbanismo sensível à água”, que pressupõe a integração de soluções baseadas na natureza à infraestrutura urbana, como telhados verdes, corredores vegetados e sistemas sustentáveis de drenagem, permanece como um ideal normativo, ainda não incorporado de maneira sistemática às legislações municipais nem aos projetos de urbanização em curso. A proposta de criação de um Sistema Regional de Áreas Verdes e protegidas, bem como a de um Observatório Metropolitano Ambiental e de consórcios intermunicipais para gestão ambiental integrada, tampouco avançou para etapas de institucionalização ou implementação concreta, o que reforça a percepção de que, apesar da

robustez técnica dos diagnósticos e da coerência estratégica das propostas, o processo de planejamento segue ancorado em um horizonte eminentemente declaratório.

Assim, permanece o desafio de transitar do discurso técnico-normativo à execução concreta de políticas públicas ambientais em escala metropolitana, especialmente no tocante à vegetação urbana, cuja preservação e expansão dependem não apenas de diretrizes bem-intencionadas, mas de arcabouços legais, financiamento contínuo e vontade política capazes de materializar as proposições delineadas no PDUI.

7. CONSIDERAÇÕES FINAIS

A análise temporal do uso e cobertura do solo na Região Metropolitana de Ribeirão Preto (RMRP), entre 1985 e 2024, revelou padrões espaciais e dinâmicas que podem oferecer contribuições para o debate sobre planejamento territorial em escala metropolitana. A expansão da vegetação arbórea, que passou de cerca de 5% para mais de 10% da área em contexto urbano, a intensificação das superfícies impermeáveis em áreas urbanas e a redução gradual das coberturas herbáceas, substituídas por áreas urbanas, sugerem transformações territoriais complexas, que envolvem tanto processos espontâneos quanto interferências institucionais, legais e normativas. A aplicação de técnicas de classificação supervisionada, aliada à validação sistemática dos dados, resultou em uma base geoespacial potencialmente útil para o acompanhamento dessas transformações ao longo do tempo.

Considerando os referenciais científicos e técnicos analisados, os dados produzidos nesta pesquisa podem fornecer subsídios importantes ao planejamento regional integrado, especialmente por sua capacidade de representar, com regularidade temporal e cobertura espacial abrangente, atributos do território que afetam diretamente políticas públicas setoriais. Autores como Carbone *et al.* (2020) argumentam que a provisão de serviços ecossistêmicos, como a oferta de água, depende de remanescentes florestais frequentemente distribuídos de forma desigual entre os municípios, o que reforça a necessidade de diagnósticos territoriais integrados. A literatura também sugere que a vegetação arbórea pode contribuir para a regulação térmica e hídrica urbana (Martins *et al.*, 2021), enquanto a expansão de áreas impermeáveis está associada ao aumento da vulnerabilidade a enchentes e à intensificação de efeitos climáticos urbanos (Mantovani *et al.*, 2024).

No contexto da RMRP, os dados de cobertura arbórea e de impermeabilização urbana obtidos por este estudo poderiam ser utilizados, por exemplo, para identificar zonas com déficit de infraestrutura verde, mapear riscos hidrológicos associados à urbanização acelerada, ou ainda apoiar propostas de corredores ecológicos entre fragmentos remanescentes. Além disso, podem dialogar com instrumentos já existentes, como o Plano de Desenvolvimento Urbano Integrado (PDUI), que embora apresente diretrizes estratégicas consistentes, ainda carece de instrumentos operacionais efetivamente implementados. Cabe destacar, no entanto, que a aplicabilidade dos dados ao planejamento público depende de fatores que vão além de sua qualidade técnica, como a existência de arranjos institucionais intermunicipais, financiamento continuado e articulação entre diferentes escalas de governo. Mesmo diante dessas limitações, o mapeamento da cobertura do solo com foco em vegetação e impermeabilização oferece

evidências úteis ao ordenamento territorial, desde que interpretado de forma crítica e associado a outras dimensões sociais, econômicas e jurídicas.

Nesse sentido, um dos principais desafios para as próximas décadas reside na articulação entre o acesso à informação e a capacidade de transformar esse conhecimento em ações concretas de planejamento e gestão territorial. Mais do que produzir dados tecnicamente qualificados, será fundamental construir instrumentos e arranjos institucionais que possibilitem sua efetiva incorporação nos processos decisórios, com transparência, e essencialmente continuidade. A efetivação do planejamento urbano integrado e seu controle exigem não apenas diagnósticos precisos, mas também mecanismos operacionais e governança articulada entre escalas e setores. O mapeamento da cobertura do solo com foco em vegetação e impermeabilização, como o realizado nesta pesquisa, representa o primeiro passo de uma cadeia de processos, ao fornecer evidências tangíveis para subsidiar políticas públicas comprometidas com a sustentabilidade e a equidade urbana.

8. REFERÊNCIAS

AGO – AUSTRALIAN GEOSPATIAL-INTELLIGENCE ORGANISATION (AUSTRALIA). **What is Geospatial Intelligence?** Disponível em: <<https://www.defence.gov.au/about/who-we-are/organisation-structure/defence-intelligence-group>>. Acesso em: 8 ago. 2023.

AMANI, M.; GHORBANIAN, A.; AHMADI, S. A.; KAKOOEI, M.; MOGHIMI, A.; MIRMAZLOUMI, S. M.; MOGHADDAM, S. H. A.; MAHDAVI, S.; GHAHREMANLOO, M.; PARSIAN, S.; WU, Q.; BRISCO, B. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 13, p. 5326–5350, 2020.

ARAI, K. Maximum Likelihood Classification based on Classified Result of Boundary Mixed Pixels for High Spatial Resolution of Satellite Images. **International Journal of Advanced Computer Science and Applications**, v. 11, n. 9, 2020. Acesso em: 22 nov. 2020.

BARBOSA, C. C. F.; NOVO, E. M. L. M.; MARTINS, V. S. **Introdução ao sensoriamento remoto de sistemas aquáticos: princípios e aplicações**. [s.l.] Instituto Nacional de Pesquisas Espaciais, 2019.

BARGOS, D. C.; MATIAS, L. F. ÁREAS VERDES URBANAS: UM ESTUDO DE REVISÃO E PROPOSTA CONCEITUAL. **Revista da Sociedade Brasileira de Arborização Urbana**, v. 6, n. 3, p. 172, 1 maio 2019. Disponível em: <<https://www.ige.unicamp.br/geoget/acervo/artigos/areas%20verdes%20urbanas%20Danubia.pdf>>.

BEATLEY, T.; NEWMAN, P. Biophilic Cities Are Sustainable, Resilient Cities. **Sustainability**, v. 5, n. 8, p. 3328–3345, 5 ago. 2013.

BROWN, C. F.; BRUMBY, S. P.; GUZDER-WILLIAMS, B.; BIRCH, T.; HYDE, S. B.; MAZZARIELLO, J.; CZERWINSKI, W.; PASQUARELLA, V. J.; HAERTEL, R.; ILYUSHCHENKO, S.; SCHWEHR, K.; WEISSE, M.; STOLLE, F.; HANSON, C.; GUINAN, O.; MOORE, R.; TAIT, A. M. Dynamic World, Near real-time global 10 m land use land cover mapping. **Scientific Data**, v. 9, n. 1, 9 jun. 2022.

CARBONE, A. S.; COUTINHO, S. M. V.; FERNANDES, V.; PHILIPPI, A. Serviços ecossistêmicos no planejamento integrado do território metropolitano: oferta, demanda e

pressões sobre a provisão de água na região metropolitana de Curitiba. **Revista Brasileira de Ciências Ambientais (Online)**, v. 55, n. 3, p. 381–400, 8 set. 2020. Acesso em: 19 set. 2021.

CAVALHEIRO, F.; DEL PICCHIA, P. C. D. ÁREAS VERDES: CONCEITOS, OBJETIVOS E DIRETRIZES PARA O PLANEJAMENTO . *In: Anais do 1º Congresso Brasileiro Sobre Arborização Urbana*. Vitória, ES: SBAU, 1992. p. 29–38.

CHAVEZ, P. S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. **Remote Sensing of Environment**, v. 24, n. 3, p. 459–479, abr. 1988. Acesso em: 9 ago. 2020.

COSTA, M. A.; FAVARÃO, C. B.; TAVARES, S.; BLANCO JUNIOR, C. Do processo de metropolização institucional à implementação do Estatuto da Metrópole: dois balanços, suas expectativas e incertezas EXPECTATIVAS E INCERTEZAS. *In: MARGUTI, B. O.; COSTA, M. A.; FAVARÃO, C. B. Brasil metropolitano em foco: desafios à implementação do Estatuto da Metrópole*. Brasília: Ipea, 2018. p. 511.

CRESSON, R. A Framework for Remote Sensing Images Processing Using Deep Learning Techniques. **IEEE Geoscience and Remote Sensing Letters**, v. 16, n. 1, p. 25–29, 1 jan. 2019. Acesso em: 28 jun. 2023.

DASH, P.; SANDERS, S. L.; PARAJULI, P.; OUYANG, Y. Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data in an Agricultural Watershed. **Remote Sensing**, v. 15, n. 16, p. 4020, 1 jan. 2023. Disponível em: <<https://www.mdpi.com/2072-4292/15/16/4020>>.

DINIZ, C. G. **Três décadas de mudanças na planície costeira brasileira: O status dos manguezais, da aquicultura e salicultura a partir de séries temporais Landsat e técnicas de aprendizado de máquina**. 2021. 92 f. Instituto de Geociências, Universidade Federal do Pará, 2021. Disponível em: <<https://repositorio.ufpa.br/jspui/handle/2011/13504>>. Acesso em: out. 2024.

ELIAS, D. Globalização, agricultura e urbanização no Brasil. **Revista ACTA Geográfica**, p. 13–32, 2013. Acesso em: 6 nov. 2020.

ELIAS, D. **Globalização e agricultura**. [s.l.] EdUSP, 2003.

FELDMAN, S. Entre o regional e o metropolitano: pensamento urbanístico e metrópole no Brasil na década de 1950. **Revista USP**, n. 102, p. 13, 1 ago. 2014. Acesso em: 13 jun. 2022.

FINEGOLD, Y.; ORTMANN, A. **Map Accuracy Assessment and Area Estimation: A Practical Guide**. Rome, Italy: FAO – Organização das Nações Unidas para a Alimentação e a Agricultura, 2016. v. 46

FIPE. **Diagnóstico da Região Metropolitana de Ribeirão Preto (P7)**. São Paulo: Fipe, nov. 2021. Disponível em: <https://rmrp.pdui.sp.gov.br/?page_id=131>. Acesso em: jun. 2024.

FIRKOWSKI, O. L. C. F. METRÓPOLES E REGIÕES METROPOLITANAS NO BRASIL: CONCILIAÇÃO OU DIVÓRCIO? In: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. C. B. DE. **Território metropolitano, políticas municipais : por soluções conjuntas de problemas urbanos no âmbito metropolitano**. Brasília: Ipea, 2013. p. 21.

FORESTI, C.; HERZ, R. **Avaliação e monitoramento ambiental da expansão urbana do setor oeste da área metropolitana de São Paulo: análise através de dados e técnicas de sensoriamento remoto**. 1987. Universidade de São Paulo, São Paulo, 1987.

GAO, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, v. 58, n. 3, p. 257–266, dez. 1996.

GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, p. 18–27, dez. 2017. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425717302900>>.

HANSEN, M. C.; POTAPOV, P. V.; MOORE, R.; HANCHER, M.; TURUBANOVA, S. A.; TYUKAVINA, A.; THAU, D.; STEHMAN, S. V.; GOETZ, S. J.; LOVELAND, T. R.; KOMMAREDDY, A.; EGOROV, A.; CHINI, L.; JUSTICE, C. O.; TOWNSHEND, J. R. G. High-Resolution Global Maps of 21st-Century Forest Cover Change. **Science**, v. 342, n. 6160, p. 850–853, 14 nov. 2013.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features for Image Classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. SMC-3, n. 6, p. 610–621, nov. 1973. Disponível em: <<https://ieeexplore.ieee.org/document/4309314/>>.

HAYASHI, C. Caracterização geográfica, política e socioeconômica da Região Metropolitana de Ribeirão Preto-SP. **Revista Brasileira de Desenvolvimento Regional**, v. 8, n. 3, 23 out. 2020. Acesso em: 4 out. 2023.

HE, D.; WANG, L. Texture Unit, Texture Spectrum, And Texture Analysis. **IEEE Transactions on Geoscience and Remote Sensing**, v. 28, n. 4, p. 509–512, jul. 1990. Acesso em: 5 maio. 2022.

HEAVISIDE, C.; MACINTYRE, H.; VARDOULAKIS, S. The Urban Heat Island: Implications for Health in a Changing Environment. **Current Environmental Health Reports**, v. 4, n. 3, p. 296–305, 10 jul. 2017.

HOURI GHOLAMREZAIE; MAHDI HASANLOU; AMANI, M.; S. MOHAMMAD MIRMAZLOUMI. Automatic Mapping of Burned Areas Using Landsat 8 Time-Series Images in Google Earth Engine: A Case Study from Iran. **Remote Sensing**, v. 14, n. 24, p. 6376–6376, 16 dez. 2022. Acesso em: 26 set. 2023.

HUANG, X.; WENG, C.; LU, Q.; FENG, T.; ZHANG, L. Automatic Labelling and Selection of Training Samples for High-Resolution Remote Sensing Image Classification over Urban Areas. **Remote Sensing**, v. 7, n. 12, p. 16024–16044, 1 dez. 2015. Acesso em: 5 nov. 2019.

HUETE, A. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. **Remote Sensing of Environment**, v. 59, n. 3, p. 440–451, mar. 1997.

HUETE, A. R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v. 25, n. 3, p. 295–309, ago. 1988.

HUETE, A. R.; HUA, G.; QI, J.; CHEHBOUNI, A.; VAN LEEUWEN, W. J. D. Normalization of multidirectional red and NIR reflectances with the SAVI. **Remote Sensing of Environment**, v. 41, n. 2-3, p. 143–154, ago. 1992.

INSTITUTO PÓLIS. **Estatuto da cidade : guia para implementação pelos municípios e cidadãos**. Brasília: Câmara Dos Deputados, Centro De Documentação E Informação, Coordenação De Publicações, 2001.

IRONS, J. R.; MARKHAM, B. L.; NELSON, R.; TOLL, D. L.; WILLIAMS, D. L.; LATTY, R. S.; STAUFFER, M. L. The effects of spatial resolution on the classification of Thematic

Mapper data. **International Journal of Remote Sensing**, v. 6, n. 8, p. 1385–1403, 1 ago. 1985. Acesso em: 17 set. 2023.

JAVED, A.; CHENG, Q.; PENG, H.; ALTAN, O.; LI, Y.; ARA, I.; HUQ, E.; ALI, Y.; SALEEM, N. Review of Spectral Indices for Urban Remote Sensing. **Photogrammetric Engineering & Remote Sensing**, v. 87, n. 7, p. 513–524, 1 jul. 2021.

JUSTINIANO, E. F.; JUNIOR, S.; MELO, B. M.; NASCIMENTO, V.; MORATO, R. G.; FANTIN, M.; PEDRASSOLI, J. C.; MARTINES, M. R.; KAWAKUBO, F. S. Proposal for an index of roads and structures for the mapping of non-vegetated urban surfaces using OSM and Sentinel-2 data. **International Journal of Applied Earth Observation and Geoinformation**, v. 109, p. 102791–102791, 30 abr. 2022. Acesso em: 12 nov. 2024.

KAWAMURA, M.; JAYAMANNA, S.; TSUJIKO, Y. QUANTITATIVE EVALUATION OF URBANIZATION IN DEVELOPING COUNTRIES USING SATELLITE DATA. **Doboku Gakkai Ronbunshu**, v. 1997, n. 580, p. 45–54, 22 nov. 1997. Acesso em: 1 fev. 2024.

LOBODA, C. R.; ANGELIS, B. L. D. de . ÁREAS VERDES PÚBLICAS URBANAS: CONCEITOS, USOS E FUNÇÕES. **AMBIÊNCIA**, v. 1, n. 1, p. 125–139, 2017. Disponível em: <<https://revistas.unicentro.br/index.php/ambiencia/article/view/157>>. Acesso em: 16 jul. 2024.

LOUPPE, G. **Undertanding Random Forest - From theory to pratice**. 2014. 223 f. Universidade de Liege, 2014. Disponível em: <<https://doi.org/10.48550/arXiv.1407.7502>>. Acesso em: 20 jul. 2024.

MANTOVANI, J.; ENNER ALCÂNTARA; PAMPUCH, L. A.; CHEILA FLÁVIA PRAGA BAIÃO; PARK, E.; MARIA SOUZA CUSTÓDIO; LUIZ FELIPPE GOZZO; CASSIANO ANTONIO BORTOLOZO. Assessing flood risks in the Taquari-Antas Basin (Southeast Brazil) during the September 2023 extreme rainfall surge. **Deleted Journal**, v. 1, n. 1, 3 jun. 2024. Disponível em: <<https://www.nature.com/articles/s44304-024-00009-8>>. Acesso em: 31 ago. 2024.

MARTINS, A. P. G.; RIBEIRO, A. P.; FERREIRA, M. L.; MARTINS, M. A. G.; NEGRI, E. M.; SCAPIN, M. A.; OLIVEIRA, A. D.; SAIKI, M.; SALDIVA, P. H. N.; LAFORTEZZA, R. Infraestrutura verde para monitorar e minimizar os impactos da poluição atmosférica. **Estudos**

Avançados, v. 35, p. 31–57, 23 ago. 2021. Disponível em: <<https://www.scielo.br/j/ea/a/jzQHmkK75XRSYNJzw58NLhR/>>. Acesso em: 18 maio. 2022.

MINISTÉRIO DA DEFESA (BRASIL). **Manual de Fundamentos – EB20-MF-10.107 – Inteligência Militar Terrestre**. 2. ed. Brasil: EXÉRCITO BRASILEIRO, 2015.

NGHIYALWA, H. S.; URBAN, M.; BAADE, J.; SMIT, I. P. J.; RAMOELO, A.; MOGONONG, B.; SCHMULLIUS, C. Spatio-Temporal Mixed Pixel Analysis of Savanna Ecosystems: A Review. **Remote Sensing**, v. 13, n. 19, p. 3870–3870, 27 set. 2021. Acesso em: 5 nov. 2024.

NORMAN, B. **Urban Planning for Climate Change**. [s.l.] Taylor & Francis, 2022.

NUCCI, J. C. **Qualidade Ambiental e Adensamento Urbano**. 2. ed. Curitiba, PR: Edição do Autor, 2008.

OLOFSSON, P.; FOODY, G. M.; HEROLD, M.; STEHMAN, S. V.; WOODCOCK, C. E.; WULDER, M. A. Good practices for estimating area and assessing accuracy of land change. **Remote Sensing of Environment**, v. 148, p. 42–57, maio 2014. Disponível em: <https://glad.umd.edu/Potapov/_Library/Oloffson.%20Good%20practices%20accuracy%20assessment.pdf>.

ROY, D. P.; WULDER, M. A.; LOVELAND, T. R.; C.E., W.; ALLEN, R. G.; ANDERSON, M. C.; HELDER, D.; IRONS, J. R.; JOHNSON, D. M.; KENNEDY, R.; SCAMBOS, T. A.; SCHAAF, C. B.; SCHOTT, J. R.; SHENG, Y.; VERMOTE, E. F.; BELWARD, A. S.; BINDSCHADLER, R.; COHEN, W. B.; GAO, F.; HIPPLE, J. D. Landsat-8: Science and product vision for terrestrial global change research. **Remote Sensing of Environment**, v. 145, p. 154–172, abr. 2014.

SAERP. **Secretaria de Água e Esgoto de Ribeirão Preto**. Disponível em: <<https://www.ribeiraopreto.sp.gov.br/portal/saerp/agua>>. Acesso em: set. 2024.

SANTOS JUNIOR, E. R. **Análise dos efeitos da mudança de cobertura da terra sobre a oferta de serviços ecossistêmicos em áreas úmidas periurbanas no estado de São Paulo**. 2023. Universidade de São Paulo, 2023. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/18/18139/tde-12042023-170518/>>. Acesso em: 4 nov. 2024.

SANTOS JUNIOR, E. R.; MONTAÑO, M.; JUSTINIANO, E. F.; MELO, B. M. de ; HIRYE, M. C. de M.; SHIMBO, J.; AZEVEDO, T.; PEDRASSOLI, J. C. **Análise da ocupação urbana em torno de corpos hídricos no Brasil**. Brasil: Mapbiomas, 2022. Disponível em: <<https://mapbiomas.org/nota-t>>. Acesso em: 4 out. 2023.

SÃO PAULO. **Região Metropolitana de Ribeirão Preto (RMRP)**. Disponível em: <https://rmrp.pdui.sp.gov.br/?page_id=127>. Acesso em: set. 2024.

SHAW, G. A.; BURKE, H. K. Spectral Imaging for Remote Sensing. **LINCOLN LABORATORY JOURNAL**, v. 14, n. 1, p. 3–28, 2003. Disponível em: <https://archive.ll.mit.edu/publications/journal/pdf/vol14_no1/14_1remotesensing.pdf>. Acesso em: 30 set. 2024.

SHEN, H.; LI, X.; CHENG, Q.; ZENG, C.; YANG, G.; LI, H.; ZHANG, L. Missing Information Reconstruction of Remote Sensing Data: A Technical Review. **IEEE Geoscience and Remote Sensing Magazine**, v. 3, n. 3, p. 61–85, 1 set. 2015. Disponível em: <<https://ieeexplore.ieee.org/document/7284768>>. Acesso em: 10 mar. 2021.

SILVEIRA, D. T.; CÓRDOVA, F. P. UNIDADE 2 – A PESQUISA CIENTÍFICA. In: GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de Pesquisa**. [s.l.] PLAGEDER, 2009.

SONG, X.-P.; HANSEN, M. C.; STEHMAN, S. V.; POTAPOV, P. V.; TYUKAVINA, A.; VERMOTE, E. F.; TOWNSHEND, J. R. Global land change from 1982 to 2016. **Nature**, v. 560, n. 7720, p. 639–643, ago. 2018. Disponível em: <<https://www.nature.com/articles/s41586-018-0411-9>>.

SOUZA, C. M.; Z. SHIMBO, J.; ROSA, M. R.; PARENTE, L. L.; A. ALENCAR, A.; RUDORFF, B. F. T.; HASENACK, H.; MATSUMOTO, M.; G. FERREIRA, L.; SOUZA-FILHO, P. W. M.; DE OLIVEIRA, S. W.; ROCHA, W. F.; FONSECA, A. V.; MARQUES, C. B.; DINIZ, C. G.; COSTA, D.; MONTEIRO, D.; ROSA, E. R.; VÉLEZ-MARTIN, E.; WEBER, E. J. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. **Remote Sensing**, v. 12, n. 17, p. 2735, 1 jan. 2020. Disponível em: <<https://www.mdpi.com/2072-4292/12/17/2735>>.

TUCKER, C. J. Radiometric resolution for monitoring vegetation How many bits are needed? **International Journal of Remote Sensing**, v. 1, n. 3, p. 241–254, 1 jul. 1980. Acesso em: 19 jul. 2025.

TUIA, D.; VOLPI, M.; COPA, L.; KANEVSKI, M.; MUNOZ-MARI, J. A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. **IEEE Journal of Selected Topics in Signal Processing**, v. 5, n. 3, p. 606–617, jun. 2011.

UNITED NATIONS HUMAN SETTLEMENTS PROGRAMME (UN-HABITAT). **World Cities Report 2022**. [s.l.] United Nations Human Settlements Programme (UN-Habitat), 2022.

VELASTEGUI-MONTOYA, A.; MONTALVÁN-BURBANO, N.; CARRIÓN-MERO, P.; RIVERA-TORRES, H.; SADECK, L.; ADAMI, M. Google Earth Engine: A Global Analysis and Future Trends. **Remote Sensing**, v. 15, n. 14, p. 3675, 1 jan. 2023. Disponível em: <<https://www.mdpi.com/2072-4292/15/14/3675>>. Acesso em: 14 set. 2023.

WANG, Y.; LI, Z.; ZENG, C.; XIA, G.-S.; SHEN, H. An Urban Water Extraction Method Combining Deep Learning and Google Earth Engine. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 13, p. 768–781, 2020.

XU, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. **International Journal of Remote Sensing**, v. 27, n. 14, p. 3025–3033, 20 jul. 2006. Disponível em: <<http://www.aari.ru/docs/pub/060804/xuh06.pdf>>.

YAO, J.; WU, J.; XIAO, C.; ZHANG, Z.; LI, J. The Classification Method Study of Crops Remote Sensing with Deep Learning, Machine Learning, and Google Earth Engine. **Remote Sensing**, v. 14, n. 12, p. 2758, 8 jun. 2022. Acesso em: 26 jun. 2022.

YOSIO EDEMIR SHIMABUKURO; FLÁVIO JORGE PONZONI. **Mistura espectral**. [s.l.] Oficina de Textos, 2017.

ZHA, Y.; GAO, J.; NI, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. **International Journal of Remote Sensing**, v. 24, n. 3, p. 583–594, jan. 2003.

ZHANG, H.; GORELICK, S. M.; ZIMBA, P. V. Extracting Impervious Surface from Aerial Imagery Using Semi-Automatic Sampling and Spectral Stability. **Remote Sensing**, v. 12, n. 3, p. 506, 4 fev. 2020. Acesso em: 14 out. 2020.

ZHAO, Q.; YU, L.; LI, X.; PENG, D.; ZHANG, Y.; GONG, P. Progress and Trends in the Application of Google Earth and Google Earth Engine. **Remote Sensing**, v. 13, n. 18, p. 3778, 21 set. 2021.

ZHONG, T.; LÜ, G.; ZHONG, X.; TANG, H.; YE, Y. Measuring Human-Scale Living Convenience through Multi-Sourced Urban Data and a Geodesign Approach: Buildings as Analytical Units. **Sustainability**, v. 12, n. 11, p. 4712, 9 jun. 2020.

ZHOU, D.; XIAO, J.; BONAFONI, S.; BERGER, C.; DEILAMI, K.; ZHOU, Y.; FROLKING, S.; YAO, R.; QIAO, Z.; SOBRINO, J. Satellite Remote Sensing of Surface Urban Heat Islands: Progress, Challenges, and Perspectives. **Remote Sensing**, v. 11, n. 1, p. 48, 29 dez. 2018.

ZHU, X. X.; TUIA, D.; MOU, L.; XIA, G.-S.; ZHANG, L.; XU, F.; FRAUNDORFER, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. **IEEE Geoscience and Remote Sensing Magazine**, v. 5, n. 4, p. 8–36, dez. 2017.

APÊNDICE – FLUXO DE CLASSIFICAÇÃO

O fluxo de trabalho da pesquisa foi desenvolvido de forma modular e automatizada, com o objetivo de realizar a classificação temporal do uso e cobertura da terra na Região Metropolitana de Ribeirão Preto (RMRP), entre os anos de 1985 e 2024. As principais etapas estão organizadas na Tabela A1.1, cada uma associada a um script ou notebook específico que compõe o pipeline de processamento.

Tabela A.1: Estruturação do fluxo de trabalho

Algoritmo	Etapa do Pipeline	Descrição Resumida
A1 – Amostragem Automática	GEE - Geração de Amostras	Cria 30.000 pontos aleatórios com dados espectrais a partir do MapBiomass/DW.
A2 – Filtro De Amostras	Colab - Seleção Automática das amostras de Classificação	Processa e estrutura amostras para posterior uso classificação com base na série de índices espectrais utilizados. São removidos <i>outliers</i> e amostras com comportamento divergente ao da classe
A3 – Classificação	GEE - Classificação Supervisionada	Treina/aplica Random Forest com mosaicos e amostras validadas.
A4 – Filtros De Classificação	GEE - Pós-processamento Temporal	Suaviza ruídos temporais e espaciais aplicando filtros à série histórica classificada.
A5 – Aplicativo de Validação Visual	Python Local - Validação Visual das Amostras de Controle	Interface para rotulagem manual com base em mosaicos Landsat (RGB, NDVI, MNDWI).

Os scripts utilizados encontram-se disponibilizados de forma publica em repositório do GitHub³ em conjunto com arquivos auxiliares. Aqueles destinados ao ambiente do Google Earth Engine encontram-se no formato JavaScript (.js), enquanto os scripts auxiliares e complementares desenvolvidos em ambiente local estão estruturados em Python, nos formatos .py e .ipynb. Em complemento, os mapas resultantes do processo podem ser acessados via Google Earth Engine Apps⁴.

³ Repositório disponibilizado em: <https://github.com/Brenomalheiros/Mestrado>

⁴ Mapas disponibilizados em: <https://brenomalheiros-ufscar.projects.earthengine.app/view/classificacao-rmrp-1985-2024>

ALGORÍTIMO A1 – Amostragem Automática

```
// =====
// SCRIPT: 1001_amostragem_auto
// ETAPA: Geração de Amostras
// OBJETIVO: Gerar amostras balanceadas para treinamento do classificador Random
// Forest,
// com base nos mapas de referência do MapBiomas e Dynamic World
// =====

// 1. Define geometria e pontos fixos
var geometry = ee.FeatureCollection("projects/brenomalheiros-
ufscar/assets/LimiteMunicipal_2021")
  .filter(ee.Filter.eq('RM','RM de Ribeirão Preto'))
  .geometry()
  .dissolve()

Map.centerObject(geometry,10)
Map.addLayer(geometry,{},'RMRP')

var seed = 20
var nSamples = 30000
var version = 7

var pts = ee.FeatureCollection.randomPoints(geometry, nSamples*0.85,
seed).set({'origem':'rmp'})
Map.addLayer(pts, {}, 'Pontos Base')

// 2. Define anos e bandas
var years = [2015,2016,2017,2018,2019,2020,2021,2022,2023]
var cloud_cover_value = 10

//var waterLines =
ee.FeatureCollection('users/spatialeanalytics/OSM_water_Brazil/water_lines_intersect_br
asil')
var waterPolygons =
ee.FeatureCollection('users/spatialeanalytics/OSM_water_Brazil/water_polygons_intersec
t_brazil')
  .filterBounds(geometry)

// Map.addLayer(waterLines,{},'water_lines')
// Map.addLayer(waterPolygons,{},'ater_polygons')

var urban_ibge = ee.FeatureCollection("projects/brenomalheiros-
ufscar/assets/AreasUrbanizadas2019_Brasil")
  .filter(ee.Filter.neq('Densidade','Loteamento vazio'))
```

```

        .filterBounds(geometry)

Map.addLayer(urban_ibge, {}, 'urban_ibge')

var pts_wt = ee.FeatureCollection.randomPoints(waterPolygons, nSamples*0.05,
seed).set({'origem':'osm_water'})
Map.addLayer(pts_wt, {}, 'pts_wt')
var pts_ua = ee.FeatureCollection.randomPoints(urban_ibge, nSamples*0.1,
seed).set({'origem':'ibge_ua2019'})
Map.addLayer(pts_ua, {}, 'pts_ua')

pts = pts.merge(pts_wt).merge(pts_ua)

// 3. Funções espectrais (como antes)
var maskcloud = function(image) {
  var qa = image.select('QA_PIXEL')
  return image.updateMask(qa.bitwiseAnd(1 << 3).eq(0))
}
var applySaturationMask = function(image) {
  var sat = image.select('QA_RADSAT')
  // Remover qualquer pixel onde qualquer banda (1–7) esteja saturada (bits 0–6)
  var mask = sat.bitwiseAnd(127).eq(0) // 127 = 01111111
  return image.updateMask(mask)
}

var getNDVI = function(image) {
  var ndvi = image.expression(
    '(nir - red) / (nir + red)',
    {
      'nir': image.select('nir'),
      'red': image.select('red')
    }
  ).rename('ndvi');
  return image.addBands(ndvi);
};

var getMNDWI = function(image) {
  var mndwi = image.expression(
    '(green - swir1) / (green + swir1)',
    {
      'green': image.select('green'),
      'swir1': image.select('swir1')
    }
  ).rename('mndwi');
  return image.addBands(mndwi);
};

```

```

var getNDBI = function(image) {
  var ndbi = image.expression(
    '(swir1 - nir) / (swir1 + nir)',
    {
      'swir1': image.select('swir1'),
      'nir': image.select('nir')
    }
  ).rename('ndbi');
  return image.addBands(ndbi);
};

var getBSI = function(image) {
  var bsi = image.expression(
    '((swir1 + red)-(nir + blue))/((swir1 + red)+(nir + blue))',
    {'swir1': image.select('swir1'), 'red': image.select('red'), 'nir': image.select('nir'), 'blue':
image.select('blue')}
  ).rename('bsi')
  return image.addBands(bsi)
}

var getSMA = function(image){
  var img = image.select('blue','green','red','nir','swir1','swir2')
  var substrate = [0.211200, 0.317455, 0.426777, 0.525177, 0.623311, 0.570544]
  var veg      = [0.093026, 0.086723, 0.049961, 0.611243, 0.219628, 0.079601]
  var dark     = [0.081085, 0.044215, 0.025971, 0.017209, 0.004792, 0.002684]
  return image.addBands(img.unmix([substrate, veg, dark], true, true).rename(['sub', 'veg',
'dark']))
}

// Normalização com base em p5–p95 usando unitScale
var normalizeByPercentile = function(image) {
  var bands = ['blue','green','red','nir','swir1','swir2']

  var percentiles = image.select(bands).reduceRegion({
    reducer: ee.Reducer.percentile([5, 95]),
    geometry: geometry,
    scale: 30,
    bestEffort: true,
  })

  // Cria bandas normalizadas usando unitScale(p5, p95)
  var normalized = ee.ImageCollection(bands.map(function(b) {
    var min = ee.Number(percentiles.get(b + '_p5'))
    var max = ee.Number(percentiles.get(b + '_p95'))
    return image.select(b).unitScale(min, max).rename(b)
  })).toBands().rename(bands.map(function(b) { return b })))

```

```

    return normalized
  }

```

```

// 4. Loop por ano: extrair dados e juntar
var enriched_pts = pts

```

```

var mb = ee.Image('projects/mapbiomas-
public/assets/brazil/lulc/collection9/mapbiomas_collection90_integration_v1')
var dw = ee.ImageCollection('GOOGLE/DYNAMICWORLD/V1').filterBounds(geometry)

```

```

years.forEach(function(year){
  var l5 = ee.ImageCollection('LANDSAT/LT05/C02/T1_TOA')
    .filterBounds(geometry)
    .filterDate(year + '-01-01', year + '-12-31')
    .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
    .filter(ee.Filter.neq('CLOUD_COVER', -1))

  .select(['B1','B2','B3','B4','B5','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

```

```

  var l7 = ee.ImageCollection('LANDSAT/LE07/C02/T1_TOA')
    .filterBounds(geometry)
    .filterDate(year + '-01-01', year + '-12-31')
    .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
    .filter(ee.Filter.neq('CLOUD_COVER', -1))

  .select(['B1','B2','B3','B4','B5','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

```

```

  var l8 = ee.ImageCollection('LANDSAT/LC08/C02/T2_TOA')
    .filterBounds(geometry)
    .filterDate(year + '-01-01', year + '-12-31')
    .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
    .filter(ee.Filter.neq('CLOUD_COVER', -1))

  .select(['B2','B3','B4','B5','B6','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

```

```

  var l9 = ee.ImageCollection("LANDSAT/LC09/C02/T1_TOA")
    .filterBounds(geometry)
    .filterDate(year + '-01-01', year + '-12-31')
    .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))

```

```

    .filter(ee.Filter.neq('CLOUD_COVER', -1))
    .select(['B2','B3','B4','B5','B6','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

```

```

var landsat = l9.merge(l8).merge(l7).merge(l5).sort('CLOUD_COVER',true).limit(30)
print(year,landsat)

```

```

landsat = landsat
    .map(maskcloud)
    .map(applySaturationMask)
    .map(getNDVI)
    .map(getMNDWI)
    .map(getNDBI)
    .map(getBSI)
    .map(getSMA)
    .map(function(img) {
        // Normaliza apenas bandas espectrais, mantendo índices intactos
        var normalized =
normalizeByPercentile(img.select(['blue','green','red','nir','swir1','swir2']))
        return img.select(['ndvi','mndwi','ndbi','bsi','sub','veg','dark'])
            .addBands(normalized)
    })

```

```

var mosaic = landsat.median().clip(geometry)
var stdImage = landsat.reduce(ee.Reducer.stdDev())

```

```

// Renomeia as bandas de desvio padrão para manter clareza
var stdBands = stdImage.bandNames().map(function(b){
    return ee.String(b).replace('_stdDev', '_std')
})
stdImage = stdImage.rename(stdBands)
print(stdImage)

```

```

Map.addLayer(mosaic,{},'mosaic_' + year, false)
Map.addLayer(stdImage,{},'std_' + year, false)

```

```

mosaic = mosaic.addBands(stdImage)

```

```

print('mosaic_final', mosaic)

```

```

var mapbiomas = mb.select('classification_' + year).rename('mb_' + year)

```

```

var dw_median = dw.filterDate(year + '-01-01', year + '-12-31')
    .select('label')
    .reduce(ee.Reducer.median())

```

```

        .rename('dw_' + year)

// Renomeia bandas com sufixo do ano
var mosaicBands = mosaic.bandNames()
print('mosaicBands',mosaicBands)

var renamedBands = mosaicBands.map(function(b) {
    return ee.String(b).cat('_').cat(ee.Number(year).format())
})
print('renamedBands',renamedBands)

var image = mosaic
    .select(mosaicBands,renamedBands)
    .addBands(mapbiomas)
    .addBands(dw_median)

print('image',image.bandNames())

var sampled = image.reduceRegions({
    collection: enriched_pts,
    reducer: ee.Reducer.first(),
    scale: 30
})
//print('sampled',sampled.propertyNames())

// Une atributos no ponto base
enriched_pts = sampled
})

// 5. Exporta asset final com todas as colunas multitemporais
Export.table.toAsset({
    collection: enriched_pts,
    description: 'ASSET-Samples_2015-2023_v' + version,
    assetId: 'projects/brenomalheiros-ufscar/assets/samples_v2/Samples_2015-2023_v' +
version
})

Export.table.toDrive({
    collection: enriched_pts,
    folder:'GEE-Samples_RMRP-v'+ version,
    description: 'Samples_2015-2023_v' + version,
    fileFormat: 'CSV'
})

```

ALGORÍTIMO A2 – Filtro De Amostras

```
# =====
# SCRIPT: filtro_amstras
# ETAPA: Seleção e Análise de Amostras de Classificação
# OBJETIVO: Processar, estruturar e analisar amostras extraídas a partir de mapas de
referência (MapBiomass e Dynamic World), convertendo geometrias, conciliando classes
temáticas, filtrando outliers e aplicando critérios estatísticos para validação espectral por
classe e ano.

# =====
from google.colab import drive
import geopandas as gpd
from shapely.geometry import Polygon
import folium
import matplotlib.pyplot as plt

import geopandas as gpd
import os
import pandas as pd
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', 100) # mostra todas as linhas

# Caminho para a pasta com os arquivos .shp/.dbf/.shx/.prj
version = 7
folder = f'/content/drive/MyDrive/GEE-Samples_RMRP-v{version}'
filename = f'Samples_2015-2023_v{version}.csv'

# Lê o shapefile
df = pd.read_csv(os.path.join(folder, filename))
df = df.rename(columns={'system:index': 'feature_id'})
# Verifica as primeiras linhas
df.head()

import pandas as pd
import geopandas as gpd
from shapely.geometry import shape
import json

def add_geometry_column(df: pd.DataFrame, geo_col: str = '.geo') -> gpd.GeoDataFrame:
    """
    Converte a coluna .geo em geometria shapely e retorna um GeoDataFrame.

    Parâmetros:
    df (pd.DataFrame): DataFrame contendo a coluna .geo.
    geo_col (str): Nome da coluna com os dados GeoJSON (default: '.geo').
    """
```

Retorna:

gpd.GeoDataFrame: GeoDataFrame com coluna 'geometry' convertida.
 """

if geo_col not in df.columns:

raise ValueError(f"Coluna '{geo_col}' não encontrada no DataFrame.")

def parse_geometry(x):

try:

if isinstance(x, str):

x = json.loads(x)

return shape(x['geometry']) if 'geometry' in x else shape(x)

except Exception as e:

print(f"Erro ao converter geometria: {e}")

return None

df = df.copy()

df['geometry'] = df[geo_col].apply(parse_geometry)

Remove entradas inválidas

df = df[df['geometry'].notnull()].reset_index(drop=True)

Retorna como GeoDataFrame

gdf = gpd.GeoDataFrame(df, geometry='geometry', crs='EPSG:4326')

return gdf

gdf = add_geometry_column(df)

print(gdf.geometry.geom_type.value_counts())

gdf.plot()

Verifica colunas por ano

[col for col in gdf.columns if '_2016' in col]

Dicionário de conciliação MB → classe temática

mapa_mb = {

1: [3, 4, 5, 6, 49, 9], # Árboreo/florestal

2: [24], # Impermeabilizado

3: [12, 15, 21], # Vegetação herbácea

4: [26, 33, 31], # Água

5: [19, 39, 20, 40, 62, 41, 36, 46,
47, 35, 48] # Agricultura

}

Dicionário de conciliação DW → classe temática

mapa_dw = {

1: [1], # Árboreo/florestal

2: [6], # Impermeabilizado

```

3: [2],    # Vegetação herbácea
4: [0],    # Água
5: [4]     # Agricultura
}

anos = [2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023]

def inverter_mapa(mapa):
    return {v: k for k, vs in mapa.items() for v in vs}

inv_mb = inverter_mapa(mapa_mb)
inv_dw = inverter_mapa(mapa_dw)

for ano in anos:
    gdf[f'mbc_{ano}'] = gdf[f'mb_{ano}'].map(inv_mb).fillna(6).astype(int)
    gdf[f'dwc_{ano}'] = gdf[f'dw_{ano}'].map(inv_dw).fillna(6).astype(int)

display(gdf[[f'mb_{ano}' for ano in anos] + [f'mbc_{ano}' for ano in anos]].head())
display(gdf[[f'dw_{ano}' for ano in anos] + [f'dwc_{ano}' for ano in anos]].head())

"""### Conciliação de Classes

"""

import matplotlib.pyplot as plt

# Lista de anos analisados
#anos = [2015, 2016, 2017, 2018, 2019, 2020, 2021, 2023]

# Nome das classes (pode ajustar se quiser usar texto)
nomes_classes = {
    1: "Árboreo/florestal",
    2: "Impermeabilizado",
    3: "Vegetação herbácea",
    4: "Água",
    5: "Agricultura",
    6: "Outros"
}

bandas_base = [
    'blue', 'green', 'red', 'nir', 'swir1', 'swir2',
    'ndvi', 'mndwi', 'ndbi', 'bsi', 'sub', 'veg', 'dark',
    'blue_std', 'green_std', 'red_std', 'nir_std', 'swir1_std', 'swir2_std',
    'ndvi_std', 'mndwi_std', 'ndbi_std', 'bsi_std', 'sub_std', 'veg_std', 'dark_std'
]

```

```

# Define subplots (2 linhas, n colunas)
fig, axs = plt.subplots(len(anos), 2, figsize=(12, len(anos)*3))
fig.suptitle('Distribuição por classe (mbc vs dwc)', fontsize=16)

for i, ano in enumerate(anos):
    # Contagens para mbc e dwc
    contagem_mbc = gdf[f'mbc_{ano}'].value_counts().sort_index()
    contagem_dwc = gdf[f'dwc_{ano}'].value_counts().sort_index()

    # Prepara rótulos e valores
    labels_mbc = [nomes_classes.get(k, str(k)) for k in contagem_mbc.index]
    labels_dwc = [nomes_classes.get(k, str(k)) for k in contagem_dwc.index]

    # Gráfico para mbc
    axs[i, 0].pie(contagem_mbc, labels=labels_mbc, autopct='%1.1f%%', startangle=90)
    axs[i, 0].set_title(f'mbc_{ano}')

    # Gráfico para dwc
    axs[i, 1].pie(contagem_dwc, labels=labels_dwc, autopct='%1.1f%%', startangle=90)
    axs[i, 1].set_title(f'dwc_{ano}')

plt.tight_layout(rect=[0, 0, 1, 0.97])
plt.show()

import pandas as pd

# Inicializa dicionários para armazenar contagens
dados_mbc = []
dados_dwc = []

for ano in anos:
    # Contagens por classe (valor numérico)
    contagem_mbc = gdf[f'mbc_{ano}'].value_counts().sort_index()
    contagem_dwc = gdf[f'dwc_{ano}'].value_counts().sort_index()

    for classe in contagem_mbc.index:
        nome_classe = nomes_classes.get(classe, f'Classe {classe}')
        dados_mbc.append({'Ano': ano, 'Classe': nome_classe, 'Total': contagem_mbc[classe]})

    for classe in contagem_dwc.index:
        nome_classe = nomes_classes.get(classe, f'Classe {classe}')
        dados_dwc.append({'Ano': ano, 'Classe': nome_classe, 'Total': contagem_dwc[classe]})

# Converte para DataFrames

```

```

df_mbc      =      pd.DataFrame(dados_mbc).pivot(index='Ano',      columns='Classe',
values='Total').fillna(0).astype(int)
df_dwc      =      pd.DataFrame(dados_dwc).pivot(index='Ano',      columns='Classe',
values='Total').fillna(0).astype(int)

# Exibe
print("Quantidade de amostras por classe (MBC):")
display(df_mbc)

print("Quantidade de amostras por classe (DWC):")
display(df_dwc)

"""### Filtro de tendencia de Classe"""

import numpy as np

for ano in anos:
    col_mbc = f'mbc_{ano}'
    col_dwc = f'dwc_{ano}'
    col_valid = f'v_y_{ano}'
    col_cc = f'cc_{ano}' # ← nova coluna de classe conciliada

    # Etapa 1: validação por conciliação de classe
    conciliadas = gdf[col_mbc] == gdf[col_dwc]
    gdf[col_valid] = conciliadas.astype(int)

    # Nova coluna: classe conciliada ou -1
    gdf[col_cc] = np.where(conciliadas, gdf[col_mbc], -1)

    # Etapa 2: refinamento estatístico por classe
    for classe in gdf[col_cc].unique():
        if pd.isna(classe) or classe == -1: continue # ignora valores inválidos

        # Filtra amostras válidas daquela classe no ano atual
        filtro = (gdf[col_valid] == 1) & (gdf[col_cc] == classe)
        subset = gdf[filtro]

        if subset.empty: continue # pula se não houver dados

        # Bandas daquele ano
        bandas_ano = [f'{b}_{ano}' for b in bandas_base]

        # Média e desvio por banda para essa classe
        medias = subset[bandas_ano].mean()
        desvios = subset[bandas_ano].std()
        print('\n')
        print(f'Ano:{ano} \nClasse:{nomes_classes[classe]}')

```

```

print(f'Média:\n{medias}\n')
print(f'stDesv.:\n{desvios}\n')

# Aplica desvio diferente se for classe 4 (Água)
fator_desvio = 1 if classe == 4 else 1.0

# Máscara: todas bandas dentro da faixa [média ± desvio]
dentro_faixa = ((gdf.loc[filtro, bandas_ano] >= (medias - fator_desvio * desvios)) &
                (gdf.loc[filtro, bandas_ano] <= (medias + fator_desvio * desvios))).all(axis=1)

# Atualiza as amostras que não passam o refinamento como inválidas
idx_invalidas = dentro_faixa[~dentro_faixa].index
gdf.loc[idx_invalidas, col_valid] = 0

gdf

# Dicionário para armazenar contagens
contagem_agua = {}

for ano in anos:
    col_cc = f'cc_{ano}'
    col_valid = f'v_y_{ano}'

    total = ((gdf[col_cc] == 4) & (gdf[col_valid] == 1)).sum()
    contagem_agua[ano] = total

# Converte para DataFrame
df_agua = pd.DataFrame.from_dict(contagem_agua, orient='index', columns=['Total Água'])
df_agua.index.name = 'Ano'

# Exibe
display(df_agua)

import pandas as pd
import matplotlib.pyplot as plt

# Preparação dos dados
dados_abs = []
dados_pct = []

for ano in anos:
    col_class = f'mbc_{ano}'
    col_valido = f'v_y_{ano}'

```

```

df_validas = gdf[gdf[col_valido] == 1]
total_ano = len(df_validas)

contagens = df_validas[col_class].value_counts().sort_index()
proporcoes = df_validas[col_class].value_counts(normalize=True).sort_index()

for classe in contagens.index:
    dados_abs.append({'ano': ano, 'classe': classe, 'total': contagens[classe]})
    dados_pct.append({'ano': ano, 'classe': classe, 'proporcao': proporcoes[classe]})

# Cria DataFrames pivotados
df_abs = pd.DataFrame(dados_abs)
df_pct = pd.DataFrame(dados_pct)

pivot_abs = df_abs.pivot(index='ano', columns='classe', values='total').fillna(0)
pivot_pct = df_pct.pivot(index='ano', columns='classe', values='proporcao').fillna(0)

# Ordena colunas e renomeia para nomes legíveis
pivot_abs = pivot_abs[sorted(pivot_abs.columns)]
pivot_pct = pivot_pct[sorted(pivot_pct.columns)]
pivot_abs.columns = [nomes_classes.get(c, str(c)) for c in pivot_abs.columns]
pivot_pct.columns = [nomes_classes.get(c, str(c)) for c in pivot_pct.columns]

# Plotar os dois gráficos lado a lado
fig, axs = plt.subplots(1, 2, figsize=(16, 6))

# Gráfico de totais absolutos
pivot_abs.plot(kind='bar', stacked=True, ax=axs[0], colormap='tab20c')
axs[0].set_title('Total de amostras válidas por classe e ano')
axs[0].set_ylabel('Quantidade de amostras')
axs[0].set_xlabel('Ano')
axs[0].grid(axis='y', linestyle='--', alpha=0.5)
axs[0].legend(title='Classe', loc='upper left', bbox_to_anchor=(1, 1))

# Gráfico de proporções
pivot_pct.plot(kind='bar', stacked=True, ax=axs[1], colormap='tab20c')
axs[1].set_title('Proporção (%) de amostras válidas por classe e ano')
axs[1].set_ylabel('Proporção')
axs[1].set_xlabel('Ano')
axs[1].set_ylim(0, 1)
axs[1].grid(axis='y', linestyle='--', alpha=0.5)
axs[1].legend(title='Classe', loc='upper left', bbox_to_anchor=(1, 1))

plt.tight_layout()
plt.show()

import pandas as pd

```

```

import plotly.graph_objects as go

# Coleta de dados
dados_abs = []
dados_pct = []

for ano in anos:
    col_class = f'mbc_{ano}'
    col_valido = f'v_y_{ano}'

    df_validas = gdf[gdf[col_valido] == 1]
    total_ano = len(df_validas)

    contagens = df_validas[col_class].value_counts().sort_index()
    proporcoes = df_validas[col_class].value_counts(normalize=True).sort_index()

    for classe in contagens.index:
        nome = nomes_classes.get(classe, str(classe))
        dados_abs.append({'ano': ano, 'classe': nome, 'total': contagens[classe]})
        dados_pct.append({'ano': ano, 'classe': nome, 'proporcao': proporcoes[classe]})

# Criação dos DataFrames
df_abs = pd.DataFrame(dados_abs)
df_pct = pd.DataFrame(dados_pct)

# Gráfico de valores absolutos (Plotly)
fig_abs = go.Figure()
for classe in df_abs['classe'].unique():
    dados = df_abs[df_abs['classe'] == classe]
    fig_abs.add_trace(go.Bar(
        x=dados['ano'],
        y=dados['total'],
        name=classe
    ))

fig_abs.update_layout(
    title='Total de amostras válidas por classe e ano (v_y_ano = 1)',
    barmode='stack',
    xaxis_title='Ano',
    yaxis_title='Quantidade de amostras',
    legend_title='Classe',
    template='plotly_white'
)
fig_abs.show()

# Gráfico de proporções (Plotly)
fig_pct = go.Figure()

```

```

for classe in df_pct['classe'].unique():
    dados = df_pct[df_pct['classe'] == classe]
    fig_pct.add_trace(go.Bar(
        x=dados['ano'],
        y=dados['proporcao'],
        name=classe
    ))

fig_pct.update_layout(
    title='Proporção de amostras válidas por classe e ano (v_y_ano = 1)',
    barmode='stack',
    xaxis_title='Ano',
    yaxis_title='Proporção',
    yaxis=dict(tickformat='.0%'),
    legend_title='Classe',
    template='plotly_white'
)
fig_pct.show()

# Constrói a matriz com anos nas linhas, classes nas colunas
tabela_abs = pd.pivot_table(
    df_abs,
    index='ano',
    columns='classe',
    values='total',
    aggfunc='sum',
    fill_value=0
)

# Mostra a tabela
display(tabela_abs)

import matplotlib.pyplot as plt

# Define a linha da amostra (ex: índice 100)
linha = gdf.iloc[100]
display(linha.head(100))

# Cria dicionário com listas dos valores por banda ao longo dos anos
series = {}
for banda in bandas_base:
    valores = [linha[f'{banda}_{ano}']] for ano in anos]
    series[banda] = valores

# Plota o gráfico
plt.figure(figsize=(12, 6))

```

```

for banda, valores in series.items():
    plt.plot(anos, valores, label=banda)

plt.title('Evolução temporal dos índices para a amostra da linha 100')
plt.xlabel('Ano')
plt.ylabel('Valor dos índices')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
plt.grid(True, linestyle='--', alpha=0.6)
plt.tight_layout()
plt.show()

import plotly.express as px
import pandas as pd

# Prepara os dados no formato longo, apenas com amostras válidas por ano
dados_plot = []
class_view = 2
for ano in anos:
    col_valido = f'v_y_{ano}'
    col_mbc = f'mbc_{ano}'
    col_dwc = f'dwc_{ano}'

    # Filtra apenas as amostras válidas naquele ano
    gdf_filtrado = gdf[(gdf[col_valido] == 1) & (gdf[col_mbc] == class_view)]

    for idx, row in gdf_filtrado.iterrows():
        for banda in bandas_base:
            dados_plot.append({
                'ano': ano,
                'banda': banda,
                'valor': row[f'{banda}_{ano}'],
                'amostra': idx,
                'mbc': nomes_classes.get(row[col_mbc], str(row[col_mbc])),
                'dwc': nomes_classes.get(row[col_dwc], str(row[col_dwc]))
            })

# Cria DataFrame longo para plotagem
df_longo = pd.DataFrame(dados_plot)

# Gráfico de dispersão por banda
fig = px.scatter(
    df_longo,
    x='ano',
    y='valor',
    facet_col='banda',
    facet_col_wrap=3,
    color='amostra', # cor por amostra (pode trocar por 'mbc' ou 'dwc' se quiser)

```

```

    hover_data=['amostra', 'mbc', 'dwc'],
    title='Valores por banda para amostras válidas em cada ano',
    template='plotly_white'
)

fig.update_layout(
    height=800,
    showlegend=False,
    yaxis=dict(range=[-1, 1]) # Limita o eixo Y de -1 a 1
)

# Limpa os títulos dos subgráficos
fig.for_each_annotation(lambda a: a.update(text=a.text.split("=")[-1]))
fig.show()

print(len(gdf))
gdf

import geopandas as gpd
import pandas as pd

#gdf = gdf.drop(columns='geometry') # evita erro no melt

# Lista de bandas e colunas a incluir por ano
colunas_base = bandas_base + ['mbc', 'dwc', 'v_y', 'cc']

# Lista que armazenará os novos DataFrames por ano
linhas = []

for ano in anos:
    # Monta dict para renomear as colunas do ano atual
    renomear = {f'{c}_{ano}': c for c in colunas_base}

    # Filtra e renomeia
    df_ano = gdf[[f'{c}_{ano}' for c in colunas_base]].rename(columns=renomear)

    # Adiciona colunas de geometria e ano
    df_ano['ano'] = ano
    df_ano['geometry'] = gdf['geometry']
    df_ano['geo'] = gdf['geo']
    df_ano['feature_id'] = gdf['feature_id']

    linhas.append(df_ano)

# Junta todos os anos num único GeoDataFrame
gdf_long = pd.concat(linhas, ignore_index=True)

```

```
# Verifica
print("Formato final:", gdf_long.shape)
display(gdf_long.head())

import os
# 2. Define caminho de destino
output_folder = folder
os.makedirs(output_folder, exist_ok=True)

# 3. Exporta o shapefile (todos os componentes: .shp, .shx, .dbf, .prj)
shapefile_name = f'Samples_2016-2023-v{version}-filtered'
gdf_long.to_csv(f'{output_folder}/{shapefile_name}.csv', index = False)#, driver='ESRI
Shapefile')
```

ALGORÍTIMO A3 – Classificação

```
// =====
// SCRIPT: 1002-classificacao_multimodelo
// ETAPA: Classificação Supervisionada
// OBJETIVO: Treinar e aplicar modelos Random Forest anuais com mosaicos Landsat e
// amostras de treinamento
// =====

/**** Start of imports. If edited, may not auto-convert in the playground. ****/
var validation_2022 = ee.FeatureCollection("projects/brenomalheiros-
ufscar/assets/validation-1700_v1");
/***** End of imports. If edited, may not auto-convert in the playground. *****/
// 1. Define geometria e pontos fixos
var geometry = ee.FeatureCollection("projects/brenomalheiros-
ufscar/assets/LimiteMunicipal_2021")
  .filter(ee.Filter.eq('RM','RM de Ribeirão Preto'))
  .geometry()
  .dissolve()

Map.addLayer(geometry,{},'RMRP')

var desc = 'Classificação com amostras automatizadas: multiprobabilidade construída com
5 modelos treinados separadamente, sem uso da classe de agricultura'

var output_version = 8
var step = 'classification_raw'
var seed = 20
var year_list =
[1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,
2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,
2015,2016,2017,2018,2019,2020,2021,2022,2023,2024]

var cloud_cover_value = 10

//Amostras
var samples = ee.FeatureCollection('projects/brenomalheiros-
ufscar/assets/samples_v2/Samples_2016-2023-v7-filtered')

var classProperty = 'cc'

var inputBands = [
  'blue', 'green', 'red', 'nir', 'swir1', 'swir2',
  'ndvi', 'mndwi', 'ndbi', 'bsi', 'sub', 'veg', 'dark',
  'blue_std', 'green_std', 'red_std', 'nir_std', 'swir1_std', 'swir2_std',
  'ndvi_std', 'mndwi_std', 'ndbi_std', 'bsi_std', 'sub_std', 'veg_std', 'dark_std']
```

```

]

/
// Treinamento de modelos
//1-arboreo, 2-imperm, 3-herbacea, 4-agua, 5-agricultura, 6-outros
var training = samples.remap([1,2,3,4,5,6],[1,2,3,4,3,6],'cc')
  .filter(ee.Filter.eq('v_y', 1))    // somente amostras válidas
  .filter(ee.Filter.lt('cc', 6));    // descarta classe 6 ("outras")

// Mapear a classe 'cc' para binária para cada classe-alvo
var training_arb = training.remap([1,2,3,4], [1,0,0,0], 'cc')
var training_imper = training.remap([1,2,3,4], [0,1,0,0], 'cc')
var training_herb = training.remap([1,2,3,4], [0,0,1,0], 'cc')
var training_water = training.remap([1,2,3,4], [0,0,0,1], 'cc')
//var training_agri = training.remap([1,2,3,4,5], [0,0,0,0,1], 'cc')

function trainBinaryRF(trainingSet) {
  return ee.Classifier.smileRandomForest({
    numberOfTrees: 200,
    seed: seed
  })
  .setOutputMode('PROBABILITY')
  .train({
    features: trainingSet,
    classProperty: classProperty,
    inputProperties: inputBands
  });
}

var clf_arb = trainBinaryRF(training_arb);
var clf_imper = trainBinaryRF(training_imper);
var clf_herb = trainBinaryRF(training_herb);
var clf_water = trainBinaryRF(training_water);
//var clf_agri = trainBinaryRF(training_agri);

/// 3. Funções espectrais
var maskcloud = function(image) {
  var qa = image.select('QA_PIXEL')
  return image.updateMask(qa.bitwiseAnd(1 << 3).eq(0))
}
var applySaturationMask = function(image) {
  var sat = image.select('QA_RADSAT')
  // Remover qualquer pixel onde qualquer banda (1–7) esteja saturada (bits 0–6)

```

```

    var mask = sat.bitwiseAnd(127).eq(0) // 127 = 01111111
    return image.updateMask(mask)
  }

  var getNDVI = function(image) {
    var ndvi = image.expression(
      '(nir - red) / (nir + red)',
      {
        'nir': image.select('nir'),
        'red': image.select('red')
      }
    ).rename('ndvi');
    return image.addBands(ndvi);
  };

  var getMNDWI = function(image) {
    var mndwi = image.expression(
      '(green - swir1) / (green + swir1)',
      {
        'green': image.select('green'),
        'swir1': image.select('swir1')
      }
    ).rename('mndwi');
    return image.addBands(mndwi);
  };

  var getNDBI = function(image) {
    var ndbi = image.expression(
      '(swir1 - nir) / (swir1 + nir)',
      {
        'swir1': image.select('swir1'),
        'nir': image.select('nir')
      }
    ).rename('ndbi');
    return image.addBands(ndbi);
  };

  var getBSI = function(image) {
    var bsi = image.expression(
      '((swir1 + red)-(nir + blue))/((swir1 + red)+(nir + blue))',
      {'swir1': image.select('swir1'), 'red': image.select('red'), 'nir': image.select('nir'), 'blue':
image.select('blue')}
    ).rename('bsi')
    return image.addBands(bsi)
  }

  var getSMA = function(image){

```

```

var img = image.select('blue','green','red','nir','swir1','swir2')
var substrate = [0.211200, 0.317455, 0.426777, 0.525177, 0.623311, 0.570544]
var veg      = [0.093026, 0.086723, 0.049961, 0.611243, 0.219628, 0.079601]
var dark     = [0.081085, 0.044215, 0.025971, 0.017209, 0.004792, 0.002684]
return image.addBands(img.unmix([substrate, veg, dark], true, true).rename(['sub', 'veg',
'dark']))
}

```

```

// Normalização com base em p5–p95 usando unitScale

```

```

var normalizeByPercentile = function(image) {
  var bands = ['blue','green','red','nir','swir1','swir2']

```

```

  var percentiles = image.select(bands).reduceRegion({
    reducer: ee.Reducer.percentile([5, 95]),
    geometry: geometry,
    scale: 30,
    bestEffort: true,
  })

```

```

  // Cria bandas normalizadas usando unitScale(p5, p95)
  var normalized = ee.ImageCollection(bands.map(function(b) {
    var min = ee.Number(percentiles.get(b + '_p5'))
    var max = ee.Number(percentiles.get(b + '_p95'))
    return image.select(b).unitScale(min, max).rename(b)
  })).toBands().rename(bands.map(function(b) { return b }))

```

```

  return normalized
}

```

```

var mosaic_dataset = ee.ImageCollection([])

```

```

var main_function = year_list.map(function(year){

```

```

  var I5 = ee.ImageCollection('LANDSAT/LT05/C02/T1_TOA')
    .filterBounds(geometry)
    .filterDate(year + '-01-01', year + '-12-31')
    .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
    .filter(ee.Filter.neq('CLOUD_COVER', -1))

```

```

  .select(['B1','B2','B3','B4','B5','B7','QA_PIXEL','QA_RADSAT'],['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

```

```

  var I7 = ee.ImageCollection('LANDSAT/LE07/C02/T1_TOA')
    .filterBounds(geometry)

```

```

.filterDate(year + '-01-01', year + '-12-31')
.filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
.filter(ee.Filter.neq('CLOUD_COVER', -1))

.select(['B1','B2','B3','B4','B5','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

var l8 = ee.ImageCollection('LANDSAT/LC08/C02/T2_TOA')
.filterBounds(geometry)
.filterDate(year + '-01-01', year + '-12-31')
.filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
.filter(ee.Filter.neq('CLOUD_COVER', -1))

.select(['B2','B3','B4','B5','B6','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

var l9 = ee.ImageCollection("LANDSAT/LC09/C02/T1_TOA")
.filterBounds(geometry)
.filterDate(year + '-01-01', year + '-12-31')
.filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
.filter(ee.Filter.neq('CLOUD_COVER', -1))

.select(['B2','B3','B4','B5','B6','B7','QA_PIXEL','QA_RADSAT'], ['blue','green','red','nir','swir1',
'swir2','QA_PIXEL','QA_RADSAT'])

var landsat = l9.merge(l8).merge(l7).merge(l5).sort('CLOUD_COVER',true).limit(30)
print(year,landsat)

landsat = landsat
.map(maskcloud)
.map(applySaturationMask)
.map(getNDVI)
.map(getMNDWI)
.map(getNDBI)
.map(getBSI)
.map(getSMA)
.map(function(img) {
  // Normaliza apenas bandas espectrais, mantendo índices intactos
  var normalized =
normalizeByPercentile(img.select(['blue','green','red','nir','swir1','swir2']))
  return img.select(['ndvi','mndwi','ndbi','bsi','sub','veg','dark'])
    .addBands(normalized)
})

```

```

var mosaic = landsat.median().clip(geometry)
var stdImage = landsat.reduce(ee.Reducer.stdDev())

// Renomeia as bandas de desvio padrão para manter clareza
var stdBands = stdImage.bandNames().map(function(b){
  return ee.String(b).replace('_stdDev', '_std')
})
stdImage = stdImage.rename(stdBands)

//Map.addLayer(mosaic,{},'mosaic_' + year, false)
//Map.addLayer(stdImage,{},'std_' + year, false)

mosaic = mosaic.addBands(stdImage)
//Map.addLayer(mosaic,{min:0,max:1.5, bands:['red','green','blue']], 'mosaic-' + year,
false)

var image = mosaic
//print('image',image.bandNames())

var prob_arb = mosaic.classify(clf_arb) .rename('prob_arb' ) .multiply(100).toInt8()
var prob_imper = mosaic.classify(clf_imper) .rename('prob_imper'
).multiply(100).toInt8()
var prob_herb = mosaic.classify(clf_herb) .rename('prob_herb' ) .multiply(100).toInt8()
var prob_water = mosaic.classify(clf_water) .rename('prob_water'
).multiply(100).toInt8()
//var prob_agri = mosaic.classify(clf_agri) .rename('prob_agri' ) .multiply(100).toInt8()

var classification_prob = prob_arb
  .addBands(prob_imper)
  .addBands(prob_herb)
  .addBands(prob_water)
  //.addBands(prob_agri);

var class_idx = classification_prob
  .toArray()
  .arrayArgmax()
  .arrayFlatten([[ 'class_idx' ]]);

var final_class = class_idx.add(1).rename('classification');

// Adiciona 1 para que os valores fiquem de 1 a 5
var final_class = class_idx.add(1).rename('classification').addBands(classification_prob)
print(final_class)

```

```

//1-arboreo, 2-imperm, 3-herbacea, 4-agua, 5-agricultura, 6-outros

//Map.addLayer(classification,{min:1,max:6,palette:['lime','red','yellow','blue','purple','gray']},'classification_'+year,false)

Map.addLayer(final_class,{min:1,max:5,bands:['classification'],palette:['lime','red','yellow','blue','purple']},'classification_'+year,false)
Map.addLayer(classification_prob, {min: 0, max: 1}, 'probabilidades_' + year, false);

/*
// ==== FEATURE IMPORTANCE ====
print('Feature Importance:', classifier.explain())

var variable_importance = ee.Feature(null,
ee.Dictionary(classifier.explain()).get('importance'));

var chart1 =
ui.Chart.feature.byProperty(variable_importance)
.setChartType('ColumnChart')
.setOptions({
title: 'RF Variable Importance - Method 1',
legend: {position: 'none'},
hAxis: {title: 'Bands'},
vAxis: {minValue:0, title: 'Importance'},
});

print(chart1, 'Relative Importance');

*/

var asset_name = 'class_'+year + '-' + output_version

Export.image.toAsset({
image:final_class.set({'year':year,'step':step,'output_ver':output_version,'description':description}),
description:asset_name,
assetId:'projects/brenomalheiros-ufscar/assets/classification_RMRP/' + asset_name,
region:geometry,
scale:30,
})

return null

```

```
})
```

```
main_function
```

ALGORÍTIMO A4 – Filtros De Classificação

```
// =====
// SCRIPT: 1003_filtro_pos-classificacao
// ETAPA: Pós-processamento Temporal
// OBJETIVO: Aplicar filtros para suavizar ruídos na série temporal classificada
// =====

// =====
// CONFIGURAÇÃO
// =====
var input_version = 8
var output_version = 10

var desc = 'Filtro espacial, com foco em remover ruído de água, e filtro de pixel solitário
para todas as classes e patches para áreas impermeáveis; uso do IRS para reduzir comissão
da classe impermeável'

var step = 'spatial and temporal filter'
var years = ee.List.sequence(1985, 2024).getInfo();
var anosCentrais = years.filter(function(y){ return y > 1985 && y < 2024 });
print(anosCentrais)
var palette = ['gray', '#006400', '#FF0000', '#FFFF00', '#0000FF', '#800080'];
var cloud_cover_value = 10;

// =====
// BASE ORIGINAL
// =====
var raw_col = ee.ImageCollection('projects/brenomalheiros-ufscar/assets/classification_RMRP')
    .filter(ee.Filter.eq('output_ver', input_version))
    .filter(ee.Filter.eq('step', 'classification_raw'))

var geometry = raw_col.geometry().bounds();

var irs_threshold = 500

// add the IRS index to the spatial filter
var irs = ee.ImageCollection('users/efjustiniano/IRS2022/IRS2022_v5_30').sum()
var irsUrb = irs.gte(irs_threshold)

// =====
// CÁLCULO DA SEGUNDA CLASSE (com WHERE)
// =====
var addSecondClassWhere = function(img) {
    // Probabilidades individuais
```

```

var prob_arb = img.select('prob_arb');
var prob_imper = img.select('prob_imper');
var prob_herb = img.select('prob_herb');
var prob_water = img.select('prob_water');
//var prob_agri = img.select('prob_agri');

// Classe predominante (1 a 5)
var main_class = img.select('classification');

// Zera a classe principal: define -999 para ignorar
prob_arb = prob_arb.where(main_class.eq(1), -999);
prob_imper = prob_imper.where(main_class.eq(2), -999);
prob_herb = prob_herb.where(main_class.eq(3), -999);
prob_water = prob_water.where(main_class.eq(4), -999);
//prob_agri = prob_agri.where(main_class.eq(5), -999);

// Inicializa com 1 (arb) e substitui com where()
var second_class = ee.Image.constant(1)
  .where(prob_imper.gt(prob_arb)
    .and(prob_imper.gte(prob_herb))
    .and(prob_imper.gte(prob_water))
    //and(prob_imper.gte(prob_agri))
    , 2)
  .where(prob_herb.gt(prob_arb)
    .and(prob_herb.gte(prob_imper))
    .and(prob_herb.gte(prob_water))
    //and(prob_herb.gte(prob_agri))
    , 3)
  .where(prob_water.gt(prob_arb)
    .and(prob_water.gte(prob_imper))
    .and(prob_water.gte(prob_herb))
    //and(prob_water.gte(prob_agri))
    , 4)
  // .where(prob_agri.gt(prob_arb)
  //   .and(prob_agri.gte(prob_imper))
  //   .and(prob_agri.gte(prob_herb))
  //   //and(prob_agri.gte(prob_water))
  //   , 5)
  .rename('second_class');

return img.addBands(second_class);
};

raw_col = raw_col.map(addSecondClassWhere)

// =====
// 1. MÁSCARA NDWI

```

```
// =====
var getMNDWI = function(image) {
  return image.expression(
    '(green - swir1) / (green + swir1)',
    {'green': image.select('green'), 'swir1': image.select('swir1')}
  ).rename('mndwi');
};

var l8 = ee.ImageCollection('LANDSAT/LC08/C02/T1_TOA')
  .filterBounds(geometry)
  .filterDate('2020-01-01', '2023-12-31')
  .filter(ee.Filter.lt('CLOUD_COVER', cloud_cover_value))
  .select(['B2','B3','B4','B5','B6','B7','QA_PIXEL',
    ['blue','green','red','nir','swir1','swir2','QA_PIXEL']]);

var ndwiMascara = l8.map(getMNDWI).median().gte(0);

// =====
// 2. FILTRO ESPACIAL PERSONALIZADO
// =====
function remove_patches(img,class_in,class_out,nPix) {
  //Map.addLayer(img, {min:0, max:5, palette: palette}, 'Original',false);
  var agri = img.eq(class_in)

  // Remove buracos internos pequenos

  var grupoPequeno = agri.selfMask()
    .connectedPixelCount()
    .reproject({crs:'EPSG:4326', scale:30})
    .lt(nPix)

  //Map.addLayer(grupoPequeno,{min:0,max:1,palette: ['white','black']},'grupoPequeno',
  false);

  img = img.where(grupoPequeno.eq(1), class_out).reproject({crs:'EPSG:4326', scale:30});
  //Map.addLayer(img, {min:0, max:5, palette: palette}, 'Corrigido (classe in -> out)',false);

  return img
}

function fs_neighborhood(img) {
  var kernel = ee.Kernel.square({radius: 1});

  // Classe 4 (água)
  var agua = img.eq(4);
  var vizinhosAgua = agua.neighborhoodToBands(kernel).reduce(ee.Reducer.sum());
}
```

```

var manterAgua = agua.and(vizinhosAgua.gte(4));
img = img.where(agua.and(manterAgua.not()), img.focal_mode());

// Classe 2 (impermeabilizado)
var imp = img.eq(2);
var vizinhosImp = imp.neighborhoodToBands(kernel).reduce(ee.Reducer.sum());
var manterImp = imp.and(vizinhosImp.gte(4));
img = img.where(imp.and(manterImp.not()), img.focal_mode());

// Classe 5 (agricultura): substitui grupos pequenos
img = remove_patches(img,2,3,50)

var filter = img.where(irsUrb.eq(0).and(img.eq(2)),3).selfMask()

return filter.rename('classification');
}

// =====
// 3. APLICA CORREÇÕES NDWI + ESPACIAL
// =====
function aplicarCorrecao(img) {

  var corrigido = img.select('classification')
    .where(img.select('classification').eq(4)
      .and(ndwiMascara.not()),img.select('second_class') ); // NDWI

  return fs_neighborhood(corrigido)
}

var collectionCorrigida = raw_col.map(aplicarCorrecao);

// =====
// 4. DICIONÁRIO DOS CORRIGIDOS
// =====
var dictCorrigido = ee.Dictionary(
  collectionCorrigida.toList(collectionCorrigida.size()).iterate(function(img, acc) {
    img = ee.Image(img);
    var ano = img.getNumber('year').int()
    return ee.Dictionary(acc).set(ano, img);
  }, ee.Dictionary({}))
);

// =====

```

```
// 5. FILTRO TEMPORAL K3 – ANOS CENTRAIS
// =====
print('dictCorrigido',dictCorrigido)
var imagensK3Centrais = anosCentrais.reverse().map(function(ano) {
  var anoNum = ee.Number(ano);
  var prev = ee.Image(dictCorrigido.get(anoNum.subtract(1)));
  var curr = ee.Image(dictCorrigido.get(anoNum));
  var next = ee.Image(dictCorrigido.get(anoNum.add(1)));

  return ee.ImageCollection([prev, curr, next])
    .reduce(ee.Reducer.mode())
    .rename('classification')
    .set('year', anoNum);
});
print('colecacaoK3Centrais',colecacaoK3Centrais)
var colecacaoK3Centrais = ee.ImageCollection.fromImages(imagensK3Centrais).sort('year');

// =====
// 6. DICIONÁRIO DOS ANOS CENTRAIS CORRIGIDOS
// =====
var dictK3 = ee.Dictionary(
  colecacaoK3Centrais.toList(colecacaoK3Centrais.size()).iterate(function(img, acc) {
    img = ee.Image(img);
    var ano = img.getNumber('year').int()
    return ee.Dictionary(acc).set(ano, img);
  }, dictCorrigido) // mantém os extremos ainda sem K3
);

// =====
// 7. FILTRO TEMPORAL K3 – EXTREMOS
// =====
var extremos = [1985, 2024];

var imagensK3Extremos = extremos.map(function(ano) {
  var anoNum = ee.Number(ano);
  var anosViz = ee.List(
    ano === 1985 ? [1985, 1986, 1987] : [2022, 2023, 2024]
  );

  var imgs = anosViz.map(function(a) {
    return ee.Image(dictK3.get(ee.Number(a)));
  });

  return ee.ImageCollection(imgs)
    .reduce(ee.Reducer.mode())

```

```

        .rename('classification')
        .set('year', anoNum);
    });

    var colecaoK3Extremos = ee.ImageCollection.fromImages(imagensK3Extremos);

    // =====
    // 8. UNIÃO FINAL
    // =====
    var colecaoFinalK3 = colecaoK3Centrais.merge(colecaoK3Extremos).sort('year');

    // =====
    // 9. VISUALIZAÇÃO
    // =====
    years.forEach(function(year){
        var raw = raw_col.filter(ee.Filter.eq('year', year)).first();
        var corrigida = collectionCorrigida.filter(ee.Filter.eq('year', year)).first();
        var k3 = colecaoFinalK3.filter(ee.Filter.eq('year', year)).first();

        print(year,raw.bandNames(),corrigida.bandNames(),k3.bandNames())

        //Map.addLayer(raw, {min:0, max:5, palette: palette, bands:['classification']}, 'Raw_' +
        year, false);
        Map.addLayer(corrigida,{min:0, max:5, palette: palette}, 'Corrigido_' + year, false);
        Map.addLayer(k3, {min:0, max:5, palette: palette}, 'K3_' + year, false);

        var asset_name = 'filter_'+year + '-' + output_version

        Export.image.toAsset({
            image:k3.set({'year':year,'step':step,'input_ver':input_version,'output_ver':output_version
            , 'description':desc}),
            description:asset_name,
            assetId:'projects/brenomalheiros-ufscar/assets/classification_RMRP/' + asset_name,
            region:geometry,
            scale:30
        })
    });

```

ALGORÍTIMO A5 – Aplicativo de Validação Visual

```
#=====
# Interface construída em para visualização
# das amostras de validação e processo de rotulagem das mesmas.
# obs: Necessário acesso ao GEE
#=====

import ee
import geemap.foliummap as geemap
import streamlit as st
import pandas as pd
import geopandas as gpd
from shapely.geometry import Point
import os
import time
import shutil

# Funções auxiliares
def timer_debug(label, start_time):
    print(f'[TIMER] {label} levou {time.time() - start_time:.2f}s')

def sanitize_dataframe(df):
    if 'id' in df.columns:
        df['id'] = pd.to_numeric(df['id'], errors='coerce')
        if df['id'].isnull().any():
            raise ValueError('Existem amostras sem ID definido!')
        df['id'] = df['id'].astype(int)
    for col in df.columns:
        if col.startswith('class_id_'):
            df[col] = pd.to_numeric(df[col], errors='coerce')
    return df

def backup_csv(original_path):
    backup_folder = os.path.join(os.path.dirname(original_path), 'backup')
    os.makedirs(backup_folder, exist_ok=True)
    timestamp = time.strftime("%Y%m%d-%H%M%S")
    backup_filename = os.path.join(backup_folder, f"backup_{timestamp}.csv")
    shutil.copy2(original_path, backup_filename)
    print(f'[DEBUG] Backup criado: {backup_filename}')

# Limpeza dos backups antigos (manter os 100 mais recentes com mais de 1h)
backups = sorted([os.path.join(backup_folder, f) for f in os.listdir(backup_folder) if
f.endswith('.csv')], key=os.path.getmtime, reverse=True)
backups_to_check = backups[100:]
now = time.time()
for bkp in backups_to_check:
```

```

    if now - os.path.getmtime(bkp) > 3600:
        os.remove(bkp)
        print(f'[DEBUG] Backup antigo removido: {bkp}')

def create_square(point, distance_m=500):
    start = time.time()
    center = ee.Geometry.Point(point.x, point.y)
    rect = center.buffer(distance_m).bounds()
    timer_debug('Create square', start)
    return rect

def get_mosaic(year):
    start = time.time()
    mosaic = ee.Image(f'projects/brenomalheiros-ufscar/assets/mosaic_RMRP/mosaic_{year}-3')
    timer_debug(f'Load mosaic {year}', start)
    return mosaic

# Inicialização do Streamlit
if 'ee_initialized' not in st.session_state:
    start = time.time()
    try:
        ee.Initialize()
    except Exception:
        ee.Authenticate()
        ee.Initialize()
    st.session_state.ee_initialized = True
    timer_debug('ee.Initialize()', start)

st.set_page_config(layout="wide")

output_folder = 'streamlit_samples_data'
os.makedirs(output_folder, exist_ok=True)

csv_filename = os.path.join(output_folder, 'amostras_validadas.csv')
gpkg_points_filename = os.path.join(output_folder, 'pontos_amostras.gpkg')
anos = [1985, 1990, 2000, 2010, 2020, 2024]

# Carrega pontos
if os.path.exists(gpkg_points_filename):
    gdf_points = gpd.read_file(gpkg_points_filename)
else:
    amostras = ee.FeatureCollection('projects/brenomalheiros-ufscar/assets/raw_val_samples-600-seed_10')
    features = amostras.getInfo()['features']
    data = [{'longitude': f['geometry']['coordinates'][0], 'latitude':
f['geometry']['coordinates'][1]} for f in features]

```

```

gdf_points = gpd.GeoDataFrame(data, geometry=gpd.points_from_xy([d['longitude'] for
d in data], [d['latitude'] for d in data]), crs="EPSG:4326")
gdf_points.to_file(gpkg_points_filename, driver="GPKG")

```

Inicializa CSV

```

if not os.path.exists(csv_filename):
    df_init = pd.DataFrame({'id': range(len(gdf_points))})
    for ano in anos:
        df_init[f'class_id_{ano}'] = pd.NA
        df_init[f'class_nm_{ano}'] = pd.NA
        df_init[f'obs_{ano}'] = pd.NA
    df_init['longitude'] = gdf_points.geometry.x
    df_init['latitude'] = gdf_points.geometry.y
    df_init['geometry'] = gdf_points.geometry
    df_init = sanitize_dataframe(df_init)
    df_init.to_csv(csv_filename, index=False)

```

Carrega CSV

```

df_samples = pd.read_csv(csv_filename, dtype=str)
df_samples = sanitize_dataframe(df_samples)
st.session_state.samples = df_samples.to_dict('records')

```

Interface Streamlit

```

st.markdown("<h4> 🌍 Rotulagem de Amostras Landsat</h4>", unsafe_allow_html=True)
tabs = st.tabs(["🚩 Classificar Amostras", "📄 Visualizar CSV"])

```

```

classe_options = {
    1: " 🌳 Arbórea",
    2: " 🏠 Impermeável",
    3: " 🌿 Herbácea",
    4: " 💧 Água",
    5: " 🚜 Agricultura",
    6: " ? Outros"
}

```

```

total_pontos = len(gdf_points)
if 'current_index' not in st.session_state:
    st.session_state.current_index = 0
if 'current_year_index' not in st.session_state:
    st.session_state.current_year_index = 0

```

Aba 1: Classificar Amostras

```

with tabs[0]:
    with st.sidebar:
        index = st.number_input(' 📍 Ponto', min_value=0, max_value=total_pontos-1,
value=int(st.session_state.current_index), step=1)

```

```

st.session_state.current_index = index

year = st.selectbox('📅 Ano', anos, index=st.session_state.current_year_index)
st.session_state.current_year_index = anos.index(year)

completed_years = []
current_id = int(index)
for ano in anos:
    for s in st.session_state.samples:
        if int(s['id']) == current_id and pd.notna(s.get(f'class_id_{ano}')):
            completed_years.append(ano)

year_status = "☑" if year in completed_years else "☐"
st.markdown(f"<b>📅 Ano atual:</b> {year} {year_status}",
unsafe_allow_html=True)

class_selection = st.selectbox("👉 Classe:", options=list(classe_options.keys()),
format_func=lambda x: f"{x} - {classe_options[x]}")
obs_text = st.text_input("📝 Observações", key="obs_text")
st.markdown("***Sugestões:** `duvida`, `borda`, `sem imagem`")

col_botoes = st.columns(2)
with col_botoes[0]:
    if st.button('⬅️ Voltar'):
        st.session_state.current_index = max(st.session_state.current_index - 1, 0)
        st.session_state.current_year_index = 0
        st.rerun()
with col_botoes[1]:
    save_clicked = st.button('💾 Salvar e Avançar')

st.markdown("<b>📅 Progresso deste ponto:</b>", unsafe_allow_html=True)
st.markdown(" ".join([f"{a} ☑" if a in completed_years else f"{a} ☐" for a in anos]))

if save_clicked:
    start = time.time()
    row = gdf_points.iloc[index]

    sample = {
        'id': int(index),
        f'class_id_{year}': int(class_selection),
        f'class_nm_{year}': classe_options[class_selection][2:].lower(),
        f'obs_{year}': obs_text,
        'longitude': row.geometry.x,
        'latitude': row.geometry.y,
        'geometry': row.geometry
    }

```

```

updated = False
for s in st.session_state.samples:
    if int(s['id']) == int(index):
        s.update(sample)
        updated = True
        break
if not updated:
    st.session_state.samples.append(sample)

backup_csv(csv_filename)

df = pd.DataFrame(st.session_state.samples)
df = sanitize_dataframe(df)
gdf = gpd.GeoDataFrame(df, geometry=gpd.points_from_xy(df.longitude, df.latitude),
crs="EPSG:4326")
gdf.to_csv(csv_filename, index=False)
timer_debug('Salvar CSV', start)

# Avança para o próximo ano ou ponto
if st.session_state.current_year_index < len(anos) - 1:
    st.session_state.current_year_index += 1
else:
    filled_all = all(
        not pd.isna(s.get(f'class_id_{a}'))
        for s in st.session_state.samples if int(s['id']) == index
        for a in anos
    )
    if filled_all:
        st.session_state.current_index = min(index + 1, total_pontos - 1)
        st.session_state.current_year_index = 0

st.rerun()

row = gdf_points.iloc[st.session_state.current_index]
point = ee.Feature(ee.Geometry.Point([row.geometry.x, row.geometry.y]))
mosaic = get_mosaic(year)
geom_1km = create_square(row.geometry, 500)
geom_5km = create_square(row.geometry, 2500)
clip_1km = mosaic.clip(geom_1km)

rgb_vis = {'bands': ['red', 'green', 'blue'], 'min': 0, 'max': 1.5}
ndvi_vis = {
    'bands': ['ndvi'],
    'min': -0.07,
    'max': 0.7,
    'palette': [

```

```

        '#654321', # solo exposto / negativo
        '#ffbb22', # vegetação rala
        '#ffff4c', # pasto
        '#aaff55', # vegetação média
        '#55aa00', # vegetação densa
        '#007700' # vegetação muito densa
    ]
}

mndwi_vis = {
    'bands': ['mndwi'],
    'min': -0.5,
    'max': 0.5,
    'palette': [
        '#ffffff', # áreas secas ou urbanas (negativo)
        '#c0c0c0', # solo úmido / transição
        '#66ccff', # água rasa
        '#0000ff' # água profunda
    ]
}

ponto_vis = {'color': 'pink'}

col1, col2, col3 = st.columns(3)

with col1:
    st.markdown("<h5> 🗺️ RGB (1km)</h5>", unsafe_allow_html=True)
    Map1 = geemap.Map()
    Map1.centerObject(point, 15)
    Map1.addLayer(clip_1km, rgb_vis, 'RGB 1km')
    Map1.addLayer(point, ponto_vis, 'Ponto')
    Map1.to_streamlit(height=250)

with col2:
    st.markdown("<h5> 🌿 NDVI (1km)</h5>", unsafe_allow_html=True)
    Map2 = geemap.Map()
    Map2.centerObject(point, 15)
    Map2.addLayer(clip_1km, ndvi_vis, 'NDVI 1km')
    Map2.addLayer(point, ponto_vis, 'Ponto')
    Map2.to_streamlit(height=250)

with col3:
    st.markdown("<h5> 💧 MNDWI (1km)</h5>", unsafe_allow_html=True)
    Map3 = geemap.Map()
    Map3.centerObject(point, 15)
    Map3.addLayer(clip_1km, mndwi_vis, 'MNDWI 1km')
    Map3.addLayer(point, ponto_vis, 'Ponto')

```

```
Map3.to_streamlit(height=250)
```

```
st.markdown("<h5> 📡 Satélite Atual (5km)</h5>", unsafe_allow_html=True)
```

```
Map4 = geemap.Map()
```

```
Map4.centerObject(point, 13)
```

```
Map4.add_basemap('SATELLITE')
```

```
Map4.addLayer(point, ponto_vis, 'Ponto')
```

```
Map4.to_streamlit(height=400)
```

```
# Aba 2: Visualizar CSV
```

```
with tabs[1]:
```

```
st.markdown("<h5> 📄 Dados Salvos até o momento</h5>", unsafe_allow_html=True)
```

```
df_display = pd.read_csv(csv_filename)
```

```
st.dataframe(df_display, use_container_width=True)
```