



**Universidade Estadual Paulista “Júlio de Mesquita Filho”
Faculdade de Ciências e Tecnologia
Engenharia de Energia**

Matheus Vinicius Giglio Brisolla dos Santos

**Mapeamento da vegetação para análise de faltas no sistema de
distribuição do município de Presidente Prudente – SP**

Rosana - SP

2024

Matheus Vinicius Giglio Brisolla dos Santos

**Mapeamento da vegetação para análise de faltas no sistema de
distribuição do Município de Presidente Prudente - SP**

Trabalho de Conclusão de Curso apresentado
ao Curso de Engenharia de Energia da
Faculdade de Engenharia e Ciências - Câmpus
de Rosana, Universidade Estadual Paulista
“Júlio de Mesquita Filho”, como parte dos
requisitos para obtenção do grau de Bacharel
em Engenharia de Energia

Orientador(a): Dra. Letícia Sabo Boschi

Rosana – SP

2024

S237m

Santos, Matheus Vinicius Giglio Brisolla dos

Mapeamento da Vegetação para análise de faltas na rede de distribuição do município de Presidente Prudente-SP / Matheus Vinicius Giglio Brisolla dos Santos. -- Rosana, 2024

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

Trabalho de conclusão de curso (Bacharelado - Engenharia de Energia) - Universidade Estadual Paulista (UNESP), Faculdade de Engenharia e Ciências, Rosana

Orientador: Leticia Sabo Boschi

1. Sensoriamento Remoto. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Engenharia e Ciências, Rosana. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

A confecção da ficha catalográfica é realizada exclusivamente pelo Serviço Técnico de Biblioteca e Documentação e deve ser inserida no lugar desta folha.

Para solicitá-la, comunique a biblioteca do Campus Experimental de Rosana.

Obs: Imprimir na parte inferior, no verso da Folha de Rosto.

Matheus Vinicius Giglio Brisolla dos Santos

Mapeamento da vegetação para análise de faltas no sistema de distribuição do
Município de Presidente Prudente - SP

Trabalho de Conclusão de Curso apresentado
ao Curso de Engenharia de Energia da
Faculdade de Engenharia e Ciências - Câmpus
de Rosana, Universidade Estadual Paulista
“Júlio de Mesquita Filho”, como parte dos
requisitos para obtenção do grau de Bacharel
em Engenharia de Energia.

Rosana, ____/____/____.

Componentes da Banca Examinadora:

Presidente e Orientador: Profa. Dra. Letícia Sabo Boschi, Câmpus de Rosana, UNESP-FEC

Membro Titular: Profa. Dra. Andréia Fátima Zanette, Câmpus de Rosana, UNESP-FEC

Membro Titular: Prof. Dr. Jose Francisco Resende da Silva, Câmpus de Rosana, UNESP-FEC

AGRADECIMENTOS

Gostaria primeiramente de agradecer a Deus por me permitir chegar onde estou hoje, pois sem Ele nada seria.

Agradeço à minha orientadora Profa. Dra. Letícia pela paciência, ajuda e compreensão nestes anos de graduação, me ensinando e instruindo cada dia mais, e acima de tudo por me dar a oportunidade de finalizar esse ciclo sob sua orientação.

Agradeço à minha família por estar ao meu lado e me prover tudo que necessitei em uma das fases mais interessantes que tive e que muito provavelmente terei em toda minha vida, que foi a graduação.

Agradeço à minha futura esposa Thais, por ter sempre me incentivado a ser melhor e buscar coisas melhores em minha vida, por seu companheirismo e amor.

Agradeço aos inúmeros amigos que fiz durante a graduação, Lucas Valle, Gabriel Khayat, Heitor Plazas, Gabriel Heimman, Mariana Kayashima, Beatriz Gonçalves, Larissa Marinho, Vinicius Miquelon, João Leme, José Mejia, entre muitos outros que fizeram parte dessa fase maravilhosa que vivi por alguns anos.

Por último gostaria de agradecer a todos os professores, funcionários e alunos com quem tive contato durante esse tempo.

“As pessoas costumam dizer que a motivação não dura para sempre. Bem, nem o efeito do banho, por isso recomenda-se diariamente.”

Zig Ziglar

RESUMO

As interrupções de energia em redes de distribuição de alta, média e baixa tensão, causadas pela vegetação arbórea são um grande problema no sistema de redes aéreas, o mais utilizado no Brasil. Visto essa problemática, um caminho para a solução deste tópico é o mapeamento da vegetação, utilizando classificadores de alta precisão. O mapeamento da vegetação, quando efetuado de maneira assertiva, fornece dados qualitativos e quantitativos para tomadas de decisões efetivas e preventivas, como por exemplo as podas de árvores com maiores riscos de interrupção na rede. Esse estudo foi realizado a partir das ferramentas QGIS e Google Earth Engine (GEE). Utilizando a ferramenta GEE, foi possível obter imagens do satélite Sentinel 2A, e classificá-las em 6 classes distintas. A precisão dos resultados se tornou possível por conta do algoritmo de classificação utilizado, o *Random Forest*, e os índices espectrais NDVI, MNDWI e NDBI. A ferramenta GEE se mostrou eficiente para o objetivo deste trabalho, uma vez que aplicado os índices de qualidade, Kappa e Acurácia Global (AC), estes retornaram com precisão acima de 91% em todas as classes de classificação. Após obter a imagem com classificação precisa, foi feito o tratamento de dados na ferramenta QGIS, uma análise dos dados de redes de distribuição que passam no município de Presidente Prudente, obtidos através do portal da ANEEL (Agência Nacional de Energia Elétrica). Com os dados das linhas de baixa, média e alta tensão, foi possível sobrepor a classificação da vegetação arbórea e indicar quais áreas têm maior risco de interrupções na rede. Após realizar essa análise, foi efetuado um levantamento com dados quantitativos das interrupções por Fator Gerador (FG), constatando que a vegetação arbórea é o quarto maior FG dentre os apresentados no relatório da ANEEL. A partir dos resultados apresentados, é possível afirmar que o mapeamento da vegetação permite que ações preventivas sejam executadas a fim de evitar interrupções na rede, como podas de árvores, uma vez que as localidades com maiores riscos são identificadas.

PALAVRAS-CHAVE: Mapeamento da vegetação. Classificação supervisionada. Rede de Distribuição.

ABSTRACT

Power interruptions in high, medium and low voltage distribution networks, caused by tree vegetation, are a major problem in the aerial network system, the most used in Brazil. Given this problem, one way to resolve this topic is vegetation mapping, using high-precision classifiers. Vegetation mapping, when carried out assertively, provides qualitative and quantitative data for effective and preventive decision-making, such as pruning trees with greater risks of interruption to the network. This study was carried out using QGIS and Google Earth Engine (GEE) tools. Using the GEE tool, it was possible to obtain images from the Sentinel 2A satellite and classify them into 6 different classes. The precision of the results became possible due to the classification algorithm used, Random Forest, and the spectral indices NDVI, MNDWI and NDBI. The GEE tool proved to be efficient for the objective of this work, since when applying the veracity, Kappa and Global Accuracy (AC) indices, these returned with accuracy above 91% in all classification classes. After obtaining the image with precise classification, data processing was carried out using the QGIS tool, an analysis of data from distribution networks that pass through the municipality of Presidente Prudente, obtained through the ANEEL (National Electric Energy Agency) portal. With data from low, medium and high voltage lines, it was possible to overlay the classification of tree vegetation and indicate which areas are at greater risk of network interruptions. After carrying out this analysis, a survey was carried out with quantitative data on interruptions by Generating Factor (FG), finding that tree vegetation is the fourth largest FG among those presented in the ANEEL report. From the results presented, it is possible to affirm that vegetation mapping allows preventive actions to be carried out in order to avoid interruptions in the network, such as tree pruning, since the locations with the greatest risks are identified.

KEYWORDS: Vegetation mapping. Supervised classification. Electrical Distribution Network.

LISTA DE ILUSTRAÇÕES

Figura 1 – Gráfico representando as ocorrências diárias durante o período de 2023 e os Fatores Geradores	15
Figura 2 – Mapa do Estado de São Paulo Destacando o Município de Presidente Prudente	20
Figura 3 – Fluxograma de Atividades Realizadas na Metodologia	23
Figura 4 – Distribuição dos dados de treinamento na área de estudos	25
Figura 5 – Código fonte da classificação com Random Forest	27
Figura 6 – Código fonte para determinação das métricas estatísticas	28
Figura 6 – Índice MNDWI	29
Figura 8 – Índice NDVI	30
Figura 9 – Índice NDBI	30
Figura 10 – Mapa de Linhas de Distribuição por Setor Censitário	31
Figura 11 – Mapa de uso e cobertura do solo	36
Figura 12 – Relatório da área total de cada uma das classes em hectares	36
Figura 13 – Soma dos recortes dos setores censitários previamente recortados	37
Figura 14 – Setores Censitários sujeitos à falta	39
Figura 15 – Setor Censitário 4	39
Figura 16 – Evidência Fotográfica 1	41
Figura 17 – Evidência Fotográfica 2	41

LISTA DE TABELAS

Tabela 1 – Valores encontrados na Matriz de Confusão	33
Tabela 2 – Índices de Acurácia Global e Kappa	34
Tabela 3 – Acurácia Produtor 34	34
Tabela 4 – Acurácia do Consumidor	35
Tabela 5 – Setores com maiores proporções arbóreas/área	38
Tabela 6 – Interrupções por nível de tensão	42

LISTA DE QUADROS

Quadro 1 – Dados Quantitativos dos Fatores Geradores	16
Quadro 2 – Representação da Matriz de Confusão	18
Quadro 3 – Desempenho do Índice Kappa	19
Quadro 4 – Características das bandas do Sentinel 2A	22
Quadro 5 – Classes do mapa de uso e cobertura do solo	24

SUMÁRIO

1.	INTRODUÇÃO	15
2.	OBJETIVO	17
2.1	Geral	17
2.2	Específico	17
3.	JUSTIFICATIVA	18
4.	LEVANTAMENTO BIBLIOGRÁFICO	20
4.1.	Sensoriamento Remoto	20
4.2.	Classificação Supervisionada	20
4.3.	Matriz de Confusão	21
4.4.	Índice Kappa	22
4.5.	Acurácia Global	23
5.	MATERIAL E MÉTODOS	24
5.1.	Área de Estudo	24
5.2.	<i>Softwares</i> Utilizados	25
5.3.	Imagens Sentinel 2	26
5.4.	Metodologia	27
5.4.1.	Pré-Processamento dos Dados	29
5.4.2.	Processamento de dados	29
5.4.3.	Índices Espectrais Utilizados	31
5.4.4.	Aplicação do Algoritmo <i>Random Forest</i>	34
5.4.5.	Geração do Índice Kappa, Acurácia Global e Matriz de Confusão	35
5.4.6.	Tratamento dos Dados	36
5.4.7.	Relatórios quantitativos	39
6.	RESULTADO E DISCUSSÃO	40

7.	CONCLUSÃO	52
	REFERÊNCIAS	53

1. INTRODUÇÃO

A ANEEL (Agência Nacional de Energia Elétrica) é responsável por regular e fiscalizar a produção, transmissão, comercialização e distribuição da energia elétrica em território nacional. À ela também é atribuída a função de conceder, autorizar ou permitir instalações e serviços de energia elétrica, implementar políticas no setor, realizar leilões e concessões, gerir contratos, estabelecer regras para o serviço de energia, criar metodologias para o cálculo de tarifas, fiscalizar o fornecimento de energia e mediar conflitos (CPFL, 2024).

O Sistema de Distribuição de Energia Elétrica (SDEE) compreende os equipamentos e instalações elétricas pertencentes à distribuidora da região de sua concessão, incluindo os sistemas de proteção (ANEEL, 2024), importantes para o bom funcionamento e segurança de um sistema elétrico de potência. Entretanto, historicamente, os sistemas de potência apresentam muitas falhas, e nem sempre possuem técnicas de detecção confiáveis, apesar da evolução dos equipamentos e novas técnicas. Dentre as potenciais falhas, estão as Faltas de Alta Impedância (FAI) (FARIAS, 2013).

As FAIs ocorrem pelo contato entre um condutor energizado rompido e superfícies de alta impedância, o que gera correntes com amplitudes reduzidas. Pelo fato dessas correntes, em grande parte, serem menores que as correntes de carga do alimentador, não chegam a sensibilizar os sistemas convencionais de proteção, propiciando riscos a bens materiais e vidas (FARIAS, 2013).

De acordo com a localização das redes, sendo elas redes aéreas ou subterrâneas, as especificidades das construções usadas na distribuição se alteram. Tendo em vista que as redes aéreas se caracterizam por serem construções mais ágeis e por serem mais eficientes na identificação de defeitos, e ainda por possuírem baixo custo de implantação, quando comparado à subterrânea esta se torna a mais comum entre as regiões do Brasil. Em contrapartida, a rede aérea apresenta maiores falhas, por estar vulnerável a agentes externos, como árvores, condições meteorológicas, animais, vandalismos e furtos, fenômenos naturais, dentre outros. A qualidade e continuidade da energia fornecida pela concessionária deve atender índices estabelecidos por leis e regulados pela ANEEL, sendo passível de perda de concessão caso os índices não sejam atendidos (LIMA et al., 2022).

Como obstáculo, as concessionárias enfrentam grandes dificuldades na distribuição da energia elétrica, o que é resultado da qualidade do serviço de distribuição de energia, entendido como continuidade de fornecimento, considerado pelas faltas que podem ser divididas em

manutenção corretiva (faltas não programadas) e manutenção preventiva (faltas programadas).

Das principais faltas não programadas no sistema, têm como destaque a influência de vegetações arbóreas em contato com as linhas de distribuição, que pode ser evitada sistematicamente com a poda preventiva dessa vegetação (TRONCHONI, 2008).

Diante do exposto o trabalho tem como objetivo analisar a relação entre a vegetação e as faltas no sistema de distribuição, com enfoque nas interferências causadas pela vegetação presente no município de Presidente Prudente. Dessa forma, se torna imprescindível fazer o mapeamento e a quantificação da vegetação, os quais são importantes para com o planejamento urbano, tomada de decisões e aferição da qualidade urbana de distintos setores do município, como uma análise das regiões que estarão sujeitas à falhas em redes de distribuição (HIRYE, 2020).

2. OBJETIVO

2.1. Geral

O objetivo geral do presente trabalho é mapear a vegetação arbórea do município de Presidente Prudente–SP e quantificar a área ocupada pela vegetação em cada setor censitário, uma vez que a vegetação está entre as variáveis locais que tornam os alimentadores de determinadas áreas mais propensos às falhas.

2.2. Específico

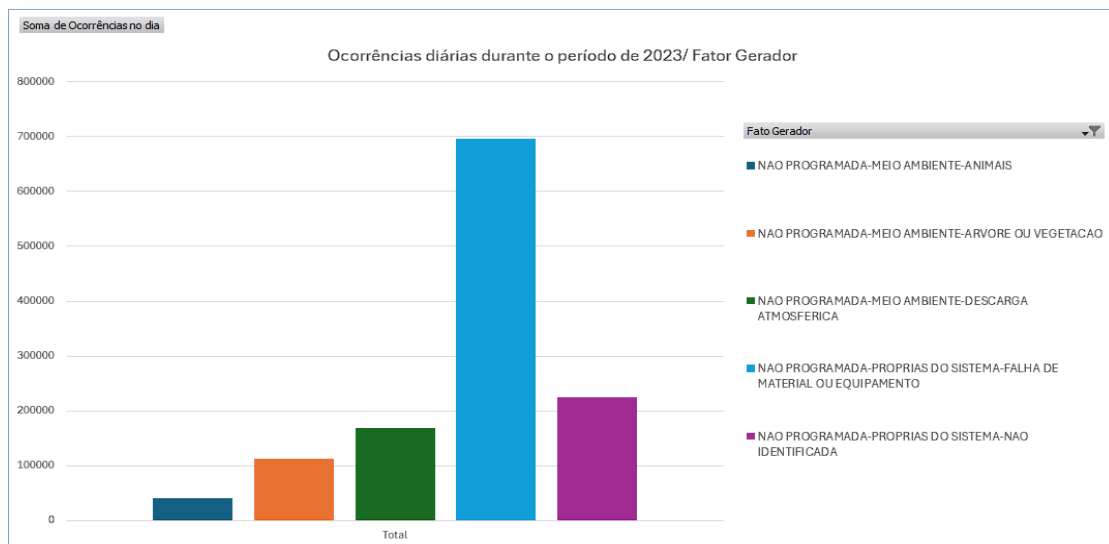
Para atingir os objetivos serão desenvolvidas as seguintes etapas:

- Mapeamento da vegetação arbórea do município a partir de imagens do satélite Sentinel 2 e do processo de extração de informações baseado em aprendizado de máquina;
- Geração de representações que permitam visualizar a distribuição espacial da vegetação no município de Presidente Prudente; e
- Quantificação da área ocupada por vegetação em cada um dos setores censitários.

3. JUSTIFICATIVA

Através da base de dados abertos da ANEEL sobre Interrupções de Energia Elétrica nas Redes de Distribuição, podemos verificar o número de cada ocorrência de acordo com o Fator Gerador (FG) e a partir das informações descritas na base, analisar as causas e suas particularidades, padrões e tendências. As ocorrências resultantes de Meio Ambiente e de Causa Própria do Sistema de distribuição representam aproximadamente 93% de todo volume de interrupções no ano de 2023, totalizando 1.323.718 ocorrências de interrupção de fornecimento de energia elétrica. O gráfico na Figura 1 indica as ocorrências diárias durante esse período para os maiores Fatores Geradores.

Figura 1: Gráfico representando as ocorrências diárias durante o período de 2023 e os Fatores Geradores



Fonte: Elaborado pelo autor.

Como exposto acima, a maior causa das faltas não programadas no período se deve a falhas próprias do sistema, falha de material ou equipamento, seguido de faltas não identificadas, descarga atmosférica, árvore ou vegetação, e animais, respectivamente.

Analisando os dados quantitativos que deram origem à Figura 1, verifica-se que a quantidade de faltas por conta da vegetação é de 112.839 durante o período de 2023. Vale ressaltar que esses dados, obtidos através da ANEEL, representam as faltas em toda a região de concessão da concessionária ENERGISA SUL-SUDESTE, a qual possui concessão no

município estudado.

No Quadro 1 são descritos os dados quantitativos para os 5 maiores Fatores Geradores de faltas.

Quadro 1: Dados Quantitativos dos Fatores Geradores

Fator Gerador	Ocorrências Diárias
MEIO AMBIENTE-ANIMAIS	40777
MEIO AMBIENTE-ARVORE OU VEGETACAO	112839
MEIO AMBIENTE-DESCARGA ATMOSFERICA	168872
PROPRIAS DO SISTEMA-FALHA DE MATERIAL OU EQUIPAMENTO	695847
PROPRIAS DO SISTEMA-NAO IDENTIFICADA	225320

Fonte: Elaborado pelo autor.

A partir desses dados é observado que a vegetação tem grande impacto nas falhas de redes de distribuição. Essas falhas podem ocorrer por falta de podas preventivas em vegetações arbóreas. Desse modo, o mapeamento de cobertura e uso do solo, principalmente de áreas urbanizadas, apresenta-se como uma informação importante para a definição de ações preventivas, para o planejamento urbano e tomadas de decisões de diversos setores de uma determinada região (LILLESAND e KIEFER, 1994).

Em grande parte dos estudos de mapeamento da vegetação, as classificações espectrais são utilizadas para caracterizá-los. Essas classificações podem ser divididas em subgrupos, sendo elas: classificação não supervisionada, supervisionada, orientada a objetos e abordagens estatísticas. A classificação supervisionada se utiliza de amostras fornecidas pelo usuário para treinar os algoritmos, assim caracterizando as classes espectrais. Os algoritmos mais comuns são o Método como Máxima Distância, Mínima Distância e *Random Forest*, o qual será utilizado no presente trabalho (HIRYE, 2020).

Nesse estudo será utilizada a classificação supervisionada, que é o procedimento mais frequentemente empregado para análise quantitativa de dados de imagens de sensoriamento remoto, a fim de fornecer, como objeto de estudo, os locais de possíveis faltas causadas por vegetação nas linhas de distribuição, tanto de alta como média e baixa tensão.

4. LEVANTAMENTO BIBLIOGRÁFICO

4.1. Sensoriamento remoto

A principal função do processo de sensoriamento remoto é de fornecer ferramentas capazes de facilitar a identificação de informações contidas na imagem, para que assim possam ser posteriormente interpretadas e analisadas.

Dessa forma, os sistemas e ferramentas são utilizados para essas análises e manutenção de imagens brutas, resultando na produção de outras imagens com especificações e informações palpáveis para um estudo. A informação de interesse é caracterizada em função das propriedades dos objetos ou padrões que compõem a imagem. Portanto, extrair informação de imagens envolve o reconhecimento de objetos ou padrões. O sistema visual humano possui uma notável capacidade de reconhecer padrões, porém raramente tem capacidade de processar volumes maiores de informações presentes nesse tipo de imagem é capaz de processar o enorme volume de informação presente numa imagem. O objetivo principal do processamento de imagens é o de remover essas barreiras, facilitando a extração de informações a partir de imagens. Deste modo, o processo de sensoriamento remoto deve ser encarado como um estágio obrigatório num contexto de análises de imagens (CRÓSTA, 2002).

4.2. Classificação Supervisionada

A classificação de imagens é de um processo de coleta e análise de dados espaciais de diferentes fontes atribuídos a um determinado fenômeno, com o objetivo de auxiliar na compreensão da cena e facilitar a tomada de decisão (CONGALTON e GREEN, 2008).

Para aprimoramento de classificações de imagens espectrais, é utilizado a técnica *Random Forest*. Esta é uma técnica de árvore de regressão que usa agregação bootstrap (meta-algoritmo de conjunto de aprendizado de máquina projetado para melhorar a estabilidade e a precisão dos algoritmos de aprendizado de máquina usados na classificação e regressão estatística) e randomização de preditores para atingir um alto grau de precisão preditiva. Os vários parâmetros de entrada da floresta aleatória são explorados (STROBL, 2009).

4.3. Matriz de Confusão

A Matriz de Confusão é a forma de representação da qualidade obtida em uma classificação digital de imagem, sendo ela definida pela relação de informações dos dados de referência (compreendido como verdadeiro) com os dados classificados (compreendidos como previstos) (PRINA, 2015).

Ela é formada pela matriz quadrada, arranjada entre linhas e colunas, dos números de pixels atribuídos a determinada classe pertencente. As classes são definidas durante o processo de classificação de imagens. Os números de pixels da diagonal representam os que foram classificados como Positivos Verdadeiros (PV), estes que eram previstos para serem classificados dentro da classe definida. O Quadro 2 mostra a disposição da MC.

Quadro 2: Representação da Matriz de Confusão

	Mata Nativa	Área Urbana	Cultura	Pastagem	Solo Exposto
Mata Nativa	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
Área Urbana	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
Cultura	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}
Pastagem	X_{41}	X_{42}	X_{43}	X_{44}	X_{45}
Solo Exposto	X_{51}	X_{52}	X_{53}	X_{54}	X_{55}

Fonte: Adaptado de Congalton e Green (2008).

Os números que se encontram fora da diagonal principal possibilitam a verificação de erros de omissão e de inclusão. Erros de omissão são aqueles onde os pixels pertencentes a uma classe de interesse são alocados a outras classes, e os erros de inclusão ocorrem quando os pixels de outras classes são associados à classe de interesse (CONGALTON e GREEN, 2008).

A partir da matriz de confusão há a possibilidade de avaliar quais classes amostrais apresentaram mais confusão e obter algumas medidas descritivas, como Acurácia Global e estimativa do Índice Kappa, os quais nos informam a qualidade da classificação.

4.4 Índice Kappa

O índice Kappa indica a qualidade da classificação, podendo ser considerado um coeficiente de veracidade dos dados obtidos por ela (CONGALTON e GREEN, 1999; VIEIRA, 2000, p. 82). Este índice é desenvolvido como medida de avaliação estatística da relação de exatidão entre os dados de referência e os dados de classificados, fornecendo um valor numérico da diferença entre esses dados.

O valor de resposta do Índice Kappa é associado à qualidade da classificação em uma escala de desempenho que varia de Péssima à Excelente, no qual Péssima equivale ao valor do índice como 0 e Excelente entre 0,81 e 1 (LANDIS e KOCH, 1997). O Quadro 3 apresenta a escala de desempenho do Índice.

Quadro 3: Desempenho do Índice Kappa

Valor do Índice Kappa	Desempenho
0,00	Péssima
0,01 a 0,20	Ruim
0,21 a 0,40	Razoável
0,41 a 0,60	Boa
0,61 a 0,80	Muito Boa
0,81 a 1,00	Excelente

Fonte: Adaptado de Landis e Koch (1997).

4.5 Acurácia Global

A Acurácia Global é uma medida específica calculada pela soma do total dos pixels classificados corretamente dividida pelo número total de pixels da MC. As acurácias individuais são calculadas através da acurácia do produtor que indica a probabilidade de um pixel ter sido corretamente classificado pela ferramenta utilizada e a acurácia do usuário calcula a probabilidade de um pixel ter sido corretamente classificado pelo usuário da ferramenta (CONGALTON, 1991).

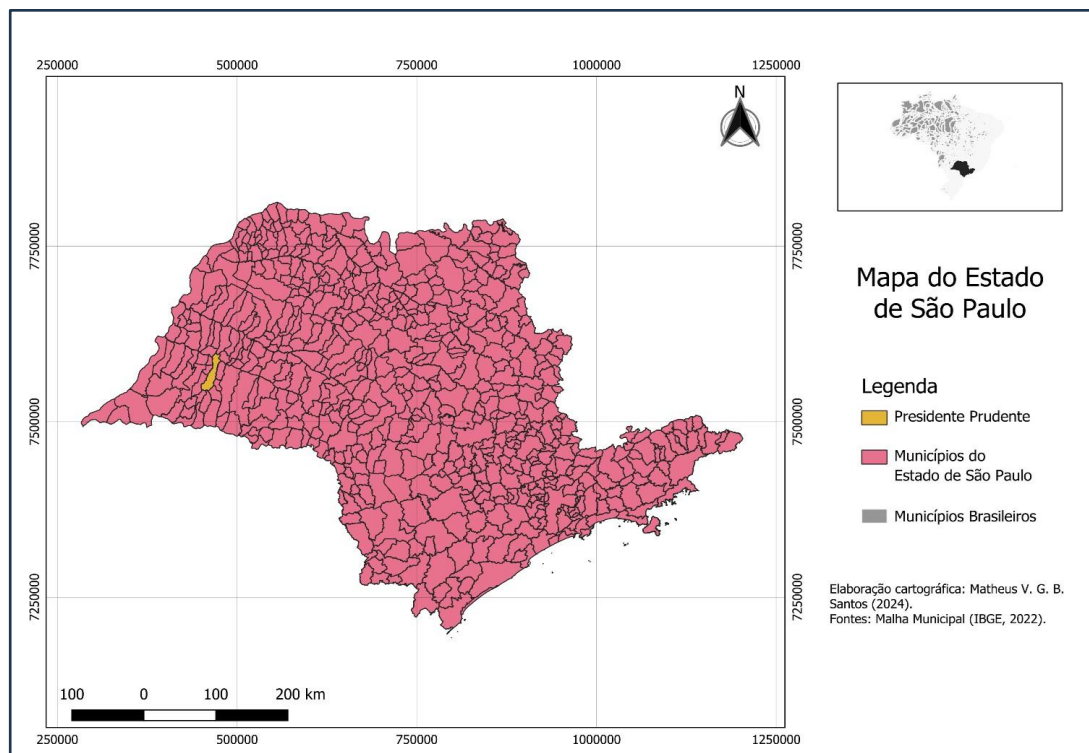
5. MATERIAL E MÉTODOS

5.1 Área de Estudo

A área selecionada para este estudo foi o município de Presidente Prudente-SP, pertencente à microrregião do Pontal do Paranapanema. De acordo com o último censo, Presidente Prudente conta com uma população residente de 225.668 habitantes (IBGE, 2022). A escolha da área em questão se dá pelo fato de Presidente Prudente ser considerada a “Capital do Oeste Paulista”. A Figura 2 apresenta a localização do município de Presidente Prudente no Estado de São Paulo.

A imagem selecionada neste estudo data de dezembro de 2023. O critério na seleção da imagem foi a porcentagem de nuvens, uma vez que as nuvens “escondem” informações, e a estação do ano na qual a vegetação apresenta boa resposta na imagem. Logo, a imagem adquirida possui 2% de nuvens, que foi o critério estabelecido para o trabalho.

Figura 2: Mapa do Estado de São Paulo destacando o município de Presidente Prudente.



Fonte: Elaborado pelo autor.

5.2 Softwares Utilizados

Para realizar o processamento das imagens do satélite Sentinel 2A a fim de produzir o mapeamento foram utilizados dois *softwares*, um para o processamento dos dados e o outro para o tratamento dos dados. São eles: Google Earth Engine (GEE) e QGIs, respectivamente.

O Google Earth Engine (GEE) é uma plataforma em nuvem desenvolvida pela Google, que tem como objetivo auxiliar processos de análises ambientais de grandes escalas. O GEE permite o geoprocessamento de dados através do *Code Editor* do Google Earth Engine, que possibilita a inserção de códigos em linguagem *JavaScript* para a criação de algoritmos classificadores, refinamento de imagens por índices espectrais, índices de qualidade, entre outros. Quando comparado a outros, o GEE apresenta vantagem, pois seu processamento de imagens é realizado em nuvem, não necessitando assim de máquinas de alto rendimento e processamento de dados (GOOGLE INC, 2021).

O Quantum GIS (QGis) é um *software* que permite a edição, visualização e análise de dados georreferenciados. É um *software* gratuito, que atua como um SIG (Sistema de Informação Geográfica) (QGIS DEVELOPMENT TEAM, 2021).

Na plataforma GEE foi realizada a classificação da imagem a partir do classificador *Random Forest* e dos índices espectrais, e gerados os índices referentes à qualidade da classificação. Já no *software* QGis foram elaborados os mapas temáticos e gerados os relatórios quantitativos das áreas de vegetação de cada setor censitário.

5.3 Imagens Sentinel 2

O satélite Sentinel teve início em 2014 como parte de um projeto da Agência Espacial Europeia (ESA), em conjunto com a Comissão Europeia, dentro do Programa Copernicus. O objetivo era dar continuidade a várias missões antigas gerenciadas pela ESA, como ERS, ENVISAT e SPOT. O satélite Sentinel 2A foi o primeiro satélite óptico da série a ser operado pela ESA e foi lançado em 2015. Ele leva a bordo o sensor multiespectral MSI, com 13 bandas espectrais, variando de 443 a 2190 nm. A resolução espacial é de 10m para as bandas do visível, 20m para o infravermelho e 60m para as bandas de correção atmosférica. Suas aplicações incluem desde o monitoramento da agricultura, águas interiores, florestas, zonas costeiras, desastres naturais, até o uso e ocupação das terras (EMBRAPA, 2020).

O Quadro 4 descreve as características das bandas presentes no Sentinel 2A, o qual será

utilizado no presente trabalho.

Quadro 4: Características das bandas do Sentinel 2A.

Banda	Nome da Banda	Comprimento de onda central (nm)	Resolução Espacial (m)
B02	Azul	490	10
B03	Verde	560	10
B04	Vermelho	665	10
B08	NIR (Infravermelho Próximo)	842	10
B05	<i>Red Edge 1</i>	705	20
B06	<i>Red Edge 2</i>	740	20
B07	<i>Red Edge 3</i>	783	20
B08A	<i>Red Edge 4</i>	865	20
B11	SWIR 1	1610	20
B12	SWIR 2	2190	20
B01	<i>Aerossol</i>	443	60
B09	<i>Water Vapor</i>	940	60
B10	<i>Cirrus</i>	1375	60

Fonte: Adaptado de ENGESAT (2024).

5.4 Linhas de Distribuição

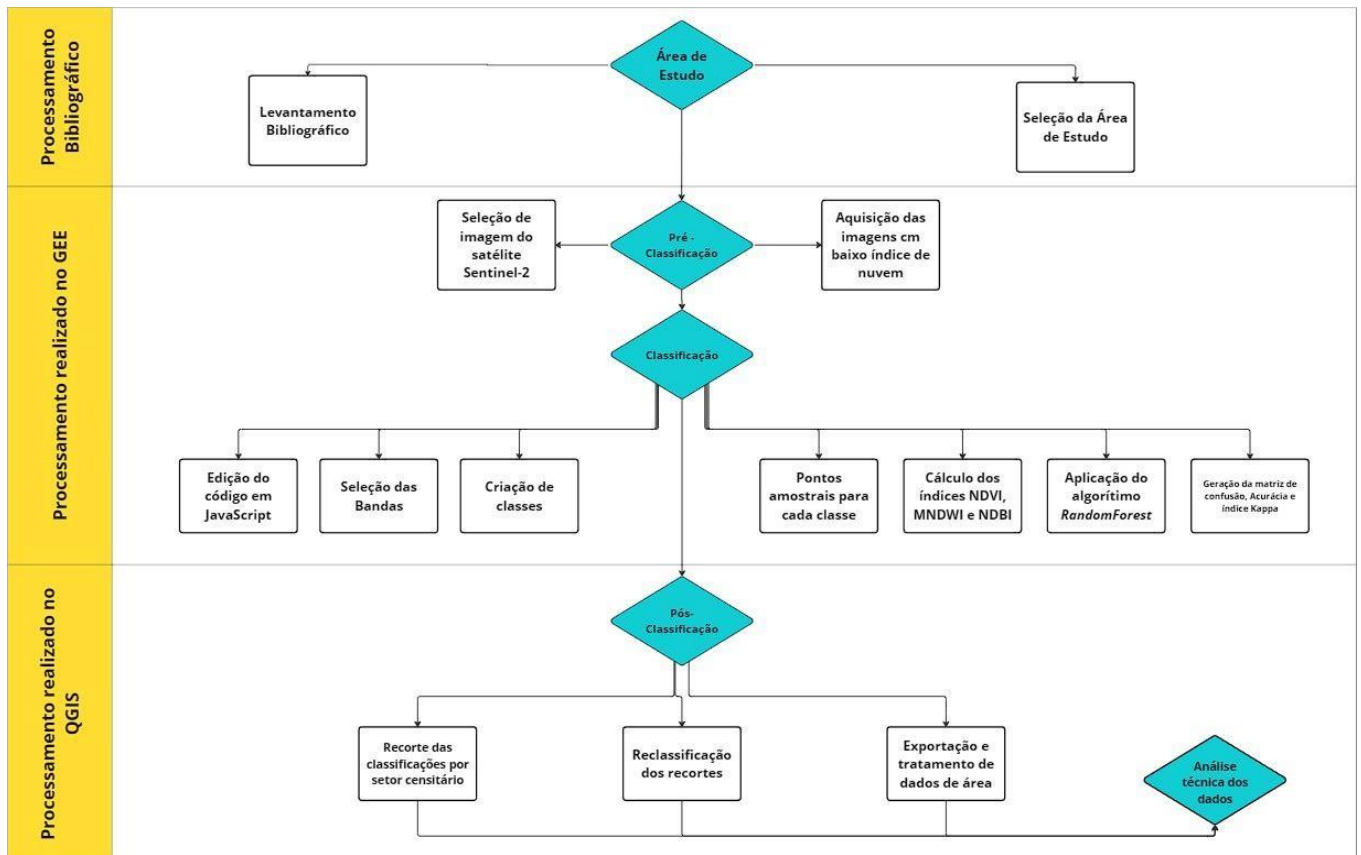
Através do portal da ANEEL, foi adquirido o *shapefile* com as linhas de distribuição de alta, média e baixa tensão, que será utilizado para análise de áreas que possuem riscos de faltas na rede em função da área coberta por vegetação arbórea no setor censitário estudado. Foi obtido também um arquivo .csv contendo todas as interrupções programadas e não programadas na rede distribuição nacional, sendo ele dividido em setores que são relevantes para a área de estudo.

5.5 METODOLOGIA

Como já citado anteriormente, o resultado da classificação e os índices de qualidade da classificação foram adquiridos através da ferramenta GEE. Os mapas visuais e os dados de área

coberta por vegetação arbórea para as análises quantitativas foram obtidos pela ferramenta QGIS com auxílio de tabelas dinâmicas produzidas no Excel. A Figura 3 indica a metodologia em um fluxograma de atividades realizadas.

Figura 3: Fluxograma de atividades realizadas na metodologia



Fonte: Elaborado pelo autor.

5.5.1 Pré-Processamento dos Dados

Nesta etapa, dentro da aba de *Code Editor* na plataforma GEE, é necessário informar a área de interesse. Essa área foi escolhida a partir de um ponto geométrico, sendo esse o centro do município de Presidente Prudente. Na sequência foi realizada a definição do satélite a ser usado como base para o estudo, neste caso, o Sentinel 2A.

Após a definição do satélite foi definida a data da imagem a ser utilizada e a cobertura de nuvem máxima, a fim de que se tenham dados mais confiáveis e de fácil análise. Neste caso foi definida uma cobertura máxima de 2% e escolhida a imagem de 13/12/2023, que possuía menos de 1% de cobertura de nuvem. A imagem gerada foi formada pelas bandas espectrais B02, B03 e B04.

5.5.2 Processamento de dados

Após a etapa de pré-processamento, em que se obteve a imagem selecionada, foi realizado o desenvolvimento do código em *JavaScript* declarando as bandas que serão utilizadas para o cálculo dos índices espectrais, sendo elas: B02, B03, B04, B05, B06, B07, B08, B11e B12.

Após a seleção das bandas foi realizada a definição das classes que serão utilizadas no processo de classificação da imagem e para a coleta dos dados de treinamento. O Quadro 5 apresenta as classes que foram definidas com as suas respectivas descrições.

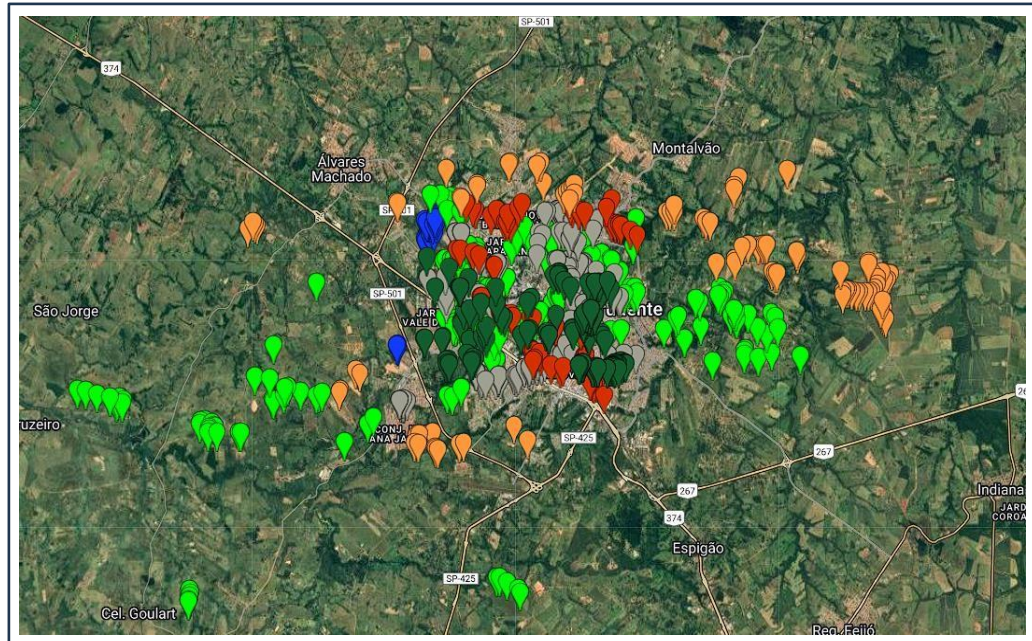
Quadro 5: Classes do mapa de uso e cobertura do solo

Classe	Definição
Corpo d'água (1)	Área de águas superficiais, onde se destacam rios, lagos, lagoas e oceanos.
Área Urbana (2)	Área com alta interferência humana.
Vias Terrestres (3)	Ruas, avenidas, logradouros, caminhos, passagens, rodovias e estradas, sendo de área urbana ou rural.
Vegetação Arbórea (4)	Áreas com presença de vegetação arbórea, sendo de área urbana ou rural.
Vegetação Rasteira / Cultura (5)	Áreas com presença de vegetação de médio porte, desde monocultura a cultura por ciclos, possuindo simetria entre os cultivos e pastagens.
Solo Exposto (6)	Área onde não há a presença de vegetação.

Fonte: Elaborado pelo autor.

Após a definição das classes foi realizada a coleta dos dados de treinamento da classificação, ou seja, foram coletadas amostras de cada uma das classes, na área de interesse que serviram como referência para que os demais pixels da imagem fossem classificados. A Figura 4 apresenta a distribuição dos dados de treinamento coletados no GEE.

Figura 4: Distribuição dos dados de treinamento na área de estudos.



Fonte: Elaborado pelo autor.

5.5.3 Índices Espectrais Utilizados

Após a criação dos pontos amostrais de cada uma das classes foram calculados os índices espectrais, escolhidos para melhorar o processo de classificação. Os índices selecionados foram o NDVI (Índice de vegetação por diferença normalizada), NDBI (Índice de Construção por Diferença Normalizada), e o MNDWI (Índice da diferença normalizada da água modificado). Os índices promovem uma maior acurácia no processo da classificação, pois suas respostas são utilizadas pelo algoritmo no momento da separação dos pixels de cada classe. A seguir são descritos os índices utilizados.

NDVI

O Índice de Vegetação por Diferença Normalizada (NDVI - Normalized Difference Vegetation Index) permite avaliar a densidade da cobertura vegetal, a partir da resposta espectral da vegetação em relação à absorção e reflectância dentro do espectro. Esse índice analisa a condição da vegetação, avaliando a intensidade da presença de clorofila associada ao vigor vegetativo na região (HUANG, 2021). O NDVI é calculado através da razão entre a

diferença da reflectância das bandas do infravermelho próximo (NIR) e do vermelho (Red), e a soma das duas reflectâncias utilizadas (ROUSE et al., 1974). No Sentinel 2A a banda do infravermelho próximo (ρ_{nir}) é equivalente à banda 08 e a banda do vermelho (ρ_{red}) equivale a banda 04, conforme a equação 1.

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (1)$$

MNDWI

O índice MNDWI (*Modified Normalized Difference Water Index*) foi proposto por Xu (2006) e é uma modificação do índice de diferença normalizada de água (NDWI – *Normalized Difference Water Index*). O MNDWI permite identificar cursos d'água, a partir da maximização da reflectância de áreas com a presença de água. No Sentinel 2A a banda do verde (ρ_{green}) é equivalente à banda 03 e a banda infravermelho próximo (ρ_{nir}) é equivalente à banda 08, como descrito na equação 2.

$$MNDWI = \frac{\rho_{green} - \rho_{NIR}}{\rho_{green} + \rho_{NIR}} \quad (2)$$

NDBI

O Índice de Construção por Diferença Normalizada (NDBI - *Normalized Difference Built-up Index*) foi baseado nos estudos de Zha et al. (2003) e possui como principal finalidade identificar áreas urbanas e impermeabilizadas através da resposta que o pixel terá durante a análise. Isso se dá pelo fato de que as áreas urbanas possuem predominância de materiais que absorvem a irradiância solar. A sua expressão matemática é descrita pela equação 3:

$$NDBI = \frac{\rho_{NIR} - \rho_{midIR}}{\rho_{NIR} + \rho_{midIR}} \quad (3)$$

No Sentinel 2A a banda infravermelho próximo (ρ_{nir}) é equivalente à banda 08 e a banda do infravermelho médio (ρ_{midIR} ou ρ_{SWIR}) equivale à banda 11.

5.5.4 Classificação a partir do Algoritmo *Random Forest*

O código fonte foi desenvolvido em *JavaScript* para atender os objetivos do presente trabalho. O código foi construído com o algoritmo *Random Forest* e foi utilizado no GEE para a classificação supervisionada da área de estudos. A Figura 5 apresenta a parte do código que trata da classificação.

Figura 5 – Código fonte da classificação com *Random Forest*.

```
var classificacao = ee.Classifier.smileRandomForest(7).train({
  features: amostra,
  classProperty: 'class',
  inputProperties: bands
});

//Imprima algumas informações sobre o classificador (especifico para CART)
var Classificado = dataset.classify(classificacao);
print ('RandomForest, explained', classificacao.explain());

//classifique o composto
var classificado = dataset.classify(classificacao);

//Display classification
var classificacao = s2img.select(bands).classify(classificacao);
Map.addLayer(classificacao, {min:0, max:5, palette: ['blue', 'grey', 'red', 'green', 'lime', 'orange']}, 'classificacao');

var validation = classificacao.sampleRegions({
  collection: amostra,
  properties: ['class'],
  scale: 10,
});
```

Fonte: Elaborado pelo autor.

Após as etapas de seleção de área e definição de índices foi aplicado o algoritmo de classificação e o treinamento foi realizado com 70% dos pixels selecionados na criação de classes amostrais e 30% reservado para testes, a fim de evitar o overfitting do modelo.

5.4.5 Geração do Índice Kappa, Acurácia Global e Matriz de Confusão

Os índices Kappa e a Acurácia Global são amplamente utilizados para determinar a qualidade do mapeamento e são obtidos a partir da matriz de confusão que descreve as classes amostrais e a quantidade de pixels que podem interferir na discriminação, em função da ambiguidade resultante de pixels semelhantes. Estes foram calculados e estimados através do código fonte apresentado na Figura 6.

Figura 6: Código fonte para determinação das métricas estatísticas

```
//Acurária
// Opcionalmente, faça alguma avaliação de precisão. Adicione uma coluna de
// uniformes aleatórios para o conjunto de dados de treinamento.
var withRandom = amostra.randomColumn('random');

//Queremos reservar alguns dos dados para teste, para evitar overfitting do modelo.
var split = 0.7; // Aproximadamente 70% treinando, 30% testando.
var trainingPartition = withRandom.filter(ee.Filter.lt('random', split));
var testingPartition = withRandom.filter(ee.Filter.gte('random', split));

// Treinado com 70% dos nossos dados..
var trainedClassifier = ee.Classifier.smileRandomForest(7).train({
  features: trainingPartition,
  classProperty: 'class',
  inputProperties: bands
});

// Classifica o teste FeatureCollection.
var validation = testingPartition.classify(trainedClassifier);

//Imprime a matriz de confusão.
var confusionMatrix = validation.errorMatrix(labelcover, 'classification');
print('Confusion Matrix', confusionMatrix);

var confMatrix = trainedClassifier.confusionMatrix()
var OA = confMatrix.accuracy()
var CA = confMatrix.consumersAccuracy()
var Kappa = confMatrix.kappa()
var PA = confMatrix.producersAccuracy()
var Order = confMatrix.order()

print(confMatrix, 'Matriz de Confusão')
print(OA, 'Acurácia geral da validação:')
print(Kappa, 'Kappa ')
print(CA, 'Acurácia do Consumidor ')
print(PA, 'Acurácia do Produtor ')
print(Order, 'Order')
```

Fonte: Elaborado pelo autor.

Seguido da determinação das métricas estatísticas, o arquivo *.tiff* contendo a imagem classificada foi exportado através do Google Drive para o *software* QGis, para tratamento e análise da classificação em questão.

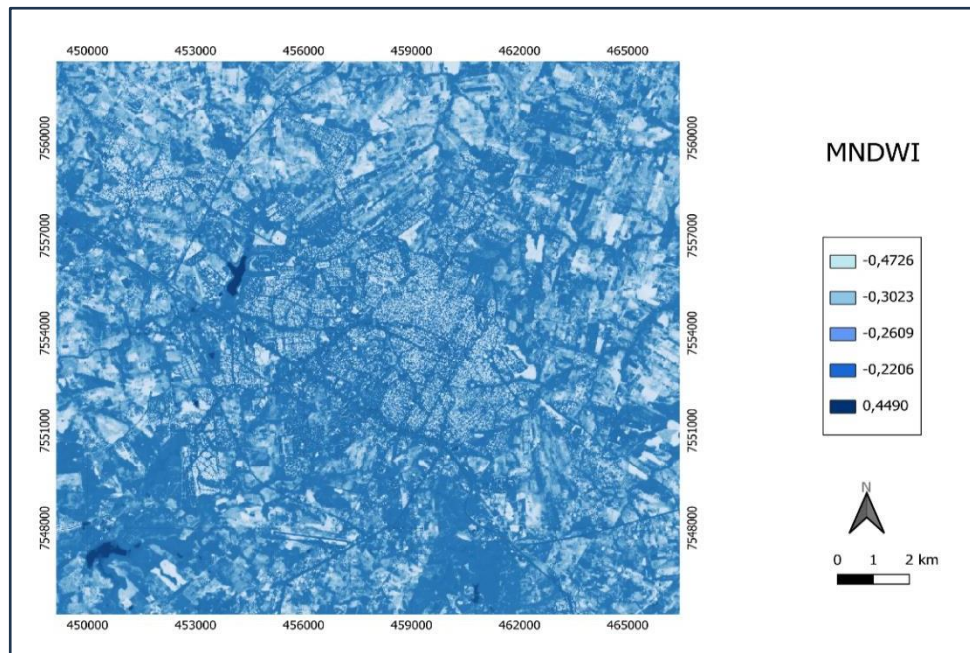
5.4.6 Tratamento de dados

Após realizado o processamento de dados, a classificação no GEE e a geração das métricas estatísticas, os resultados foram exportados através da nuvem para a ferramenta QGis, para que fossem elaborados os mapas temáticos, relatórios e análises classificatórias

quantitativas.

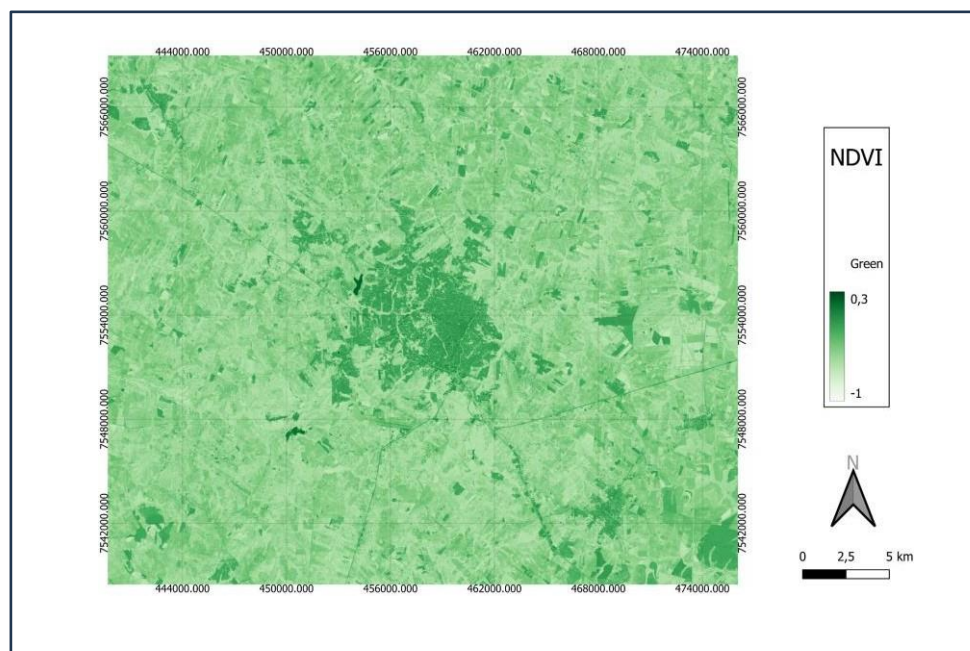
Utilizando a mesma imagem com a qual foi realizada a classificação na plataforma GEE, foram gerados os índices, já calculados no GEE, no QGis, a fim de apresentar visualmente cada um deles. Sendo eles: NDVI, NDBI e MNDWI, apresentados nas Figuras 7, 8 e 9, respectivamente.

Figura 7: Índice MNDWI



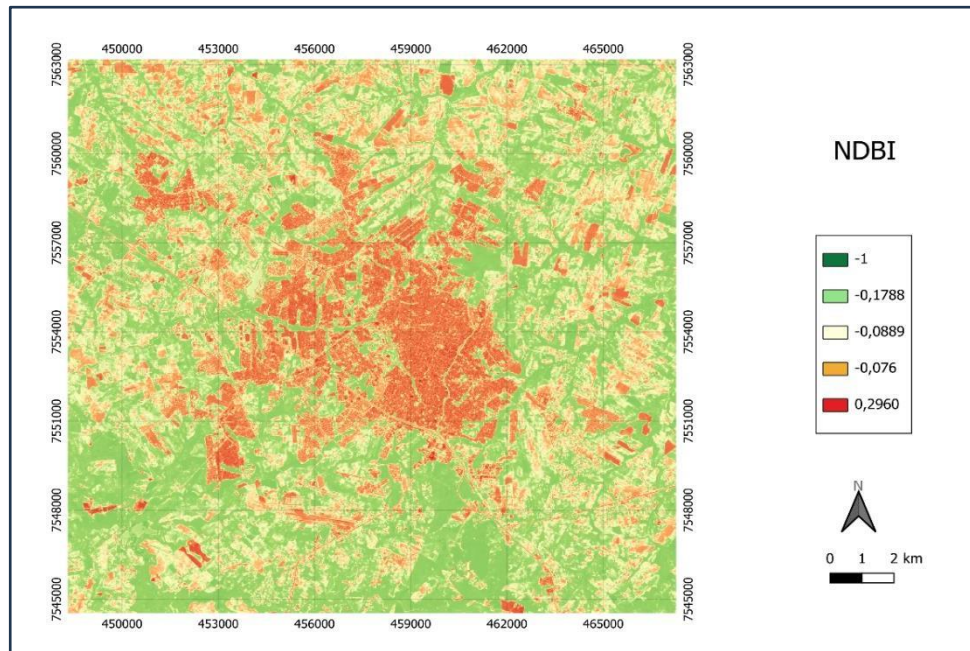
Fonte: Elaborado pelo autor.

Figura 8: Índice NDVI



Fonte: Elaborado pelo autor.

Figura 9: Índice NDBI

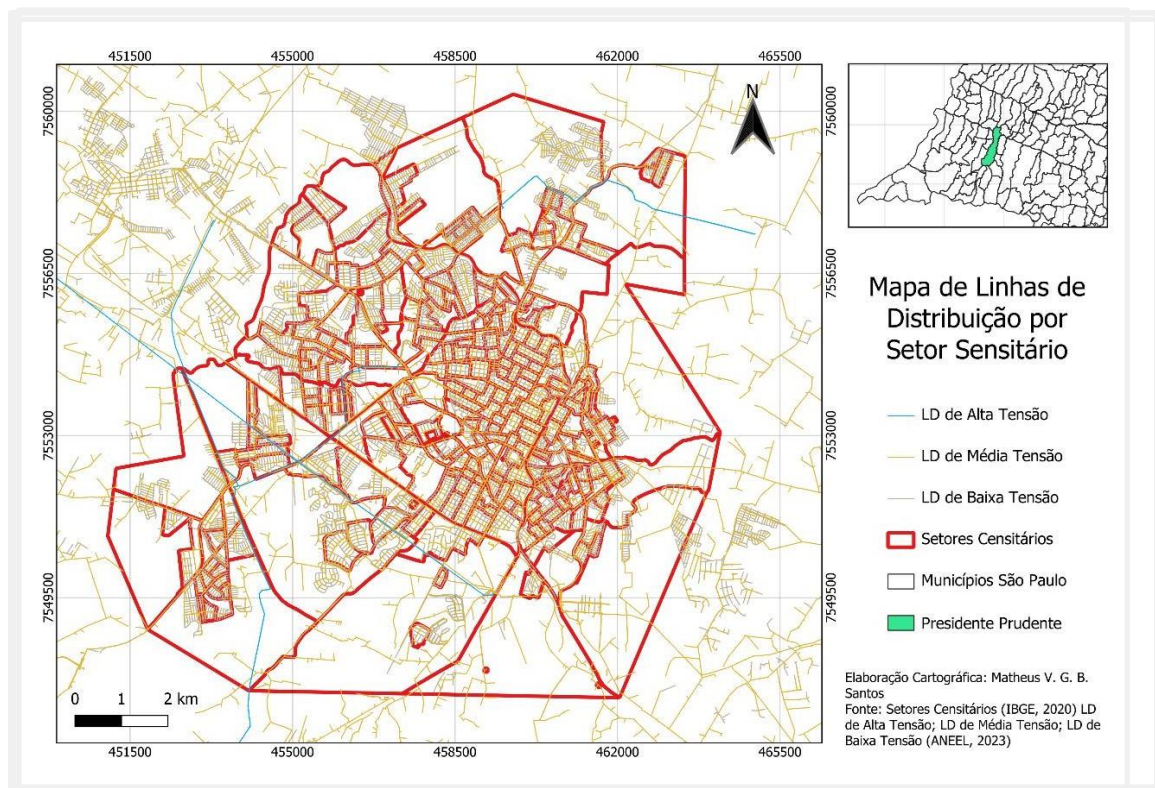


Fonte: Elaborado pelo autor.

Em seguida foram sobrepostos na imagem os arquivos no formato *shapefile*. O primeiro arquivo, disponibilizado pelo IBGE através de seu portal, contém os setores censitários do município de Presidente Prudente – SP. Já o segundo arquivo, disponibilizado pela ANEEL, contém as linhas de distribuição Baixa Tensão (BT), Média Tensão (MT) e Alta Tensão (AT), presentes em todo o território nacional. O objetivo dessa sobreposição é analisar a informação da cobertura de vegetação arbórea de cada setor censitário em conjunto com as linhas de distribuição e verificar se as áreas de vegetação estão próximas às linhas de distribuição. A Figura 10 permite a visualização das linhas de transmissão em cada setor censitário.

A partir do arquivo contendo os setores censitários foi efetuado o recorte da classificação, no *software* QGIS, separando cada um dos setores censitários, representados numericamente de 0 a 303. Esse passo foi necessário para que posteriormente fosse possível calcular a porcentagem de vegetação presente em cada um dos setores censitários, para que em um segundo momento fossem analisados os setores com maior porcentagem de vegetação em conjunto com as linhas de distribuição.

Figura 10: Mapa de Linhas de Distribuição por Setor Censitário



Fonte: Elaborado pelo autor.

5.4.7 Relatórios quantitativos

Após o recorte da área a partir dos setores censitários, ainda no QGIS, utilizou-se a ferramenta de processamento *processing r.report* presente já nativamente no software, para quantificar a área de cada uma das classes, em cada setor censitário. Nesse relatório foi possível observar a divisão de áreas para cada classe presente no setor analisado, descrita como um valor numérico inteiro (0-Corpos d'Água; 1-Área Urbana; 2-Vias Terrestres; 3-Vegetação Arbórea; 4-Vegetação Rasteira/Cultura; 5-Solo Exposto). É observado que o relatório retorna valores somente para as classes presentes no setor censitário.

Após esta etapa, os dados obtidos por cada relatório foram exportados para uma planilha no Excel para tratamento da informação e análises quantitativas.

Com as informações das áreas das classes de cada setor, alinhadas e dispostas em formato de tabela, foi feito o cálculo de proporção de vegetação arbórea para o setor correspondente, obtendo assim as áreas com maior probabilidade de faltas na rede por conta de

vegetação.

Em seguida foram identificados na sobreposição do resultado da classificação, dos setores censitários e das linhas de distribuição, os setores censitários com maior probabilidade de interrupções na rede, ou seja, os setores com maiores proporções de vegetação arbórea por área.

Foi realizado um levantamento do período de novembro de 2023, através do portal da distribuidora do município (ENERGISA-SUL-SUDESTE - ESS), com o objetivo de determinar para esse período: a quantidade de interrupções na rede, de objetos retirados, clientes afetados e a duração média das interrupções.

É também utilizado para a análise um conjunto de dados, obtidos pelo portal da ANEEL, que informa a quantidade de interrupções na rede de distribuição, sendo ele dividido por interrupções programadas e não programadas. Com esse conjunto de dados e as informações levantadas, pode-se levar em conta o risco gerado pela falta de manutenção e poda preventiva em vegetações arbóreas.

6. RESULTADOS OBTIDOS

Este tópico é majoritariamente apresentado em formato de mapas, figuras e tabelas. Através do desenvolvimento metodológico foi possível gerar uma representação visual do uso e cobertura do solo nos setores censitários do município de Presidente Prudente-SP, produzida a partir de etapas da classificação supervisionada através do classificador *Random Forest*, e índices espectrais. A validação do resultado foi realizada por meio de índices como Matriz de Confusão, Kappa e Acurácia Global.

Após a realização da classificação, a Matriz de Confusão, o índice Kappa e a Acurácia Global foram obtidos, ainda no GEE, para que a qualidade da classificação pudesse ser verificada. A Tabela 1 apresenta os valores encontrados na Matriz de Confusão.

Tabela 1: Valores encontrados na Matriz de Confusão

Report	Corpo D'água	Área Urbana	Via terrestre	Vegetação Arbórea	Vegetação Rasteira/Cultura	Solo Exposto	Total	Erro de Inclusão (%)
Corpo D'água	46	0	0	0	0	0	46	0
Área Urbana	0	156	1	2	0	1	160	2,56
Via terrestre	0	3	52	0	0	0	55	5,77
Vegetação Arbórea	0	6	0	126	0	0	132	4,76
Vegetação Rasteira/Cultura	0	0	0	1	67	0	68	1,47
Solo Exposto	0	2	0	0	0	61	63	3,17
Total	46	167	53	129	57	62	524	
Erro de Omissão (%)	0	7,05	1,92	2,38	0	1,63		

Fonte: Elaborado pelo autor.

Analisando a Tabela 1 é possível verificar que as classes que apresentaram maior confusão foram as de “Vegetação Arbórea” e “Área Urbana”. Isso se dá pelo fato de haver vegetação na área urbana. Ainda é possível verificar que para as amostras obtidas, os valores são significativamente baixos, tendo como resultado uma boa classificação. É possível observar que as classes “Via Terrestre” e “Área Urbana” geraram maior confusão na amostra. Isso se dá em função de ambas possuírem respostas espectrais semelhantes. Essas confusões tiveram baixo impacto na qualidade da classificação, visto que representam apenas 3,8% e 1,52% de pixels, respectivamente.

O fato das classes amostrais serem definidas por pontos geométricos e não polígonos,

permite que as classificações atinjam maiores índices de acurácia, uma vez que é definido pixel por pixel.

A partir da Matriz de Confusão, obtiveram-se os dados do índice Kappa e Acurácia Global, e os resultados são apresentados na Tabela 2. Ambos os índices têm como objetivo representar o grau de confiabilidade dos dados classificados, logo, quanto mais próximo de 1 ou 100%, maior concordância os dados classificados têm com os dados de referência.

Tabela 2: Índices de Acurácia Global e Kappa

Índices	
Acurácia Global	0,969052243713723
Kappa	0,960315861182272

Fonte: Elaborado pelo autor.

Analisando a Tabela 2 identificamos que os valores obtidos tanto para Acurácia Global quanto para Kappa estão acima de 96% e de acordo com o Quadro 3, o resultado obtido no índice Kappa é classificado como “Excelente”, revelando que houve pouca confusão de pixels.

De posse dos dados da Matriz de confusão foram obtidos os valores de Acurácia do Produtor e Acurácia do Usuário. Na Tabela 3 são apresentados os valores obtidos para ambos.

Tabela 3: Acurácia Produtor

Acurácia Produtor	
Corpo D'Água	1
Áre Urbana	0.9831460674157303
Via Terrestre	0.9107142857142857
Vegetação Arbórea	0.9661016949152542
Vegetação Rasteira/Cultura	0,982758621
Solo Exposto	0,955223881

Fonte: Elaborado pelo autor.

Os valores de acurácia para cada classe, com exceção da “Via Terrestre”, foram superiores a 95%. A classe “Corpo d’Água” foi a mais bem identificada, apresentando valores de acurácia de 100%, tanto para a acurácia do produtor quanto para a acurácia do consumidor/usuário, como mostra a Tabela 4. A acurácia do produtor indica a precisão com

que a ferramenta, neste caso GEE, conseguiu identificar corretamente as classes nos dados analisados.

Segundo os dados apresentados, com exceção da classe “Via Terrestre”, a classificação atingiu valores de alta excelência para todas as classes, apresentando acurácia superior a 95%. O valor de 91% de acurácia para a classe “Via Terrestre” deve-se à confusão de pixels presente na matriz de confusão. Essa confusão se deve à presença de cidades de pequeno porte na região em que foi realizada a amostragem, o que dificulta a coleta de um grande número de pontos amostrais e também a presença de estradas não pavimentadas, que apresentam respostas espectrais semelhantes ao solo exposto. Pequenas cidades e áreas urbanas tendem a apresentar um valor menor de acurácia devido à sua natureza heterogênea, contendo desde vegetação até áreas de solo exposto, em uma área relativamente pequena.

A Tabela 4 apresenta os valores de Acurácia do Consumidor, que indicam uma resposta elevada e validam a excelente classificação obtida. Esses elevados valores de acurácia são atribuídos à qualidade da coleta de pontos amostrais realizada pelo operador.

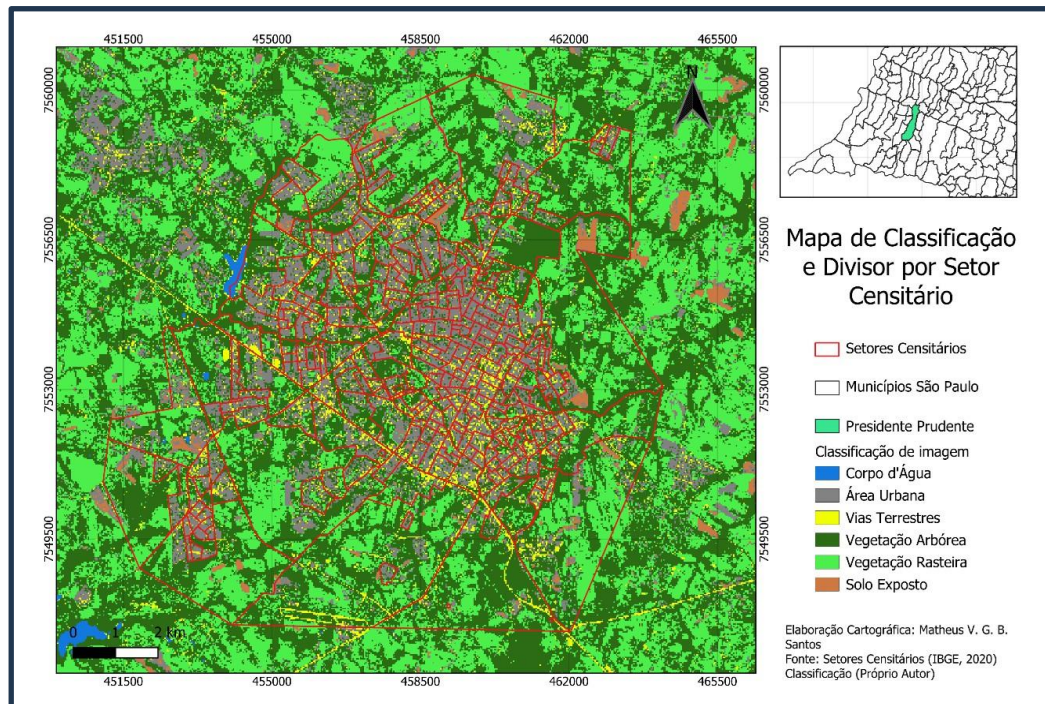
Tabela 4: Acurácia do Consumidor

Acurácia Consumidor	
Corpo D'Água	1
Áre Urbana	0.9510869565217391
Via Terrestre	0.9444444444444444
Vegetação Arbórea	0.9743589743589743
Vegetação Rasteira/Cultura	1
Solo Exposto	0.9846153846153847

Fonte: Elaborado pelo autor.

Após a etapa de validação da classificação foi exportado o resultado da classificação para o *software* QGIS. A Figura 11 apresenta o resultado da classificação supervisionada produzida no GEE, a partir do classificador *Random Forest* e das bandas espectrais B2, B3, B4, B5, B6, B7, B8, B11, B12 do satélite Sentinel 2A, do dia 13 de dezembro de 2023, e posteriormente tratada no QGis.

Figura 11: Mapa de uso e cobertura do solo.



Fonte: Elaborado pelo autor.

Ao analisarmos a Figura 11, observamos que grande parte da área total é coberta por vegetação arbórea. Dado este, que foi obtido precisamente através do relatório quantitativo do *r.report* da área total, conforme descrito na Figura 12.

Figura 12: Relatório da área total de cada uma das classes em hectares

RASTER MAP CATEGORY REPORT		
LOCATION: temp_location	Wed Jun 19 00:11:43 2024	
REGION	north: 7560340 south: 7547350 res: 30	east: 464190 west: 451050 res: 30
MASK: none		
MAP: (untitled) (rast_66724c6e6a1f75 in PERMANENT)		
Category Information		
#	description	hectares
-1	.	45.540
5	.	243.630
2	.	1010.880
1	.	3236.490
4	.	4797.720
3	.	7734.600
TOTAL		17,068.860

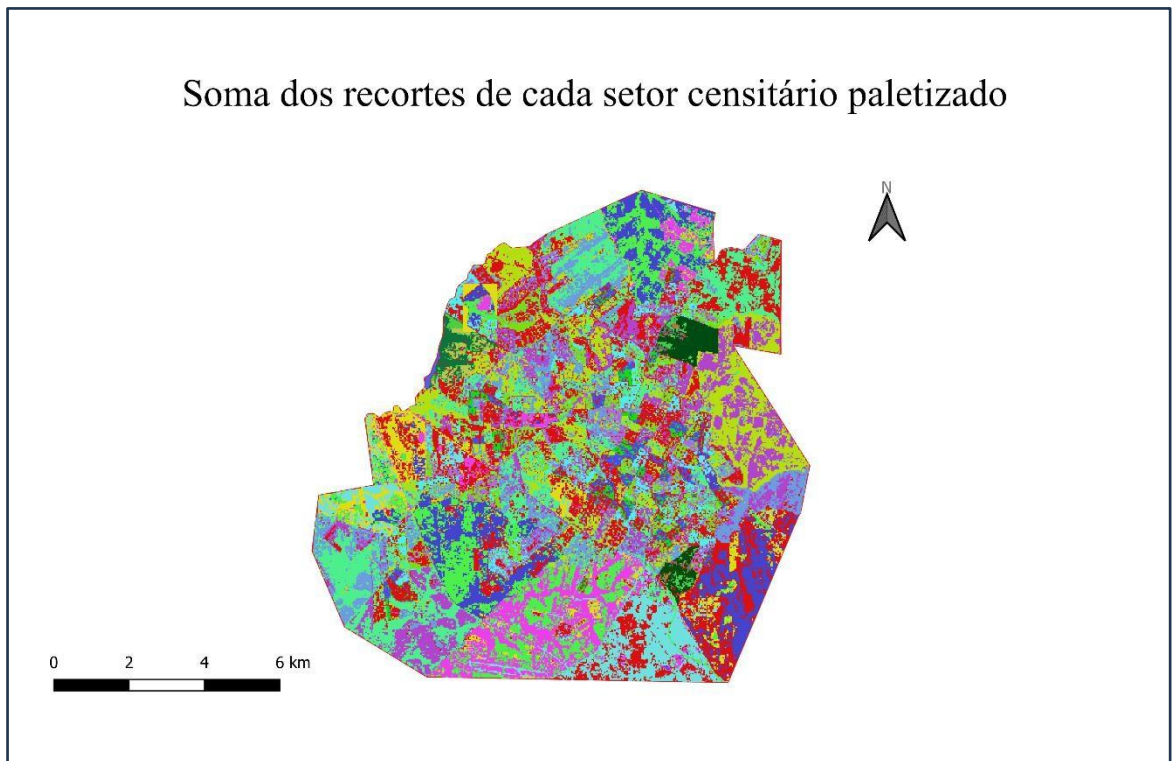
Fonte: Elaborado pelo autor.

Com base nos resultados apresentados na Figura 12 é possível observar a área total para classe de interesse, destacando as 3 maiores de forma decrescente:

- 3-Vegetação Arbórea;
- 4-Vegetação Rasteira/Cultura;
- 1-Área Urbana.

A seguinte etapa se dá pelo recorte e paletização, ou seja, a utilização de cores associadas a cada uma das classes, para cada um dos setores censitários, a fim de que seja possível analisar separadamente as regiões que mais estão sujeitas às interrupções. Na Figura 13 é apresentada a soma de todos os setores censitários que foram recortados anteriormente.

Figura 13: Soma dos recortes dos setores censitários previamente recortados



Fonte: Produzido pelo autor.

É observado que a paletização não segue um padrão de cores. Isso se dá pelo fato de que em setores censitários que não possuem todas as classes de classificação, o conjunto da paleta de cores se altera automaticamente.

Como a imagem já foi previamente classificada, é descartada a necessidade de alterar as cores das classes, colocando paletas específicas para todos os 300 setores censitários. A geração dos relatórios técnicos não sofre alteração alguma por conta da paleta de cores utilizada. Dessa forma, se faz entendido, por questões visuais, indicar a mesma coloração apenas para os 5 setores censitários que apresentaram maior proporção de vegetação arbórea por área total.

Após a separação e alteração da paleta de cores dos 5 setores censitários, foram emitidos os relatórios em arquivos *.txt* de todas as 300 áreas. Na sequência foram exportados os dados contidos em cada relatório para uma planilha do Excel, facilitando a análise de proporções.

Seguido do tratamento dos dados foi realizado o cálculo de proporção de vegetação arbórea por área total, a partir da razão entre a área de vegetação pela área total do setor censitário.

Com esses dados foi possível realizar a separação dos 5 setores, com ao menos 100.000 m² de área, que indicaram as maiores proporções de vegetação. A Tabela 5 indica os 5 setores que apresentaram esses resultados.

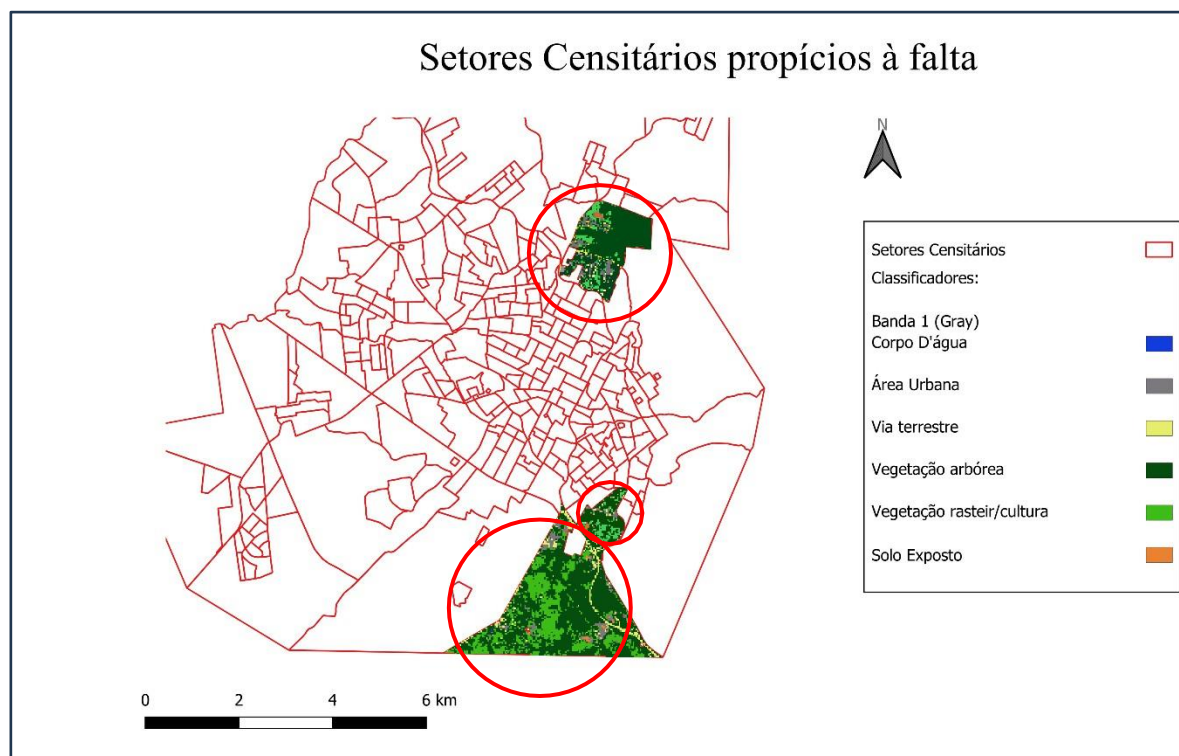
Tabela 5: Setores com maiores proporções arbóreas/área

Report	Corpo D'água	Área Urbana	Via terrestre	Vegetação Arbórea	Vegetação Rasteira/Cultura	Solo Exposto	Área do Setor	Porporção Arbórea/Área
Report 13		105.300	30.600	1.214.100	158.400	15.300	1.523.700	0,7968
Report 157		24.300	13.500	133.200	22.500		193.500	0,6884
Report 22		63.900	50.400	569.700	146.700	11.700	842.400	0,6763
Report 125		77.400	41.400	331.200	73.800	1.800	525.600	0,6301
Report 4	1.800	319.500	326.700	3.741.300	1.636.200	30.600	6.056.100	0,6178

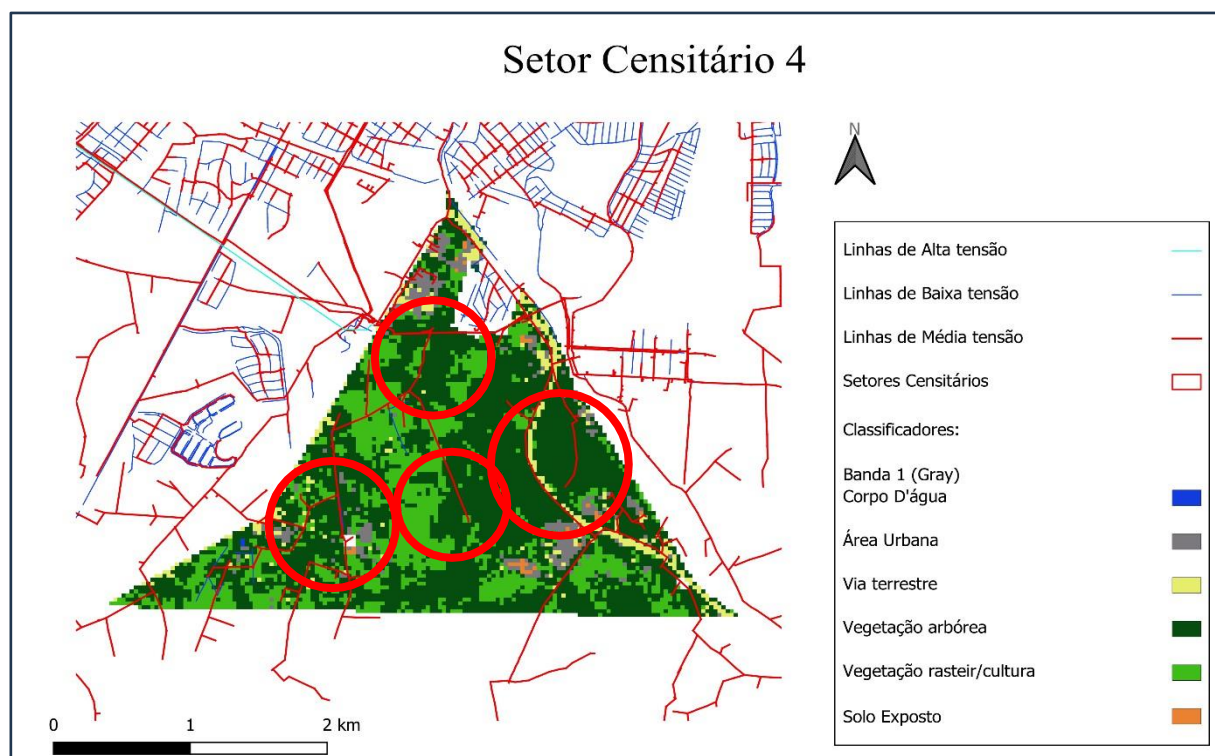
Fonte: Elaborado pelo autor.

A Figura 14 a seguir apresenta a localização dos 05 setores censitários (Tabela 6) com maior proporção de vegetação.

Após a análise dos setores destacados foi utilizado o arquivo das linhas de distribuição para que fossem verificadas se as áreas que possuem vasta vegetação arbórea e continham linhas passando por elas. A Figura 15 traz o resultado da sobreposição da classificação e das linhas de distribuição, utilizando o setor censitário 4 como exemplo. Este setor foi selecionado por possuir maior área entre os analisados, porém, o estudo foi feito para todas as amostras.

Figura 14: Setores Censitários sujeitos à falta

Fonte: Elaborado pelo autor.

Figura 15: Setor Censitário 4

Fonte: Elaborado pelo autor.

Como destacado na imagem (Figura 15), há diversas linhas de média tensão que passam por uma vasta vegetação densa, podendo causar faltas não programadas na rede, caso não haja o controle e manutenção dessa vegetação.

Realizada a análise e identificados os pontos de atenção, foi feito um levantamento do período de novembro de 2023, no portal da distribuidora do município (ENERGISA-SUL-SUDESTE - ESS), das interrupções que foram relatadas e notícias de reportagem indicando as causas de cada desligamento na rede em toda a região que a concessionária possui concessão, dentre elas Presidente Prudente. Esses dados nos informam que foram registrados:

- Retirada de 284 objetos estranhos da rede, dentre eles galhos de árvores, árvores entre outros;
- Reparo de 150 cabos partidos (MT/BT);
- Manutenção de 18 postes quebrados ou tombados.

Por conta desses danos, ocorreram 134.563 interrupções no período e 92.747 clientes afetados. Essas interrupções ocasionaram 418,46 minutos, em média, de falta de energia para esses clientes. As Figuras 16 e 17 trazem um recorte do relatório mensal no período, contendo evidências das faltas.

Figura 16: Evidência Fotográfica 1



Fonte: Fotos Coletadas Durante o Evento pela Distribuidora

Reportagens da imprensa local:



Fonte: Energisa Sul Sudeste (2023).

Figura 17: Evidência Fotográfica 2



Fonte: Energisa Sul Sudeste (2023).

Essas informações reforçam a necessidade do controle sobre a vegetação, assim também

apontando os riscos que podem ser causados pela falta do controle.

Através de um arquivo .csv referente às faltas de energia em território nacional, foram filtrados dados quantitativos que revelam a proporção desses desligamentos no município de Presidente Prudente durante o ano de 2023. Na Tabela 6 é indicado o número de interrupções separadas pelo nível de tensão.

Tabela 6: Interrupções por nível de tensão

Tensão (V)	Número de Interrupções
ESS	11G80
110	40
115	12
120	9
127	6119
220	3495
230	52
254	24
380	11
440	1
11400	2164
13800	28
34500	17
69000	1
138000	7
Total Geral	11G80

Fonte: Elaborado pelo autor.

É verificado que em linhas de média tensão (entre 600V e 69kV), ocorreram 2.192 interrupções, e para linhas de baixa tensão (até 600V), 9.763 interrupções. O total de faltas durante o período, apenas para Presidente Prudente, é de 11.980, sendo esse representado por faltas programadas e não programadas. Dentro de tais faltas apresentadas, não é informada a quantidade de interrupções por conta da vegetação arbórea ou quaisquer outros fatores, mas é possível estimar, além dos dados quantitativos já obtidos através da ANEEL com base nas evidências fotográficas, que uma parcela significativa ocorre devido tal vegetação.

7. CONCLUSÃO

A partir dos resultados obtidos da classificação das imagens espectrais e visualização das linhas de distribuição de alta, média e baixa tensão, e análises quantitativas referentes às interrupções na rede de distribuição da região de concessão da distribuidora Energisa Sul-Sudeste, esta que inclui Presidente Prudente, é possível afirmar que a vegetação arbórea é um Fator Gerador de grande impacto para situações de faltas elétricas na rede. Em função disso se faz necessário o mapeamento da vegetação arbórea para planejamento urbano e ações efetivas e preventivas no sistema de distribuição.

A partir dos resultados obtidos pode-se afirmar que a ferramenta Google Earth Engine (GEE) obteve um ótimo desempenho. O algoritmo de classificação utilizado (*Random Forest*) e os índices espectrais, permitiram que a classificação obtivesse maiores níveis de precisão ao selecionar os pixels. Através dos índices de qualidade (Kappa e Acurácia Global), gerados a partir da Matriz de Confusão, foi possível estimar quantitativamente a qualidade da classificação usada para esse estudo. Esses obtiveram valores acima de 96%, que segundo a literatura, classificam o mapeamento realizado com “excelente”.

Foi também exposto que as interrupções não programadas representam um total de 93% do volume total de faltas no sistema, sendo um de seus principais fatores geradores, causados por influência da vegetação arbórea, o qual afeta milhares de consumidores. Segundo o dado presente neste trabalho, 623.811 consumidores no período de 2023 sofreram por interrupções não programadas causadas pela vegetação arbórea na região de concessão da Energisa Sul-Sudeste, sendo este o quarto maior fator gerador de interrupções no período. O destaque vale também para Falhas em Equipamentos, que por disparidade se torna o fator com maior volume de interrupções.

O fato das interrupções não programadas causadas pela vegetação se apresentarem como o quarto maior gerador, reforça a importância de estudos a partir de imagens de sensoriamento remoto, classificadores e índices espectrais voltados a esse tema, uma vez que o processamento desses dados permite espacializar as áreas de vegetação arbórea e desse modo verificar a distribuição dessas regiões dentro da área de interesse. Auxiliando assim a tomada de decisões dos órgãos competentes.

Desta forma, é possível concluir que este trabalho serve, também, como material de estudo para monitoramento e controle da vegetação arbórea, sendo de extrema importância para o planejamento e desenvolvimento de determinada região.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Interrupções de energia elétrica 2022**. 2022. Disponível em: <<https://dadosabertos.aneel.gov.br/dataset/interruptoes-de-energia-eletrica-nas-redes-de-distribuicao/resource/7d081751-3f4c-4ede-96b5-1bf2e46b61ca>>. Acesso em: 02 mai. 2024.

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote sensing of environment**, v. 37, n. 1, p. 35-46, 1991.

CRÓSTA, A. P. **Processamento digital de imagens de sensoriamento remoto**. UNICAMP/Instituto de Geociências, 1999.

ENERGISA. **Relatório situação de emergência novembro/2023**. 2023. Disponível em: <https://www.energisa.com.br/Institucional/relatoriosituacaoemergencia/ISE_ESS_8-112023.pdf> . Acesso em 2 de maio de 2024.

FARIAS, A. P. S.; FONTANA, M. E.; MORAIS, D. C. Modelo de sistema de informação e decisão para intervenções de reabilitação em redes de distribuição de água. **RBRH - Revista Brasileira de Recursos Hídricos**, v. 18, n. 2, p. 55-65, 2013.

HELOU, L. F.; NOVAES, B. C.. Utilização da matriz de confusão na indicação de aparelho de amplificação sonora individual. **Distúrbios da Comunicação**, v. 17, n. 2, 2005.

HIRYE, M. C. M. **Estudos e técnicas de mapeamento de vegetação urbana**. São José dos Campos: INPE, 2020. 22 p. Disponível em: <<http://urlib.net/ibi/8JMKD3MGP3W34R/3UN44CH>>. Acesso em 16 de julho de 2024.

HUANG, S. et al . A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. **Journal of Forestry Research**, v. 32, n.1, p.1-6. 2021. Disponível em: https://www.researchgate.net/publication/341782370_A_commentary_review_on_the_use_of_normalized_difference_vegetation_index_NDVI_in_the_era_of_popular_remote_sensing#fullTextFileContent. Acesso em 16 de julho de 2024.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 01, p. 159-174, 1977.

LILLESAND, T. M.; KIEFER, R. W. **Remote sensing and image interpretation**. (S.L) Jhon Wiley and Sons, Inc, 1994.

LIMA, N. A.; NOGUEIRA, F. J.. Melhorias nos sistemas de distribuição de energia elétrica. **Caderno de Estudos em Engenharia Elétrica**, v. 4, n. 1, 2022.

TINNEY, L. R. Remote sensing and geographic information system data integration- Error sources and research issues. **Photogrammetric engineering and remote sensing**, v.57, n.6, p.677-687, 1991.

PRINA, B. Z.; TRENTIN, R. GMC: Geração de Matriz de Confusão a partir de uma classificação digital de imagem do ArcGIS®. In: SIMPÓSIO BRASILEIRO DE

SENSORIAMENTO REMOTO, 17., 2015, São José dos Campos. **Anais [...]**. São José dos Campos: SBSR, 2015. p. 137.

RIGATTI, S. J. Floresta Aleatória. **J Insur Med.** 2017; v. 47 n. 1, p. 31–39, doi: <https://doi.org/10.17849/insm-47-01-31-39.1>

ROUSE, J. H. **Monitoring Vegetation Systems in the Great Plains with ERTS. In: Proceedings.** Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, 1974. 1974. p. 3010- 3017. (modelo evento parte monografia)

SANTOS, W. C. dos. **Identificação de faltas de alta impedância em sistemas de distribuição.** 117 f. Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Centro de Engenharia Elétrica e informática, Universidade Federal de Campina Grande, Paraíba, Brasil, 2016.

SOUZA, E.. **Índices de confiabilidade devido a vegetação e planejamento de podas de árvores em redes de distribuição.** 2013. 225. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho Faculdade de Engenharia de Ilha Solteira, 2013.

SOUZA, JOSIANE APARECIDA CARDOSO et al. Análise de medidas de acurácia em processos de classificação baseadas na matriz de confusão. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 2019, Limeira. **Anais eletrônicos...** Campinas: Galoá, 2019. Disponível em: <<https://proceedings.science/sbpo/sbpo-2019/trabalhos/analise-de-medidas-de-acuracia-em-processos-de-classificacao-baseadas-na-matriz?lang=pt-br>>. Acesso em: 16 jul. 2024.

STROBL, C.; MALLEY, J.; TUTZ, G.. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. **Psychological methods**, v. 14, n. 4, p. 323, 2009.

TRONCHONI, A. B.. **Identificação de causas de desligamentos não programados em redes de distribuição.** Dissertação de Mestrado. Pontifícia Universidade Católica do Rio Grande do Sul, 2016.

VILELA JUNIOR, G. B. et al. **Métricas utilizadas para avaliar a eficiência de classificadores em algoritmos inteligentes.** Centro de Pesquisas Avançadas em Qualidade de Vida. 14. 1. 10.36692/v14n2-01R. Disponível em: <https://www.researchgate.net/publication/359541310_metricas_utilizadas_para_avaluar_a_eficiencia_de_classificadores_em_algoritmos_inteligentes>. Acesso em 16 de julho de 2024.

XU, H. Modification of normalized 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, 2006.

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, 2003.