



**Universidade Estadual Paulista “Júlio de Mesquita Filho”  
Campus Experimental de Rosana  
Engenharia de Energia**

**Victor Siqueira Ferreira**

**Mapeamento do uso e cobertura do solo de uma região do Pontal do Paranapanema  
para estimativa da produção de biocombustível**

Rosana - SP  
2022

**Victor Siqueira Ferreira**

**Mapeamento do uso e cobertura do solo de uma região do Pontal do Paranapanema  
para estimativa da produção de biocombustível**

Trabalho de Conclusão de Curso apresentado à  
Coordenadoria de Curso de Engenharia de  
Energia do Campus Experimental de Rosana,  
Universidade Estadual Paulista, como parte dos  
requisitos para obtenção do diploma de  
Graduação em Engenharia de Energia.

Orientador(a): Dra. Letícia Sabo Boschi

Rosana – SP  
2022

F383m           Ferreira, Victor Siqueira  
                  Mapeamento do uso e cobertura do solo de uma região do Pontal do  
Paranapanema para estimativa da produção de biocombustível / Victor  
Siqueira Ferreira. -- Rosana, 2022  
                  57 p.

Trabalho de conclusão de curso (Bacharelado - Engenharia de  
Energia) - Universidade Estadual Paulista (Unesp), Faculdade de  
Engenharia e Ciências, Rosana  
Orientadora: Letícia Sabo Boschi

1. Mapeamento do uso e cobertura do solo. 2. Google Earth Engine.  
3. Biocombustíveis. I. Título.

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

Essa ficha não pode ser modificada.



**Universidade Estadual Paulista “Júlio de Mesquita Filho”**

**Campus Experimental de Rosana**

**Engenharia de Energia**

**Victor Siqueira Ferreira**

ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO  
PARTE DO REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE  
“GRADUADO EM ENGENHARIA DE ENERGIA”

APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE  
GRADUAÇÃO EM ENGENHARIA DE ENERGIA

Prof. Dr. José Francisco Resende da Silva  
Coordenador

**BANCA EXAMINADORA:**

Profª. Dra. Leticia Sabo Boschi  
Orientadora/UNESP-FEC

Profª. Dra. Andréia Fátima Zanette  
UNESP-FEC

Prof. Dr. Guilherme Henrique Barros de Souza  
UNESP-FCT

Janeiro de 2022

## **AGRADECIMENTOS**

Em primeiro lugar agradeço a minha família, pelo incentivo durante toda a graduação. Obrigado por me ampararem até aqui e por me ajudarem a formar a pessoa que sou hoje. Sou grato a tudo que fazem e já fizeram por mim.

Agradeço à minha orientadora Prof.<sup>a</sup> Dr.<sup>a</sup> Leticia pela ajuda, paciência e aprendizado durante os anos em que trabalhamos juntos. Por todo suporte e incentivo em adquirir conhecimentos na área de cartografia, agradeço a oportunidade de trabalharmos juntos tanto neste trabalho como em outros trabalhos que realizamos.

Agradeço principalmente aos meus amigos Beatriz Resende, Renan Mendes, Patrick Allan, Júlia Vilalva, Eduardo Brunhara, Julia Alejandra, Thalita Botuem, Samantha Nogueira, Vinicius Leal, Mariana Valentim, Alice Gomes e Vitória Santos, obrigado por estarem comigo sempre ao meu lado durante toda minha graduação, por todo acolhimento e suporte emocional que me deram e todos os momentos que tivemos. Com certeza a minha graduação teria sido muito diferente sem o suporte que criamos uns com os outros, aprendi muito com todos vocês compartilhando os momentos bons e ruins dessa fase.

Agradeço também a todos os amigos, professores e funcionários da UNESP que fizeram parte da minha graduação, por me auxiliarem em toda a evolução e aprendizado durante todos esses anos.

“Cada sonho que você deixa pra trás, é uma parte do seu futuro que deixa de existir.”

Steve Jobs

## RESUMO

Atualmente o etanol vem despontando como uma fonte de geração de energia no Brasil e no mundo. Devido a isso, desde o início do programa Proálcool há incentivo à expansão de lavouras de cana-de-açúcar pelo estado de São Paulo. Nessa expansão, a região do Pontal do Paranapanema vem presenciando uma crescente alteração do uso e cobertura do solo, ocorrendo assim, a substituição de áreas de pastagem por plantação de cana-de-açúcar. Diante do exposto, o objetivo do trabalho é estimar a produção de biocombustíveis, em específico o etanol de cana-de-açúcar, para uma pequena região do Pontal do Paranapanema. A estimativa foi realizada a partir de áreas de pastagens extraídas de imagens multiespectrais do satélite Sentinel 2A através do processo de classificação supervisionada. Para realizar a estimativa foi produzido um mapa de uso e cobertura do solo da região a partir da classificação supervisionada realizada no *Google Earth Engine* (GEE), através do algoritmo *Random Forest* (RF) e dos índices NDVI, NDBI, MNDWI e RS, que auxiliaram na melhora da classificação. O GEE demonstrou ser uma ferramenta versátil na análise de dados para identificação do uso e cobertura do solo e a classificação produzida teve uma acurácia excelente, com valores de Índice Kappa de 99,7% e Acurácia Global de 99,78%. Verificada a qualidade do mapeamento foram calculados os valores aproximados de área em hectare para cada uma das classes. A partir do valor de área estimado para a classe de pastagem, da relação da Conab que diz que 1 tonelada de cana-de-açúcar produz de 80 a 90 litros de etanol e levando em conta as perdas inerentes ao processo, foi estimado o potencial aproximado de produção de biocombustível que esse cenário de substituição da pastagem por cultura de cana-de-açúcar poderia gerar na região. Ao atribuir o valor de 196.416,71 hectares de área de pastagem, com uma perda de 14,14%, prevê-se um potencial de produção entre 944.402.968,4 a 1.062.453.339 litros de etanol. Segundo os resultados obtidos é possível afirmar que a área apresenta um enorme potencial de produção de biocombustíveis. A região ainda se encontra pouco explorada e pesquisas como essa trazem comprovações plausíveis de atrativa para investimentos no setor, favorecendo a economia e a população da dos municípios.

**PALAVRAS-CHAVE:** Mapeamento do uso e cobertura do solo. Google Earth Engine. Biocombustíveis.

## ABSTRACT

Currently, ethanol is emerging as a source of energy generation in Brazil and in the world. Because of this, since the beginning of the Proálcool program, there has been an incentive to expand sugarcane plantations throughout the state of São Paulo. In this expansion, the Pontal do Paranapanema region has been witnessing a growing change in land use and cover, thus resulting in the replacement of pasture areas by sugarcane plantations. Given the above, the objective of this work is to estimate the production of biofuels, specifically sugarcane ethanol, for a small region of Pontal do Paranapanema. The estimation was carried out from pasture areas extracted from multispectral images from the Sentinel 2A satellite through a supervised classification process. To perform the estimate, a land use and land cover map of the region was produced from the supervised classification carried out in Google Earth Engine (GEE), through the Random Forest (RF) algorithm and the NDVI, NDBI, MNDWI and RS indices, which helped to improve the rating. The GHG proved to be a versatile tool in data analysis to identify land use and land cover and the classification produced had excellent accuracy, with Kappa Index values of 99.7% and Global Accuracy of 99.78%. After checking the quality of the mapping, the approximate values of area in hectare were calculated for each of the classes. Based on the estimated area value for the pasture class, the Conab list that says that 1 ton of sugarcane produces 80 to 90 liters of ethanol and considering the losses inherent to the process, the potential approximate amount of biofuel production that this scenario of replacing pasture with sugarcane could generate in the region. By assigning the value of 196,416.71 hectares of pasture area, with a loss of 14.14%, a production potential of between 944,402,968.4 and 1,062,453,339 liters of ethanol is predicted. According to the results obtained, it is possible to affirm that the area has an enormous potential to produce biofuels. The region is still little explored and research like this brings plausible evidence of attractiveness for investments in the sector, favoring the economy and population of the municipalities.

**KEYWORDS:** Land Use Classification. Google Earth Engine. Biofuels.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Preço do barril de Petróleo Mundial em Dólares (US\$)	16
Figura 2 – Mapa com a localização das usinas dos municípios	22
Figura 3 – Espectro Eletromagnético	24
Figura 4 – Mapa do Estado de São Paulo destacando os municípios escolhidos	28
Figura 5 – Fluxograma das Atividades Realizadas na Metodologia	32
Figura 6 – Recorte da área de estudo com os polígonos da coleta de pontos amostrais	35
Figura 7 – Fluxograma de operação do <i>Random Forest</i>	38
Figura 8 – Exemplificação da Matriz de Confusão	40
Figura 9 – Fluxograma do processo de produção do etanol com as perdas elencadas	44
Figura 10 – Mapa de uso e cobertura do solo	47

## LISTA DE TABELAS

Tabela 1 – Usinas dos Municípios Utilizados	21
Tabela 2 – Índices Espectrais	26
Tabela 3 –Características das Bandas do Sentinel 2A	31
Tabela 4 –Classes Utilizadas na Elaboração da Classificação	34
Tabela 5 – Associação de qualidade de classificação com Valores de Kappa	41
Tabela 6 – Principais Perdas Consideradas	46
Tabela 7 –Matriz de Confusão	48
Tabela 8 – Índices Kappa e Acurácia Global	49
Tabela 9 – Valores de Acurácia do Produtor (ferramenta)	49
Tabela 10 – Valores de Acurácia do Usuário	50
Tabela 11 –Valores em Hectare Encontrados	51

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO</b>	15
<b>2.</b>	<b>OBJETIVO</b>	19
<b>3.</b>	<b>JUSTIFICATIVA</b>	20
3.1.	Indústrias na Região	21
<b>4.</b>	<b>LEVANTAMENTO BIBLIOGRÁFICO</b>	23
4.1.	Sensoriamento Remoto	23
4.2.	Classificação	25
4.3.	Índices Espectrais	25
4.4.	Etanol	26
<b>5.</b>	<b>MATERIAL UTILIZADO</b>	28
5.1.	Área de Estudo	28
5.2.	Softwares Utilizados	29
5.3.	Imagens Sentinel 2	30
5.4.	Metodologia	31
5.4.1.	Pré-Processamento dos Dados	33
5.4.2.	Processamento	33
5.4.3.	Índices Espectrais Utilizados	35
5.4.4.	Aplicação do Algoritmo <i>Random Forest</i>	38
5.4.5.	Geração do Índice Kappa, Acurácia Global e Matriz de Confusão	39
5.4.6.	Tratamento dos Dados no QGis	42
5.4.7.	Estimativa de Área	43
<b>6.</b>	<b>RESULTADO E DISCUSSÃO</b>	47
<b>7.</b>	<b>CONCLUSÃO</b>	52
	<b>REFERÊNCIAS</b>	53

## 1 . INTRODUÇÃO

O Brasil possui uma forte presença da cana-de-açúcar em sua história, iniciando no século XVI com o intuito de produzir açúcar. Com o passar dos séculos, o país se tornou um grande produtor de açúcar, e esteve por muitos anos entre os maiores produtores de açúcar para exportação, promovendo assim, a expansão de fazendas por todo o país (GROFALO; PACKER; RAMOS; KONDO; FOLEGATTI; CABRAL, 2020).

Durante a década de 30 já havia a produção do etanol, mas sem representatividade expressiva na matriz energética brasileira, pois o etanol anidro produzido era destinado ao uso de aditivo da gasolina. Um enorme marco do século XX ocorreu em 1973, quando houve um aumento do preço do barril de petróleo devido a um embargo econômico causado pelos países do Golfo Pérsico como forma de protesto contra apoiadores da Nação de Israel. Com esse embargo o mundo sofreu com o aumento do preço do barril que foi de US \$3,00 para US \$12,00 (valores da época). Em função desse aumento a economia mundial sofreu uma enorme crise econômica (SZMRECSÁNYI, 1979).

Depois desse episódio, as nações mundiais começaram a investir em novas tecnologias no campo de combustíveis. Como resposta à crise mundial do petróleo da década de 70, o governo passou a incentivar a produção de veículos movidos a etanol, com o objetivo de diminuir a dependência do petróleo importado, que representava aproximadamente 80% do petróleo consumido no país. O objetivo principal do PNA (Programa Nacional do Álcool) era diminuir a dependência de combustíveis importados, favorecer a criação e o fortalecimento de uma indústria nacional de produção de combustíveis. Em 1975 passam a operar os incentivos à produção de etanol com a intenção de substituir derivados de petróleo por um combustível produzido internamente. Esses incentivos modificaram a infraestrutura industrial do país atingindo diretamente o setor de combustíveis e com isso também a indústria automotiva. Em 1978 os primeiros automóveis movidos 100% a etanol passaram a surgir no mercado, que viu durante o programa Proálcool, uma mudança na frota nacional. Em 1979 0,45% dos veículos eram movidos a etanol, e em 1986 a participação do álcool hidratado como combustível principal passou a ser de 76,1% (KOHLHEPP, 2010).

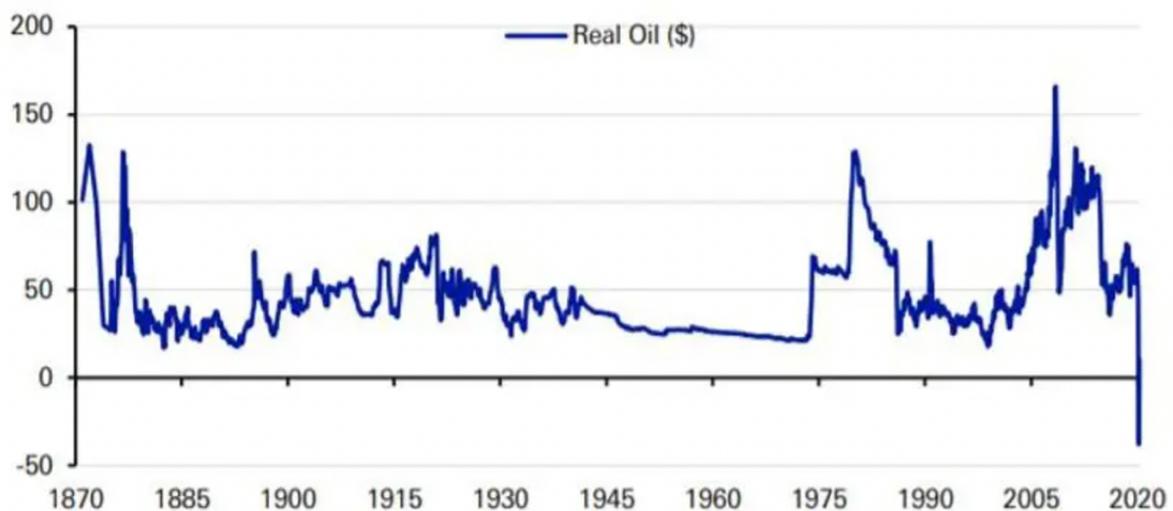
Durante as décadas de 80 e 90 o etanol encontrou uma grande aceitação no país. A produção alcooleira cresceu de 600 milhões de litros no ano de 1976 para 3,4 bilhões de litros no ano de 1980. Nos anos 80 devido a uma segunda crise petrolífera mundial, os valores do barril de petróleo dispararam novamente, nesse momento, o petróleo importado representava aproximadamente 46% de participação de mercado. Em 1987 o Brasil viu sua produção de

etanol atingir um dos maiores valores da indústria nacional, com uma produção de 12,3 bilhões de litros de etanol no ano (SOARES; ZUKOWSKI JUNIOR., 2021).

A partir de 1990 o preço do barril do petróleo passou a ter uma maior estabilidade de mercado, com isso houve um aumento do uso de motores a combustão na frota nacional, mas a partir de 2003 com o advento da tecnologia *Flex Fuel Vehicles* nos veículos nacionais, os quais utilizavam os dois combustíveis para o funcionamento do motor, houve a necessidade de se manter alta a produção desse biocombustível. (COSTA, 2019).

De acordo com as informações divulgadas pelo *Deutsche Bank* (2020) e apresentadas na Figura 1, é possível notar a volatilidade do preço do barril no mercado ao longo das últimas décadas.

**Figura 1: Preço do barril de petróleo mundial em dólares (US\$).**



Fonte: Deutsche Bank, Global Financial (2020).

Globalmente falando o principal combustível principal que compõe a matriz energética mundial são os combustíveis fósseis, provindos do carvão mineral, gás natural e derivados do petróleo, o uso desses combustíveis vem despertando uma preocupação global devido a alta emissão de gases poluentes e a degradação ambiental que a atividade de extração causa. Hoje em dia há uma preocupação mundial referente às mudanças climáticas causadas por interferência humana no meio ambiente, onde o aumento de temperatura que em grande parte vem sendo ocasionado pelo efeito estufa e o aumento expressivo da temperatura global, causados majoritariamente pelo uso de combustíveis fósseis como carvão mineral, petróleo e o gás natural, devido a isso estão ocorrendo diversas movimentações e tentativas da redução da emissão desses gases (THE ECONOMIST, 2021).

Segundo a EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária) (2020), o etanol provindo de processo destilação do caldo de cana-de-açúcar se torna uma abordagem promissora dentro do cenário ambiental mundial, por se tratar de uma fonte renovável e em termos de emissão de gases poluentes, quando comparado com a gasolina, se torna um excelente combustível, devido a sua baixa emissão. Com a substituição do combustível provindo do refino do óleo bruto para um biocombustível, e o uso de veículos movidos a álcool, as emissões de  $CO_2$  na atmosfera foram reduzidas em 73%, quando comparados à emissão dos veículos a combustão (GROFALO; PACKER; RAMOS; KONDO; FOLEGATTI; CABRAL, 2020).

Com o enrijecimento das políticas de proteção ambiental a fim de diminuir a emissão dos gases de efeito estufa, o incentivo e a procura por combustíveis menos poluentes vem aumentando. Em 2021, entre os dias 31 de outubro e 12 de novembro, ocorreu em Glasgow a COP26 (*Conference of the Parties*). A Conferência das Partes da Convenção-Quadro das Nações Unidas de Mudanças Climáticas reuniu líderes mundiais com o objetivo de discutir problemas globais atuais, com enfoque na diminuição em 45% das emissões globais de  $CO_2$  até 2030, buscando manter a meta de limitar o aquecimento global em 1,5 °C até 2100, proposto pelo acordo de Paris. Durante a conferência foram discutidas políticas globais de diminuição e políticas com o objetivo de zerar as emissões de dióxido de carbono até 2050, a fim de evitar o aumento da temperatura global. Devido a essas políticas globais estima-se que os biocombustíveis terão mais incentivos e investimentos para substituição do petróleo como principal combustível nas frotas de veículos mundial (THE ECONOMIST, 2021).

As empresas sucroalcooleiras estão expandindo ou migrando para regiões mais interioranas do estado de São Paulo devido a incentivos fiscais para criação de empregos e industrialização de regiões majoritariamente agrárias, devido a crescente industrialização dos principais polo sucroalcooleiros como Campinas e Ribeirão Preto, a indústria do etanol passou a realizar uma expansão para o oeste do estado (ANDRADE, CARVALHO E SOUZA, 2009). O Pontal do Paranapanema historicamente possui uma característica predominante de atividades de agropecuária de corte, pois na década de 40 passou por uma ampla expansão da indústria de agropecuária de gado de corte, ocorrendo uma expansão do desmatamento da região, substituindo a vegetação nativa por fazendas de corte para suprir o mercado de carnes que começava a crescer no bioma do cerrado brasileiro, tendo uma interiorização dentro do território nacional. Junto a isso houve a expansão de regiões de cultura, principalmente de algodão, mandioca e amendoim, durante as primeiras décadas de expansão agrícola nos municípios (BENTO, 2020).

No intervalo compreendido entre os anos de 2002 e 2012, a região do Pontal do Paranapanema apresentou mudanças expressivas na paisagem, frente à disseminação de áreas canavieiras. As incorporações e implantação de novas agroindústrias, e o crescimento das culturas da cana-de-açúcar, foram os principais responsáveis por estas modificações. Nos municípios em que as terras se destinavam a pastagens/pecuária ou a outras lavouras como milho, soja, amendoim, e outras, passaram a ser dedicadas à produção da cultura canvieira (BARRETO e THOMAZ JR, 2012).

Atualmente, a região vem presenciando uma expansão da ocupação do solo por plantações de cana-de-açúcar, devido o bioma cerrado ser um ambiente adaptável ao plantio, favorecendo a migração de empresas sucroalcooleiras na região (BENTO, 2020).

No mercado agropecuário a cana-de-açúcar apresenta um alto retorno financeiro devido ao processo de produção de açúcar e álcool ser relativamente baixo em valores de produção, chegando a valores inferiores a US \$200/toneladas. E na produção de energia é considerada atualmente uma das fontes de produção com um amplo potencial de crescimento. De acordo com o Balanço Energético Nacional (BEM) a energia provinda de biocombustíveis possui uma capacidade instalada nacional de 10,8 GW compondo em 8,2% de participação na matriz energética brasileira (EMPRESA DE PESQUISA ENERGÉTICA, 2021).

## 2 . OBJETIVO

O objetivo geral do trabalho é estimar a produção de biocombustíveis, em específico o etanol de cana-de-açúcar, para uma pequena região do Pontal do Paranapanema. A estimativa foi realizada a partir de áreas de pastagens extraídas de imagens multiespectrais do satélite Sentinel 2A através do processo de classificação supervisionada. Com base nesse objetivo, foram delineados os seguintes objetivos específicos:

Utilizando da classificação supervisionada extrair áreas de pastagens de imagens multiespectrais do satélite sentinel 2A.

- Produzir um mapa de uso e cobertura da área a partir de uma classificação supervisionada realizada no Google Earth Engine, utilizando o algoritmo Random Forest para levantamento de valores em hectare de cada classe;
- Estimar a produção aproximada de biocombustível nas áreas de pastagens identificadas através da classificação, que seriam possíveis áreas de interesse para a expansão da indústria sucroalcooleira.

### 3 . JUSTIFICATIVA

Quando se fala em biocombustíveis, o Brasil é uma das referências mundiais no assunto, principalmente em relação à cana-de-açúcar, por isso o setor vem incentivando cada vez mais a indústria sucroalcooleira. Devido aos grandes investimentos das últimas décadas, as fazendas produtoras de cana-de-açúcar se espalharam por todo o país, aumentando o número de hectares destinados ao plantio. Devido a importância que o país vê em investir nesse combustível, conseguimos atingir a marca de 30% da produção mundial de etanol (GROFALO; PACKER; RAMOS; KONDO; FOLEGATTI; CABRAL, 2020).

Os quatro municípios da região do Pontal do Paranapanema foram selecionados devido a presença de usinas sucroalcooleiras na região, que está sofrendo uma expansão canavieira intensa, junto com o interesse das indústrias da região em encurtar as distâncias e transporte da matéria-prima para as fábricas (REZENDE, 2011).

Essa expansão profunda e intensificada pela rapidez da adesão de investidores na região do Oeste Paulista trouxeram a necessidade de um plantio em grande escala que é atraído por muitos fatores. Dentre eles:

- O relevo propício à mecanização da produção, pois possui um relevo pouco acidentado, podendo ser facilmente corrigido através de curvas de nível, tornando a área favorável ao plantio;
- O tipo de solo que é propício à produção de uma cana-de-açúcar com alto teor de sacarose, adequado à produção de açúcar;
- O valor do metro quadrado da região é baixo quando comparado com outras regiões do estado;
- A logística favorável que a região possui, seja para transporte de insumos e maquinários, utilizados principalmente para processos produtivos. A região possui rodovias (através da SP-270/Raposo Tavares), ferrovias e hidrovias (o potencial navegável de ambos os rios Paraná e Paranapanema).

Além desses aspectos físicos, algumas medidas políticas incentivam, ainda que indiretamente, a produção se estender ao Pontal, inclusive medidas legislativas (Lei Estadual 11.241 / 02) que proíbe as queimadas nos canaviais, levando à mecanização, favorecendo a região a adotar a expansão e ser tornar alvo de investimentos (BARRETO e THOMAZ JR, 2012).

### 3.1. Indústrias na Região

Para essa monografia foram selecionadas 4 usinas de produção de açúcar e etanol, localizadas na área de estudo selecionada. A Tabela 1 apresenta cada uma das usinas, os respectivos municípios e os grupos investidores e a Figura 2 apresenta a distribuição dessas usinas na região.

**Tabela 1:** Usinas dos municípios que compõem a área de estudos.

Nome da Usina	Município	Grupo Investidor Responsável
Alcídia	Teodoro Sampaio – SP	ATVOS (Antiga Odebrecht)
Conquista do Pontal	Mirante do Paranapanema -SP	ATVOS (Antiga Odebrecht)
Paranapanema II	Sandovalina – SP	UMOE BioEnergy
Usina Decasa	Marabá Paulista – SP	Grupo Olival Tenório

Fonte: Elaborado pelo autor.

**A Destilaria Alcídia S/A** localizada em Teodoro Sampaio – SP sob a responsabilidade da empresa ATVOS pertencente ao Grupo Novonor (antigo Grupo Odebrecht Agroindustrial), tem enfoque na produção de Açúcar VHP (Açúcar *Very High Polarization*), Energia Elétrica, Etanol de Cana Hidratado. Essa usina foi inaugurada durante o período do Governo Militar do Presidente Ernesto Geisel, sendo fruto da primeira fase do Programa Proálcool. Em 1978 iniciou a produção do etanol. Com uma capacidade máxima de moagem de 4 milhões de toneladas de cana por ano (ATVOS, 2021).

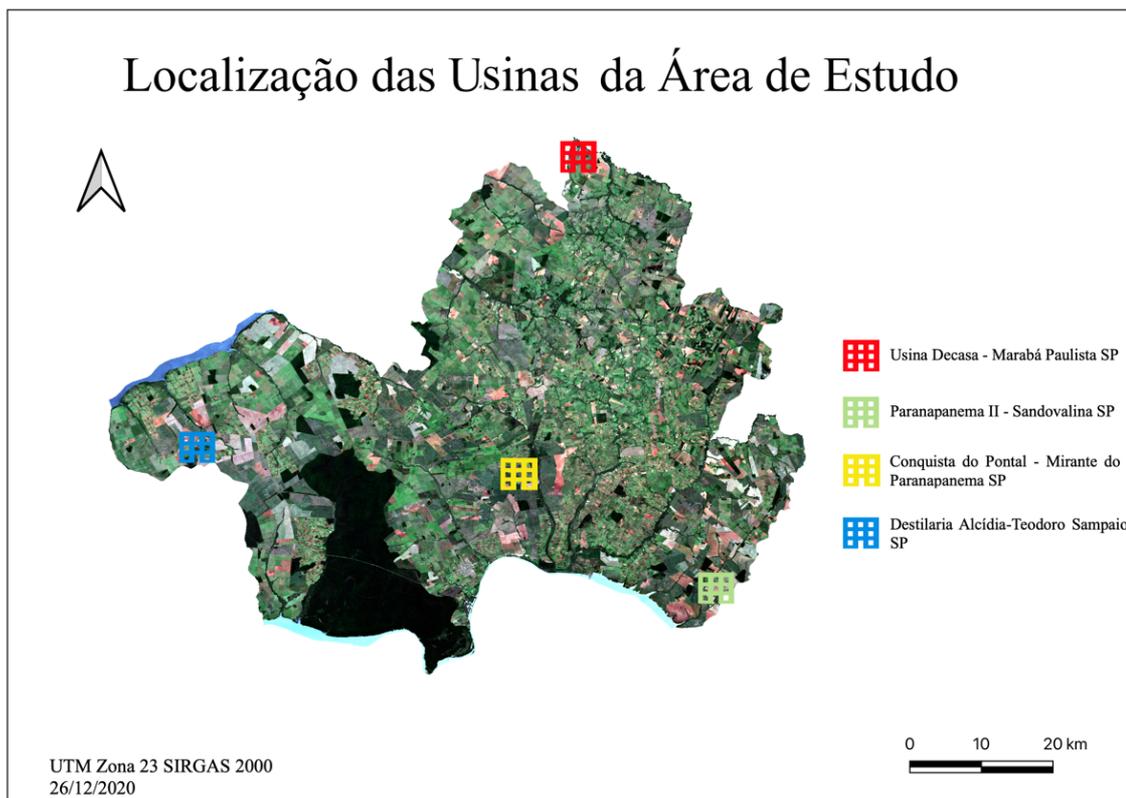
**A Usina Conquista do Pontal S/A** localizada no município de Mirante do Paranapanema – SP, sob a responsabilidade da empresa ATVOS pertencente ao Grupo Novonor (antigo Grupo Odebrecht Agroindustrial). Com enfoque na produção de Açúcar VHP (Açúcar *Very High Polarization*), Energia Elétrica, Etanol de Cana Anidro, Etanol de Cana Hidratado. Fundada em 2004, a Conquista do Pontal pertence ao grupo de usinas que vieram para suprir a expansão recente do de etanol devido a popularização dos veículos Flex Fuel. Possuindo uma capacidade máxima de moagem de 4 milhões de toneladas de cana por ano (ATVOS, 2021).

**A Destilaria Paranapanema II** localizada no município de Sandovalina – SP, sob direção da *Umo BioEnergy* pertencente ao Grupo norueguês Umo AS, fundada em 2006 com enfoque em suprir a demanda de produção de etanol hidratado e anidro. A Umo possui faz uso de 100% de colheita mecanizada em todas as suas áreas cultiváveis que compreendem 44 mil

hectares de área destinada a plantio. A planta de Sandovalina possui uma capacidade de moagem de 2,8 milhões de toneladas de cana por ano (UMOEBIOENERGY, 2021).

A **Destilaria Decasa Açúcar e Álcool** localizada no município de Marabá Paulista, pertencente atualmente ao Grupo Olival Tenório. Foi inaugurada em 1982 durante os investimentos do Projeto Governamental do Proálcool sob controle do Grupo Jacinto e Associados, em 1998 foi declarado falência ficando inativa por 2 anos. Em 2000 passa para o grupo pernambucano Arraes que opera a planta até 2004 e vende para o Grupo Olival Tenório. Em março de 2021 houve a interrupção da produção devido ao decreto de falência e a usina foi colocada à venda, em outubro do mesmo ano houve o início das novas operações. A planta de Marabá possui uma capacidade de moagem de 1,6 milhões de toneladas de cana por ano (NOVACANA, 2021).

**Figura 2:** Mapa com a localização das usinas dos municípios.



Fonte: Elaborado pelo próprio autor.

## 4. LEVANTAMENTO BIBLIOGRÁFICO

### 4.1. Sensoriamento Remoto

O sensoriamento remoto atua como uma das principais ferramentas para monitoramento e estudos da superfície terrestre, auxiliando às diversas pesquisas nos campos das geociências, trazendo como principal vantagem a rápida execução e aquisição de dados, pelo menor custo atrelado ao processo se comparado aos métodos com observação *in loco*. Através de sensores há a possibilidade de se observar e analisar a superfície e a atmosfera terrestre, a partir dos dados numéricos adquiridos. Os sensores embarcados nos satélites são capazes de coletar os valores da energia refletida ao atingir um objeto, convertendo-a em um sinal passível de ser registrado e convertê-lo para uma informação numérica e visível (JENSEN, 2009; FORMAGGIO, 2017).

A visão humana não está capacitada a detectar diferenças sutis nos objetos quando se trata de comprimento de onda, devido a isso houve a necessidade da criação de ferramentas para investigar valores de energia magnética e eletromagnética, presente da reflexão da energia propagada do espaço e atingir a superfície. A energia produzida pelo Sol chega à Terra na forma de um campo elétrico magnético variável com raios perpendiculares entre si. Uma parte dessa radiação eletromagnética que chega à Terra é visível ao olho humano, que é a chamada luz visível. Porém, dentro do espectro eletromagnético existem outras formas de radiação eletromagnética, como a radiação infravermelha, ultravioleta, raio-x, ondas de rádio e micro-ondas as quais não é possível a visualização a olho nu (SEARS; YOUNG; ZEMANSKY, 2016).

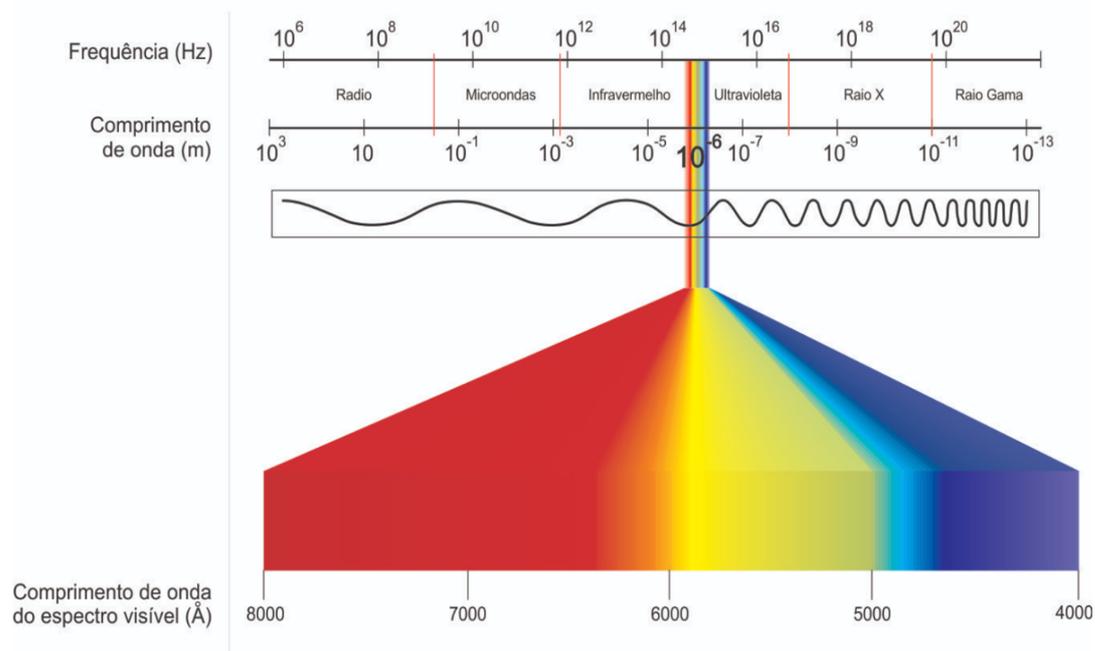
Essa energia que chega ao planeta através de ondas é avaliada numericamente a partir da verificação do comprimento e frequência da onda. O comprimento de onda refere-se à distância de um pico de onda outro e é medido em metros. Porém, por compreender valores extremamente baixos se utiliza dos prefixos micro e nano variando  $10^{-9}$  m (1 nm) a  $10^{-6}$  m (1  $\mu$ m). A frequência é medida pelo número de ondas que passam por um ponto em um determinado espaço de tempo, utilizando da unidade de medida em Hertz (Hz). O comprimento de onda e a frequência são parâmetros inversamente proporcionais entre si. Desse modo, ondas com comprimento elevado possuem uma baixa frequência (SEARS; YOUNG; ZEMANSKY, 2016).

O espectro eletromagnético é a representação gráfica ordenada e contínua entre o comprimento de onda em função da frequência, dividido em intervalos referentes ao comprimento de onda, tendo como divisões principais a modalidade de ondas a que se referem. O intervalo entre dois comprimentos de ondas é denominado banda espectral. A denominação

de banda é amplamente utilizada pelos satélites para agrupar respostas espectrais captadas pelos sensores, com a finalidade de auxiliar na elaboração da coleta e uso desses dados. Como demonstrado na Figura 3, o espectro visível o olho humano compreende apenas o intervalo de  $4,10^{-7}$  m (violeta) ao  $7,10^{-7}$  m (vermelho) (LILLESAND, 1995).

A **Figura 3** abaixo mostra a segmentação do espectro eletromagnético e os valores de comprimento de cada classe.

**Figura 3:** Espectro Eletromagnético.



Fonte: MORAES e FIORIO (2002).

Os sensores embarcados nos satélites geram como produto uma imagem em escala de cinza, onde cada pixel dessa imagem é a representação dos dados provindos de uma matriz de números digitais. Entre esses sensores podemos considerar Landsat, Sentinel, CBERS, ASTER, MODIS entre outros. A formação de uma imagem por pixels depende da resolução em que o sensor opera podendo variar o tamanho, isso acaba por impactar o resultado do produto, pois a resolução das imagens dos satélites é um dos fatores fundamentais para a acurácia do mapeamento. Esses dados possuem diversas aplicações de estimação de variáveis geofísicas e biofísicas, dentre elas, modelagem e análise de processos naturais ou causados por interferências humanas (desertificação, desmatamento, mudanças climáticas, poluição, expansão urbana etc.) (ANDRADE, 2019).

## 4.2. Classificação

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).

Primeiramente há a observação qualitativa, depois há a medição de valores quantitativos dos fenômenos espaciais com as ferramentas e instrumentos disponíveis, resultando na obtenção descritiva, que são os atributos sobre o alvo estudado. A classificação associada ao sensoriamento remoto possui uma ampla aplicação na elaboração de mapas temáticos que trazem uma visão facilitada dos resultados analisados. Esses resultados necessitam de uma atribuição de classes de uma forma clara e coesa, pois leva-se em consideração o nível de detalhamento oferecido pelos dados coletados (CONGALTON e GREEN, 2008).

Segundo Congalton e Green (2008) a classificação é uma representação abstrata da situação observada em campo utilizando critérios de diagnósticos bem definidos, realizando o agrupamento de informações em classes ou grupos com base na sua resposta. A melhor forma de visualização é a utilização de mapas temáticos generalizando as feições da superfície terrestre categorizando e agrupando informações com respostas semelhantes. Os esquemas de classificação buscam organizar informações espaciais de uma maneira lógica e ordenada, que possibilita a caracterização das feições da paisagem, facilitando para o usuário reconhecê-las.

## 4.3. Índices Espectrais

Os índices espectrais têm como principal utilidade fazer uso das respostas espectrais de cada banda, ou do agrupamento de várias, realçando um alvo específico ou uma classe durante a classificação por operações aritméticas entre as bandas espectrais compreendendo as faixas de reflectância, facilitando assim a visualização para a diferenciação visual dos alvos (JENSEN, 2009).

Segundo a literatura há diversos índices desenvolvidos para análise de classes específicas. Dentre eles, os índices para detecção de uso e cobertura do solo como o NDVI (*Normalized Difference Vegetation Index*) e RS (Razão Simples) que acordo com Rouse et al., (1974) e Campbell e Wynne (2011), altos valores indicam pixels com vegetação saudável. Para destacar as áreas com construção civil se faz uso do NDBI (*Normalized Difference Built-up Index*) gerado com base na metodologia de Zha (2003); para o realce de áreas com uma presença de umidade no solo há o MNDWI (*Modified Normalized Difference Water Index*) proposto por Xu (2006), que se mostra eficiente pois ele parte da modificação do índice NDWI (*Normalized*

*Difference Water Index*) que auxilia na minimização de ruídos de vegetação, solo impermeabilizado e solo exposto presente próximo a corpos d'água.

Na **Tabela 2**, segundo Jensen (2009), são apresentados os índices que foram utilizados na elaboração da classificação.

**Tabela 2:** Índices Espectrais.

Índices	Equações	Referência
NDVI	$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$	Rouse et al., (1974)
RS	$RS = \frac{\rho_{red}}{\rho_{nir}}$	Birth; McVey, 1968
MNDWI	$MNDWI = \frac{\rho_{green} - \rho_{MidIR}}{\rho_{green} + \rho_{MidIR}}$	Xu (2006)
NDBI	$B_u = \frac{\rho_{nir} - \rho_{MidIR}}{\rho_{nir} + \rho_{MidIR}}$	Zha et al., (2003)
	$NDBI = B_u - NDVI$	

Fonte: Jensen (2009).

#### 4.4. Etanol

Como resposta à crise do petróleo ocorrida na década de 70, o Brasil em 1975 iniciou o incentivo à criação do programa Proálcool, buscando evitar uma dependência do preço do petróleo internacional. Esse incentivo ao programa trouxe resultados à nação em 2006, quando o Brasil se tornou energeticamente autossuficiente na produção de etanol. A participação desse biocombustível na matriz energética brasileira trouxe diversos benefícios ambientais, primeiramente por se tratar de uma fonte renovável de energia e por ter baixas emissões de CO<sub>2</sub>, contribuindo assim para uma menor degradação do meio ambiente, evitando a emissão de gases do efeito estufa. Essa substituição do petróleo pelo biocombustível pode trazer uma redução de 73% da emissão de gases de poluentes na atmosfera (MACEDO, 2008).

A variedade de cana-de-açúcar comumente utilizada nas fazendas de cana-de-açúcar é a espécie *Saccharum officinarum* L, essa espécie possui como principal característica um alto teor de açúcar com uma produção média de 480 milhões de toneladas anual. Ela se destaca no cenário global devido ao seu alto valor energético atribuído à concentração de açúcar, que é

convertido em álcool através da fermentação. A produção de etanol ocorre por meio da fermentação alcoólica, processo anaeróbio catabólico, onde há a degradação de moléculas de sacarose, por leveduras, que metabolizam a glicose em moléculas de  $CO_2$ , etanol e energia (FAO, 2021).

No Brasil a produção de etanol, como um biocombustível renovável, tem sido intensamente estimulada principalmente para suprir a demanda do setor de transportes, pois a frota brasileira conta com a tecnologia *flex-fuel*, na qual os veículos funcionam 100% com o etanol hidratado ou em diferentes misturas de etanol e gasolina. Portanto, esta tecnologia vem estimulando o aumento da produção de cana-de-açúcar no país para substituir os combustíveis fósseis por biocombustíveis (COSTA, 2019).

Em 2021 o mundo presenciou a COP26, Conferência das Partes da Convenção-Quadro das Nações Unidas de Mudanças Climáticas, que foi uma resposta de elaboração e atualização de políticas ambientais mundiais para frear as altas emissões de Dióxido de Carbono ( $CO_2$ ), Metano ( $CH_4$ ), Óxido Nitroso ( $N_2O$ ), Hexafluoruro de Enxofre ( $SF_6$ ) entre outros gases prejudiciais ao meio ambiente. Vários países iniciaram seus próprios programas de produção e uso de etanol como combustível para reduzir a dependência do petróleo e as emissões de gases de efeito estufa (GEE) (THE ECONOMIST, 2021).

## 5. MATERIAL E MÉTODO

### 5.1. Área de Estudo

A região utilizada para testes compreende os municípios da macrorregião do Pontal do Paranapanema no Oeste Paulista. Os municípios escolhidos foram Teodoro Sampaio, Marabá Paulista, Mirante do Paranapanema e Sandovalina com uma população de 46.956 habitantes (IBGE, 2020). Devido os quatro municípios possuírem usinas sucroalcooleiras, há um interesse em avaliar o potencial dessas regiões.

De acordo com a classificação Köppen e Geiger, a região do Pontal do Paranapanema é classificada como clima tropical com estação seca (Aw), possuindo temperatura média anual de 21,9 °C com uma média de precipitação de 1.370mm durante o ano. A vegetação predominante é do tipo Cerrado com resquícios de Mata Atlântica e zonas de transição entre elas (M.C. Peel, B.L. Finlayson e T.A. McMahon, 2007).

A data escolhida para a aquisição das imagens foi de 26 de dezembro de 2020, pois encontra-se entre a safra e entressafra do cultivo de cana-de-açúcar da região e foi a única data que atendeu a exigência máxima de 10% de nuvens estipulada para o trabalho.

A Figura 4 demonstra a localização dos municípios selecionados em destaque:

**Figura 4:** Mapa do Estado de São Paulo destacando os municípios escolhidos.



Fonte: Elaborado pelo autor.

A área foi escolhida a partir da presença de quatro usinas sucroalcooleiras descritas na Tabela 1 e optou-se por trabalhar com essa região, assim delimitando a área de estudo a ser utilizada.

## 5.2. Softwares Utilizados

Durante o processamento das imagens do satélite Sentinel 2A 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 em 2016 pela Google, com a finalidade de ajudar em processos de análises ambientais em grande escala, realizando atividades de geoprocessamento através da seção *Google Earth Engine Code Editor*, que permite a implantação de algoritmos em *JavaScript* ou em Python para a implementação de análises geoespaciais em um ambiente em nuvem, facilitando processamentos em larga escala, devido ao fato de não ser necessário efetuar download das imagens de satélite. O GEE disponibiliza em seu catálogo o acesso direto a imagens dos sensores SENTINEL 1 e 2; LANDSAT 5, 7 e 8; ASTER, MODIS e NOAA (GOOGLE INC, 2021).

**O Quantum GIS (QGis)** é um software gratuito desenvolvido pela *QGIS Development Team* que é um SIG (Sistema de Informação Geográfica) projetado pela *Open Source Geospatial Foundation (OSGeo)*, que permite realizar análises de dados espaciais, com um sistema em código aberto programado em C++ e Python edição e análise de dados georreferenciados. Tem como característica um amplo desenvolvimento de uma arquitetura de plug-ins que podem ser baixados para executar análises específicas. O QGis foi lançado em 2002, tem como característica, a liberdade de escolha de download de sensores presentes no mercado, mas há a necessidade de fazer o download as bandas na máquina para efetuar o processamento (QGIS DEVELOPMENT TEAM, 2021).

A principal ferramenta utilizada foi o Google Earth Engine para o processamento das imagens e o software QGis foi utilizado para a elaboração dos mapas temáticos e geração das análises numéricas presentes nesse trabalho.

### 5.3. Imagens Sentinel 2

O programa Copernicus foi desenvolvido pela Agência Espacial Européia (*ESA – European Spatial Agency*), em parceria com a Organização Européia para Exploração de Satélites Meteorológicos. (*EUMETSAT - European Organisation for the Exploitation of Meteorological Satellites*), Centro Europeu de Previsões Meteorológicas de Médio Prazo, Agências da UE e *Mercator Océan*. Estabelecido em 2014 pela regulamentação (EU) N° 377/2014, nomeado em homenagem ao cientista Nicolaus Copernicus, teórico do modelo heliocentrismo durante o século 16.

O programa foi desenvolvido com o objetivo de atingir uma capacidade de observação global da Terra com alta qualidade, em larga escala e de forma precisa. O acesso às imagens é gratuito e permite que sejam utilizadas para estudos de mitigação ambiental, impactos de mudanças climáticas, estudos no campo da geociência abordando os campos de monitoramento atmosférico, estudo do uso e cobertura do solo, ecossistema marinho, mudanças climáticas, manejo de situações emergenciais e segurança para organizações europeias e não europeias. O programa Copernicus conta com 2 satélites, o Sentinel 2A e o 2B na mesma órbita, localizados a uma altitude de 786 km, cobrindo 290 km de largura e com uma diferença de fase de 180°. Devido à geometria de operação dos satélites é possível uma revisão da coleta de imagens a cada 5 dias (*ESA – EUROPEAN SPATIAL AGENCY, 2021*).

O satélite Sentinel 2 desempenha a função da aquisição sistemática de imagens de alta resolução espacial com 10m, 20m a 60m de densidade de pixel. Dispõe de um sensor MSI (*MultiSpectral Instrument*) com 13 bandas multiespectrais dispostas entre os espectros visível (VIS), infravermelho de ondas curtas (SWIR) e infravermelho próximo (NIR) (*ESA – EUROPEAN SPATIAL AGENCY, 2021*).

Durante a elaboração dessa pesquisa optou-se por utilizar imagens providas do Copernicus devido a imagem Sentinel 2A possuir uma resolução espacial de 10m x 10m no pixel, sendo considerado uma boa resolução para se trabalhar, além de ser disponibilizada gratuitamente (<https://earthexplorer.usgs.gov/>). A Tabela 3 abaixo descreve as características das bandas presentes no Sentinel 2A. É importante destacar que o presente trabalho utiliza de imagens providas de satélites e não de veículos aero tripulados (drones), pois esses veículos possuem uma abrangência de área bem menor quando comparado com as imagens providas de sensores. Devido a essa baixa abrangência de área o uso de drones se torna inviável para análises de áreas extensas.

**Tabela 3:** Características das bandas do Sentinel 2A.

<b>Banda</b>	<b>Nome da Banda</b>	<b>Comprimento de onda central (nm)</b>	<b>Resolução Espacial (m)</b>
<b>B01</b>	<i>Coastal Aerosol</i>	443	60
<b>B02</b>	<i>Blue</i>	490	10
<b>B03</b>	<i>Green</i>	560	10
<b>B04</b>	<i>Red</i>	665	10
<b>B05</b>	<i>Vegetation Red Edge</i>	705	20
<b>B06</b>	<i>Vegetation Red Edge</i>	740	20
<b>B07</b>	<i>Vegetation Red Edge</i>	783	20
<b>B08</b>	<i>Near Infrared</i>	842	10
<b>B08A</b>	<i>Near Infrared Narrow</i>	865	20
<b>B09</b>	<i>Water Vapour</i>	945	60
<b>B10</b>	<i>Shortwave Infrared Cirrus</i>	1380	60
<b>B11</b>	<i>Shortwave Infrared</i>	1610	20
<b>B12</b>	<i>Shortwave Infrared</i>	2190	20

Fonte: (ESA – EUROPEAN SPATIAL AGENCY, 2021) e adaptado pelo autor.

#### 5.4. METODOLOGIA

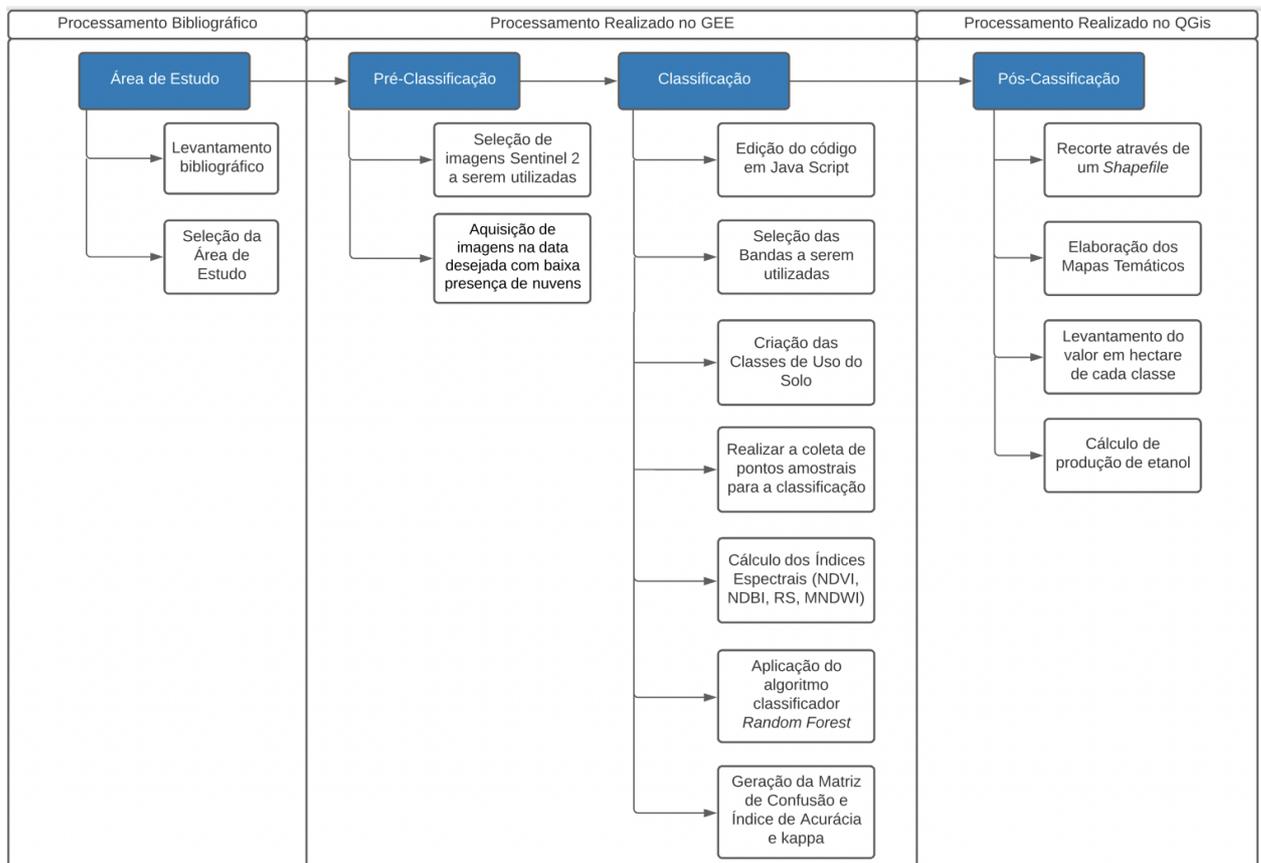
Os resultados desse trabalho foram gerados através da combinação do uso de duas ferramentas, GEE e o QGIS, como explicitado anteriormente. Durante o processo de definição da metodologia foram avaliadas algumas questões e chegou-se à conclusão de que a combinação pretendida inicialmente, que abordaria o uso dos softwares ENVI e o WEKA, não seria tão interessante quando comparada à combinação do GEE e o software QGIS. Os pontos favoráveis a esta combinação foram: alto poder de processamento do GEE, o fato de não haver necessidade de download das imagens e de uma máquina de alto desempenho, pois os processamentos são realizados em uma nuvem. Além disso, a facilidade do uso das ferramentas, o tempo de processamento muito reduzido, quando comparado à combinação pretendida inicialmente. No caso da primeira opção, as imagens teriam que ser selecionadas uma a uma, dentro do período estipulado e da porcentagem de nuvens definida, o que onera bastante o processo, posteriormente seria realizado o download das imagens selecionadas, o que exige uma internet de qualidade, espaço na máquina, o processamento dos dados, que exigiria uma máquina com poder de processamento adequado. Ao utilizarmos o GEE e o QGIS para produzir

os mapas e os resultados numéricos, a qualidade do processamento foi mantida e o tempo de desenvolvimento, que poderia levar dias foi reduzido a pouco mais de um dia.

Segundo Cho *et. al* (2021), Vale (2019) e Alencar *et al.* (2020) o desempenho da ferramenta da Google para a classificação de uso e cobertura do solo se mostrou eficiente e apresentou resultados excelentes de classificação. Isso reforça que a metodologia definida para o processo de classificação é eficiente.

O processo metodológico desenvolvido no presente trabalho é descrito no fluxograma e na sequência são descritas cada uma das etapas.

**Figura 5:** Fluxograma de atividades realizadas na metodologia.



Fonte: Elaborado pelo próprio autor.

#### 5.4.1. Pré-Processamento dos Dados

Durante essa etapa de pré-processamento, já atuando dentro da aba principal na plataforma em nuvem na aba *Code Editor* do GEE no link (<https://code.earthengine.google.com/>), a primeira etapa é definir a área de interesse a partir de um polígono desenhado na imagem e na sequência é definido o satélite a partir do qual serão extraídas as imagens multiespectrais. Optou-se por utilizar imagens do satélite Sentinel 2A, por possuir uma resolução espacial de 10m, a qual é considerada uma adequada para o objetivo do trabalho.

Em seguida realizou-se a escolha das imagens no período de 01/12/2020 a 31/12/2020 e definiu-se a porcentagem máxima de nuvem em 10%. Optou-se pelo mês de dezembro pois há a uma baixa presença de áreas que foram aradas. A aquisição das imagens se deu através da aba principal na plataforma em nuvem na aba *Code Editor* do GEE no link (<https://code.earthengine.google.com/>).

#### 5.4.2. Processamento

Após o pré-processamento, onde ocorreu a seleção da imagem a ser utilizada, levando em consideração a resolução espacial de 10m, a porcentagem de nuvem máxima dentro do intervalo selecionado, houve a seleção das bandas a serem utilizadas, que foram: bandas B2, B3, B4, B5, B6, B7, B8, B11, B12.

Após a escolha das bandas foram definidas as 6 classes a serem mapeadas. A Tabela 4 descreve cada uma das classes.

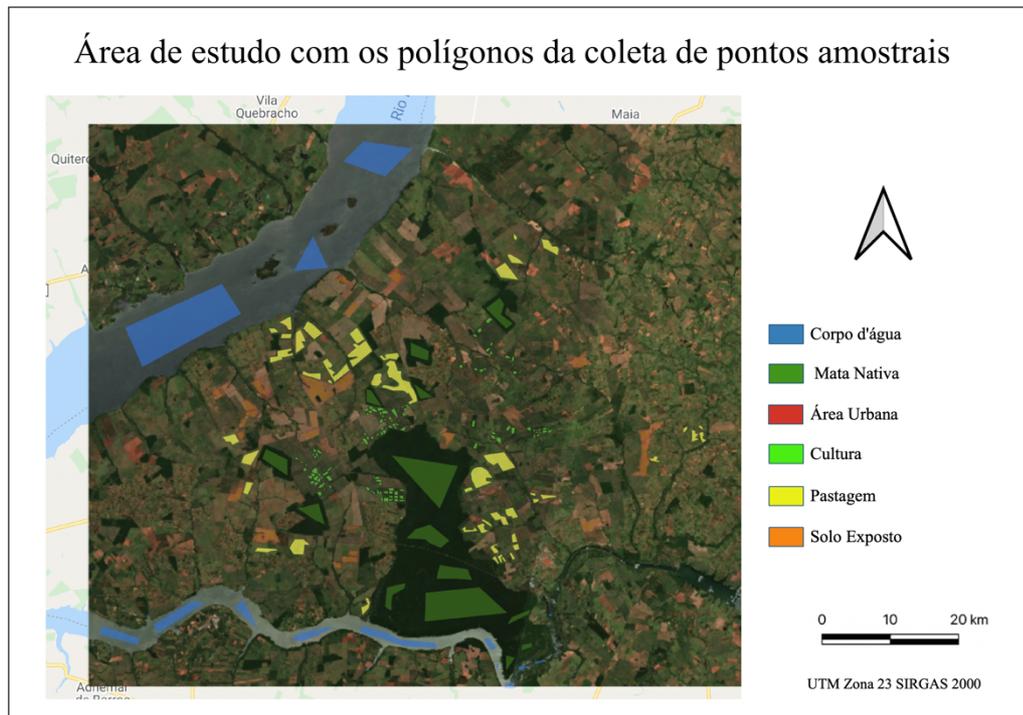
**Tabela 4:** Classes do mapa de uso e cobertura do solo

Classe	Definição
Corpo d'água	Área de águas superficiais, onde se destacam rios, lagos, lagoas e oceanos.
Mata Nativa	Área de vegetação densa, não alterada, compondo a vegetação nativa da região.
Área Urbana	Área com alta interferência humana de impermeabilização de solo.
Cultura	Áreas com presença de vegetação, desde monocultura ou cultura por ciclos, vegetação de porte médio e possuindo simetria de espaçamento entre uma planta e outra.
Pastagem	Áreas com presença de vegetação rasteira, ou regiões que ocorreu a retirada da vegetação original para o desenvolvimento da atividade pastoril.
Solo Exposto	Área onde não há a presença de vegetação.

Fonte: Elaborado pelo próprio autor.

Após a definição das classes foi realizada a coleta das amostras de cada uma das classes, necessárias para a etapa da classificação supervisionada. As amostras foram adquiridas a partir de polígonos referenciando as informações contidas naquele pixel ou grupo de pixels que representa um tipo específico de cobertura terrestre (LILLESAND e KIEFER, 1995). A Figura 6 apresenta a coleta de pontos amostrais realizada no GEE.

**Figura 6:** Recorte da área de estudo com os polígonos da coleta de pontos amostrais.



Fonte: Elaborado pelo próprio autor.

### 5.4.3. Índices Espectrais Utilizados

Posteriormente à criação dos polígonos amostrais de cada uma das classes foram calculados os índices espectrais selecionados para comporem e melhorarem o desempenho da classificação. Os índices utilizados foram NDVI (Índice de vegetação por diferença normalizada), NDBI (Índice de Construção por Diferença Normalizada), MNDWI (Índice da diferença normalizada da água modificado) e a RS (Razão simples) descritos no tópico anterior. Os índices promovem uma maior acurácia, pois suas respostas serão utilizadas pelo algoritmo no momento da separação dos pixels de cada classe. Os índices são descritos a seguir.

#### NDVI

O Índice de Vegetação por Diferença Normalizada (NDVI - *Normalized Difference Vegetation Index*) desenvolvido por Rouse et al. (1974) e aprimorado posteriormente por Campbell e Wynne (2011), tem a finalidade de avaliar a densidade da cobertura vegetal, levando em consideração a resposta espectral da vegetação em relação a 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, esse índice é muito utilizado em averiguação de desmatamento, monitoramento de safras, expansão ou encolhimento de áreas vegetativas. O NDVI é calculado através da diferença entre a reflectância das bandas do infravermelho próximo (NIR) e do vermelho (*Red*), dividido pela soma das duas reflectâncias utilizadas e é expresso matematicamente como:

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

No Sentinel 2 a banda do infravermelho próximo ( $\rho_{nir}$ ) é equivalente à banda 08 e a banda do vermelho ( $\rho_{red}$ ) equivale a banda 04. A formulação resulta numa resposta numérica entre -1 e 1. Quanto mais próximo de 1, maior o vigor vegetativo no local representado pelo pixel e valores mais próximos de -1 compreendem áreas com uma menor ou nenhuma incidência de matéria composta por clorofila.

### Razão Simples

O Índice de Vegetação da Razão Simples (RS – Simple Ratio) desenvolvido primeiramente por Birth e McVey (1968) é um índice físico com intuito de avaliar, identificar e realçar através da análise de dados de reflectância, regiões com a presença de vegetação e exibe boa relação com características das florestas. A RS se assemelha muito com a análise do NDVI, pois também leva em consideração a resposta de absorção da vegetação na faixa do visível, a diferença está no fato de ser uma avaliação mais simplificada. A RS é um índice calculado pela divisão da banda vermelha (*red*), representada pela banda 04 no Sentinel 2, pela banda do infravermelho próximo (*NIR*), representada pela banda 08 no Sentinel 2. A razão simples é expressa pela fórmula:

$$RS = \frac{\rho_{red}}{\rho_{nir}} \quad (2)$$

### MNDWI

O índice MNDWI (*Modified Normalized Difference Water Index*) é fundamentado pela modificação do índice de diferença normalizada de água (NDWI – *Normalized Difference Water Index*). Proposto por McFeeters (1996) foi desenvolvido para aplicação de estudos ambientais a fim de identificar cursos d'água, visando maximizar a reflectância de áreas com a presença de água. Esse índice tem como principais aplicações a avaliação e manejo ambiental

de recursos hídricos, realizando estudos de cursos de rios, monitoramento de áreas inundadas e gestão de recursos hídricos. O NDWI é concebido pela relação matemática entre as bandas verde (Green) e a banda infravermelha próxima (NIR), no Sentinel 2 a banda verde ( $\rho_{green}$ ) representado pela banda 03 e a infravermelha próxima ( $\rho_{NIR}$ ) representado pela banda 08. Calculado pela equação abaixo:

$$NDWI = \frac{\rho_{green} - \rho_{NIR}}{\rho_{green} + \rho_{NIR}} \quad (3)$$

O MNDWI foi desenvolvido para suprir a deficiência do índice NDWI proposto por Xu (2006), uma vez que este apresenta limitações devido a ruídos provenientes de áreas construídas, que acabam interferindo na resposta espectral dos pixels. Devido a isso, o MNDWI tem uma aplicação no monitoramento de corpos hídricos que possam ter interferências devido a áreas impermeabilizadas. O cálculo do MNDWI difere do NDWI, pois nesse índice considera-se a relação entre a banda verde (*green*) e a infravermelho médio (*midIR*), descrito pela equação 4:

$$MNDWI = \frac{\rho_{green} - \rho_{MidNIR}}{\rho_{green} + \rho_{MidNIR}} \quad (4)$$

No Sentinel 2 o infravermelho médio ( $\rho_{midIR}$  ou  $\rho_{SWIR}$ ) equivalente a banda 11 e a verde ( $\rho_{green}$ ) equivale a banda 03.

## **NDBI**

O Índice de Construção por Diferença Normalizada (NDBI - *Normalized Difference Built-up Index*) é baseado na abordagem dos estudos de Zha et al. (2003). O NDBI possui como principal finalidade identificar áreas construídas e impermeabilizadas através da resposta que o pixel terá durante a análise, pois solos impermeabilizados de áreas urbanas tem a predominância de materiais que absorvem a irradiância solar (asfalto, concreto, metal, entre outros), diferindo assim, da resposta de áreas com presença de vegetação. O NDBI é determinado através de equações aritméticas da diferença entre a reflectância das bandas do infravermelho próximo (*NIR*) e do infravermelho médio (*midIR*), dividido pela soma das duas reflectâncias utilizadas, conforme a Equação 5:

$$NDBI = \frac{\rho_{nir} - \rho_{midIR}}{\rho_{nir} + \rho_{midIR}} \quad (5)$$

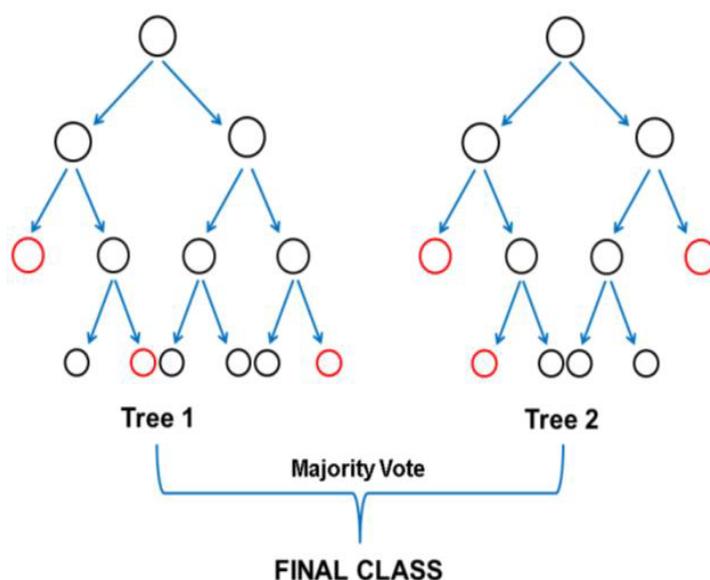
No Sentinel 2 o infravermelho próximo ( $\rho_{nir}$ ) é equivalente a banda 08 e a infravermelho médio ( $\rho_{midIR}$  ou  $\rho_{SWIR}$ ) equivale a banda 11. O resultado expresso pelo índice produz uma imagem com valores de -1 a 1. Espera-se que o valor mais alto indique os pixels incluídos na área construída e o valor negativo indica os pixels presentes na área não construída.

#### 5.4.4. Aplicação do Algoritmo *Random Forest*

O código fonte desenvolvido em *JavaScript* foi disponibilizado por Vale (2019) e alterado para atender os objetivos do presente trabalho. O código foi desenvolvido com o algoritmo *Random Forest* e foi utilizado no GEE para a classificação supervisionada da área de estudos.

O *Random Forest* (RF) é um algoritmo de classificação supervisionada, baseado no conjunto de árvores de decisão gerados de uma forma randômica, a partir do treinamento com as classes amostrais coletadas no processo de classificação. É constituída pela reamostragem aleatória do treinamento inicial, o RF funciona pela classificação de um vetor com uma característica específica, submetendo o vetor à árvore da floresta de decisão, sendo o algoritmo que toma uma decisão do vetor com a melhor resposta através de votação. Na Figura 7 é apresentado o fluxograma de como o algoritmo opera na tomada de decisão (KAGGLE, 2019).

**Figura 7** – Fluxograma de operação do *Random Forest*.



Fonte: Kaggle, (2019).

O *Random Forest* como algoritmo para classificação tem se mostrado uma excelente ferramenta quando comparado com outros modelos. Isso se dá devido a rapidez na etapa de treinamento e a minimização do efeito de ajustamento do modelo apenas para os dados de treinamento. Uma desvantagem do RF é a necessidade de uma enorme quantidade de dados de treinamento, a geração de possíveis lacunas de dados e uma possível lentidão no processamento devido a necessidade da utilização de diversas árvores de decisão para se chegar em um resultado mais coeso com o esperado (KAGGLE, 2019).

Ao trabalhar com o GEE, esta desvantagem do RF é minimizada, em função do processamento ser realizado em nuvem. A utilização desse algoritmo no *script* alinhado com os índices espectrais (NDVI, NDBI, MNDWI, entre outros) torna ainda mais eficiente a classificação, trazendo uma acurácia maior para o processamento e dados mais condizentes com a verdade terrestre da região (BELGIU & DRAGUT, 2016; SCHOTT, 2019).

Segundo Cho *et. al* (2021), dentre diversos classificadores, esse algoritmo se destaca na detecção de vegetação e possui um desempenho considerado excelente para ser utilizado como ferramenta de classificação, apontando como vantagens um bom resultado de acurácia, desempenho ágil de processamento quando combinado a uma ferramenta como o GEE, mostrando uma classificação consistente com a verdade terrestre da região estudada.

Após as etapas de seleção de área e definição de índices foi realizada a aplicação do algoritmo e o treinamento do classificador foi realizado com 70% dos pixels selecionados na criação de classes amostrais.

Na sequência foram calculados os índices para verificação da qualidade da classificação.

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

A veracidade e desempenho da classificação das imagens e do mapa do uso e cobertura do solo produzido foram avaliados a partir de alguns índices. Na cartografia, os índices Kappa e a Acurácia Global são amplamente utilizados para determinar a exatidão do mapeamento, junto a isso se destaca a Matriz de Confusão que descreve as classes amostrais e a quantidade de pixels que podem interferir numa boa resposta, devido à confusão gerada por pixels semelhantes.

## Matriz de Confusão

A Matriz de Confusão ou Matriz de Erro é uma das melhores formas para avaliar a acurácia de uma classificação. A matriz de confusão (Figura 8) é constituída pela formação de matriz quadrada contendo números arranjados em linhas e colunas, que apresentam o número de pixels atribuídos a uma classe específica em relação à verdadeira classe a qual o pixel pertence. As linhas representam as classes definidas através do processo de classificação e as colunas os dados de referência. Os números de pixels da diagonal representam os pixels que foram classificados de forma correta (CONGALTON; GREEN, 2008).

Além disso, as informações marginais das linhas e colunas, possibilita fazer uma distinção entre 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; GREEN, 2008).

**Figura 8:** Exemplificação 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: (CONGALTON; GREEN, 2008) e adaptado pelo próprio autor.

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.

## Índice Kappa

O Índice Kappa 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. Esse índice é indicado para avaliar a legitimidade de uma classificação, avaliando as interações realizadas e utiliza como base os dados da matriz de

confusão para se estimar a exatidão do processo. Proposto por Cohen (1960) a validação do cálculo desse índice pode ser expressa pela Equação 6:

$$K = \frac{\sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+}n_{+i}}{n^2 - \sum_{i=1}^k n_{i+}n_{+i}} \quad (6)$$

Na qual o  $k$  refere-se o número de linhas e colunas presentes na matriz,  $n_{ii}$  é o número de observações na coluna  $i$  e linha  $i$ ,  $n$  se refere ao número total de observações e  $n_{i+}n_{+i}$  são os valores totais da linha e coluna  $i$  (COHEN, 1960).

Segundo Landis e Koch (1997) o valor de resposta do Índice Kappa é atribuído à qualidade da classificação variando de Péssima a Excelente, sendo tabelada por valores atribuindo uma associação qualitativa de processo. A Tabela 5 representa a divisão e classificação:

**Tabela 5:** Associação de Qualidade de Classificação com Valores de 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: (LANDIS; KOCH, 1997) e adaptado do próprio autor.

### Acurácia Global

A Acurácia Global (AG) ou Exatidão Global se refere a uma das muitas medidas de análise de confiabilidade de classificação. É uma medida simples relacionada aos valores presentes na diagonal principal da matriz de confusão, relacionada com o total de pontos amostrados na classificação. De acordo com Congalton e Green (2008), a Acurácia Global (AG)

ou Exatidão Global é calculada através da divisão da somatória da diagonal principal da matriz de confusão  $n_{ii}$ , pelo número de amostras totais coletadas  $n$  como descrita pela Equação 7:

$$AG = \sum_{i=1}^K \frac{n_{ii}}{n} \quad (7)$$

A acurácia pode ser dividida em dois tipos específicos para a avaliação dos resultados, são elas: a Acurácia do Usuário (AU) e a Acurácia do produtor (AP). A AU se refere à divisão entre o número de pixels classificados corretamente em cada classe e o total de pixels classificados na classe (total da linha). A AP é a divisão do número de pixels corretamente classificados em cada classe pelo número de pixels do conjunto de dados de treinamento usado nas classes (total da coluna) (CONGALTON; GREEN, 2008). As Equações 8 e 9 exemplificam as duas acurácias:

$$AP = \frac{n_{ii}}{n_{+j}} \quad (8)$$

$$AU = \frac{n_{ii}}{n_{j+}} \quad (9)$$

Após a determinação dos índices foi realizado o tratamento dos dados no QGIS a fim de obter a espacialização dos resultados obtidos no GEE.

#### 5.4.6 Tratamento dos dados no QGIS

Efetuada todo o processamento dentro do GEE os resultados foram exportados para o Google Drive e baixados para a elaboração dos mapas temáticos no software QGIS. O objetivo dessa etapa é apresentar os resultados obtidos na classificação supervisionada no GEE em forma de um mapa, para uma melhor visualização dos resultados.

Após baixar o arquivo *Geotiff*. gerado no GEE foi aplicada uma máscara *shapefile* dos municípios escolhidos para desenvolver o trabalho. A máscara utilizada é composta por 4 municípios e é apresentada na Figura 4 (em amarelo). Através desta máscara é feito o recorte da área e a elaboração do mapa temático para uma melhor visualização dos resultados.

Após a elaboração do mapa de uso e ocupação do solo foi realizada a estimativa das áreas de cada uma das classes e a partir das áreas estimadas foi calculada a produção de etanol.

#### 5.4.7 Estimativa de Área das Classes e Cálculo da Produção do Etanol

Realizado o recorte da área aplicando o *shapefile*, dentro das ferramentas do QGis utilizou-se da ferramenta de processamento (*processing r.report*) presente já nativamente no software, gerou-se um relatório estatístico de área para levantar os valores em Hectare de cada classe presente na classificação. Realizado isso, os dados são disponibilizados e basta apenas tabelar os resultados encontrados.

A partir dos valores em hectare para cada uma das classes foi possível realizar a estimativa do potencial de produção de biocombustível. A estimativa visou avaliar o cenário de substituição das áreas de pastagem da região por cana-de-açúcar. Devido a expansão crescente na região do Oeste Paulista, esse cenário de substituição vem se tornando uma realidade.

Segundo a Conab (2021) um hectare de plantação de cana equivale a 70 toneladas de cana-de-açúcar (equação 10). Com essa informação foi estimado o valor de produção em tonelada de cana, substituindo toda a área de pastagem da região por cultura. O valor encontrado na classificação foi aplicado nesse cálculo para assim dar continuidade aos cálculos de estimativa de produção.

$$1 \text{ hectare de plantação} \rightarrow \text{Aproximadamente } 70 \text{ toneladas de cana de açúcar} \quad (10)$$

O processo de produção adotado mundialmente para produção é centralizado no método de fermentação. A fermentação alcoólica é um processo químico de transformação de açúcares fermentescíveis (glicose e frutose) em uma cadeia alcoólica como o etanol através da ação de leveduras que facilitam o processo fermentativo, acelerando a reação. Esse processo pode ocorrer como uma reação anaeróbica, sem a presença de oxigênio. A primeira etapa é o processo de glicólise, onde fosfatos são incorporados pelas moléculas de glicose facilitando a sua quebra em 2 moléculas de ácido pirúvico. A segunda etapa consiste na fermentação alcoólica realizada na presença das leveduras, nessa reação o ácido pirúvico ( $C_3$ ) é descarboxilado gerando o acetaldeído. Através do NADH ocorre a redução do acetaldeído, gerando etanol ( $C_2H_6O$ ) e dióxido de carbono ( $CO_2$ ) (DE GÓES-FAVONI, 2018).

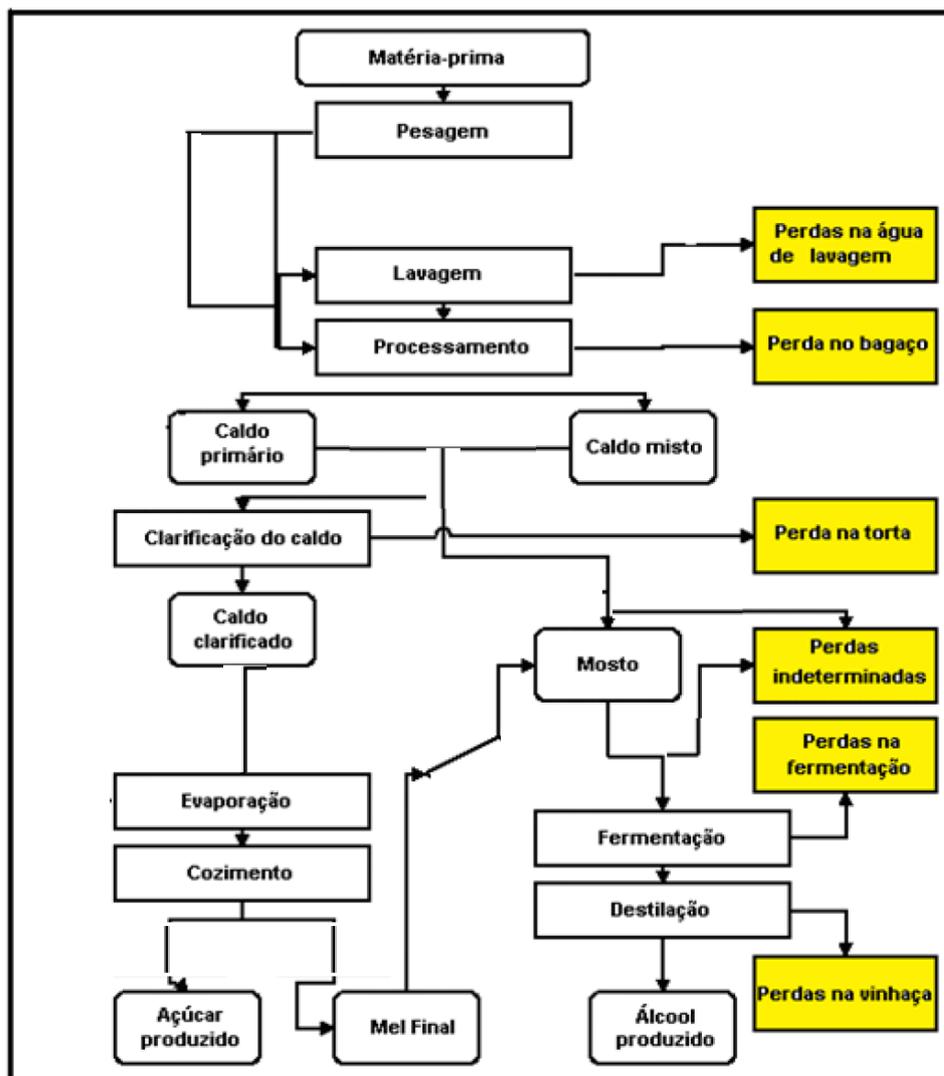
Durante a modelagem desse cenário de substituição foram avaliadas as principais perdas presentes no processo de produção de biocombustíveis. É sabido que durante o processo de produção de etanol, ocorrem algumas perdas, dessa forma, foram avaliadas as principais através de dados de referências, mediante o uso de cálculos aritméticos foi realizada a análise da produção do combustível e suas principais perdas de processo. As perdas a serem consideradas

para a estimação, segundo Santos (2014) são: Na etapa de lavagem, de moagem, de decantação, de destilação, de fermentação e as outras perdas principais. Essas perdas estão relacionadas aos processos industriais utilizados atualmente e são levadas em consideração para os cálculos de estimação o ART (Açúcares redutores totais fermentescíveis) (DELGADO; CESAR; SILVA, 2019; RAÍZEN ENERGIA, 2021).

As perdas de processos são catalogadas como perdas determinadas, ou seja, são as que podem ser quantificadas, mensuradas e monitoradas durante a produção, podendo corrigi-las a fim de diminuí-las. E as perdas indeterminadas são consideradas as que não há possibilidade de prever (vazamentos, erro de amostragem, transbordamento nos tonéis) (NOVACANA, 2018).

O fluxograma da Figura 9 demonstra o processo de produção de etanol, juntamente com a exemplificação de onde ocorrem as principais perdas no processo.

**Figura 9:** Fluxograma do processo de produção do etanol com as perdas elencadas.



Fonte: (DELGADO; CESAR; SILVA, 2019). Adaptado pelo próprio autor.

Durante a produção do etanol a matéria-prima retirada da lavoura é pesada e encaminhada para a planta, pode-se citar como a primeira perda determinada a perda por lavagem. Nesse processo ocorre a remoção de impurezas (palha, terra, areia, materiais em geral, canas danificadas e material que possa ter caído no transporte) com uso de jatos de água para a lavagem, segundo o estudo de caso desenvolvido pela Novacana (2018) foi considerado o percentual de perda dessa etapa como 0,47%, sendo enquadrado dentro dos percentuais estipulados de perda.

A segunda perda considerada foi a perda na etapa de moagem, é resultado do processo de moagem para extração do caldo presente na cana. Essa perda consiste na eficiência de processamento da moenda da indústria e, segundo a modelagem proposta pela Novacana (2018) o valor de perda nessa etapa consistiu em 3,73% estando dentro dos parâmetros de perda de açúcares (ART) que considera de 3% a 4% nessa etapa (NOVACANA, 2018).

Após o processo de moagem o caldo é encaminhado para o processo de tratamento químico para remoção de impurezas que restaram junto com as correções de pH, sendo assim preparado para a decantação por clarificação. As impurezas (lodo) contidas na mistura são decantadas e passam por filtros para a recuperação de açúcares residuais. Devido a essa recuperação há a perda durante a filtragem sendo considerada a terceira perda principal que é a perda por torta, onde na modelagem utilizada dessa etapa consiste em 0,54% (NOVACANA, 2018).

O caldo é encaminhado para realizar a evaporação eliminando 90% da água presente. Nessa etapa há a formação do mosto que é encaminhado para o processo de fermentação, que consiste na reação da transformação da glicose em álcool pela reação de fermentação. Após a fermentação o mosto se transforma no vinho fermentado (líquido composto de água, álcool e leveduras). Esse vinho é encaminhado para uma centrífuga separando as leveduras do resto da mistura, nesta etapa ocorre a quarta principal perda, que é a perda relacionada a separação das leveduras do resto da mistura, sendo relacionada a qualidade e eficiência do tratamento do caldo. Nos dados de Novacana (2018) essa etapa compreende como um dos maiores percentuais de perda, em torno de 5,17% devido a problemas relacionados a qualidade e remoção de impurezas da matéria-prima e da limpeza da moenda.

O vinho centrifugado e sem a presença das leveduras é encaminhado para a destilaria onde ocorre a separação da água e do álcool. Nesse processo busca-se atingir o percentual mínimo de 94,5% de etanol. A formação de vinhaça ocorre nessa etapa. Ocorre também a formação de um composto que é descartado após o processo de decantação, em resultado disso a perda correspondeu a 0,18% no experimento (NOVACANA, 2018).

Dentro do processo há as perdas indeterminadas compostas por vazamentos, transbordamentos, eventuais manutenções, acidentes, entre outras situações não mensuráveis e impossíveis de prever que ocorram. Essa perda compreende 4,05% de perdas no processo (NOVACANA, 2018). A Tabela 6 descreve os valores em porcentagem de toneladas de perdas analisados.

**Tabela 6:** Principais perdas consideradas.

Principais Perdas (%)	
Etapa de Lavagem	0,47
Etapa de Moagem	3,73
Etapa de Decantação (Torta)	0,54
Etapa de Fermentação	5,17
Etapa de Destilação (Vinhaça)	0,18
Perdas Indeterminadas	4,05
<b>Total</b>	<b>14,14</b>

Fonte: (NOVACANA, 2018). Adaptação do próprio autor.

Realizados os levantamentos de formulação e porcentagens nas análises de perdas foi aplicado o valor em área encontrado a partir das representações produzidas nas formulações e porcentagens dos cálculos de produção se realizando o levantamento final de valores de etanol em usinas sucroalcooleiras.

Como cita Delgado (2019) e Conab (2021) considera-se a eficácia de produção de etanol por tonelada de cana-de-açúcar, atribuindo a relação: cada 1 tonelada de cana processada resultará em uma produção de 80 litros a 90 litros de etanol hidratado (eq. 11).

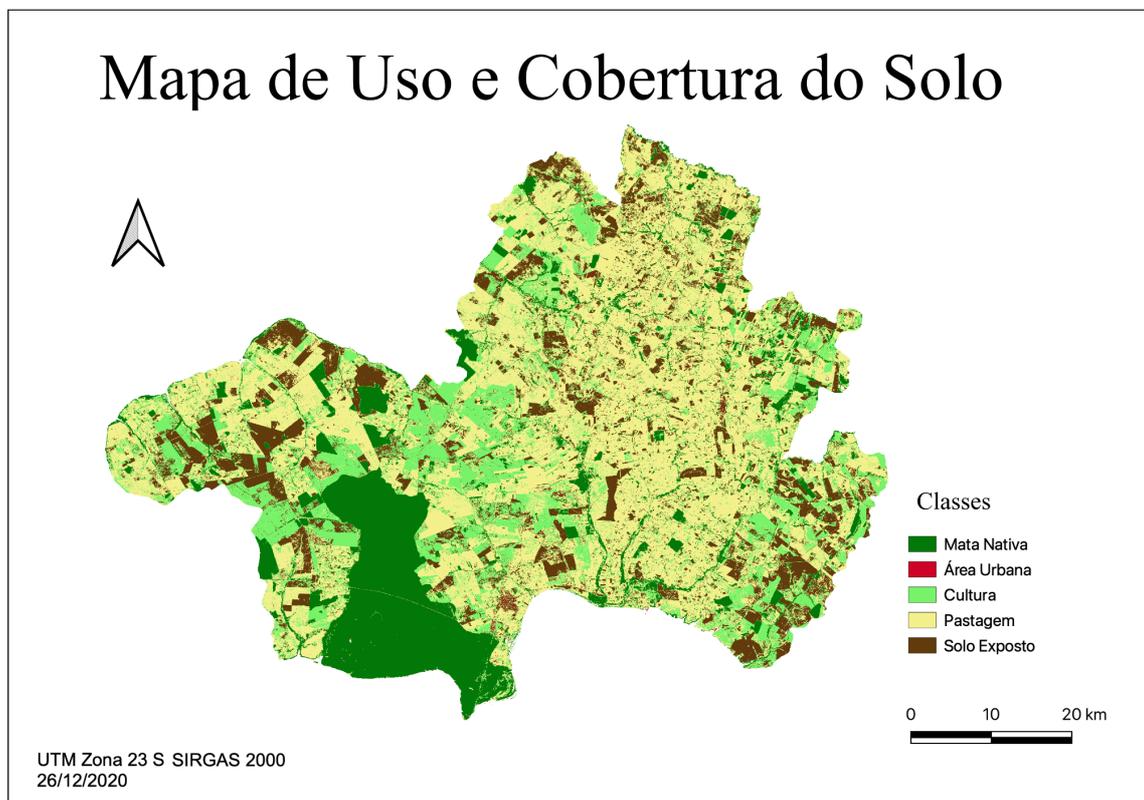
$$1 \text{ tonelada de cana} \rightarrow \text{de 80 a 90 litros de etanol hidratado} \quad (11)$$

A partir dos valores de área da classe de pastagem obtidos do mapeamento de uso e cobertura do solo e do levantamento de perdas e eficiência de produção foi possível estimar o potencial em litros de etanol na região.

## 6. RESULTADOS E DISCUSSÕES

A partir do desenvolvimento metodológico foi possível gerar uma representação visual do uso e cobertura do solo da área estudada juntamente com a construção de um cenário de estimativa do potencial de produção de etanol da região estudada. Neste trabalho, os resultados estão dispostos em um mapa e tabelas para facilitar a análise. A Figura 10 apresenta o mapa de uso e cobertura do solo da área de estudos resultante da classificação supervisionada produzida no GEE, através do algoritmo *Random Forest* e das bandas B2, B3, B4, B5, B6, B7, B8, B11, B12 do satélite Sentinel 2 para a data de 26 de dezembro de 2020. Na legenda são apresentadas as classes que compõem o mapa.

**Figura 10:** Mapa de uso e cobertura do solo.



Fonte: Elaborado pelo próprio autor.

Ao analisarmos a Figura 10 verificamos uma maior concentração das áreas de cultura na porção sul e à leste da área de estudos. De acordo com Thomaz Jr. (2009), esse fato se deve à expansão da monocultura de cana-de-açúcar na região em função da implantação de usinas canavieiras em vários municípios, intensificando o cultivo.

Ainda de acordo com o mapa de uso e cobertura do solo é possível notar que a região possui grande presença de áreas de pastagem. Como descrito por Bento (2020), a região possui uma indústria agropecuária forte, mas que vem sendo substituída pelo plantio de cana devido a sua maior rentabilidade e incentivos municipais. Devido a isso, futuramente a região pode experimentar o aumento da classe de cultura em seu território.

A fim de validar o mapa gerado pela classificação das imagens Sentinel 2 foi produzida a matriz de confusão (Tabela 7). A matriz de confusão mostra em sua diagonal o número de pixels de verdade terrestre corretamente rotulados pelo classificador e, fora da diagonal, o número de pixels que foi classificado erroneamente para cada classe.

**Tabela 7:** Matriz de Confusão

	Mata Nativa	Área Urbana	Cultura	Pastagem	Solo Exposto	Total	Erro de Inclusão (%)
Mata Nativa	74.880	0	6	1	0	74.887	0,01
Área Urbana	0	629	3	8	105	745	15,57
Cultura	3	1	27.832	259	22	28.117	1,01
Pastagem	3	2	140	29.520	14	29.679	0,54
Solo Exposto	0	14	13	4	28.021	28.052	0,11
Total	74.886	646	27.994	29.792	28.162	161.480	
Erro de Omissão (%)	0,01	2,63	0,58	0,91	0,50		

Fonte: Elaborado pelo próprio autor.

Analisando a matriz de confusão nota-se que as classes que apresentaram maior confusão foram as de Cultura e Pastagem, essa confusão se dá pela proximidade da resposta espectral dessas classes entre si. Mas os valores de confusão representam uma interferência baixa no resto da classificação, representando 0,24% de pixels na classificação. As classes de Área Urbana e Solo Exposto também foram as classes que apresentaram maior confusão, também em função das respostas espectrais serem semelhantes, mas essa confusão gera uma baixíssima interferência no resultado, representando 0,07% dos pixels na classificação. As demais confusões apresentaram valores bem inferiores aos apresentados por essas classes.

Com base nos valores que estão fora da diagonal principal da matriz de confusão foram determinados os erros de omissão e inclusão (Tabela 7). Os erros de omissão foram de: 0,01%, 15,57%, 1,01%, 0,54% e 0,1%, respectivamente para as classes Mata Nativa, Área Urbana, Cultura, Pastagem e Solo Exposto e os erros de inclusão encontrados foram: 0,01%, 2,63%, 0,58%, 0,91% e 0,5% para as classes Mata Nativa, Área Urbana, Cultura, Pastagem e Solo Exposto, respectivamente. Tanto o erro de omissão quanto o de inclusão foram bastante baixos,

com exceção da área urbana que apresentou um erro de inclusão de 15,57%, em função de parte dos pixels da classe de solo exposto (105 pixels) terem sido classificados como área urbana.

A classificação supervisionada da qual resultou esse mapa temático mostrou resultados bastante satisfatórios, ou seja, condizentes com a verdade terrestre da região, o que pode ser verificado através do índice Kappa e da Acurácia Global, determinados com base na matriz de confusão. Ambos têm o objetivo de representar o grau de confiabilidade dos dados classificados, de modo que quanto mais próximo de 1 ou 100%, maior concordância os dados classificados têm com os dados de referência. Os valores de Índice de Acurácia Global e Índice Kappa foram de 99,78% e 99,70%, respectivamente (Tabela 8). O valor do Índice Kappa, segundo Landis e Koch (1997), foi considerado “excelente”.

**Tabela 8:** Índices Kappa e Acurácia Global.

<b>Índices</b>	
<b>Acurácia</b>	0.9977905647154292
<b>Kappa</b>	0.9969595919883522

Fonte: Elaborado pelo próprio autor.

Devido aos valores altos encontrados é possível afirmar que a GEE atuou como um excelente classificador, o índice Kappa com 99,7% de precisão demonstra que houve poucas confusões de pixel, a AG mostrou valores considerados excelentes de classificação.

A partir dos dados da matriz de confusão foram determinados os valores de Acurácia do Produtor e Acurácia do Usuário. Na Tabela 9 são apresentados os valores de Acurácia do Produtor.

**Tabela 9:** Valores de Acurácia do Produtor (ferramenta).

<b>Acurácia Produtor</b>	
<b>Mata Nativa</b>	0.9999065258322539
<b>Área Urbana</b>	0.8442953020134228
<b>Cultura</b>	0.9898637834761888
<b>Pastagem</b>	0.9946426766400486
<b>Solo Exposto</b>	0.9988593020354329

Fonte: Elaborado pelo próprio autor.

Os valores de acurácia para cada classe, com exceção da área urbana, foram acima de 0,97. A classe Mata Nativa foi a mais bem identificada, com valores de acurácia iguais a 0,999, tanto para a acurácia do produtor quanto do usuário. A acurácia do produtor demonstra a veracidade da resposta da ferramenta, neste caso o GEE. Segundo os dados apresentados, a classificação atingiu valores de alta excelência, a classe de área urbana é a única que se encontra inferior a 90%, esse valor de 84,43% deve-se muito provavelmente a confusão de pixels presente na matriz. Essa confusão se deve a presença de cidades de pequeno porte na região do Pontal, que gera uma impossibilidade de uma coleta alta de pontos amostrais e a presença de estradas não pavimentadas que possuem resposta espectral semelhante a solo exposto. Pequenas cidades/áreas urbanas tendem a apresentar um valor menor de acurácia devido a ser uma classe muito mista (tendo a presença desde vegetação até áreas de solo exposto em uma área pequena).

A Tabela 10 apresenta os valores de Acurácia do Usuário, que apresentou valores elevados de resposta, legitimando uma excelente classificação. Os valores elevados de acurácia devem-se à boa coleta de pontos amostrais feita pelo operador.

**Tabela 10:** Valores de Acurácia do Usuário.

<b>Acurácia Consumidor/Usuário</b>	
<b>Mata Nativa</b>	0.9999198782148866
<b>Área Urbana</b>	0.9736842105263158
<b>Cultura</b>	0.9942130456526399
<b>Pastagem</b>	0.9908700322234156
<b>Solo Exposto</b>	0.9949932533200767

Fonte: Elaborado pelo próprio autor.

Após a análise da qualidade do mapeamento, na etapa de pós-processamento foi realizada no QGis a determinação das áreas em hectare, de cada uma das classes do mapa de uso e cobertura do solo (Tabela 11).

**Tabela 11:** Valores em Hectare Encontrados.

<b>Classes</b>	<b>Valor Área (hectare)</b>	<b>Valor Área (%)</b>
<b>Mata Nativa</b>	68.015,19	16,65
<b>Área Urbana</b>	696,71	0,17
<b>Cultura</b>	77.227,65	18,90
<b>Pastagem</b>	196.416,71	48,08
<b>Solo Exposto</b>	66.160,98	16,20
<b>Total</b>	408.517,24	100%

Fonte: Elaborado pelo próprio autor.

Ao quantificarmos as áreas em função do tipo de uso e cobertura (Tabela 11), verificamos que a classe pastagem ocupou mais de 48,08% da área de estudo. A classe cultura foi a segunda maior em extensão territorial e na sequência as áreas de mata nativa e solo exposto.

Através desses valores de áreas calculados foi possível fazer uma análise da estimativa da substituição da pastagem por plantações de cana-de-açúcar, com a finalidade da produção do etanol. Através do valor encontrado de 196.416,71 hectares de área destinada a pastagem e as estimativas de eficiência e perdas propostas pela literatura levantada na metodologia chegou-se na seguinte relação: “1 hectare de plantio de cana-de-açúcar possui potencial de produção de 70 toneladas de cana”. Nesse caso a região teria um potencial de produção de 13.749.169,7 de toneladas (CONAB, 2021).

Durante o processo de produção há as principais perdas, descritas na Tabela 6. Aplicando os valores de perda no resultado do potencial de produção em toneladas, segundo o cenário proposto por Novacana (2018), o somatório das perdas de processo equivale a 14,14% por tonelada. Na situação criada para esse trabalho, o valor de toneladas perdido durante o processo de produção encontrado foi de 1.944.132,60 toneladas, que equivale a 14,14% de 13.749.169,7 toneladas de cana-de-açúcar. Logo o que será utilizado para a produção do etanol corresponde a 11.805.037,10 toneladas.

Para Delgado (2019) e Conab (2021) a estimativa de produção de etanol hidratado pode ser relacionada da seguinte forma, segundo as normas de produção nacional, “1 tonelada de cana processada produz de 80 a 90 litros de etanol hidratado”. Utilizando a norma de produção nacional, a área de estudos composta por 4 municípios do Pontal do Paranapanema possui um potencial futuro de produção que varia entre 944.402.968,4 a 1.062.453.339 de litros de etanol

para ser utilizado no setor de transportes do país, isso no caso de substituição das áreas de pastagens determinadas no mapa de uso e cobertura do solo por cultura de cana-de-açúcar.

## 7. CONCLUSÃO

A partir dos resultados obtidos com a classificação de uso e cobertura do solo no Pontal do Paranapanema foi observado que a região possui um enorme potencial futuro de produção de biocombustíveis, podendo atrair investimentos de indústrias interessadas na expansão sucroalcooleira na região do Oeste Paulista. Embora o potencial agrícola da região ainda seja pouco explorado, a área possui solo e relevo propícios para a expansão de lavouras.

Em relação à ferramenta utilizada, o Google Earth Engine se mostrou uma excelente opção, possuindo como principais vantagens, a eficiência de processos, a agilidade de processamento e a execução de todas as análises na nuvem, evitando assim a necessidade de download de bancos de dados de imagens pesados. Essas vantagens contribuíram para bons resultados durante esse trabalho, como os valores do índice Kappa e da Acurácia Global que tiveram valores de precisão de 99,78% e 99,70%, respectivamente. Esses valores, segundo a literatura, legitimaram a classificação como excelente, comprovando que essa ferramenta vinculada com a metodologia aplicada foi considerada eficaz em gerar resultados coerentes com a verdade terrestre. Essa eficiência está atrelada ao uso do algoritmo Random Forest que alinhado com os índices espectrais trouxe uma boa resposta de classificação.

Foi constatado através da realização do estudo de caso proposto por esse trabalho que a região tem um potencial de produção entre 944.402.968,4 a 1.062.453.339 litros de etanol. Devido a esse elevado potencial e os incentivos mundiais para a adoção de biocombustíveis e fontes de baixa emissão de gases causadores de efeito estufa, a região nos próximos anos pode enfrentar uma mudança na utilização do solo, pois a substituição de terras destinadas à pecuária por lavouras vem se tornando algo recorrente nos últimos anos, principalmente e em função do maior retorno financeiro que o arrendamento para produção de cana-de-açúcar traz para os proprietários de terra.

Segundo as análises efetuadas com os resultados obtidos na pesquisa, verificamos que há um enorme potencial econômico viável de produção de biocombustíveis. A região ainda se encontra pouco explorada, pesquisas como essa trazem comprovações plausíveis de que a região é atrativa para investimentos no setor, favorecendo assim a economia e a população da região.

## REFERÊNCIAS

- ALENCAR, Ane *et al.* Mapping Three Decades of Changes in the Brazilian Savanna Native Vegetation Using Landsat Data Processed in the Google Earth Engine Platform. **Mdpi - Multidisciplinary Digital Publishing Institute**, Basel, Switzerland, v. 12, n. 6, p. 1-23, mar. 2020.
- ANDRADE, Ana Beatriz Silva de. **UTILIZAÇÃO DOS ÍNDICES NDWI E MNDWI NA DETECÇÃO DE CORPOS HÍDRICOS EM IMAGENS SENTINEL-2 NA BACIA HIDROGRÁFICA DO RIO TRAIPU – ALAGOAS**. 2019. TCC (Graduação) - Curso de Engenharia de Agrimensura, Centro de Ciências Agrárias-Ceca, Universidade Federal de Alagoas- Ufal Centro de Ciências Agrárias-Ceca Curso de Engenharia de Agrimensura, Rio Largo - Alagoas, 2019.
- ANDRADE, Ednilton Tavares de; CARVALHO, Sergio Roberto Garcia de; SOUZA, Lucas Fernandes de. PROGRAMA DO PROÁLCOOL E O ETANOL NO BRASIL. **Engevista**, Rio de Janeiro, v. 11, n. 2, p. 127-136, dez. 2009. Disponível em: <http://umbu.uft.edu.br/handle/11612/3398>. Acesso em: 06 dez. 2021.
- ATVOS (Brasil). Grupo Novonor (emp.). **Sobre a Empresa**. 2021. Disponível em: <https://atvos.com/a-empresa/sobre-a-empresa/>. Acesso em: 25 nov. 2021.
- BANK, Deutsche (org.). **Covid-19 and Commodities**. 2020. Disponível em: <https://flow.db.com/trade-finance/covid-19-and-commodities>. Acesso em: 05 nov. 2021.
- BARRETO, M. J. THOMAZ JR, A. **O cenário do agronegócio canavieiro na região do Pontal do Paranapanema-SP**. In: XXI Encontro Nacional de Geografia Agrária, UFU – MG, 2012.
- BELGIU, M.; DRĂGUT, L. Random Forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, v 114, p. 24-31, abr. 2016.
- BENTO, Fredi dos Santos. Os impactos da expansão do agronegócio canavieiro para as relações campo-cidade no Pontal do Paranapanema (SP). **Revista Geografia em Atos (Geo Atos online)**, v. 03, n. 18, p. 181-205, maio-ago, 2020. DOI: 10.35416/geoatos.v3i18.7219
- CAMPBELL, J.B.; WYNNE, R.H. **Introduction to Remote Sensing**. 4.ed.2011.
- CHO, David Fernando *et al.* Desempenho do Algoritmo de Classificação de Imagens Random Forest para Mapeamento do Uso e Cobertura do Solo no Cerrado Brasileiro. **Anuário do Instituto de Geociências**, Rio de Janeiro, v. 44, mar. 2021.
- COHEN, J. A. Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, v. 20, n. 01, p. 37-46, 1960.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO (Brasil). Governo Federal. **INFORMAÇÕES AGROPECUÁRIAS CUSTOS DE PRODUÇÃO**. 2021. Disponível em: <https://www.conab.gov.br/info-agro/custos-de-producao>. Acesso em: 29 nov. 2021.

CONGALTON, R.G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. CRC press, 2008.

COSTA, Juliana Arouche. **A Indústria Sucroalcooleira: uma análise da competitividade do setor no período 2003 a 2018**. 2019. Monografia (TCC) - Curso de Ciências Econômicas, Universidade Federal do Maranhão, São Luís, 2019.

DELGADO, Afrânio Antônio; CESAR, Marco Antônio Azeredo; SILVA, Fábio Cesar da. **Elementos de Tecnologia e Engenharia da Produção do Açúcar, Etanol e Energia**. Piracicaba: Fealq, 2019. 984 p.

DE GÓES-FAVONI, Silvana Pedrosa et al. Fermentação alcoólica na produção de etanol e os fatores determinantes do rendimento. **Revista Ibero-Americana de Ciências Ambientais**, v. 9, n. 4, p. 285-296, 2018.

Empresa de Pesquisa Energética. Epe. Ministério de Minas e Energia. **Balço Energético Nacional 2021: Ano base 2020**. Rio de Janeiro: Empresa de Pesquisa Energética, 2021

ESA – EUROPEAN SPATIAL AGENCY (Paris França). **Satellite Constellation**. 2021. Disponível: [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Sentinel-2/Satellite\\_constellation](https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Satellite_constellation). Acesso em: 01 dez. 2021.

FOOD AND AGRICULTURE ORGANIZATION (FAO). **Farming**. Roma: 2021. Disponível em: <https://www.fao.org/faostat/en/#data/QC>. Acesso em: 19 dez. 2021.

FERNANDES, Antonio Carlos. **CÁLCULOS NA AGROINDÚSTRIA DA CANA-DE-AÇÚCAR**. 3. ed. Piracicaba: Stab, 2011. 416 p.

FORMAGGIO, Antonio Roberto; SANCHES, Ieda Del'Arco. **Sensoriamento Remoto em Agricultura**. São Paulo: Oficina de Textos, 2017.

GROFALO, Danilo Francisco Trovo; PACKER, Ana Paula Contador; RAMOS, Nilza Patricia; KONDO, Vitor Yukio; FOLEGATTI, Marília Ieda da Silveira; CABRAL, Osvaldo Machado Rodrigues. **Dinâmica do cultivo de cana-de-açúcar no Brasil – 1990 a 2018**. 2020. Empresa Brasileira de Pesquisa Agropecuária Embrapa Meio Ambiente e Ministério da Agricultura, Pecuária e Abastecimento.

IBGE - Instituto Brasileiro de Geografia e Estatística - Rio de Janeiro - RJ). Instituto Nacional de Estatística. **Pesquisa do Censo 2020**. 2020. Disponível em: <https://www.ibge.gov.br/ibge/est%C3%ADstica/popula%C3%A7%C3%A3o/censo2000>. Acesso em: 127 out. 2021.

JENSEN, J. R. **Sensoriamento remoto do ambiente: Uma perspectiva em recursos terrestres**. Parêntese, São José dos Campos, 598 p, 2009.

KOHLHEPP, G. Análise da situação da produção de etanol e biodiesel no Brasil. **Estudos Avançados**, v.24, p.223-253, 2010. DOI: <https://doi.org/10.1590/S0103-40142010000100017>.

SOARES, André Augusto; ZUKOWSKI JUNIOR, Joel Carlos. O Brasil como Grande Player no Mercado Mundial de Etanol. **Revista de Política Agrícola**, [s. l], v. 30, n. 3, p. 57-71, jul. 2021.

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**. New York: John Wiley e Sons Inc., 1995, 708p.

MACEDO, I.C.; SEABRA, J. E. A.; SILVA, J. E. **Emissões de gases de efeito estufa na produção e uso de etanol de cana-de-açúcar no Brasil: as médias de 2005/2006 e uma previsão para 2020**. Bioenergia de Biomassa. 2008; 32: 4.

M.C. Peel; B.L. Finlayson; T.A. McMahon. **Updated world map of the Köppen-Geiger climate classification**. Hydrol. Earth Syst. Sci., v. 11, p. 1633-1644, 2007.

MORAES, Elizabeth Caria; FIORIO, Peterson Ricardo. **Fundamentos do Sensoriamento Remoto**. São José dos Campos: Inpe, 2002. 43 slides, color. Disponível em: [http://www.leb.esalq.usp.br/leb/disciplinas/Topo/LEB5838/Peterson/Fundamentos\\_energia\\_pos.pdf](http://www.leb.esalq.usp.br/leb/disciplinas/Topo/LEB5838/Peterson/Fundamentos_energia_pos.pdf). Acesso em: 04 nov. 2021.

NOVACANA (Marabá Paulista - SP). **Usina Decasa Marabá Paulista**. 2021. Disponível em: [https://www.novacana.com/usinas\\_brasil/fabrica/decasa-maraba-paulista](https://www.novacana.com/usinas_brasil/fabrica/decasa-maraba-paulista). Acesso em: 09 out. 2021.

NOVACANA. **Possíveis melhorias nas usinas de produção de etanol**. 2018. Disponível em: <https://www.novacana.com/etanol/melhorias-usinas-producao-etanol>. Acesso em: 29 nov. 2021.

OLIVEIRA, Edilaine Cristina de *et al.* **Mapeamento de uso e cobertura da terra utilizando a plataforma Google Earth Engine**. 2021. 89 f. TCC (Graduação) - Curso de Geografia, Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina, Universidade Federal de Santa Catarina, Florianópolis - Sc, 2021.

PAREDES, D. A. C; SOUZA, V; VOLPATO, M. M. **VIABILIDADE DA PLATAFORMA GOOGLE EARTH ENGINE EM PESQUISAS CAFEIEIRAS NO SUL DE MINAS GERAIS.** In: **ANAIS DO XIX Simpósio BRASILEIRO DE SENSORIAMENTO REMOTO**, 2019, Santos. Anais eletrônicos... São José dos Campos, INPE, 2019.

PAYNE, J. H. **Operações unitárias na produção de açúcar de cana.** São Paulo: Nobel/Stab, 1989.

QGIS DEVELOPMENT TEAM (Estados Unidos das Américas). Open-Source Geospatial Foundation. **Quantum GIS.** 2021. Disponível em: <https://www.qgis.org/en/site/>. Acesso em: 28 out. 2021.

RAÍZEN ENERGIA (Piracicaba). **Relatório Anual 2020/2021.** 2021. Disponível em: <https://www.raizen.com.br/relatorioanual/2021/pdf/raizen-rs2021-pt.pdf>. Acesso em: 25 nov. 2021.

REZENDE, Wallace Silva. **O SETOR SUCROALCOOLEIRO NO PONTAL DO PARANAPANEMA: UMA ANÁLISE A PARTIR DO LICENCIAMENTO AMBIENTAL.** 2011. 128 f. TCC (Graduação) - Curso de Engenharia Ambiental, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista Júlio de Mesquita Filho - Unesp, Presidente Prudente, 2011.

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.

SANTOS, Samuel César Silva dos. **ESTUDO DE OTIMIZAÇÃO PARA O PROCESSO DE PRODUÇÃO DO ETANOL EM UMA EMPRESA NA REGIÃO DE OURINHOS/SP.** 2014. TCC (Graduação) - Curso de Engenharia de Produção, Univem Curso de Engenharia de Produção, Fundação de Ensino “Eurípides Soares da Rocha” Centro Universitário Eurípides de Marília – Univem Curso de Engenharia de Produção, Marília, 2014.

SEARS, F; YOUNG, Hii Diii; ZEMANSKY, Mi Wi. **Física IV: Ótica e Física Moderna.** 14. ed. São Paulo: Pearson Education do Brasil, 2016. 534 p. Colaborador A. Lewis Ford; tradução Daniel Vieira; revisão técnica Adir Moysés Luiz.

SCHOTT, Madison. **Random Forest Algorithm for Machine Learning:** Part 4 of a Series on Introductory Machine Learning Algorithms. 2019.

SZMRECSÁNYI, T. **O planejamento da agroindústria canavieira no Brasil (1930-1975).** São Paulo: HUCITEC, Universidade Estadual de Campinas (UNICAMP), 1979.

THE ECONOMIST. **Was COP26 in Glasgow a success? The Economist.** Glasgow Scotland, p. 109-110. 14 nov. 2021. Disponível em:

<https://www.economist.com/international/2021/11/14/was-cop26-in-glasgow-a-success>. Acesso em: 17 nov. 2021.

THOMAZ JÚNIOR, A. **Dinâmica Geográfica do Trabalho no Século XXI. (Limites Explicativos, Autocrítica e Desafios Teóricos)**. p.941, Livre Docência - Universidade Estadual Paulista. Presidente Prudente. 2009.

UMOE BIOENERGY (Brasil). Umoe As. **Sobre a Empresa**. 2021. Disponível em: <http://umoebioenergy.com/empresa/>. Acesso em: 25 nov. 2021.

VALE, Jones Remo Barbosa. **ANÁLISE DA DINÂMICA DO USO E COBERTURA DA TERRA NAS ÁREAS DESFLORESTADAS DO ESTADO DO PARÁ POR MEIO DA PLATAFORMA GOOGLE EARTH ENGINE**. 2019. 69 f. TCC (Graduação) - Curso de Engenharia Cartográfica e de Agrimensura, Universidade Federal Rural da Amazônia, Belém - PA, 2019.

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.