

MIQUEIAS LIMA DUARTE

**PREVISÃO DA SUSCETIBILIDADE À INCÊNDIOS E QUEIMADAS
UTILIZANDO UM MODELO BASEADO EM INTELIGÊNCIA
ARTIFICIAL E SISTEMA DE INFERÊNCIA FUZZY**

Sorocaba
2022

PROGRAMA DE PÓS-GRADUAÇÃO em

*ciências
ambientais*

MIQUEIAS LIMA DUARTE

**PREVISÃO DA SUSCETIBILIDADE À INCÊNDIOS E QUEIMADAS
UTILIZANDO UM MODELO BASEADO EM INTELIGÊNCIA
ARTIFICIAL E SISTEMA DE INFERÊNCIA FUZZY**

Tese de doutorado apresentada como requisito para a obtenção do título de Doutor em Ciências Ambientais da Universidade Estadual Paulista "Júlio de Mesquita Filho" na Área de Concentração Diagnóstico, Tratamento e Recuperação Ambiental

Orientador: Prof. Dr. Roberto Wagner Lourenço

Sorocaba
2022

PROGRAMA DE PÓS-GRADUAÇÃO em

ciências
ambientais



D812p

Duarte, Miqueias Lima

Previsão da suscetibilidade à incêndios e queimadas utilizando um modelo baseado em inteligência artificial e sistema de inferência fuzzy / Miqueias Lima Duarte. -- Sorocaba, 2022

130 p. : il., tabs., mapas

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia, Sorocaba

Orientador: Roberto Wagner Lourenço

1. Sensoriamento Remoto. 2. Método Boruta. 3. Aprendizagem de máquina. 4. lógica fuzzy. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Ciência e Tecnologia, Sorocaba. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

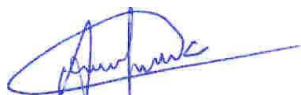
CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: Previsão da suscetibilidade à incêndios e queimadas utilizando um modelo baseado em inteligência artificial e sistema de inferência fuzzy

AUTOR: MIQUEIAS LIMA DUARTE

ORIENTADOR: ROBERTO WAGNER LOURENÇO

Aprovado como parte das exigências para obtenção do Título de Doutor em CIÊNCIAS AMBIENTAIS, área: Diagnóstico, Tratamento e Recuperação Ambiental pela Comissão Examinadora:



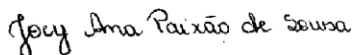
Prof. Dr. ROBERTO WAGNER LOURENÇO

Departamento de Engenharia Ambiental / Instituto de Ciência e Tecnologia Campus de Sorocaba Unesp



Prof. Dr. JOSÉ CARLOS DE SOUZA

Departamento de Geografia / Universidade Estadual de Goiás - Campus Minaçu



Dr^a. JOCY ANA PAIXÃO DE SOUSA

Universidade Estadual Paulista (Unesp)



Prof. Dr. ADMILSON IRIO RIBEIRO

Engenharia Ambiental / Unesp - ICT Sorocaba



Prof. Dr. Elfany Reis do Nascimento Lopes

Universidade Federal do Sul da Bahia (UFSB) - Porto Seguro

Sorocaba, 10 de fevereiro de 2022

AGRADECIMENTOS

Agradeço a Deus por ter me concedido vida, saúde, paz, alegria, motivação e força em continuar caminhando na jornada da vida, sem essa força divina, nenhuma conquista seria possível.

A toda minha família, em especial minha mãe, Maria e meu pai Jaílson pelo apoio incondicional, pelo exemplo de vida, educação e incentivo para vencer na vida.

A minha esposa Tatiana pelo apoio, paciência e incentivo irrestrito e fundamental durante estes anos, em especial durante a quarentena da Covid-19.

Ao meu orientador Prof. Dr. Roberto Wagner Lourenço, pela orientação, ajuda e incentivo que refletiram no meu crescimento pessoal e profissional.

Ao Laboratório de Geoprocessamento e Modelagem Matemática Ambiental (LABGEMM) do qual fiz parte nos últimos anos, principalmente aos amigos(as) Jocy, Bruna, Bruna II, Hetiany, Leticia, Camille, Rita, Amazonino. À Giovana pelo suporte técnico.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES - pela concessão de bolsa de estudo.

Ao Programa de Pós-Graduação em Ciências Ambientais da UNESP pela oportunidade de cursar o Doutorado.

Aos professores integrantes do Programa de Pós-Graduação em Ciências Ambientais da UNESP pelo profissionalismo e alto nível de conhecimento compartilhado nas disciplinas que cursei.

Em especial, aos meus amigos e professores da graduação, mestrado e doutorado, e a todos aqueles que de algum modo contribuíram direta ou indiretamente para a consolidação deste trabalho.

Muito obrigado.

EPIGRAFE

“Não venci todas batalhas que lutei, mas perdi todas que deixei de lutar”
Cecília Meireles

“O sucesso é ir de fracasso em fracasso sem perder o entusiasmo”
Winston Churchill

*“Pois quanto maior a sabedoria, maior o enfado; e
quanto maior o conhecimento, maior o desgosto”*
Eclesiastes 1:18

RESUMO

Os incêndios florestais são eventos globais que causam perdas imensuráveis para o homem e para o meio ambiente. A previsão e mapeamento desses eventos pode ser uma importante medida, uma vez que possibilita desenvolver estratégias com vistas ao controle ou a prevenção. Nesse contexto, esta pesquisa teve por objetivo desenvolver uma metodologia para a previsão da Suscetibilidade à Incêndios e Queimadas (SIQ), a partir de dados de sensoriamento remoto, com uso de um método misto de aprendizagem de máquina, tendo como unidade de pesquisa a bacia hidrográfica do rio Sorocabaçu, no município de Ibiúna/São Paulo, Brasil. Foram utilizadas 14 variáveis potencialmente influenciadoras, sendo quatro fatores climáticos, dois antrópicos, quatro topográficos e quatro fatores relacionados às características da vegetação. Os fatores mais importantes foram selecionados com uso do algoritmo Boruta e, posteriormente, estes foram utilizados para compor um modelo de classificação com uso do Sistema de Inferência Fuzzy Híbrido (HFIS). O sistema foi implementado com base em dados de treinamento (70%) e avaliado com base em dados independentes (30%). Além disso, o sistema também foi implementado em dois períodos climáticos extremos (2020 e 2018). Os resultados obtidos mostraram que os fatores mais importantes estão relacionados às características da vegetação, seguido pelos fatores climáticos e antrópicos. A implementação do sistema HFIS mostrou que a função de associação fuzzy gaussiana com particionamento em cinco valores linguísticos apresentou os melhores ajustes. Considerando o conjunto de dados independentes, o modelo HFIS apresentou bom desempenho na predição para o ano de 2019 (Acurácia de 0,93 e índice Kappa de 0,86), apresentando um valor de AUC de 93,3%, o que representa boa capacidade de replicação do modelo proposto, uma vez que nos períodos climáticos considerados extremos, o modelo apresentou valores de AUC superiores a 90%. A aplicação do modelo HFIS mostrou que em 2019 cerca de 65,93% da área foi classificada com muito baixo e baixo SIQ, enquanto as áreas com maior SIQ foram minoritárias (16,18%) e, apesar disso, é neste último onde ocorrem cerca de 76,69% dos eventos de incêndios e queimadas. Estes resultados confirmam a eficiência do modelo proposto e sua capacidade em modelar problemas não lineares complexos. A implementação do sistema HFIS mostrou-se adequado para a previsão da SIQ, e os resultados obtidos podem ser utilizados para auxiliar os gestores públicos no seu planejamento, com vista à prevenção e mitigação dos eventos.

Palavras-chave: Sensoriamento Remoto, Método Boruta, Aprendizagem de máquina, lógica fuzzy.

ABSTRACT

Forest fires are global events that cause immeasurable losses to man and the environment. The prediction and mapping of these events can be an important measure, since they allow us to develop strategies for either control or prevention. In this context, this research aimed to develop a method for the prediction of Susceptibility to Fires and Burnings (SFB) via remote sensing data, and using a mixed method of machine learning, for which the research unit was the basin of the Sorocabuçu River, in the municipality of Ibiúna, São Paulo, Brazil. We used fourteen variables of potential influence, which were four climatic factors, two anthropic factors, four topographic factors and four factors related to vegetation characteristics. The most important factors were selected using the Boruta algorithm, and these were subsequently used to compose a classification model using the Hybrid Fuzzy Inference System (HFIS). The system was implemented on the basis of training data (70%) and evaluated on the basis of independent data (30%). In addition, the system was implemented in two periods of extreme weather (2020 and 2018). The results showed that the most important factors are related to vegetation characteristics, followed by climatic and anthropic factors. The implementation of the HFIS system showed that the fuzzy Gaussian association function with partitioning in five linguistic values presented the best adjustments. Considering the set of independent data, the HFIS model showed good performance in the prediction for the year 2019 (accuracy of 0.93 and Kappa index of 0.86), and presented an AUC value of 93.3%, which represents good replication capacity for the proposed model, since in the climatic periods that are considered extreme, the model presented AUC values greater than 90%. The application of the HFIS model showed that in 2019 about 65.93% of the area was classified as having very low or low SFB, while the areas with higher SFB were the minority (16.18%) and, despite this, it is in the latter where about 76.69% of the events of fires and burnings occur. These results that confirm the efficiency of the proposed method and its ability to model complex nonlinear problems. The implementation of the HFIS model proved to be adequate for the prediction of SFB, and the results obtained can be used to assist public managers in their planning, with a view to prevention and mitigating events.

Keywords: Remote sensing, Boruta method, Machine Learning, Fuzzy logic.

LISTA DE FIGURAS

Figura 1. Localização da bacia hidrográfica do rio Sorocabuçu.	31
Figura 2. Etapas seguidas para identificação e mapeamento das cicatrizes de incêndios. .	35
Figura 3. Mapa índice das Cartas Topográficas utilizadas.	42
Figura 4. Fluxograma das etapas seguidas para regularização dos dados.	47
Figura 5. Mapa de pontos balanceado de ocorrência e não ocorrência de focos de incêndios e queimadas.	50
Figura 6. Algoritmo de seleção de recursos Boruta implementado.	51
Figura 7. Diagrama esquemático do método HFIS implementado para o mapeamento da suscetibilidade de incêndios e queimadas (SIQ).	53
Figura 8. Matriz de confusão para classificação.	55
Figura 9. Fluxograma da metodologia adotada.	57
Figura 10. Número de focos de incêndios e precipitação anual (a) e média mensal (b) observado entre 2002 a 2020 na bacia hidrográfica do rio Sorocabuçu.	58
Figura 11. Padrões temporais do SPI (a) e de focos de incêndio (b) padronizado (z-score) para a série histórica entre 2002 a 2020.	60
Figura 12. Semivariograma ajustado para a precipitação (a), umidade do ar (b), temperatura (c) e umidade do solo (d).	62
Figura 13. Mapas dos valores médios da precipitação (a), umidade do ar (b), temperatura (c) e umidade do solo (d).	63
Figura 14. Mapa hipsométrico (a), declividade (b), orientação de vertentes (c) e TWI (d).	65
Figura 15. Mapa de cobertura e uso da terra (a), NDVI (a), NDWI (b) e SAVI (c) para o ano de 2019.	67
Figura 16. Mapa de distância (a) e densidade de rodovias (b) na área de estudo.	69
Figura 17. Exemplo de áreas identificadas como cicatrizes de incêndios e queimadas em 2019.	71
Figura 18. Interseção entre focos de incêndios e queimadas e classes de uso e ocupação do solo.	72
Figura 19. Gráficos boxplot e valores do teste Wilconxon para os atributos avaliados em áreas de ocorrência e não ocorrência de incêndios e queimadas.	73
Figura 20. Correlação de Pearson entre as variáveis potenciais de previsão de incêndios e queimadas.	75

Figura 21. Importância relativa das variáveis potenciais para explicação dos eventos pelo método Boruta.	77
Figura 22. Valores linguísticos adotado para as variáveis de entrada.	79
Figura 23. Curva ROC e AUC para os dados de treinamento (a) e teste (b) para previsão de ocorrência de incêndios e queimadas.....	81
Figura 24. Mapa de SIQ obtido pelo método HFIS para o ano de 2019.	83
Figura 25. SIQ obtido pelo método HFIS para as sub-bacias do rio Sorocabuçu.	85
Figura 26. Curva ROC e AUC para os dados de treinamento e teste para previsão de ocorrência de incêndios e queimadas.	88
Figura 27. Mapa de SIQ obtido pelo método HFIS para os anos de 2018 (a) e 2020 (b)..	89

LISTA DE TABELAS

Tabela 1. Características dos satélites e sensores utilizados no monitoramento de focos de calor.	33
Tabela 2. Principais características de dados de umidade do solo estimadas pelo SMAP/Sentinel.	39
Tabela 3. Características principais e bandas do Sentinel-2.	40
Tabela 4. Classes de uso da terra na área.	44
Tabela 5. Descrição das características gerais do conjunto de variáveis utilizadas.	48
Tabela 6. Classes do Índice Padronizado de Precipitação.	56
Tabela 7. Estatística descritiva e geoestatística para as variáveis regularizadas.	61
Tabela 8. Resultado do algoritmo Boruta para selecionar os fatores mais importantes com potencial de explicação dos eventos de incêndios e queimadas.	76
Tabela 9. Métricas de desempenho para os dados de treinamento e teste.	82
Tabela 10. Área e número de focos de incêndios e queimadas por classe de Suscetibilidade de Incêndios e Queimadas-SIQ em 2019.	84
Tabela 11. Área e número de focos de incêndios e queimadas por classe de SIQ nas sub-bacias.	86
Tabela 12. Métricas de desempenho para os dados de treinamento e teste.	88
Tabela 13. Área (hectares e percentual) e número de focos por classe de Suscetibilidade de Incêndios e Queimadas-SIQ para os anos de 2018 e 2020.	89

LISTA DE ABREVIATURAS E SIGLAS

ACC – Acurácia

AUC - *Area Under the Curve*

CF - Certeza Fuzzy

CIIAGRO - Centro Integrado de Informações Agrometeorológicas

DAAC - *Distributed Active Archive Center*

FAO - *Food and Agriculture Organization*

FLDAS - *Famine Land Data Assimilation System*

FN - Falso Negativo

FP - Falso Positivo

GAM - Modelos Aditivos Generalizados

GDE - Grau de Dependência Espacial

GPM - *Global Precipitation Measurement*

HFIS - Sistema de Inferência Fuzzy Híbrido

IBGE – Instituto Brasileiro de Geografia e Estatística

ICMBio - Instituto Chico Mendes de Conservação da Biodiversidade

IG - *Information Gain*

IGC - Instituto Geográfico e Cartográfico

INPE – Instituto Nacional de Pesquisas Espaciais

IR - Importância Relativa

KO - Krigagem Ordinária

MCUT - Mapeamento de Cobertura e Uso da Terra

MDE - Modelo Digital de Elevação

MODIS - *Moderate-Resolution Imaging Spectroradiometer*

NASA - Administração Nacional da Aeronáutica e Espaço

NB - Naive *Bayes*

NBR - Normalized Burn Ratio

NDVI – Índice de vegetação por Diferença Normalizada

NDWI - Índice de Água por Diferença Normalizada

NIR – Infravermelho Próximo

ORQ - *Ordered Quantile*

OSM - *OpenStreetMap*

PMCRMAI - Plano Municipal de Conservação e Recuperação da Mata Atlântica do

Município de Ibiúna

RF - Random Forest

RGB – Red Green Blue

RMSE - *Root-Mean-Square Error*

RNA - Redes Neurais Artificiais

ROC - *Receiver-Operating Characteristic*

SAVI - *Soil Adjusted Vegetation Index*

SIG – Sistema de Informação de Geográfica

SIQ - Suscetibilidade de Incêndios e Queimadas

SMAP - *Soil Moisture Active Passive*

SPF - Especificidade

SPI - *Standardized Precipitation Index*

SR – Sensoriamento Remoto

SST – Sensitividade

SVM - Suporte Vetor Machines

SWIR – Infravermelho de Ondas Curtas

TIN - *Triangulated Irregular Network*

TWI - *Topographic Wetness Index*

USGS - *United States Geological Survey*

UTM - Universal Transversa de Mercator

VN - Verdadeiro Negativo

VP - Verdadeiro Positivo

WMS - *Web Map Service*

Sumário

1. INTRODUÇÃO	15
2. JUSTIFICATIVA.....	17
3. OBJETIVOS	18
4. REVISÃO BIBLIOGRÁFICA.....	19
4.1. QUEIMADAS E INCÊNDIOS	19
4.2. RISCO E PERIGO DE INCÊNDIOS.....	19
4.2. FATORES INFLUENCIADORES.....	21
4.2.1. Aspectos climáticos	21
4.2.2. Aspectos antrópicos.....	22
4.2.3. Fatores topográficos	23
4.2.4. Características da vegetação	24
4.3. SENSORIAMENTO REMOTO E SIG APLICADO À DETECÇÃO DE INCÊNDIOS E QUEIMADAS	25
4.4. TÉCNICAS DE MACHINE LEARNING APLICADO AO MAPEAMENTO DA SIQ.....	27
4.5. SELEÇÃO DE ATRIBUTOS RELEVANTES	29
5. MATERIAIS E MÉTODOS.....	31
5.1. CARACTERIZAÇÃO DA ÁREA DE ESTUDO	31
5.2. IDENTIFICAÇÃO DE FOCOS DE INCÊNDIOS E QUEIMADAS E MAPEAMENTO DE CICATRIZES ..	32
5.3. AQUISIÇÃO E PROCESSAMENTO DE DADOS	35
5.3.1. Precipitação	36
5.3.2. Temperatura próximo a superfície (LST)	37
5.3.3. Umidade do ar	37
5.3.4. Umidade do solo.....	38
5.3.5. Índices de vegetação.....	40
5.3.6. Dados planialtimétricos	42
5.3.7. Cobertura e Uso da Terra	44
5.4. REGULARIZAÇÃO DE DADOS DE ENTRADA.....	45
5.5. PREVISÃO DA SUSCETIBILIDADE DE INCÊNDIO E QUEIMADAS	48
5.6. ANÁLISE DE QUALIDADE DOS MODELOS	54
6. RESULTADOS E DISCUSSÃO	58
6.2. SAZONALIDADE DE INCÊNDIOS E QUEIMADAS	58
6.3. DESCRIÇÃO DAS VARIÁVEIS DE ENTRADA	61
6.4. ANÁLISE EXPLORATÓRIA DOS DADOS	70
6.5. IMPORTÂNCIA DAS VARIÁVEIS	74
6.6. PREVISÃO DA SUSCETIBILIDADE DE INCÊNDIOS E QUEIMADAS (SQI)	78
7. CONCLUSÕES E RECOMENDAÇÕES	91
REFERÊNCIAS BIBLIOGRÁFICAS	94
ANEXOS	109

1. INTRODUÇÃO

Os problemas causados por incêndios e queimadas são imensuráveis. Além de danos econômicos e os problemas relacionados à saúde pública que são comumente observados de imediato, esses eventos também contribuem com alterações no clima e ambiente em função de grandes quantidades de CO₂ que são emitidos para atmosfera, contribuindo com o aumento do efeito estufa, distúrbios ecológicos, o que tem afetado o ciclo hidrológico, e o aumento da erosão do solo (MACHADO; LOPES, 2014; ARAGÃO et al., 2018; VENKATESH et al., 2020).

No passado, a ocorrência de incêndios e queimadas esteve naturalmente relacionada a flutuações climáticas, tais como alterações na temperatura e precipitação. Nas últimas décadas, a atuação antrópica causou grandes mudanças no regime desses eventos (CHUVIECO et al., 2019), pois as alterações no uso e ocupação do solo associado às mudanças climáticas podem aumentar sua frequência e a gravidade (AQUILUÉ et al., 2020). Entretanto, a compreensão de sua distribuição espacial e temporal não é trivial (MACHADO; LOPES, 2014), pois inclui um conjunto de fatores dinâmicos, tais como a precipitação, temperatura, umidade do ar, uso e ocupação da terra (ARAGÃO et al. 2018; MOTA et al., 2019; POURGHASEMI et al., 2020).

Os dados da *Fire Information for Resource Management System* - FIRMS indicam que entre 2000 a 2018 ocorreram cerca de 7,27 milhões de focos de incêndio e queimadas na América do Sul (NASA, 2020). No Brasil, o número de focos de incêndio e queimadas aumentaram significativamente nos últimos anos devido a uma série de fatores, tais como o desmatamento, atividades agropastoris e queima descontrolada (CAÚLA et al., 2015; BARLOW et al., 2019). Assim como na Amazônia e Cerrado, o bioma Mata Atlântica também experimentou um aumento (cerca de 60%) no número de incêndios e queimadas na última década (INPE, 2020).

Para o gerenciamento efetivo desses eventos, é necessário um planejamento centralizado que envolvem diversas ações, dentre os quais o mapeamento da suscetibilidade e risco de incêndio e queimadas é um passo importante, pois podem ser utilizados para a prevenção e o gerenciamento de áreas críticas (EUGENIO, 2016; NÓBREGA et al., 2018), além de servir de base para o desenvolvimento de sistemas de alerta de incêndio com maior precisão, bem como para o desenvolvimento de um programa institucional consistente (ADAB et al., 2013; WHITE et al, 2016; BARLOW et al, 2019).

Os métodos usualmente utilizados para o mapeamento da suscetibilidade e risco de incêndios e queimadas descritos na literatura incluem a integração de dados de

Sensoriamento Remoto e técnicas estatísticas em Sistema de Informações Geográficas-SIG (JAISWAL et al., 2002; ADAB et al., 2013; MOTA et al., 2019; POURGHASEM et al., 2020). Esses métodos, apesar de serem muito efetivos, exigem uma grande quantidade de variáveis de entrada que em muitos casos nem sempre estão disponíveis ou são de difícil aquisição, além de possuírem graus subjetivos de importância que normalmente são dependentes do contexto geográfico em que são utilizados (CARMO et al., 2011; GRALEWICZ et al., 2012).

Na última década, os métodos de aprendizagem de máquina acompanhado pelo avanço tecnológico do sensoriamento remoto (ADIRI et al., 2020), tem sido utilizado com sucesso para previsão de suscetibilidade e riscos naturais, tais como deslizamentos de terra, inundações, apresentando resultados promissores e mostrando o potencial que essas abordagens podem oferecer nas aplicações de mapeamento de incêndios e queimadas, já que provou seu potencial em outros zoneamentos de suscetibilidade (TERMEH et al., 2018; CAO et al., 2020; NAPOLI et al., 2020; JAIN et al., 2020).

Entretanto, como a modelagem de incêndios e queimadas requer a coleta de fatores influenciadores que possuem diferentes resoluções espaço/temporal, tais como aspectos climáticos, topográficos, antrópicos e uso da terra (POURGHASEM et al., 2020), neste caso, a dificuldade é trabalhar com as incertezas e imprecisões dos dados (BUI et al., 2017). Como resultado, esses modelos têm dificuldade em processar dados com imprecisões de informações em SIG devido a discrepâncias espaciais decorrentes de diferentes resoluções (BUI et al., 2017; MOAYEDI et al., 2020). Dessa forma, o desenvolvimento de novos modelos que possibilitem lidar com incertezas e imprecisões, melhorando a capacidade da previsão de incêndios e queimadas são necessários.

Partindo desse pressuposto, essa pesquisa tem por premissa que é possível prever áreas com suscetibilidade de incêndios e queimadas de forma contínua dentro de um território, a partir de um modelo derivado de dados obtidos por sensoriamento remoto e técnicas de geoprocessamento, com uso de um sistema de inferência fuzzy integrado a aprendizagem de máquina.

2. JUSTIFICATIVA

A ocorrência de incêndios e queimadas no bioma Mata Atlântica tornou-se comum e crescente nas últimas décadas (MENEZES et al., 2017; LOPES et al., 2018; GUEDES et al., 2020; SOUSA, 2021; INPE, 2021). Esses eventos causam perdas econômicas e emitem grandes quantidades de CO₂ para atmosfera, contribuindo com o aumento do efeito estufa (ARAGÃO et al., 2018; VENKATESH et al., 2020), e com a fragmentação de habitats, constituindo um dos principais causadores do declínio da biodiversidade (LOPES et al., 2018; DRISCOLL et al., 2021; SOUSA, 2021).

Nesse contexto, o conhecimento prévio da suscetibilidade e risco de incêndios e queimadas é de fundamental importância, pois possibilita direcionar medidas de gerenciamento em áreas mais suscetíveis, auxiliar na gestão e alocação de equipes de brigadistas, bem como propor práticas de manejo do solo e organização territorial, com vistas a minimizar os danos, caso esses eventos ocorram (WHITE et al., 2016; CORRÊA, 2020; TONINI et al., 2020).

Dessa forma, esta pesquisa propõe um método para previsão da suscetibilidade de incêndios e queimadas que leva em consideração os fatores influenciadores intrínsecos da região, esses dados são processados com base em técnicas de geoprocessamento, em seguida, a previsão da suscetibilidade de incêndios e queimadas é realizada com emprego de um Sistema de Inferência Fuzzy Híbrido (HFIS). A integração de inteligência artificial e lógica fuzzy possibilita aprender os padrões do conjunto de dados, e ao mesmo tempo, lidar com a complexidade e subjetividade em que são intrínsecos (LOURENÇO et al., 2015; BRESSANE et al., 2020; LOPES et al., 2021). Os resultados obtidos são apresentados em forma de mapas que podem subsidiar o manejo constante de medidas de fiscalização nos locais mais suscetíveis.

A definição de uma bacia hidrográfica como unidade de pesquisa deu-se em função de serem unidades básicas de planejamento do uso, da conservação e da recuperação dos recursos naturais, considerada unidade de conservação e planejamento ambiental, conforme definido pela Lei Federal nº 9.433/97. Dessa forma, a bacia do rio Sorocabuçu foi escolhida por fazer parte da Área de Proteção Ambiental Itupararanga (APA Itupararanga), que por sua vez, foi criada com objetivo principal de preservar, conservar e recuperar os recursos naturais, em especial os recursos hídricos e remanescentes florestais (BERNADI et al., 2020). Entretanto, essa unidade apresenta uma intensificação de uso antrópico nos últimos anos, causando a fragmentação florestal, bem como ocorrência de incêndios e queimadas (LOPES et al., 2021; SOUSA, 2021; TAKIKAWA, 2021).

3. OBJETIVOS

3.1. Geral

Desenvolver uma metodologia para a previsão de áreas com suscetibilidade a incêndios e queimadas, a partir de técnicas de aprendizagem de máquina, lógica fuzzy e dados de sensoriamento remoto, na Bacia Hidrográfica do Rio Sorocabuçu.

3.2. Específicos

- Determinar as variáveis potencialmente explicativas de incêndios e queimadas com base em dados de Sensoriamento Remoto;
- Identificar as variáveis mais relevantes para a previsão da suscetibilidade de incêndios e queimadas;
- Implementar um sistema de Inferência Fuzzy Híbrido (HFIS) para prever a suscetibilidade de incêndios e queimadas na área de estudo;
- Mapear a suscetibilidade de incêndios e queimadas na área de estudo com base no modelo HFIS ajustado.

4. REVISÃO BIBLIOGRÁFICA

4.1. Queimadas e incêndios

No que se refere a classificação dos eventos com presença de fogo em áreas rurais e/ou terrenos baldios no perímetro urbano, existem duas definições distintas na literatura as quais podem se enquadrar: as queimadas e os incêndios. Essas definições são encontradas de forma ampla na literatura, e em alguns casos utilizadas de forma generalizada.

De acordo com Manual para Formação de Brigadista de Prevenção e Combate aos Incêndios Florestais do Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio, 2010), o termo queimada controlada pode ser entendido como sendo uma prática agrícola ou florestal que utiliza o fogo de forma racional, isto é, com o controle de sua intensidade e limitado a uma área predeterminada; o termo incêndio florestal (também utilizado para definir incêndios em outros tipos de vegetação tais como capoeiras, campos e pradarias) é entendido como sendo a ocorrência de fogo sem controle que incide sobre qualquer forma de vegetação, podendo ser provocado de forma natural, intencional ou por negligência.

Dessa forma, o termo incêndio pode ser entendido como sendo a ocorrência de queima não controlada, eventos que fazem parte da temática dos desastres naturais que ocorrem nas escalas local e regional, e podem ser originadas de forma natural (por meio de raios), acidental (fogos de artifício, estradas de ferro, fogueira de acampamento, queda de balões, rompimentos de cabos de alta tensão, dentre outros) ou criminosa (ANDRADE; FERREIRA, 2019; BERLINCK; BATISTA, 2020). Por outro lado, as queimadas são associadas a práticas de manejo, ou seja, práticas de queima controlada realizadas pela ação antrópica (DIAS, 2009; SALES et al., 2019).

De todo modo, é conhecido que em ambos os casos ocorrem danos à biodiversidade, modificação da paisagem, afetando diretamente o equilíbrio dos ecossistemas (COSTAFREDA-AUMEDES et al., 2017; LOPES et al., 2018), além de emitir grandes quantidade de gás carbônico (CO₂) para a atmosfera, contribuindo para as mudanças climáticas (BERLINCK; BATISTA, 2020; VENKATESH et al., 2020).

4.2. Risco e perigo de incêndios

O termo risco e perigo muitas vezes são tratados como sinônimos no Brasil, entretanto, possuem significados distintos (CORRÊA, 2020). Para diferenciá-los, este estudo toma como base as definições adotadas pelo manual de terminologia sobre Redução de Riscos e Desastres da *United Nations Organization* - ONU (ISDR, 2009), e o Guia para

Gerenciamento de Incêndios da *Food and Agriculture Organization of the United Nations* (FAO, 2009).

Segundo a definição do ISDR (2009), o termo perigo pode ser entendido como um fenômeno, evento ou atividade humana potencialmente danosa que pode causar perda de vidas, ferimentos ou outros impactos à saúde, danos materiais, perdas de meios de subsistência e serviços, perturbação social e econômica ou dano ambiental. Por outro lado, risco é definido como sendo a combinação da probabilidade de que um evento possa ocorrer e suas consequências negativas associadas.

Para a FAO (2009), o termo perigo de incêndio é entendido como sendo a presença de material combustível que, através de sua natureza, localização, condição de arranjo, ou qualquer combinação desses fatores, pode ser queimado e, caso ocorra, pode criar um risco à vida humana, propriedade/benfeitorias e ao ambiente. Já o risco de incêndio é definido como sendo uma real ou potencial probabilidade de ocorrência de incêndio, ou outra combustão orgânica ou inorgânica de materiais que pode pôr em perigo a vida humana, propriedade/benfeitorias e o ambiente.

Em termos gerais, o perigo de incêndio é utilizado especificamente para representar a predisposição/suscetibilidade que uma determinada área possui para ocorrência desse evento (HARDY, 2005; MICHAEL et al., 2021), enquanto o risco de incêndio refere-se a chance/probabilidade de um incêndio ocorrer em função da sua predisposição (FINNEY, 2005; CORRÊA, 2020; POURGHASEMI et al., 2020).

Dessa forma, Corrêa (2020) cita que a incidência espacial do perigo é representada pela suscetibilidade, sendo avaliado através da predisposição para ocorrência dos processos ou ações, de forma a indicar a propensão de uma área a determinado evento perigoso, ou seja, a ocorrência de incêndios e queimadas.

O mapeamento da suscetibilidade de incêndios e queimadas (SIQ) possibilita identificar os locais com predisposição de ocorrência do evento, permitindo direcionar medidas de gerenciamento estruturais, tais como a construção de aceiros preventivos, construção de estradas de acesso em pontos estratégicos, reorganização de práticas de manejo do solo, delineamento de áreas prioritárias para a intensificação de fiscalização, alocação de equipes de brigadistas, bem como a organização de ações no combate (RIBEIRO et al., 2012; SILVEIRA et al., 2013; WHITE et al., 2016).

4.2. Fatores influenciadores

A ocorrência e disseminação espacial de incêndios e queimadas é um processo não linear e complexo impulsionada pela interação de processos bióticos e abióticos que são dependentes da escala geográfica (CHUVIECO et al., 2019). Aponte et al. (2016) e Bui et al. (2017) descrevem que a ocorrência e distribuição desses eventos são influenciados por quatro fatores principais, sendo eles: aspectos climáticos e antrópicos, fatores topográficos e características da vegetação. A seguir, é apresentado uma descrição dos principais aspectos influenciadores de incêndios e queimadas citados na literatura.

4.4.1. Aspectos climáticos

Os aspectos climáticos são reconhecidos como o principal determinante dos padrões do regime de incêndios e queimadas (MARLON et al., 2013), não por acaso que o relatório do Painel Intergovernamental sobre Mudanças Climáticas (IPCC, 2018) sugere que até 2040 as mudanças climáticas irão aumentar a frequência e severidade dos incêndios e queimadas em todo o mundo, principalmente devido ao aumento na temperatura associado a oscilações nas taxas de precipitação, fatores importantes no comportamento do fogo.

Os fatores climáticos são importantes pois determinam a umidade do combustível morto (serapilheira, gramíneas secas e material lenhoso não vivo), que por sua vez irá representar o potencial de ignição e propagação do fogo. As condições que influenciam na umidade do combustível incluem a temperatura, precipitação, velocidade do vento, umidade do ar e umidade do solo (MARLON et al., 2013).

De modo geral, a temperatura, precipitação e a umidade do ar podem afetar de forma direta a umidade do combustível, aumentando sua inflamabilidade, e a longo prazo, influenciam diretamente a disponibilidade da vegetação, estágios fenológicos e estresse hídrico (CHUVIECO et al., 2019; POURGHASEMI et al., 2020; VENKATESH et al. 2020), e da mesma forma, a umidade do solo que regula a umidade da vegetação em longo prazo (RAKHMATULINA et al., 2021).

A relação entre esses parâmetros também é notória, por exemplo: a velocidade do vento influencia diretamente na propagação do fogo, além de modular a umidade do combustível de forma indiretamente por meio da evaporação do vapor de água (DONG et al., 2021), e com uma menor umidade do ar pode ocasionar o déficit de formação de nuvens, favorecendo uma maior incidência da radiação solar na superfície, e consequentemente, elevando a temperatura (MACHADO et al., 2014).

A sazonalidade também é um fator importante. Períodos de seca prolongadas, por exemplo, possuem forte correlação com incêndios e queimadas, porém, se a precipitação for uniforme durante o ano, sem uma estação seca bem definida, a suscetibilidade de incêndios é menor (FLANNIGAN et al., 2013; HANES et al., 2020), fazendo com que a precipitação seja um fator fundamental na identificação do início e término da estação de incêndios (FERREIRA, 2017).

Chen et al. (2014) citam que as variações climáticas de longo prazo (escala decenal ou mais longas) podem ter grandes impactos sobre os tipos de cargas de combustíveis (produção e armazenamento de biomassa), e as variações climáticas interanuais provavelmente influenciam os regimes de incêndios através de seus efeitos na dessecação da vegetação. Em suma, a variabilidade climática influencia diretamente na ocorrência de incêndios e queimadas, visto que regula a disponibilidade de água no sistema solo/planta/atmosfera, fatores determinantes da acumulação e predisposição de material que pode ser queimado (ANDRADE et al., 2020; ABRAM et al., 2021).

4.2.2. Aspectos antrópicos

De fato, as atividades humanas são um dos fatores básicos que afetam as ocorrências de incêndios, e a intensificação dessas atividades propicia uma maior possibilidade de ignição, seja acidental ou criminosa, o que configura maior risco de incêndio (ADAB et al., 2013). Dentre os aspectos antrópicos citados na literatura, podemos destacar a distância e densidade de rodovias e estradas vicinais e a densidade populacional.

Estudos apontam que regiões próximas às atividades humanas, como estradas e assentamentos apresentam alto risco a incêndios (CAÚLA et al., 2015; NÓBREGA et al., 2018; VENKATESH et al., 2020). Entretanto, Marlon et al. (2013), descrevem que, de forma global, existe uma tendência de incêndios mais frequente em níveis intermediários de densidade populacional, ao passo que a ocorrência desses eventos é mais rara em áreas densamente povoadas (áreas urbanas) visto que a vegetação é mais limitada, com paisagem altamente fragmentada, e em áreas com baixa densidade populacional em função de condições climáticas adversas e baixa cobertura vegetal (áreas desérticas e polares), por outro lado, observa-se um aumento em 10% a 20% na frequência de incêndios apenas para áreas com até 0,1 pessoas por km² (KNORR et al., 2014).

A densidade de rodovias e estradas vicinais, por sua vez, apresentam importante contribuição na ocorrência de incêndios em áreas com baixa acessibilidade como na região

Amazônica, pois contribuem na acessibilidade da área, abertura de novas áreas de exploração agropastoril e retirada de madeira, o que facilita a ocorrência de incêndios (RIBEIRO et al., 2012). Nessa região, o desmatamento-fogo ativo é correlacionado (ARAGÃO et al., 2018), o que contribui para que a densidade de rodovias apresente relação com os focos de incêndios (JÚNIOR et al., 2019; MOTA et al., 2019).

Entretanto, com a dissociação do desmatamento-fogo ativo (ARAGÃO et al., 2018) é provável que a relação incêndios/queimadas/densidade de rodovias torne-se não linear, assim como observado por Zumbrunnen et al. (2011), ou seja, o aumento no número de habitantes e a expansão de áreas urbanizadas, e maior densidade de rodovias e estradas vicinais não necessariamente podem resultar em mais ocorrência de incêndios e queimadas, exceto se o aumento populacional e densidade de rodovias se der em áreas menos habitadas (ZUMBRUNNEN et al., 2011; TORRES et al., 2017).

4.2.3. Fatores topográficos

As características da topografia representadas pela altitude, declividade e exposição, constituem um dos fatores importantes, pois conseguem modular aspectos da região, como as condições meteorológicas médias e a variabilidade espaço-temporal de aspectos climáticos, tais como a temperatura do ar, precipitação e radiação solar (CAÚLA et al., 2015; LEUENBERGER et al., 2018). Esses fatores controlam o ciclo de vida da cobertura vegetal e do uso e ocupações do solo, e conseqüentemente, possuindo influência na incidência de incêndios e queimadas (PARENTE; PEREIRA, 2016).

A elevação, por exemplo, é uma variável fisiográfica crucial que modula a temperatura, umidade e velocidade do vento, tendo papel importante na propagação de incêndios, pois esse fator influencia na estrutura da vegetação, a umidade do combustível e umidade do ar (ADAB et al., 2013), já a declividade influencia a taxa de propagação do fogo, sendo que os incêndios se movem mais rapidamente para cima, e menos rapidamente para baixo (WHITE et al., 2016), e também tendem a ocorrer mais rapidamente em relevos mais ondulados do que aqueles com menor declividade (ADAB et al., 2013; CAMARGO et al., 2019).

A orientação do terreno refere-se à direção para a qual as encostas do terreno estão voltadas. Este fator influencia na radiação solar incidente no terreno, e por consequência, interfere na umidade do solo, temperatura do ar e velocidade do vento, pois quanto mais direta é a incidência de radiação solar, maior será a temperatura do ar e menor a umidade,

influenciando diretamente no tipo e ressecamento do material vegetal que pode ser queimado (WHITE et al., 2016; EUGENIO et al., 2016; VENKATESH et al., 2020).

O Índice de umidade topográfico (*Topographic Wetness Index-TWI*) também é um indicador citado na literatura que pode ser associado a ocorrência de incêndios, ele refere-se à tendência de água que se acumula na bacia hidrográfica devido à força gravitacional que move a água a jusante. Dessa forma, valores mais elevados de TWI indicam maior formação de fluxo de água na superfície e, consequentemente, maior saturação nestes ambientes (LEI et al., 2016; NÓBREGA et al., 2018).

4.2.4. Características da vegetação

A dinâmica de uso e ocupação do solo também podem afetar a frequência e severidade dos incêndios e queimadas de uma região (MARLON et al., 2013), pois transformam os elementos da paisagem, alterando a disponibilidade da massa vegetal, grau de ressecamento, e consequentemente, sua inflamabilidade (FONSECA et al., 2019; GOMES et al., 2020).

Um levantamento histórico desenvolvido por Badia et al. (2019) na Catalunha-Espanha mostraram que os incêndios ocorreram exclusivamente onde houve alteração da cobertura florestal, sendo que o ponto de ignição estava associado aos campos agrícolas. No bioma Cerrado no Brasil, Gomes et al. (2020) observaram que a taxa de propagação de incêndios, sua intensidade, calor liberado e consumo de combustível fino foram maiores em áreas ocupadas por pastagens e savanas do que em áreas ocupadas por floresta.

De modo geral, a estrutura da paisagem e a forma com que os padrões de uso e ocupação da terra se organizam interferem na frequência e severidade dos incêndios, pois, além de refletir a carga de combustível disponível que pode ser queimada, a cobertura vegetal exerce influência no microclima local (AQUILUÉ et al., 2020).

Em áreas naturais, como uma floresta densa ocorre a interceptação da radiação solar, o que reduz a temperatura do ar e do material combustível, além de servir como uma barreira, reduzindo a velocidade do vento em seu interior, diminuindo as taxas de evaporação, e consequentemente, dificultando a secagem do material combustível (NUNES et al., 2008; GOMES et al., 2020). Por outro lado, em áreas antropizadas, além de ocorrer a remodelagem de forma contínua da carga de combustível, as taxas de evaporação, temperatura de superfície são alteradas, e por consequência, a suscetibilidade de incêndio (AQUILUÉ et al., 2020; GOMES et al., 2020).

Dessa forma, em estudos relacionados à suscetibilidade de incêndios e queimadas, é comum o uso de índices de vegetação para a análise do teor de umidade do combustível, bem como mapas de uso e cobertura da terra que possibilitam avaliar as alterações da cobertura vegetal, fatores potencialmente explicativos da ocorrência de incêndios e queimadas (CHUVIECO et al., 2019; POURGHASEMI et al., 2020).

4.3. Sensoriamento Remoto e SIG aplicado à detecção de incêndios e queimadas

Na literatura, existe uma descrição de várias formas de detectar/monitorar incêndios e queimadas, tais como por meio de postos de observação, vigilância terrestre, patrulhamento aéreo e/ou monitoramento por imagens de sensoriamento remoto. Entretanto, em áreas de grande extensão territorial, algumas dessas técnicas se tornam inviáveis (SZPAKOWSKI; JENSEN 2019; BARMPOUTIS et al., 2020).

Devido a capacidade de obtenção de informações de forma remota (sensores acoplados em aeronaves ou satélites), o sensoriamento remoto constitui uma vantagem sobre as demais técnicas. Dessa forma, com seu advento, o uso de imagens obtidas por sensoriamento remoto passou a representar uma das formas mais viáveis para o monitoramento de incêndios e queimadas, devido sua rapidez, eficiência, periodicidade, baixo custo e uma visão de diversos aspectos, possibilitando o monitoramento em diversas escalas (BARMPOUTIS et al., 2020).

Segundo Szpakowski e Jensen (2019), a utilização de imagens obtidas por sensoriamento remoto possibilita detectar focos de incêndio ativo, bem como estimar a área queimada e sua severidade, além do mapeamento do material combustível e demais características da superfície terrestre, viabilizando o desenvolvimento de estudos mais específicos, como o mapeamento da suscetibilidade e a previsão do risco de incêndios.

Como os eventos de incêndios e queimadas produzem uma elevação de temperatura acima do normal observado no ambiente, esses eventos podem ser detectados por sensores remotos. Segundo Chuvieco et al. (2020), várias bandas do espectro eletromagnético possibilitam uma diferenciação entre um ponto quente (incêndio) e um ponto frio circundante em imagens, dessa forma, para detecção desses eventos comumente são utilizadas as bandas do infravermelho médio e térmicas.

O Programa Queimadas do Instituto Nacional de Pesquisas Espaciais-INPE, utiliza sensores que operam na faixa do infravermelho (0,75 a 1.000 μm) para detecção de incêndios e queimadas. O INPE denomina esses eventos como “focos de calor”, sendo detectados pela

quantificação de energia (de fótons) que chega até o sensor (temperatura de brilho), obtidas a partir da irradiância espectral (INPE, 2021), possibilitando detectar os focos de calor ativo. Para fins de simplificação, neste estudo adotou-se o termo “incêndios e queimadas” para os casos da ocorrência de focos de calor identificados a partir de sensores remotos provenientes dos dados disponibilizados pela base do INPE, dado que, ao menos *a priori*, não é possível identificar a fonte causadora de ignição.

Trabalhos de validação de campo realizados pelo INPE indicam que os satélites em órbita polar (NOAA, TERRA e AQUA) utilizados no programa DBQueimadas possibilitam detectar uma frente de fogo com área superior a 30 m², já os satélites geoestacionários (GOES e MSG-3), a frente de fogo precisa ter o dobro de tamanho para ser detectado, enquanto os satélites mais recentes (NPP-SOUMI e NOAA-20) conseguem detectar áreas bem menores (INPE, 2021).

Vale ressaltar que a relação foco de calor/queimada não é direta, pois a existência de um foco de calor indica o evento em um pixel, que por sua vez, varia em função do satélite (que pode variar de 375x375 m até 5x4 km), ou seja, neste pixel pode haver uma ou várias ocorrências de queimadas ou incêndios ativo distintas, mesmo assim, será detectado um único foco de calor em função da limitação de resolução (INPE, 2021).

Dessa forma, mapear a área de queimada e sua severidade é de suma importância, pois fornecem uma representação espacial precisa da extensão e perímetro do incêndio, bem como sua severidade (SZPAKOWSKI; JENSEN, 2019; CHUVIECO et al., 2020). Segundo Chuvieco et al. (2020), existem vários métodos para o mapeamento de área queimada, esses métodos se baseiam na análise das alterações de refletância causada pelo incêndio, tais como o GEMI-Burn scar (Global Environment Monitoring Index Burn scar), NBR (*Normalized Burn Ratio*) que avaliam a proporção de refletância do infravermelho próximo e infravermelho de ondas curtas.

Os dados obtidos por sensores remotos são processados com uso de Sistemas de Informações Geográficas (SIG). O SIG é um instrumento bastante adequado para a gestão desses eventos, pois permite capturar, armazenar, editar, consultar e analisar dados espacialmente georreferenciados, de modo que possibilita a integração de diferentes bases de dados, bem como a realização de modelagem espaço-temporal, além do mais, os SIGs podem ser utilizados para geração de mapas estratégicos para auxiliar no combate de incêndios, bem como para geração de mapas de risco de incêndios, indicando as regiões com maiores probabilidade de ocorrência dos eventos, permitindo respostas imediatas que otimizam os resultados de ações (CORRÊA, 2020; CHUVIECO et al., 2020).

4.4. Técnicas de Machine Learning aplicado ao mapeamento da SIQ

A possibilidade de prever quais áreas são suscetíveis a um determinado tipo de desastre, incluindo deslizamentos de terra, alagamentos ou incêndios florestais é imprescindível para gestão desses eventos. Dessa forma, vários métodos para modelar espacialmente a ocorrência desses desastres foram desenvolvidos nas últimas décadas (SACHDEVA et al., 2018; CHUVIECO et al., 2019; POURGHASEMI et al., 2020).

Estimar a probabilidade de ocorrência de incêndios e queimadas em uma área sob determinadas condições ambientais e antrópicas é uma ferramenta moderna para apoiar os planos de proteção e reduzir as consequências desses eventos (JAIN et al., 2020). Segundo Costafreda-Aumedes et al. (2017), os primeiros modelos utilizados para determinação da ocorrência desses eventos foram construídos com uso de regressão linear, e a partir da metade da década de 1980, os modelos de regressão logística e de Poisson foram introduzidos para esse fim, sendo ambos os métodos aplicados com frequência até os dias atuais devido sua facilidade de aplicação e simplicidade de compreensão.

Nos anos subsequentes, os modelos evoluíram em paralelo às aplicações matemáticas, técnicas de análise e modelagem, poder computacional e maior disponibilidade de conjunto de dados espaciais em decorrência do advento do Sensoriamento Remoto, aumentando assim o número de estudos voltados ao mapeamento da suscetibilidade de ocorrência de incêndios (COSTAFREDA-AUMEDES et al., 2017).

As técnicas mais complexas como Random Forest (RF), Redes Neurais Artificiais (RNA), Support Vector Machines (SVM) ou Modelos Aditivos Generalizados (GAM) foram introduzidos como alternativa aos métodos estatísticos tradicionais, especialmente quando se trata de grandes bancos de dados, padrões não lineares e variáveis que são altamente correlacionadas ou não distribuídas normalmente (COSTAFREDA-AUMEDES et al., 2017; JAIN et al., 2020).

De modo geral, os resultados mais ou menos sofisticados são aplicados para combinar as variáveis predisponentes em um SIG, produzindo os mapas de suscetibilidade dos eventos (CHUVIECO et al., 2019). No caso de algoritmos de aprendizagem de máquina, os modelos produzem mapas de suscetibilidade baseados em dados de entrada (variáveis) sem a necessidade de conhecimento *a priori* dos fenômenos investigados, mas simplesmente aprendendo com a experiência, visto que o modelo é ajustado conforme os dados de

treinamento, permitindo gerar previsões sobre toda a área de estudo (COSTAFREDA-AUMEDES et al., 2017; LEUENBERGER et al., 2018).

Vários estudos foram implementados utilizando aprendizagem de máquina associado a SIG (ZHENG et al., 2017; SACHDEVA et al., 2018; BUI et al., 2019; TANG et al., 2020; TONINI et al., 2020), com destaque ao desenvolvido por Pourghasemi et al. (2020) que utilizaram o Random Forest para o mapeamento da suscetibilidade de riscos múltiplos (inundação, deslizamento e incêndios florestais) em uma província no Irã. Segundo os autores, o resultado obtido no mapeamento da suscetibilidade de incêndio foi mais efetivo que para os demais riscos, pois atingiu uma taxa de acerto de 0,943%. Li et al. (2020) utilizaram Redes Neurais Artificiais e Máquina de Vetores de Suporte para o mapeamento da suscetibilidade de incêndios e queimadas na Região Autônoma de Guangxi, na China. Os resultados obtidos pelos autores mostram uma taxa de acerto de 92,16% para o método de Redes Neurais Artificiais, enquanto o método de Máquina de Vetores de Suporte apresentou cerca de 89,89%.

Recentemente, a implementação de métodos mistos tem mostrado bons resultados quando comparado aos obtidos até então, como destacado por Eskandari et al. (2021) que utilizaram o método combinado GAM-MARS-SVM (*Generalized Additive Model - Multivariate Adaptive Regression Spline - Support Vector Machine*) para o mapeamento da suscetibilidade de incêndios na região norte do Irã, obtendo resultados consideravelmente melhores. Mohajane et al. (2021) utilizaram vários modelos combinados para prever o risco de incêndio no norte de Marrocos, os autores constataram que o modelo conjunto RF-FR (*Random Forest-Frequency Ratio*) apresentou o melhor desempenho dentre os métodos empregados.

Entretanto, Bui et al. (2017) e Moayedi et al. (2020) destacam que poucos estudos se concentraram na exploração de métodos de aprendizagem de máquina otimizados, especialmente os baseados em lógica fuzzy. Para os autores, os métodos de aprendizagem de máquina integrados a lógica fuzzy podem ser úteis nessa abordagem, visto que os fatores influenciadores são múltiplos, além disso, não são lineares, e possuem graus de subjetividade que estão relacionados a distribuição espacial dos eventos, bem como incerteza em função da resolução espacial e temporal dos dados.

Em suma, a utilização de lógica fuzzy associado a métodos de aprendizagem de máquina, tais como *Random Forest* (RF), *Support Vector Machines* (SVM), *Artificial Neural Networks* (ANN) *Naive Bayes* (NB), *Genetic Algorithms* (GA), dentre outros, para a otimização de modelos podem apresentar bons resultados, porém, no contexto de

mapeamento de suscetibilidade de incêndios e queimadas essas aplicações ainda são incipientes (JAIN et al., 2020; MOAYEDI et al., 2020).

4.5. Seleção de atributos relevantes

Seja qual for o mecanismo de classificação da suscetibilidade e/ou risco de incêndios e queimadas, ele se materializa com base na observação de um conjunto de atributos do espaço geográfico. Nesse caso, um dos desafios é selecionar os atributos que são facilmente mensuráveis e que, ao mesmo tempo, possuam relação com o fenômeno.

No que se refere à disponibilidade de dados, a facilidade de aquisição desses atributos contribui para a implementação do modelo, principalmente em regiões com baixa densidade de estações meteorológicas de superfície (JESUS et al., 2020), nesse caso, as informações obtidas por sensoriamento remoto constituem excelentes fontes de dados.

Por outro lado, identificar uma quantidade de atributos que tenha contribuição significativa ao modelo, e, ao mesmo tempo, descartar os irrelevantes, minimiza os erros de previsão, pois os atributos não fidedignos ao evento podem se tornar uma fonte de ruído (HOSSEINI et al., 2020). Dessa forma, identificar um subconjunto de dados relevantes dentro do conjunto de dados é um dos grandes desafios na mineração de dados e aprendizagem de máquina (KUHN; JOHNSON, 2013; URBANOWICZ et al., 2018).

Para cumprir esse propósito, existem uma grande variedade de metodologias disponíveis em *softwares* gratuitos e proprietários. Kuhn e Johnson (2013) classificam essas metodologias em dois grupos, os métodos de *Wrapper* e métodos de filtros. Os métodos de *Wrapper* avaliam vários modelos usando procedimentos que adicionam e/ou removem atributos para encontrar a combinação ideal que maximiza o desempenho do modelo. Esses métodos correspondem a algoritmos de pesquisa que tratam os atributos como as entradas e usam o desempenho do modelo como saída a ser otimizada (KOHAVI; JOHN, 1997). Por outro lado, os métodos de filtro avaliam a relevância dos preditores fora dos modelos preditivos, e posteriormente, consideram apenas os preditores que apresentam relevância (KUHN; JOHNSON, 2013).

Como exemplo para os métodos de filtros podemos citar o método de ganho de informação (*Information Gain* - IG), qui-quadrado e *Relief* (URBANOWICZ et al., 2018). Já o método de *Wrapper* englobam uma gama de métodos em que empregam conjunto de dados de teste e validação, o que possibilita realizar pontuações quanto a importância dos atributos através de medidas de desempenho como *Root-Mean-Square Error*-RMSE,

Receiver-Operating Characteristic-ROC e *Akaike Information Criterion* (KUHN; JOHNSON, 2013; URBANOWICZ et al., 2018).

Dentre estes métodos de *Wrapper*, Pourghasemi et al. (2020), Gholami et al. (2021) e Prasad et al. (2021) destacam o algoritmo Boruta por sua funcionalidade, adaptação às principais características de dados, sendo um método robusto que avalia os impactos das variáveis de forma combinada para encontrar as variáveis mais relevantes e ao mesmo tempo, identificar as variáveis não-redundantes, além de estar disponível em softwares gratuitos especializados.

O algoritmo Boruta (nome originário de um deus da floresta na mitologia eslava) é construído com uso de classificação por meio de *Random Forest* (RF), ou seja, obtém-se a medida de importância de cada variável independente com base na média e desvio padrão para todas as árvores de decisão, sendo calculado o *z-score* de cada variável levando em consideração atributos “sombra” (réplicas embaralhadas das variáveis originais) gerados de forma aleatória (KURSA; RUDNICKI, 2010).

De modo geral, no primeiro passo o algoritmo Boruta duplica e embaralha o conjunto de dados, no qual são denominados atributo “sombra” em seguida, ele treina um classificador de RF para verificar a importância de cada atributo, considerando as métricas de desempenho (*z-score*), em seguida, o algoritmo verifica os recursos de maior importância, ou seja, aqueles com *z-score* mais elevado do que o máximo *z-score* dos atributos “sombra”, de modo que a cada interação, o algoritmo compara os valores das cópias embaralhadas dos atributos e os seus respectivos originais (KURSA; RUDNICKI, 2010; PRASAD et al., 2021).

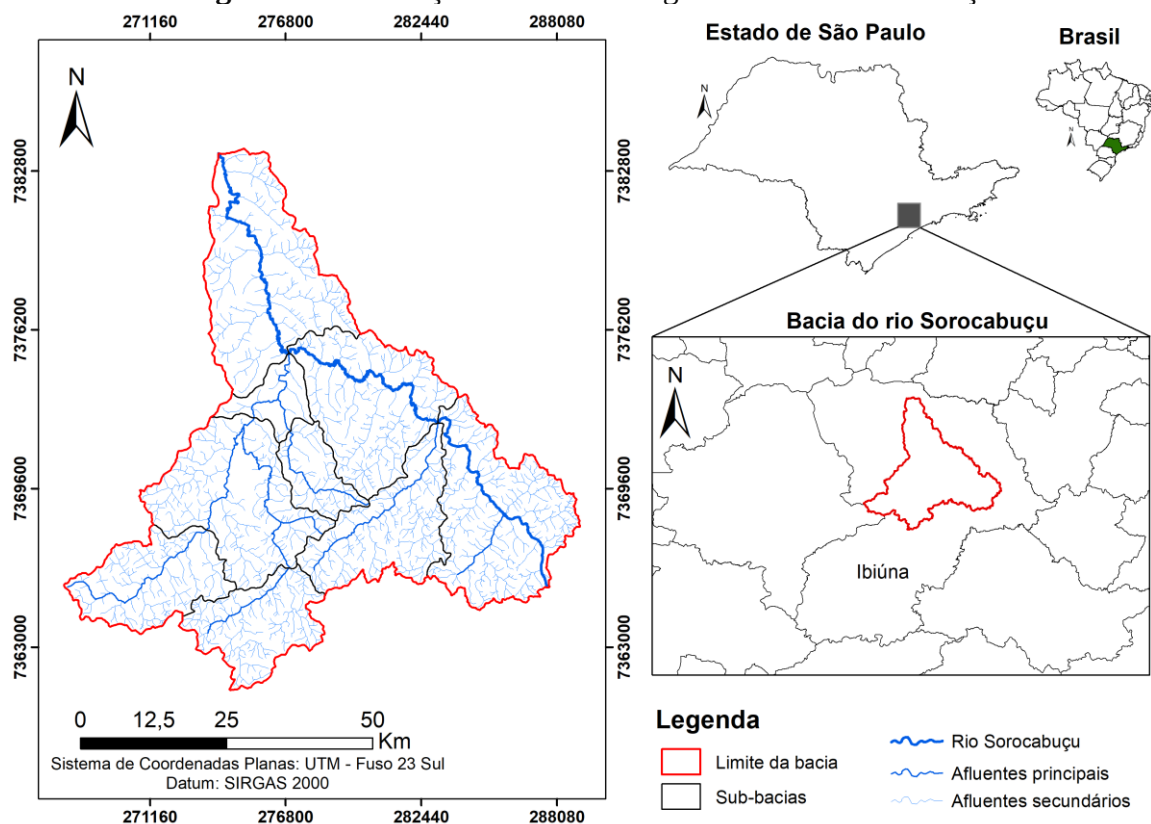
A implementação do método Boruta para seleção de atributos relevantes é ampla em vários campos da ciência, dentre vários trabalhos, destaca-se os desenvolvidos por Candanedo et al. (2017) que utilizaram o método para identificar as variáveis mais importantes para previsão de consumo de energia elétrica; Grainger et al. (2018) utilizaram para selecionar os atributos mais importantes para explicar os fatores responsáveis pelo desperdício alimentar; Amiri et al. (2019) utilizaram para identificar os fatores mais importantes responsáveis por erosão do solo, além de aplicações em estudos voltados ao mapeamento da suscetibilidade de múltiplos riscos (POURGHASEMI et al., 2020; GHOLAMI et al., 2021; PRASAD et al., 2021).

5. MATERIAIS E MÉTODOS

5.1. Caracterização da área de estudo

Este estudo foi realizado na Bacia Hidrográfica do Rio Sorocabuçu, localizado no município de Ibiúna, região sudeste do Estado de São Paulo. A bacia se localiza entre as coordenadas UTM de 274802 a 276156 m E, 7373205 a 7383470 m S, com área aproximada de 202 km² (Figura 1). O município de Ibiúna pertence a região metropolitana de Sorocaba, e possui uma população de 76.867 habitantes (SEADE, 2021).

Figura 1. Localização da bacia hidrográfica do rio Sorocabuçu.



Segundo a classificação climática de Köppen, o clima da região é do tipo subtropical de altitude (Cwb), com inverno seco e verão chuvoso. A temperatura média anual é de 20°C, com mínima de 10°C no mês de junho e máximas de 25°C em fevereiro. A precipitação média anual é de 1.493 mm, sendo janeiro o mês mais chuvoso com precipitações médias de 248 mm, e agosto o mês mais seco com precipitações médias de 43 mm (CIIAGRO, 2021).

Conforme descrito no Plano Municipal de Conservação e Recuperação da Mata Atlântica do Município de Ibiúna (PMCRMAI, 2021), a vegetação original da área era ocupada por Florestas Ombrófilas Densas, Florestas Ombrófilas Mistas e Florestas

Estacionais Semidecíduais, fitofisionomias características do bioma Mata Atlântica. Entretanto, as atividades agrícolas desenvolvidas de forma intensiva correspondem ao principal tipo de uso e ocupação do solo na área, que se desenvolve por meio de lavouras temporárias, semiperenes e perenes, seguido por capoeira, silvicultura e pastagem (SILVA et al., 2016; SOUSA, 2021).

A formação geológica predominante do município é o Grupo São Roque, seguido pelo Complexo Magmático Embu. A litologia do Grupo São Roque é constituída por metarritmitos, caracterizados por intercalações de metarenitos, filitos e pequenos corpos de metassiltitos, enquanto o Complexo Migmatítico Embu é constituído por metassedimentos com estrutura bandada, com intercalações de gnaisses, micaxistos, anfibolitos e quartzitos (ZAI, 2016).

As principais classes de solos correspondem aos Argissolos Vermelho-Amarelo e Latossolos Vermelho-Amarelo, sendo que os Argissolos apresentam caráter distrófico, com horizonte superficial moderado, textura média/argilosa e argilosa. Do mesmo modo, os Latossolos também apresentam caráter distrófico, possuindo horizonte superficial moderado, com textura predominantemente argilosa (SOUZA, 2017).

5.2. Identificação de focos de incêndios e queimadas e mapeamento de cicatrizes

Foram obtidos dados em formato *shapefile* sobre focos de incêndios e queimadas entre 2002 a 2020 junto ao Banco de Dados de Queimadas (BDQUEIMADAS) do Instituto Nacional de Pesquisas Espaciais (INPE, 2021). Atualmente, o INPE processa e distribui dados de focos de incêndios ativos de observação adquirido por dez satélites (satélites polares NOAA-18, NOAA-19, NOAA-20, METOP-B, METOP-C, MODIS-AQUA, MODIS-TERRA, NPP, e geoestacionários GOES-16 e MSG-3) que possuem sensores ópticos operando na faixa termal-média de 4 μm (INPE, 2021).

Cada satélite polar produz ao menos dois conjuntos de imagens por dia, enquanto os geoestacionários geram seis imagens por hora, sendo que no total, são processadas mais de 200 imagens por dia para detectar focos de calor, as recepções são feitas nas estações de Cachoeira Paulista, em São Paulo, e em Cuiabá, no Mato Grosso (INPE, 2021). A Tabela 1 apresenta os principais satélites (inativos e ativos) utilizados para o monitoramento de incêndios e queimadas do programa do INPE.

Tabela 1. Características dos satélites e sensores utilizados no monitoramento de focos de calor.

Satélite	Sensor	Data Lançamento	Situação	Resolução Espacial	Resolução Temporal
NOAA-12	AVHRR/2	14/05/1991	Inativo	1.000x1.000m	Diário
GOES-08	I-M	13/04/1994	Inativo	4.000x4.000m	40 minutos
NOAA-14	AVHRR/2	30/12/1994	Inativo	1.000x1.000m	Diário
GOES-10	I-M	25/04/1997	Inativo	4.000x4.000m	15 minutos
NOAA-15	AVHRR/3	13/05/1998	Ativo	1.000x1.000m	Diário
TERRA	MODIS	18/12/1999	ativo	1.000x1.000m	1-2 dias
NOAA-16	AVHRR/3	21/09/2000	ativo	1.000x1.000m	Diário
GOES-12	I-M	23/07/2001	Ativo	5.000x5.000m	30 minutos
AQUA	MODIS	04/05/2002	Ativo	1.000x1.000m	1-2 dias
NOAA-17	AVHRR/3	24/06/2002	Inativo	1.000x1.000m	Diário
NOAA-18	AVHRR/3	20/05/2005	Ativo	1.000x1.000m	Diário
MSG-02	SEVIRI	25/12/2005	Inativo	3.000x3.000m	15 minutos
GOES-13	I-M	24/05/2006	Ativo	6.000x6.000m	30 minutos
NOAA-19	AVHRR/3	06/02/2009	Ativo	1.000x1.000m	Diário
NPP-375	VIIRS	28/10/2011	Ativo	375x375m	Diário
NPP-750	VIIRS	28/10/2011	Ativo	750x750m	Diário
MSG-03	SEVIRI	05/07/2012	Ativo	3.000x3.000m	15 minutos
METOP-B	AVHRR/3	17/09/2012	Ativo	1.000x1.000m	Diário
GOES-16	I-M	19/11/2016	Ativo	2.000x2.000m	15 minutos
NOAA-20	AVHRR/3	18/11/2017	Ativo	1.000x1.000m	Diário
METOP-C	AVHRR/3	11/07/2018	Ativo	1.000x1.000m	Diário

Fonte: Elaborado pelo autor a partir de dados do INPE (2021), NASA (2021).

Os dados sobre focos pontuais de incêndios e queimadas disponibilizados pelo INPE serviram de base para a identificação e mapeamento das áreas de ocorrência dos eventos (cicatrizes de incêndios e queimadas). A partir dos dados pontuais, obteve-se a localização espacial dos incêndios e queimadas, e com base na Taxa de Queima Normalizada (*Normalized Burn Ratio-NBR*), foi realizado o mapeamento das cicatrizes de áreas onde esses eventos ocorreram. O NBR foi obtido utilizando a Equação 1 proposta por Key e Benson (2006).

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \quad (1)$$

Onde NIR refere-se a banda do infravermelho, e SWIR a banda do infravermelho de ondas curtas, sendo as bandas B8 e B12 do Sentinel-2A, respectivamente. O NBR foi obtido com resolução espacial de 30 m, dessa forma, as bandas do Sentinel-2A com resolução espacial de 10 e 20 metros foram reamostradas para 30 metros com uso do *reduceResolution* no Google Earth Engine.

Valores altos de NBR ($> 0,5$) refletem áreas cobertas por vegetação saudável, enquanto valores baixos ($< -0,5$) indicam solo descoberto e áreas recentemente queimadas (MILLER; THODE, 2007), dessa forma, para diferenciar as áreas queimadas de solo descoberto, foi utilizado uma composição cor verdadeira e falsa cor com bandas do visível, infravermelho próximo e infravermelho de ondas curtas do Sentinel-2A (RGB – B04/B03/B02 - cor verdadeira, RGB - B12/B8A/B04 - falsa cor) com resolução espacial de 30m, neste caso, as bandas com resolução espacial divergente foram reamostradas, esse processo foi realizado com uso da plataforma *Earth Observation Browser* (EO Browser, 2021), por possibilitar realizar o processo em nuvem, e exportar apenas a área de interesse, em seguida, as cicatrizes identificadas foram vetorizadas com o software ArcGIS 10.5 (ESRI, 2016).

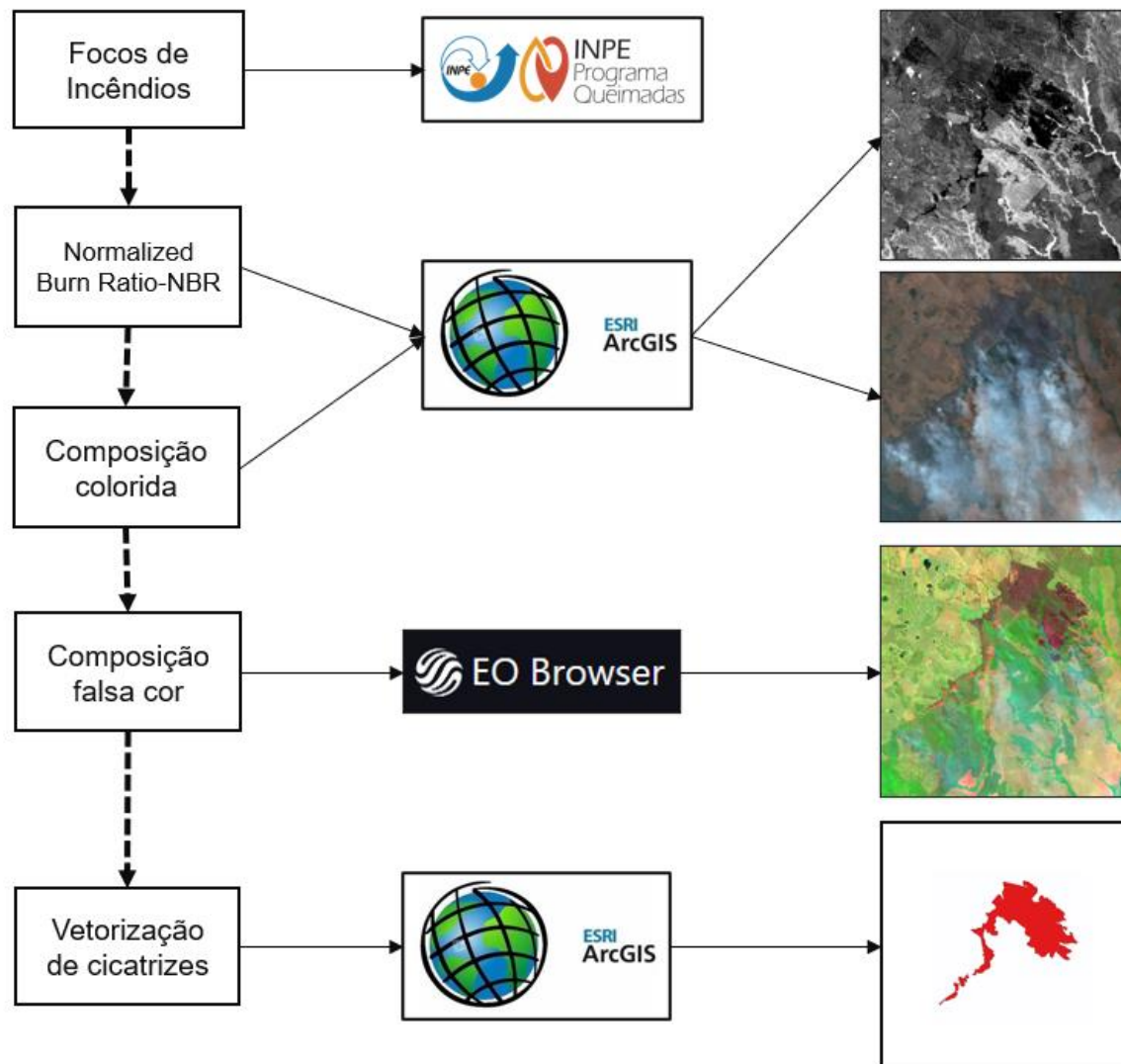
A etapa de delimitação das cicatrizes fez-se necessário em virtude da relação entre foco de calor e os incêndios e queimadas não ser direta, pois um foco de calor identificado pelo satélite indica a existência de fogo em um pixel, que por sua vez, possui resolução espacial variando de 375 m a 6 km, ou seja, em um foco de calor identificado pelo INPE podem ocorrer vários incêndios e queimadas de pequeno porte, bem como em vários pixels podem detectar um único grande incêndio ou queimada (INPE, 2021).

Além disso, como a localização do foco de calor se dá pelo ponto central do pixel (centro geométrico) onde ocorreu os incêndios, nesse caso, o ponto de indicação da ocorrência do evento pode apresentar um deslocamento em relação a sua ocorrência (deslocamento médios de 400 m em focos detectados com satélites com resolução superior a 1 km), conforme descrito em trabalhos de validação realizados pelo INPE (2021). Dessa forma, o mapeamento das cicatrizes corresponde uma etapa de suma importância, pois retifica a localização do evento, além de possibilitar identificar sua extensão.

A Figura 2 apresenta a metodologia adotada para a identificação e delimitação das cicatrizes de incêndios e queimadas empregada neste estudo. Com a identificação das

cicatrizes de incêndios e queimadas, as áreas foram vetorizadas e contabilizadas como eventos de verdadeiro positivo para ocorrência de fogo.

Figura 2. Etapas seguidas para identificação e mapeamento das cicatrizes de incêndios.



5.3. Aquisição e processamento de dados

Visando a possibilidade de replicação deste estudo em áreas com baixa densidade de dados, foi utilizado variáveis obtidas por sensoriamento remoto disponibilizado em plataformas gratuitas. Dessa forma, dados como a precipitação da região foram obtidos a partir da missão de recuperação integrada de múltiplos satélites para a medição de precipitação global (GPM - *Global Precipitation Measurement*); enquanto os dados de temperatura de superfície foram obtidos a partir do MODIS (*Moderate-Resolution Imaging Spectroradiometer*); umidade do ar foi obtida a partir de dados do FLDAS (*Famine Land*

Data Assimilation System); umidade do solo pelo SMAP/Sentinel-1 (*Soil Moisture Active Passive*); NDVI (*Normalized Difference Vegetation Index*), NDWI (*Normalized Difference Water Index*), SAVI (*Soil Adjusted Vegetation Index*) e cobertura e uso da Terra com base em dados Sentinel-2; e a elevação, declividade, orientação das vertentes, Índice de Umidade Topográfica (*Topographic Wetness Index* - TWI), distância entre as rodovias e densidade da malha viária obtida com base em cartas topográficas (IGC, 1979), retificadas manualmente no Google Earth Pro (Versão 7.3.4). A seguir é apresentada a descrição detalhada e tratamento de cada fonte de dados utilizada.

5.3.1. Precipitação

A precipitação da área de estudo foi obtida com base em dados da missão *Global Precipitation Measurement* - GPM disponibilizado pela *National Aeronautics and Space Administration* - NASA. A missão GPM foi desenvolvida em conjunto pela NASA e a *Japan Aerospace Exploration Agency*-JAXA. O projeto foi lançado em fevereiro de 2014, e fornece informações sobre chuva e neve a nível global por meio do algoritmo *Integrated Multi-satellite Retrievals for GPM* - IMERG (HUFFMAN et al., 2020).

Os dados do GPM possuem resolução espacial de ~ 10 km ($0,1^\circ$), e resolução temporal de meia hora. O algoritmo IMERG é executado duas vezes em tempo quase real gerando três produtos: o multi-satélite “inicial” (*Early Run*) disponibilizado aproximadamente 4 horas após a observação, o produto multi-satélite “atrasado” (*Early Late*) disponibilizado aproximadamente 14 horas após a observação, e por fim, o produto “final” (*Final Run*) disponibilizado aproximadamente 3,5 meses após a observação, o que permite incorporar uma análise de medidor de superfície para melhorar as estimativas do satélite (TANG et al., 2016; HUFFMAN et al., 2019).

Nesse estudo foram obtidos dados de precipitação em mm.mês^{-1} do produto *Final Run* para o ano de 2019 por meio da plataforma *Google Earth Engine* exportados em formato *GeoTIFF* com resolução espacial original (~ 10 km). A abrangência da área de interesse foi definida por polígono em forma de retângulo cobrindo a área da bacia do rio Sorocabaçu com buffer de raio de 20 km. O exemplo do *script* utilizado para exportação dos dados do GPM com uso do *Google Earth Engine* encontra-se disponível em anexo (Anexo A1). Esse método foi adotado com vista à simplificação e padronização das etapas de aquisição e subsequente processamento de dados.

5.3.2. Temperatura próximo a superfície (LST)

Para a espacialização da temperatura de superfície na área de estudo, foram utilizados dados médios mensais do *Moderate-Resolution Imaging Spectroradiometer* - MODIS (MOD11A1 – *Day time Land Surface Temperature - LST_Day_1km*), disponibilizado pela *United States Geological Survey* – USGS. O instrumento MODIS11 é operado a bordo dos satélites Terra e Aqua, possui uma largura de faixa de visualização de 2.330 km com capacidade de revista a cada um ou dois dias, resolução espectral de variando entre 0,405 a 14,385 μm distribuídas em 36 bandas, e resolução espacial variando de 250, 500 e 1.000m (WAN et al., 2015).

Os dados médios mensais para o ano de 2019 do MOD11 foram exportados via *Google Earth Engine* em formato GeoTIFF com resolução espacial original (1 km), com abrangência da área de interesse definida por polígono em forma de retângulo cobrindo a área da bacia do rio Sorocabaçu tendo por base um buffer de raio de 20 km. Esses dados são disponibilizados em número digital com escala de valores variando de 7.500 a 65.535. Para conversão dos números digitais em temperatura de superfície em $^{\circ}\text{C}$ foi utilizado a Equação 2.

$$TS = (N * 0,02) - 273,15 \quad (2)$$

Onde TS corresponde a temperatura de superfície, N o valor médio mensal dos valores digitais das imagens obtidas pelo sensor, 0,02 a constante de conversão do número digital para Kelvin, e 273,15 a constante de conversão de Kelvin para Celsius, conforme descrito por Wan et al. (2015) e Padovanni et al. (2018).

O exemplo do *script* utilizado para conversão de número digital em Kelvin, e de Kelvin para Celsius, seguido da exportação dos dados mensais do MOD11 com uso do *Google Earth Engine* encontra-se disponível em anexo (Anexo A2).

5.3.3. Umidade do ar

Para a espacialização da umidade relativa do ar próximo à superfície, foram utilizados dados médios mensais provenientes do Sistema de Assimilação de Dados Terrestres Globais *Famine Early Warning Systems Network* (FEWS NET) *Land Data Assimilation System* (FLDAS). O FLDAS é um sistema de assimilação de dados terrestres proveniente de uma combinação do produto do *Modern-Era Retrospective analysis for Research and*

Applications, Version 2 MERRA-2, e dados do *Climate Hazards Group InfraRed Precipitation with Station data* – CHIRPS (MCNALLY, 2018)

Este conjunto de dados contém uma série de parâmetros da superfície terrestre simulados pelo modelo Noah 3.6.1 do FLDAS (EK et al., 2003). Os dados possuem resolução espacial de ~10 km (0,1°) com cobertura global, e resolução temporal mensal de janeiro de 1982 até o presente (MCNALLY, 2018).

A umidade específica do ar se refere à massa de vapor em uma dada região em relação à massa total de ar, em kg.kg⁻¹. Já a umidade relativa do ar comumente utilizada em boletins climáticos é definida como a razão entre a pressão parcial do vapor de água no ar e a pressão do vapor de saturação de água sob a mesma temperatura, expressa em porcentagem (LAWRENCE, 2005), podendo ser determinada pela equação 3.

$$UR = 100 * \left(\frac{W}{W_s} \right) \quad (3)$$

Onde *UR* refere-se à umidade relativa do ar, *W* a pressão parcial de vapor de água e *W_s* a pressão de saturação do vapor de água. Nesse estudo, foram obtidos dados mensais de umidade específica (em kg.kg⁻¹) para o ano de 2019, o exemplo do *script* utilizado para recuperação da umidade do ar na base de dados do FLDAS, seguido da exportação dos dados mensais com uso do Google Earth Engine encontra-se disponível em anexo (Anexo A3).

5.3.4. Umidade do solo

A umidade do solo para a área de estudo foi obtida com base em dados da missão *Soil Moisture Active Passive* (SMAP) e do Radar de Abertura Sintética-SAR Sentinel-1A/B (Radiômetro/Radar) disponibilizada pela *National Aeronautics and Space Administration* - NASA. O projeto foi lançado em janeiro de 2015, e fornece informações a nível global sobre a umidade específica do solo (em uma profundidade aproximada de 0-5 cm) em cm³.cm⁻³ derivado de um modelo de transferência radiativa com base em dados de temperaturas de brilho obtidas pela banda L do radiômetro do SMAP e medições de retroespalhamento do SAR, banda C do Sentinel-1 (DAS et al., 2019).

Os dados de umidade específica do SMAP/Sentinel-1 é derivado de um modelo de transferência radiativa proposto por Njoku e Entekhabi (1996) e implementado por Das et al. (2018), sendo disponibilizados pelo *Distributed Active Archive Center* (DAAC) da NASA. O monitoramento teve início a partir de 15 de abril e está em operação até os dias

atuais, possuem resolução espacial de 1 km e 3 km, com resolução temporal de 6 a 12 dias (NASA, 2021). O algoritmo SMAP ativo-passivo implementado por Das et al. (2014) pode ser representado pela Equação 4.

$$T_{B_p}(M_j) = T_{B_p}(C) + \beta(C) * \{[\sigma_{pp}(M_j) - \sigma_{pp}(C)] + \Gamma * [\sigma_{pq}(C) - \rho_{pq}(M_j)]\} \quad (4)$$

Onde $T_{B_p}(C)$ corresponde a temperatura de brilho baseada no radiômetro, $\sigma_{pp}(C)$ e $\sigma_{pq}(C)$ referem-se ao retroespalhamento em dB agregado à resolução grosseira em *co-pol* (Vertical/Vertical - VV ou Horizontal/Horizontal - HH) e *cross-pol* (Vertical/Horizontal - VH ou Horizontal/Vertical - HV) respectivamente ($\sim 36\text{km}$), $\sigma_{pp}(M_j)$ e $\rho_{pq}(M_j)$ correspondem ao retroespalhamento do radar em alta resolução (1 km ou 3 km), $\beta(C)$ em K/dB e Γ correspondem a parâmetros do algoritmos em dB/dB, sendo que o primeiro refere-se a covariação entre $T_{B_p}(C)$ e $\sigma_{pp}(C)$ do radiômetro SMAP e observações de radar, e o segundo parâmetro representa a heterogeneidade induzida pela vegetação dentro das células do radiômetro de resolução grosseira detectado por $\sigma_{pp}(M)$ e $\sigma_{pq}(M)$ respectivamente (ENTEKHABI et al., 2014). A Tabela 2 apresenta as características gerais dos dados obtidos.

Tabela 2. Principais características de dados de umidade do solo estimadas pelo SMAP/Sentinel.

Data	Sat/Mode/Pol	Número de Cena	Resolução espacial
22/01/2019	1BIWDV	046W24S	1 e 3 km
26/02/2019	1AIWDV	048W24S	1 e 3 km
23/03/2019	1BIWDV	046W24S	1 e 3 km
27/04/2019	1AIWDV	048W24S	1 e 3 km
09/05/2019	1BIWDV	046W24S	1 e 3 km
02/06/2019	1AIWDV	048W24S	1 e 3 km
31/07/2019	1AIWDV	048W24S	1 e 3 km
26/08/2019	1BIWDV	046W24S	1 e 3 km
30/09/2019	1AIWDV	048W24S	1 e 3 km
25/10/2019	1BIWDV	046W24S	1 e 3 km
28/11/2019	1AIWDV	048W24S	1 e 3 km
22/12/2019	1AIWDV	048W24S	1 e 3 km

“**Sat/Mode/Pol**” - Satélite/Modo/Polarização: 1A (Sentinel-1A); 1B (Sentinel-1B); IW (Interferometric Wide-swath); DV (Dual-polarization – Vertical/Vertical-VV ou Vertical/Horizontal - VH).

A base de dados SMAP já foi incorporada ao catálogo Google Earth Engine, entretanto, o produto do SMAP/Sentinel-1 ainda não foi contemplado, dessa forma, os dados foram obtidos diretamente no portal EARTHDATA da NASA (NASA, 2021), sendo obtido

uma cena em cada mês para o ano de 2019 do produto SMAP_L2_SM_SP no formato HDF (*Hierarchical Data Format*).

De posse dos dados SMAP_L2_SM_SP no formato HDF, foi selecionado o produto “*Soil_Moisture_Retrieval_Data_1km*” com resolução espacial de 1.000 m e em seguida, foram convertidas para o formato GeoTIFF e reprojetoado para o Sistema de Projeções Geográfica e Datum SIRGAS 2000, Fuso 23 Sul com uso do software ArcGIS 10.5.

5.3.5. Índices de vegetação

Nesse estudo, foram obtidos o Índice de Vegetação por Diferença Normalizada (*Normalized Difference Vegetation Index*-NDVI), Índice de Vegetação Ajustado ao Solo (*Soil Adjusted Vegetation Index* - SAVI) e Índice de Água por Diferença Normalizada (*Normalized Difference Water Index*-NDWI) com base em dados do Sentinel-2.

A missão Sentinel-2 é composta por uma constelação de dois sensores (Sentinel-2 A e B), sendo que o Sentinel-2A foi lançado pela ESA em junho 2015, e o Sentinel-2B foi lançado em março de 2017, cada sensor possui um ciclo de 10 dias, que quando combinada, possibilita uma revisita de cinco dias (ESA, 2021). A Tabela 3 apresenta as especificações técnicas do Sensor Sentinel-2 utilizadas nesse estudo.

Tabela 3. Características principais e bandas do Sentinel-2.

Data de Lançamento		Sentinel-2A: 23 de junho de 2015			
Tipo de Órbita		Sentinel-2B: 7 de março de 2017			
Tempo de revista		Órbita sincrônica ao solo			
Altitude Orbital		10 dias cada sensor			
Complemento do sensor		~786 km			
		MSI (Multi Spectral Instrument)			
Banda	Nome	Comprimento central de onda S2A/S2B (nm)	Largura de banda S2A/S2B (nm)	Relação sinal ruído - Lref	Resolução (m)
B1	Aerossol	442,7/442,3	21/21	129	60
B2	Blue (Azul)	492,4/492,1	66/66	154	10
B3	Green (verde)	559,8/559,0	36/36	168	10
B4	Red (Vermelho)	664,5/665,0	31/31	142	10
B5	Red Edge 1	704,1/703,8	15/16	117	20
B6	Red Edge 2	740,4/739,1	15/15	89	20
B7	Red Edge 3	782,8/779,7	20/20	105	20
B8	NIR	832,8/833,0	106/106	174	10
B8A	Red Edge 4	864,7/864,0	21/22	72	20
B9	Water vapor	945,1/943,2	20/21	114	60
B10	Cirrus	1373,5/1376,9	31/30	50	60
B11	SWIR 1	1613,7/1610,4	91/94	100	20
B12	SWIR 2	2202,4/2185,7	175/185	100	20

Fonte: Adaptado de SUHET (2013) e ESA (2021).

O Sentinel-2 é equipado por um Instrumento Multiespectral (*MultiSpectral Instrument* - MSI) que realiza a captura de dados em 13 bandas espectrais que vão desde o visível e infravermelho próximo (*Visible and Near-Infra-Red* - VNIR) até o infravermelho de ondas curtas (*Short Wave Infra-Red* - SWIR). A resolução espacial do satélite varia de 10 a 60 metros, sendo que quatro bandas apresentam resolução espacial em 10 metros (B2, B3, B4 e B8), seis bandas com 20 metros (B5, B6, B7, B8A, B11 e B12) e três bandas com 60 metros (B1, B9 e B10), apresentando resolução radiométrica de 12 bits (ESA, 2021).

Os dados do Sentinel-2 disponibilizados pela ESA foram processados com uso da plataforma *Google Earth Engine*, sendo o NDVI e NDWI calculados utilizando as Equação 5 e 6, respectivamente, conforme proposto por Rouse et al. (1973) e GAO (1996), esse cálculo foi realizado com uso da função “*image.normalizedDifference*”, e para a determinação do SAVI, foi utilizado a Equação 7, conforme proposto por Huete (1998). Ambos os índices foram obtidos utilizando cenas selecionadas com uso do filtro “CLOUDY_PIXEL_PERCENTAGE” com 20% de tolerância para cobertura de nuvens.

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (5)$$

$$NDWI = \frac{\rho_{Green} - \rho_{NIR}}{\rho_{Green} + \rho_{NIR}} \quad (6)$$

$$SAVI = \frac{(\rho_{NIR} - \rho_{Red}) * (1 + L)}{\rho_{NIR} + \rho_{Red} + L} \quad (7)$$

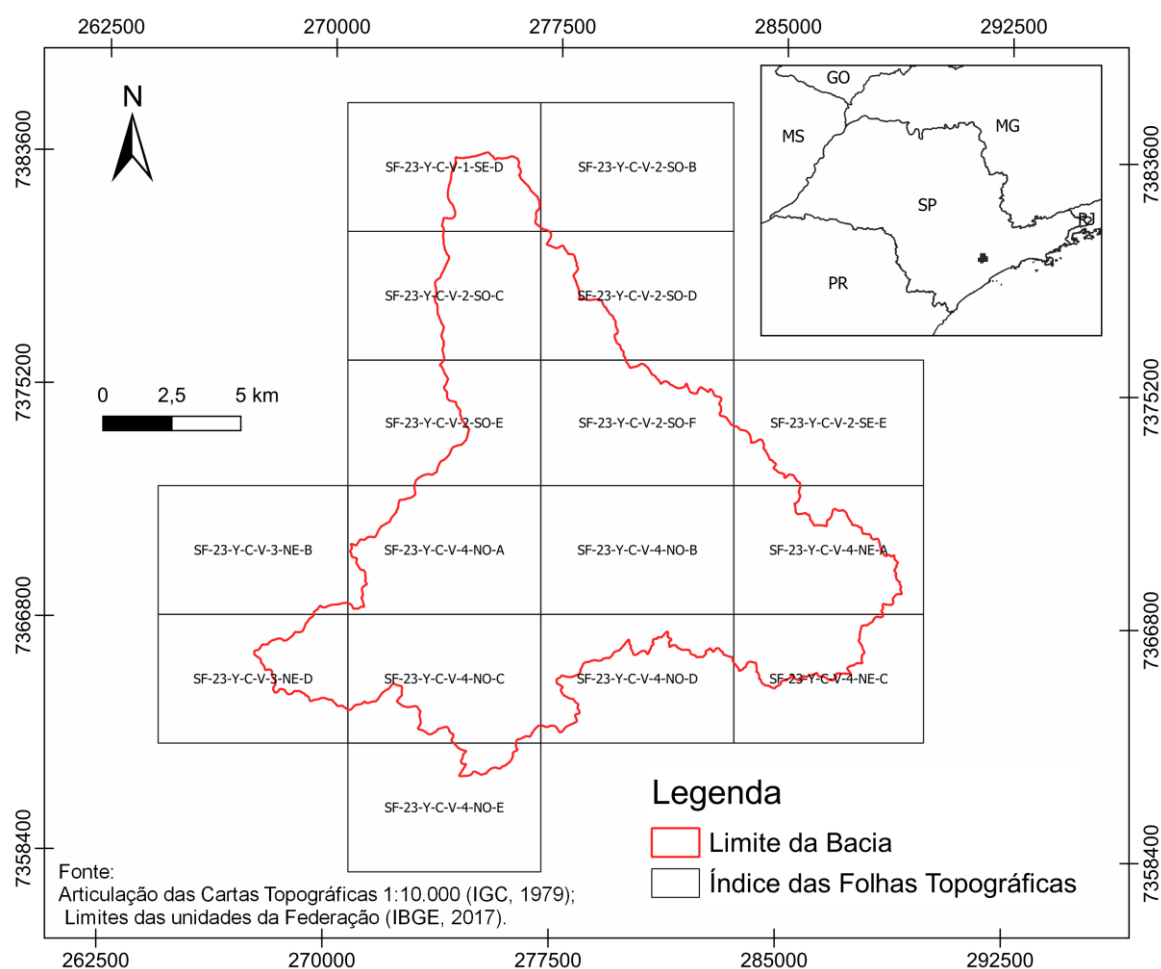
Onde ρ_{NIR} é a reflectância da banda do infravermelho próximo, ρ_{Red} refere-se a reflectância da banda do vermelho e ρ_{Green} a reflectância da banda verde, sendo a banda 8, banda 4 e banda 3 do Sentinel-2 respectivamente. Já L corresponde o fator de ajuste, sendo adotado valor 0,5, visto que é o mais comumente utilizado em função da característica de densidade da vegetação da região (AMORIM et al., 2019; MARTINS; ROSA, 2019).

Os dados sobre NDVI, SAVI e NDWI mensal para os meses de janeiro a dezembro de 2019 foram exportados em formato GeoTIFF com resolução espacial de 30 metros, neste caso, os índices obtidos foram reamostrados para 30 metros com uso do *reduceResolution* no Google Earth Engine, em seguida, foram exportados levando em consideração abrangência da área de interesse definida por polígono em forma de retângulo cobrindo a área da bacia do rio Sorocabaçu. O *script* utilizado para obtenção do NDVI, NDWI e SAVI encontra-se no Anexo A4, A5 e A6 respectivamente.

5.3.6. Dados planialtimétricos

As informações planialtimétricas utilizadas neste estudo, tais como pontos cotados, curvas de nível, rodovias e estradas vicinais, cursos hídricos (unifilar e bifilar), foram obtidas a partir de dados de um Mosaico de Cartas Topográficas em escala 1:10.000 proveniente do levantamento sistemático realizado pelo Instituto Geográfico e Cartográfico (IGC) do Estado de São Paulo (IGC, 1979). A Figura 3 apresenta a articulação de folhas e suas respectivas nomenclaturas das 16 folhas utilizadas neste estudo.

Figura 3. Mapa índice das Cartas Topográficas utilizadas.



Fonte: Elaborado pelo autor.

As Cartas Topográficas são disponibilizadas em formato padronizado *Web Map Service* (WMS) na base de dados de Infraestrutura de Dados Espaciais Ambientais do Estado de São Paulo (IDEA-SP), o Sistema Ambiental Paulista (DataGEO). Dessa base, foram

utilizadas ao total 16 Cartas Topográficas que cobrem a área de estudo, sendo realizado a vetorização em ambiente SIG de forma manual com uso do software ArcGIS 10.5.

A partir das Cartas Topográficas disponibilizadas pelo IGC (IGC, 1979), obteve-se pontos cotados e curvas de nível que serviram de base para obtenção do Modelo Digital de Elevação (MDE), e a partir deste, foram obtidos a declividade do terreno, orientação das vertentes e o Índice de Umidade Topográfica (*Topographic Wetness Index* - TWI). As rodovias e estradas vicinais serviram de base para determinação da distância direta das áreas potenciais de incêndios e queimadas até a rede viária, bem como a densidade da rede viária.

O MDE foi obtido através de interpolação por *Triangulated Irregular Network* (TIN) disponível no módulo *3D Analyst Tools*. O método TIN foi empregado por ser adequado para representação espacial da superfície do terreno, sendo amplamente citado na literatura (TUCKER et al., 2001). A partir do MDE, foi obtido a declividade da área em percentual com uso da ferramenta *slope* disponível no módulo *Spatial Analyst Tools* no ArcGIS 10.5. E com base no MDE, foi obtido a orientação de vertentes com uso da ferramenta *Aspect* disponível no módulo *Spatial Analyst Tools* no ArcGIS 10.5.

O TWI foi obtido pelo método TOPMODEL com uso da ferramenta *Topographic Wetness Index* - TWI disponível no módulo *Terrain Analysis/Hydrology* do software SAGA GIS 2.3.2. Levou-se em consideração a declividade e área de captação (*Catchment Area*) da bacia, sendo que a área de captação foi determinada pelo método *Multiple Triangular Flow Direction* com base no MDE (SEIBERT; MCGLYNN, 2007).

Devido o ano de levantamento de rodovias e estradas vicinais das Cartas Topográficas, foi realizado um ajuste da malha viária com base em imagens do Google Earth (GOOGLE LLC, 2020) e dados mapeamento voluntário da *OpenStreetMap* (OSM Foundation, 2020). A partir da malha viária ajustada, foi determinada a distância euclidiana de rodovias e estradas vicinais na área de estudo com uso da ferramenta *Euclidean Distance* disponível no módulo *Spatial Analyst Tools*.

Em seguida, obteve-se a densidade de rodovias e estradas vicinais com uso da ferramenta *Line Density* que calcula a densidade simples (densidade de linhas em km dentro de uma determinada área em km²) de feições lineares na vizinhança de cada célula do *raster* de saída, determinada por km/km² (ESRI, 2021). Ambos os produtos foram definidos com resolução espacial de saída de 30m.

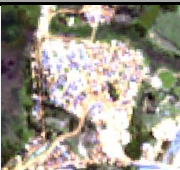
5.3.7. Cobertura e Uso da Terra

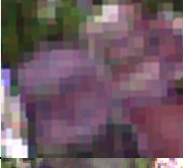
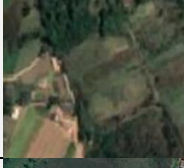
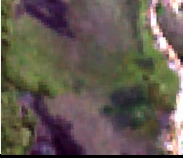
O Mapa da Cobertura e Uso da Terra na área de estudo para o ano de 2019 foi obtido por meio da classificação com base em imagens do Sentinel-2A, considerando o Sistema Básico de Classificação da Cobertura e do Uso da Terra (SCUT) da 3ª Edição do Manual Técnicos de Uso da Terra do Instituto Brasileiro de Geografia e Estatística (IBGE, 2013).

Essa classificação foi realizada tendo como base o mapeamento de uso da terra para o ano de 2010 obtido a partir de imagens de alta resolução (0,45 metro) do sensor *Vexcel Ultracam* disponibilizadas pelo IGC (IGC, 2010). Para a atualização do mapeamento de 2019 considerou as alterações no uso observadas com imagens do Sentinel-2 avaliadas em composição colorida (4R/3G/2B) e falsa cor (8R/4G/3B) sob a cena número T22KHU com data de imageamento em 08/08/2019. A atualização/edição foi realizada utilizando o conjunto de ferramentas de edição de polígonos com uso do software ArcGIS 10.5 (ESRI, 2016), sendo que essa classificação foi realizada através de interpretação visual.

Levando em consideração as classes representativas identificadas na área com base nas imagens do Sentinel-2, foram consideradas neste estudo dez classes de uso da terra, contemplando uma adaptação da Subclasse II do SCUT, o que representa a um mapeamento da cobertura e uso em escala regional (IBGE, 2013). Em seguida, as classes de uso identificadas foram retificadas com base em imagens do Google Earth Pro. A Tabela 4 apresenta as características identificadas em cada unidade mapeada.

Tabela 4. Classes de uso da terra na área.

Cobertura e Uso da Terra	4R/3G/2B – Sentinel-2	RGB - Google Earth
Áreas urbanizadas: Áreas ocupadas por edificações e sistema viário, onde predominam as superfícies artificiais não agrícolas. Incluem nessa categoria as cidades, vilas, áreas de rodovias, áreas industriais e comerciais.		
Área de Mineração: Áreas ocupadas por atividade de exploração ou extração de substância minerais comumente representadas por atividades de lavra e garimpo.		
Culturas temporárias: Áreas ocupadas por atividades agrícolas de curta duração, geralmente com ciclo vegetativo inferior a um ano, por exemplo: milho, feijão, trigo, soja, batata-inglesa, hortícolas e floríferas.		

Culturas Permanentes: Áreas ocupadas por atividade de cultivo vegetativo de longa duração, que permitem colheitas sucessivas, sem necessidade de novo plantio, por exemplo: maçã, uva, laranja, café.		
Pastagens: Áreas destinadas à atividade pecuária de pequeno e grande porte, com solo normalmente coberto por vegetação de gramíneas ou leguminosas.		
Silvicultura: Área ocupada por atividade ligada a ações de composição, trato e cultivo de povoamento florestais (Eucalipto, pinus e outras espécies) para fins comerciais.		
Área Florestal: Áreas ocupadas por vegetação arbórea com altura superior a 5 metros constituídas por diversas fitofisionomias, incluindo remanescentes primários em estágios evolutivos de recomposição florestal.		
Corpos d'água: Áreas ocupadas por corpos de água naturais e artificiais representados por rios, canais, lagos, lagoas, represas e açudes.		
Área vegetacional perturbada: Áreas com processo de antropização ou abandonadas, ocupadas por vegetação herbácea rasteira e/ou alta podendo estar associado a cobertura arbustiva densa e/ou esparsa.		
Área de várzea: Áreas ocupadas por vegetação de porte variado sob influência fluvial e topográfica.		

Fonte: Classes adaptadas do SCUT (IBGE, 2013), imagens Sentinel-2 e Google Earth Pro.

5.4. Regularização de dados de entrada

Os dados obtidos por sensores, tais como a precipitação proveniente do GPM, a temperatura do ar próximo à superfície obtida pelo produto MOD11, a umidade do ar estimada pelo Noah 3.6.1 do FLDAS e a umidade do solo determinada pelo modelo SMAP/Sentinel-1 foram processados para se tornar uma grade com resolução espacial regular conforme as demais variáveis utilizadas, tais como o uso e ocupação da terra, distância de rodovias e estradas vicinais que foram originalmente obtidos com resolução espacial de 30 m.

Dessa forma, os dados de precipitação e umidade do ar com resolução espacial original de ~10 km, bem como a temperatura do ar e umidade do solo com resolução espacial original de 1.000 m foram padronizados para 30 m, esse processo fez-se necessário para possibilitar a integração e comparação de conjuntos de dados com escala diferente. Nesse caso, foram considerados os valores do pixel central, resultando em uma grade de pontos com distribuição espacial coincidente com a resolução da fonte de dado de origem, ou seja, ~1 km e ~10 km respectivamente.

A regularização da base de dados foi realizada com uso de krigagem ordinária. Dessa forma, os dados foram avaliados por meio de análise estatística descritiva e geoestatística com uso do software R (R Core Team, 2020). Para a realização da krigagem ordinária, foi testado a normalidade dos dados por meio do teste Kolmogorov-Smirnov com nível de significância de 5%, não observando a normalidade, os dados foram transformados pelo método *Ordered Quantile* (ORQ). Em seguida, foi realizada uma análise de dependência espacial através de semivariograma visando identificar a existência de autocorrelação espacial entre os pontos, conforme o semivariograma experimental estimado pela Equação (8).

$$y^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(X_i + h) - Z(X_i)]^2 \quad (8)$$

Onde (h) corresponde o valor da semivariância para uma dada distância h ; $N(h)$ o número de pares envolvidos no cálculo da semivariância; $Z(X_i)$ o valor do atributo Z na posição X_i ; $Z(X_i + h)$ o valor do atributo Z separado da localidade X_i pela distância h (YAMAMOTO, 2020).

Com o ajuste do modelo matemático calculado são definidos os coeficientes do modelo teórico para o semivariograma representados pelo efeito pepita (C_0), variância (C_1), patamar ($C_0 + C_1$) e alcance (a). Além disso, foi verificado a existência de anisotropia por meio da observação dos semivariogramas obtidos para diferentes direções, considerando as convenções direcionais de 0°, 45° 90° e 135°, conforme descrito por Yamamoto (2020).

A escolha do semivariograma experimental (teórico) mais bem ajustado (por exemplo, esférico, exponencial e gaussiano), levou-se em consideração o menor valor de RMSE (*Root Mean Squared Error*) e o melhor valor de coeficiente de determinação (R^2) por

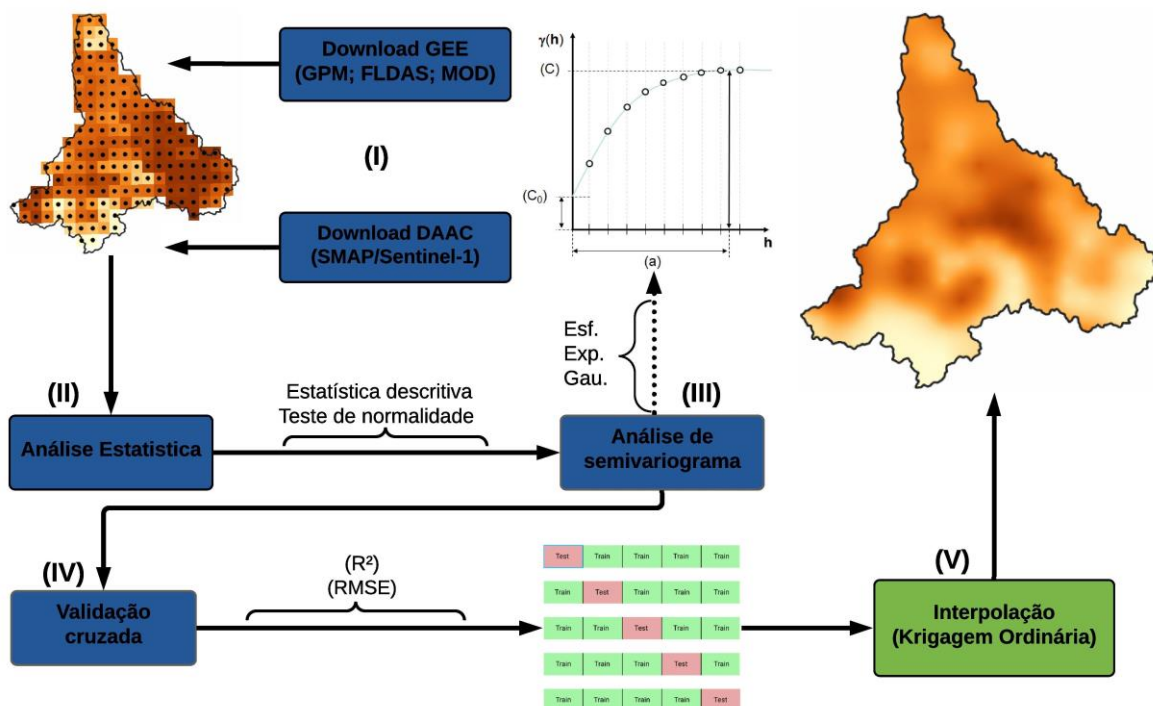
meio da validação cruzada *leave-one-out*. Após o ajuste do modelo matemático, foi realizada a interpolação dos dados por krigagem ordinária.

O Grau de Dependência Espacial (GDE) do semivariograma ajustado foi avaliado pelo efeito pepita relativo determinado pela Equação 9, sendo considerada forte dependência espacial $GDE < 0,25\%$, moderado para valores entre $0,25 \leq GDE < 0,75$, e fraca dependência espacial se $GDE \geq 0,75$, conforme classificação proposta por Cambardella et al. (1994).

$$GDE (\%) = \frac{C_0}{C_0 + C_1} \times 100 \quad (9)$$

Onde GDE refere-se ao grau de dependência espacial do modelo, C_0 refere-se ao efeito pepita e $C_0 + C_1$ o patamar respectivamente. A Figura 4 apresenta um fluxograma das etapas seguidas para a regularização do conjunto de dados do GPM, FLDAS, MODIS e SMAP por krigagem ordinária.

Figura 4. Fluxograma das etapas seguidas para regularização dos dados.



*(I): aquisição da base de dados; (II): análise estatística; (III): análise geoestatística; (IV): validação cruzada e (V): krigagem ordinária.

O ajuste dos semivariogramas foram realizados com uso do software R, sendo utilizados os pacotes *raster* (HIJMANS, 2020) e *rgdal* (BIVAND et al., 2020) para leitura e

especialização dos dados, e o pacote *gstat* (PEBESMA, 2004) para a modelagem dos semivariogramas. O exemplo do *script* utilizado para regularização encontra-se disponível em anexo (Anexo B1).

A Tabela 5 apresenta a relação das variáveis utilizados neste estudo, a respectiva fonte de aquisição, resolução espacial original e forma de processamento.

Tabela 5. Descrição das características gerais do conjunto de variáveis utilizadas.

Variáveis utilizadas	Fonte/provedor	Resolução	Regularização
Precipitação (mm)	GEE/NASA	~10km	KO (30m)
Temperatura do ar (°C)	GEE/NASA	1.000m	KO (30m)
Umidade do ar (kg.kg ⁻¹)	GEE/NASA	~10km	KO (30m)
Umidade do solo (cm ³ .cm ⁻³)	NASA/ESA	1.000m	KO (30m)
NDVI	GEE/ESA	30m	-
NDWI	GEE/ESA	30m	-
SAVI	GEE/ESA	30m	-
Elevação (m)	Carta topográfica	30m	-
Declividade (%)	MDE	30m	-
Orientação			
TWI	Carta Topográfica/ OpenStreetMap	30m	-
Distância entre vias (m)			
Densidade de vias (k/km ²)	Sentinel-2/Google Earth Pro	30m	-
Cobertura e uso da Terra			
Nº de Focos de incêndios	INPE	-	Mapeamento de cicatrizes (30m)

***NDVI**: Normalized Difference Vegetation Index; **TWI**: Topographic Wetness Index; **GEE**: Google Earth Engine; **KO**: Krigagem Ordinária; **DK**: Densidade de Kernel.

5.5. Previsão da Suscetibilidade de Incêndio e Queimadas

Na literatura, existem vários estudos que propõe a associação de fatores para indicar os locais com Suscetibilidade de Incêndios e Queimadas (SIQ) (JAISWAL et al., 2002; BONAZOUNTAS et al., 2005; PARENTE; PEREIRA, 2016; MOTA et al., 2019; POURGHASEMI et al., 2020). Entretanto, associar vários fatores para compor um método de previsão pode torna-se muito oneroso, além de apresentar graus de subjetividade regionais associados aos fatores climáticos e antrópicos (CARMO et al., 2011; GRALEWICZ et al., 2012), bem como as incertezas relacionadas a diferentes resoluções espaciais e temporais

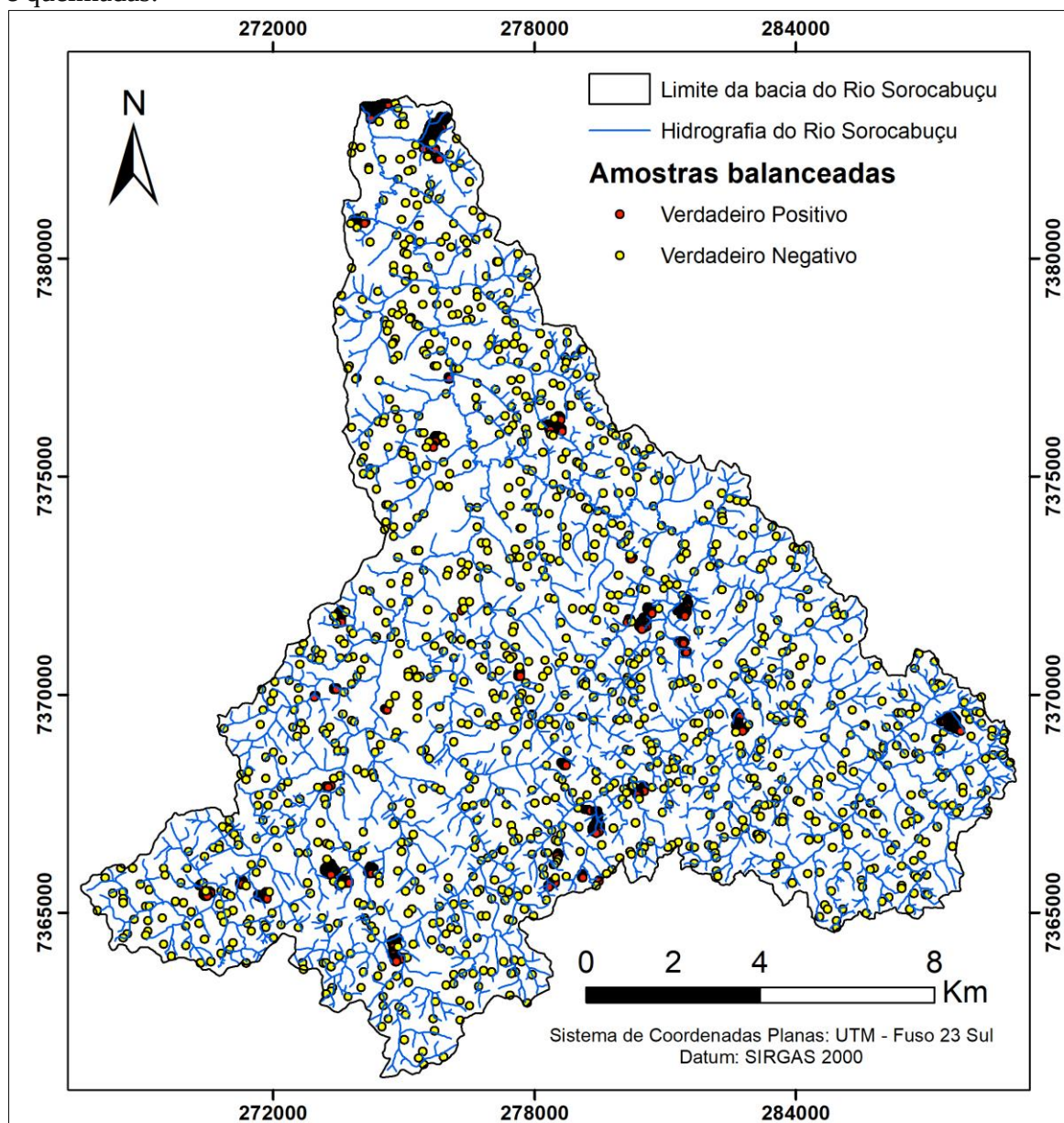
(BUI et al., 2017; MOAYEDI et al., 2020), o que impossibilita a construção de um modelo com boa precisão.

Neste estudo, foram reunidos os principais fatores citados na literatura que podem ser obtidos de forma gratuita com uso de técnicas de geoprocessamento e sensoriamento remoto. No total, 14 variáveis com potencial de explicar a ocorrência dos incêndios e queimadas na área de estudo foram obtidas, sendo elas: precipitação, temperatura, umidade do ar, umidade do solo, NDVI, NDWI, SAVI, elevação, declividade, orientação de encostas, distância das rodovias e estradas vicinais, densidade de rodovias e estradas vicinais, TWI, cobertura e uso da Terra.

A classe alvo utilizada para identificação do conjunto de variáveis mais importantes foi a ocorrência e não ocorrência de incêndios e queimadas observada na área de estudo, sendo as ocorrências de incêndios e queimadas (verdadeiro positivo) representado pelos pontos identificados no mapeamento das cicatrizes, e os de não ocorrência (verdadeiro negativo) foram selecionados de forma randômica com base no conjunto de dados com resolução espacial de 30m.

Dessa forma, foram identificadas 1.055 amostras de verdadeiro positivo, e selecionados de forma randômica sem reposição (cada ponto da grade 30m só pode ser selecionado uma vez para compor a amostra) o mesmo número de observações para a classe de verdadeiro negativo com uso da ferramenta *Create Random Points*, em *Data Management* no software ArcGIS. Esse processo fez-se necessário para equilibrar o conjunto de dados de modo a não prejudicar o desempenho dos modelos de classificação, conforme descrito por Lever et al. (2016), Pourghasemi et al., (2020). A Figura 5 apresenta a distribuição espacial do conjunto de dados balanceado de pontos verdadeiro positivo e verdadeiro negativo.

Figura 5. Mapa de pontos balanceado de ocorrência e não ocorrência de focos de incêndios e queimadas.



Com base no conjunto de dados balanceado, a abordagem adotada para selecionar os fatores mais importantes engloba a identificação das variáveis redundantes por meio da correlação de Pearson, seguido da identificação da Importância Relativa (IR) pelo algoritmo Boruta. Dessa forma, para os casos em que foram observados redundância entre as variáveis (correlação superior a 0,7, considerada neste estudo), manteve-se o atributo que apresentou maior IR pelo método de seleção com o algoritmo Boruta, e removidas as variáveis com menor IR.

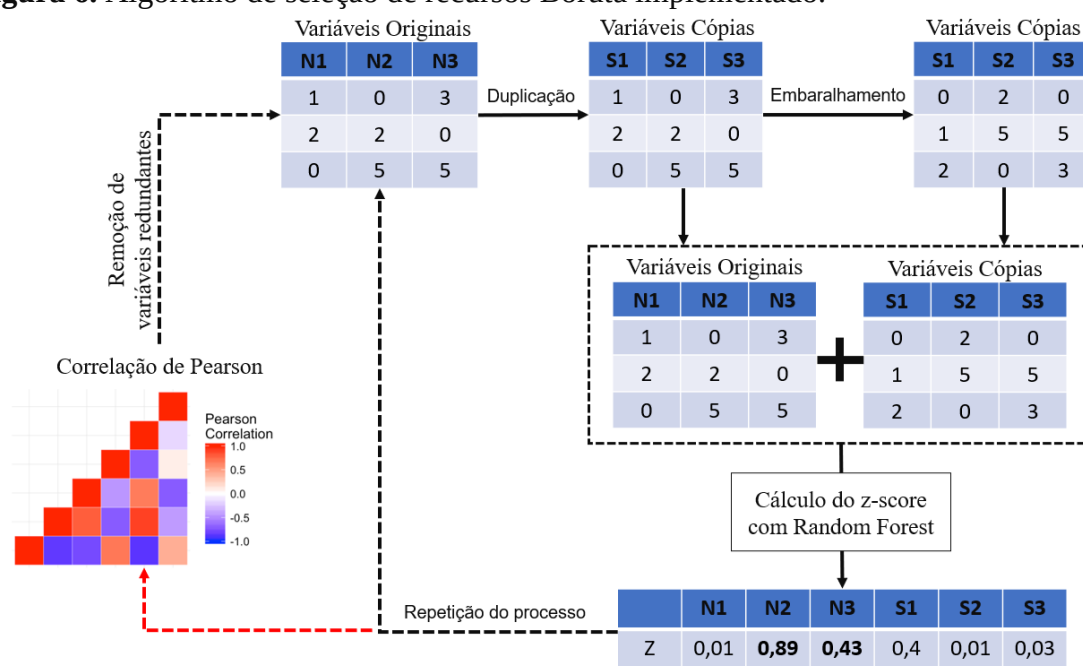
O algoritmo foi implementado pelo pacote *Boruta* (KURSA; RUDNICKI, 2010) no software RStudio. O método foi escolhido para identificar as variáveis mais importantes por

possibilitar remoção de redundância das variáveis, evitar o sobreajuste (*overfitting*) e simplificar os modelos, possibilitando uma melhor capacidade de generalização (KURSA; RUDNICKI, 2010), além de possuir capacidade de detectar padrões complexos nos dados (AMIRI et al., 2019).

O algoritmo Boruta implementado adiciona aleatoriedade à variável original, criando cópias embaralhadas de todas as variáveis (“sombras”), em seguida, ele executa um classificador RF no conjunto de dados com variáveis originais e “sombras” e avalia a pontuação média (*z-score*) na precisão para avaliar a importância de cada variável, sendo as médias mais altas as mais importantes (KURSA; RUDNICKI, 2010). Em cada iteração é realizada uma comparação entre as pontuações da variável original e “sombra”, sendo que para as variáveis com pontuações inferiores a suas respectivas “sombras” são consideradas sem importância, caso a pontuação seja maior, ela é registrada, e ao final das interações os resultados são apresentados por meio da pontuação *z-score*. O algoritmo é executado até que todas as variáveis sejam confirmadas ou rejeitadas, nesse estudo foi adotado um limite de execuções de máximo de 500 árvores, conforme descrito por Grainger et al. (2018).

A Figura 6 apresenta um fluxograma do método de seleção de atributos utilizado. O exemplo do *script* utilizado para selecionar as variáveis mais importantes encontra-se disponível em anexo (Anexo B2).

Figura 6. Algoritmo de seleção de recursos Boruta implementado.



As variáveis identificadas como mais importantes foram utilizadas para composição de um Sistema de Inferência Fuzzy Híbrido (HFIS) para o mapeamento da suscetibilidade à ocorrência de incêndios e queimadas. A abordagem HFIS foi utilizada pois consiste em um poderoso método de inteligência computacional baseada em lógica fuzzy capaz de lidar com vários problemas que incluem graus de incerteza, imprecisão e não linearidade (RIZA et al., 2015; JIMÉNEZ et al., 2017), atributos intrínsecos dos eventos de incêndios e queimadas (APONTE et al., 2016; BUI et al., 2017; CHUVIECO et al., 2019), e ao conjunto de dados utilizado (BUI et al., 2017; MOAYEDI et al., 2020).

O método HFIS consiste na combinação de lógica fuzzy e inteligência artificial baseado no método de Ishibuchi e Nakashima (2001) que gera regras de tomada de decisão considerando os graus de certeza (denominado peso). O sistema HFIS é semelhante ao proposto por Mamdani (MAMDANI; ASSILIAN, 1975), possuindo quatro partes principais: processo de aprendizagem (I), fuzzificação das variáveis de entrada (II), inferência fuzzy (IV) e defuzzificação (V).

O processo de aprendizagem é realizado por meio da técnica de partição do espaço, e ocorre de forma automática em duas etapas: a identificação de estruturas e estimativas de parâmetros, sendo que, na primeira é feito a identificação da estrutura, e determinado uma base de regra correspondente a pares de variáveis de entrada e saída, sendo otimizada a estrutura e o número das regras. Então, os parâmetros da função de pertinência são otimizados na etapa de estimativa do parâmetro. Esse processo pode ser realizado sequencialmente ou de forma simultânea (RIZA et al., 2015; JIMÉNEZ et al., 2017).

Na fuzzificação das variáveis de entrada, os valores numéricos (*crisp*) são normalizados em número real no intervalo de unidade $[0,1]$, em seguida, são convertidos em variáveis linguísticas, sendo atribuído valores linguísticos muito baixo, baixo, moderado e alto, e muito alto, que podem ser representados pelas funções de pertinência triangulares, trapezoidais e/ou gaussianas, que se sobrepõe formando regiões fuzzy, fazendo com que um dado possa pertencer a mais de um conjunto (COCCONELLO et al., 2014; ROMÁN-FLORES et al., 2020).

A base de regras é formada por um conjunto de regras que associa as entradas formando as saídas do sistema fuzzy, do tipo SE-ENTÃO, onde a parte SE é chamado de antecedente, enquanto a parte do ENTÃO é chamada de consequente, como por exemplo:

$$\text{Se } X1 \text{ é } A1 \text{ e } \dots \text{ e } Xn \text{ é } An \text{ ENTÃO } Y \text{ é } B$$

Para classificação, a base de regras é gerada, em seguida, as partes consequentes são substituídas por classes (ISHIBUCHI; NAKASHIMA, 2001), como por exemplo:

Se X_1 é A_1 e . . . e X_n é A_n ENTÃO C_j com $CF_j, j=1,2 \dots N$

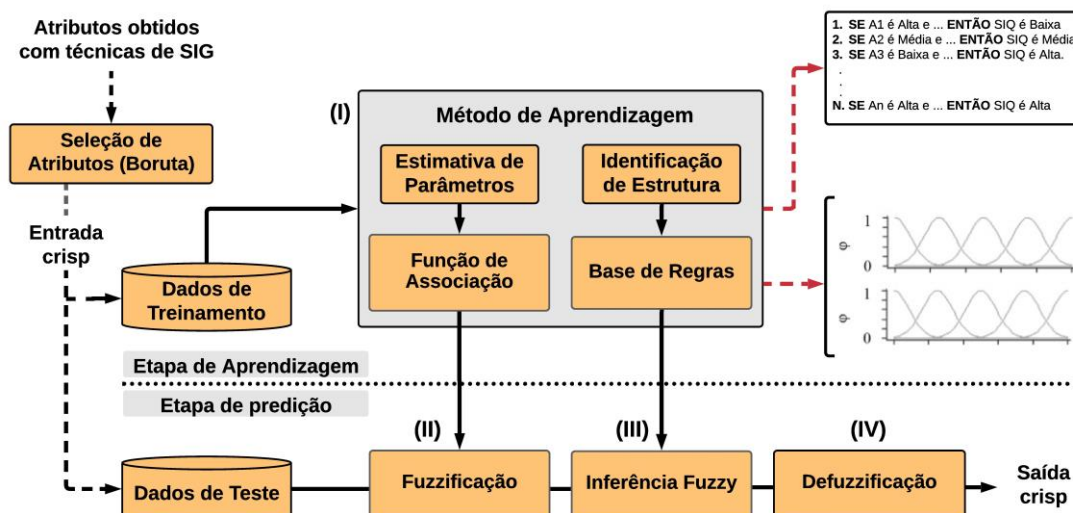
Onde X é um vetor padrão dimensional, A um valor linguístico antecedente, como pequeno, médio e grande, C_j a classe consequente, CF a certeza fuzzy da regra difusa, e N o número de regras difusas SE-ENTÃO (ISHIBUCHI; NAKASHIMA, 2001). Neste caso, a classe consequente é definida como a classe dominante no subespaço fuzzy correspondente aos antecedentes de cada regra fuzzy SE-ENTÃO. Dessa forma, a classe de uma nova instância pode ser determinada considerando a classe consequente da regra pelo produto máximo de seus graus de compatibilidade e CF , conforme definido por:

$$\mu_j(X_p) * CF_j = \max\{\mu_j(X_p) * CF_j: j = 1, 2, \dots, N\}$$

Onde $\mu_j(X_p)$ corresponde o grau de compatibilidade da regra difusa (SE-ENTÃO), e CF é a certeza fuzzy, sendo um número real no intervalo de unidade $[0,1]$ (ISHIBUCHI; NAKASHIMA, 2001). O grau de compatibilidade é determinado pela agregação de graus da função de pertinência dos antecedentes, enquanto o grau CF é calculado a partir da razão entre a classe subsequente (ISHIBUCHI; NAKASHIMA, 2001; RIZA et al., 2015).

A Figura 7 apresenta um diagrama esquemático da metodologia proposta para implementação do método HFIS de mapeamento da suscetibilidade à ocorrência de incêndios e queimadas. O sistema foi implementado no software RStudio (R Core Team, 2020) com uso do pacote “*frbs*” utilizando o método “FRBCS.W” (RIZA et al., 2015) considerando as variáveis identificadas como mais importantes. O exemplo do script desenvolvido para o mapeamento da suscetibilidade de incêndios e queimadas (SIQ) encontra-se disponível em anexo (Anexo B3).

Figura 7. Diagrama esquemático do método HFIS implementado para o mapeamento da suscetibilidade de incêndios e queimadas (SIQ).



Foram utilizados o conjunto de dados balanceados, conforme recomendações descritas por Pourghasemi et al. (2020). O modelo foi construído levando em consideração o particionamento do conjunto de dados em 70% para treinamento e 30% para teste para verificar o potencial de replicação do método. O percentual das partições para treinamento e teste adotado neste estudo apesar de não serem absolutos, é o mais comumente encontrado na literatura (LIU; COCEA, 2017).

Após ajuste do modelo HFIS, foi realizado a previsão dos eventos, em seguida, os dados de ocorrência de eventos de incêndios e queimadas previsto pelo modelo foram espacializados pelo método de densidade de Kernel considerando o número de ocorrências por km² (NASCIMENTO et al., 2017), em seguida, foi realizado uma classificação considerando o nível de densidade observada por km², sendo discretizadas nas classes Muito Baixo, Baixo, Moderado, Alto e Muito Alto utilizando o método de classificação de quebra natural (Natural Breaks - Jenks), representando o mapa de suscetibilidade de ocorrência de incêndios e queimadas, conforme descrito por Pourghasemi et al. (2020).

5.6. Análise de qualidade dos modelos

A qualidade dos classificadores foi avaliada considerando métricas de desempenho amplamente utilizados na literatura, tais como a especificidade (SPF), sensibilidade (SST), acurácia (ACC) e índice Kappa obtidos por meio da matriz de confusão de 2x2. De modo geral, a matriz de confusão fornece quatro tipos de resultados de classificação: verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN),

respectivamente (KUHN; JOHNSON, 2013; CIABURRO; VENKATESWARAN, 2017), conforme ilustrado na Figura 8 abaixo.

Figura 8. Matriz de confusão para classificação.

		Classe Prevista			
		K0	K1		
Classe Real	K0	VP	FN	VP	Verdadeiro Positivo
	K1	FP	VN	VN	Verdadeiro Negativo
				FP	Falso Positivo
				FN	Falso Negativo

Fonte: Adaptado de Lever et al. (2016).

Sendo VP para os casos em que o valor real é positivo sendo previsto um valor positivo pelo modelo, VN um valor negativo sendo previsto um valor negativo, FP e FN correspondem a valores previstos falsamente, onde o primeiro caso o valor real era negativo, mas o modelo previu um valor positivo, e no segundo caso o valor real era positivo, mas o modelo previu um valor negativo (KUHN & JOHNSON, 2013). Dessa forma, a especificidade (SPF) é dado pela Equação 10, a sensibilidade (SST) pela Equação 11, acurácia (ACC) pela Equação 12 e índice Kappa pela Equação 13 (KUHN; JOHNSON, 2013; CIABURRO; VENKATESWARAN, 2017).

$$SPF = \frac{VP}{VP + FP} \quad (10)$$

$$SST = \frac{FP}{VN + FP} \quad (11)$$

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (12)$$

$$Kappa = \frac{(VP + VN) - ((VP + FN)(VP + FP) + (FP + VN)(FN + VN))}{1 - ((VP + FN)(VP + FP) + (FP + VN)(FN + VP))} \quad (13)$$

As métricas de desempenho foram obtidas considerando o particionamento do conjunto de dados em 70% para treinamento e 30% para teste, conforme descrito

anteriormente, o que possibilitou verificar a habilidade dos diferentes métodos de aprendizagem utilizados para a previsão da densidade de ocorrência de incêndios e queimadas na área, ou seja, a capacidade de generalização dos métodos utilizados.

Após validação e identificação do modelo ideal de mapeamento da suscetibilidade de incêndios e queimadas, foi realizada a espacialização da suscetibilidade para a área de estudo. Para avaliar a eficiência do modelo HFIS, o sistema foi testado levando em consideração as classes de suscetibilidade mapeadas e o número de ocorrências de incêndios observado na região em 2019.

O mesmo processo foi realizado para dois eventos climáticos extremos (evento seco e chuvoso) com vista a verificar a eficácia do modelo proposto sob condições climáticas adversas. Para identificação dos eventos climáticos extremos foi utilizado o Índice Padronizado de Precipitação (*Standardized Precipitation Index* - SPI), conforme proposto por McKee et al. (1993), levando em consideração dados de precipitação observados entre 1990 a 2020 de duas estações pluviométricas localizada no município de Ibiúna, sob coordenadas métricas UTM 7382765 m N e 273893 m E (geocódigo DAAE nº E4-047), e 7375536 m N e 284208 m E (geocódigo DAAE nº E4-131) disponibilizada pelo Departamento de Águas e Energia Elétrica - DAEE (DAEE, 2021). Neste estudo, considerou-se como eventos extremos os anos que apresentaram valores de SPI $\pm 1,00$. Além do mais, considerando os dados de precipitação observados nessas estações, obteve-se a precipitação acumulada anual e a precipitação média mensal para a série histórica.

O SPI foi obtido no software R com uso do pacote *SPEI* (BEGUERÍA; SERRANO, 2017), sendo determinado a probabilidade de distribuição de frequência da precipitação por meio da distribuição gama, conforme descrito por Wilks (2006), em seguida, foi construído um gráfico contemplando os intervalos do SPI, sendo a classificação dos valores realizada conforme adaptação das classes propostas por McKee et al. (1993), conforme apresentado na Tabela 6.

Tabela 6. Classes do Índice Padronizado de Precipitação.

Valores de SPI	Classes de SPI
SPI < -1,50	Extremamente Seco
> -1,49 a -0,5	Moderadamente Seco
> -0,49 a 0,50	Quase Normal
> 0,5 a 1,50	Moderadamente Chuvoso
SPI > 1,50	Extremamente Chuvoso

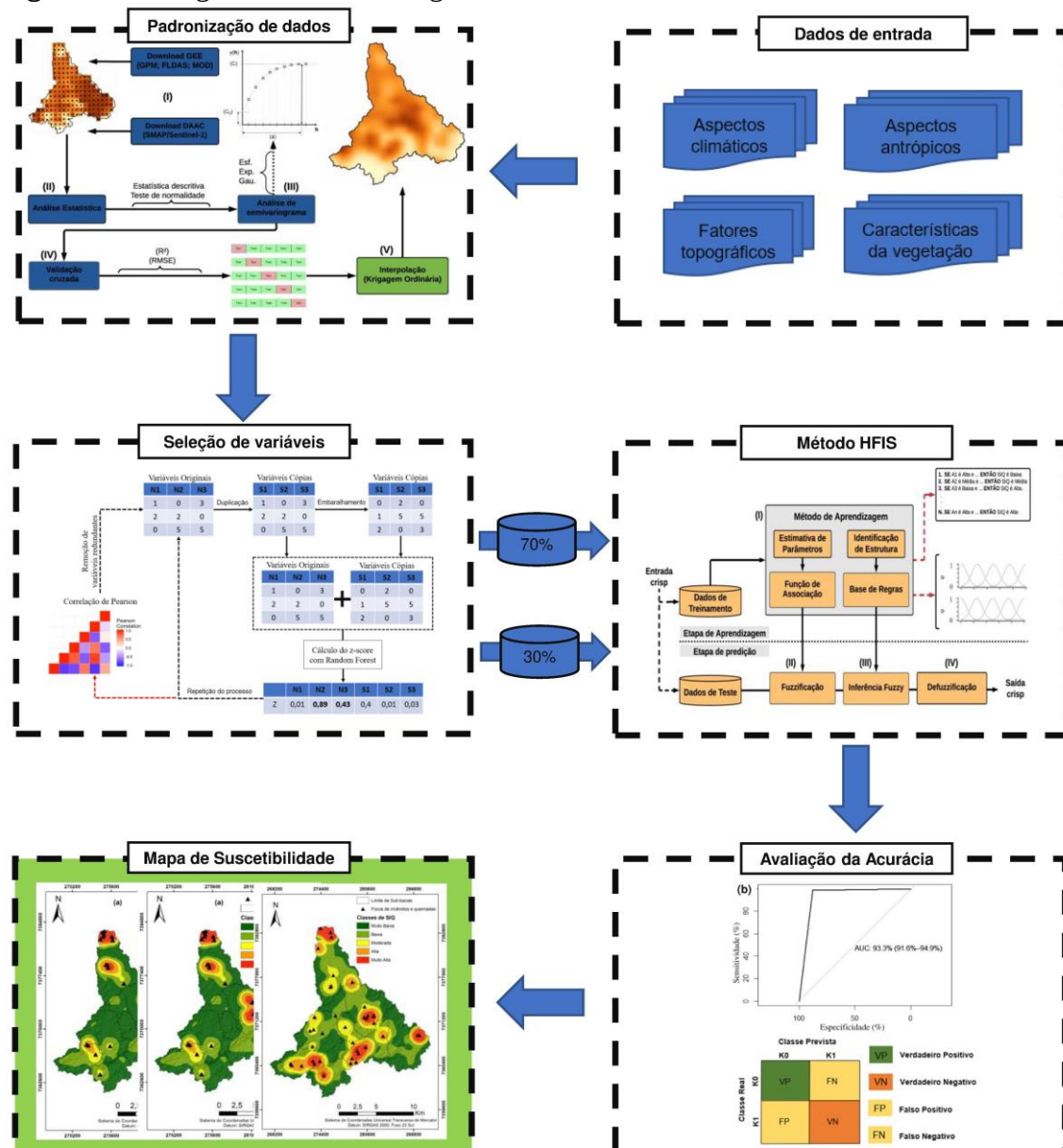
Adaptado de McKee et al. (1993).

Os mapas finais de suscetibilidade de incêndios e queimadas foram validados, conforme descrito por Venkatesh et al. (2020), sendo considerado o número de ocorrência de focos de incêndios observada nas áreas classificadas com alto e muito alta suscetibilidade pelo método HFIS, conforme apresentado na Equação 14.

$$Acurácia = \frac{\text{Focos nas classes alta e muito alta}}{\text{Total de focos}} \times 100 \quad (14)$$

A Figura 9 apresenta um fluxograma da metodologia adotada para a previsão da Suscetibilidade de Incêndios e Queimadas (SIQ) adotado neste estudo.

Figura 9. Fluxograma da metodologia adotada.

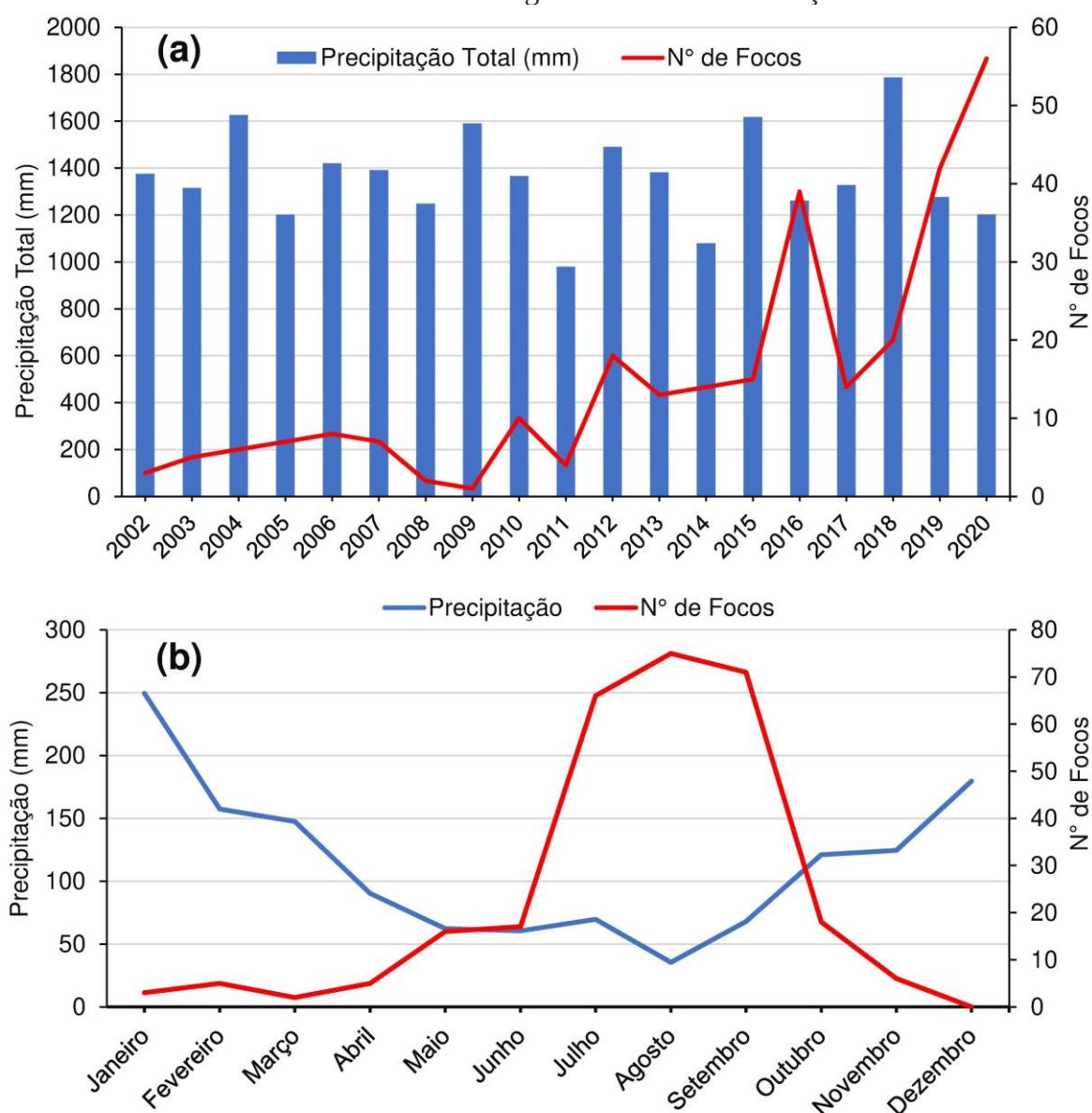


6. RESULTADOS E DISCUSSÃO

6.2. Sazonalidade de incêndios e queimadas

A Figura 10 apresenta o número de focos de incêndios, queimadas e a precipitação anual acumulada (Figura 10a), bem como a precipitação média mensal e o número de focos de incêndio acumulado (Figura 10b) entre 2002 a 2020 na Bacia Hidrográfica do Rio Sorocabaçu.

Figura 10. Número de focos de incêndios e precipitação anual (a) e média mensal (b) observado entre 2002 a 2020 na bacia hidrográfica do rio Sorocabaçu.

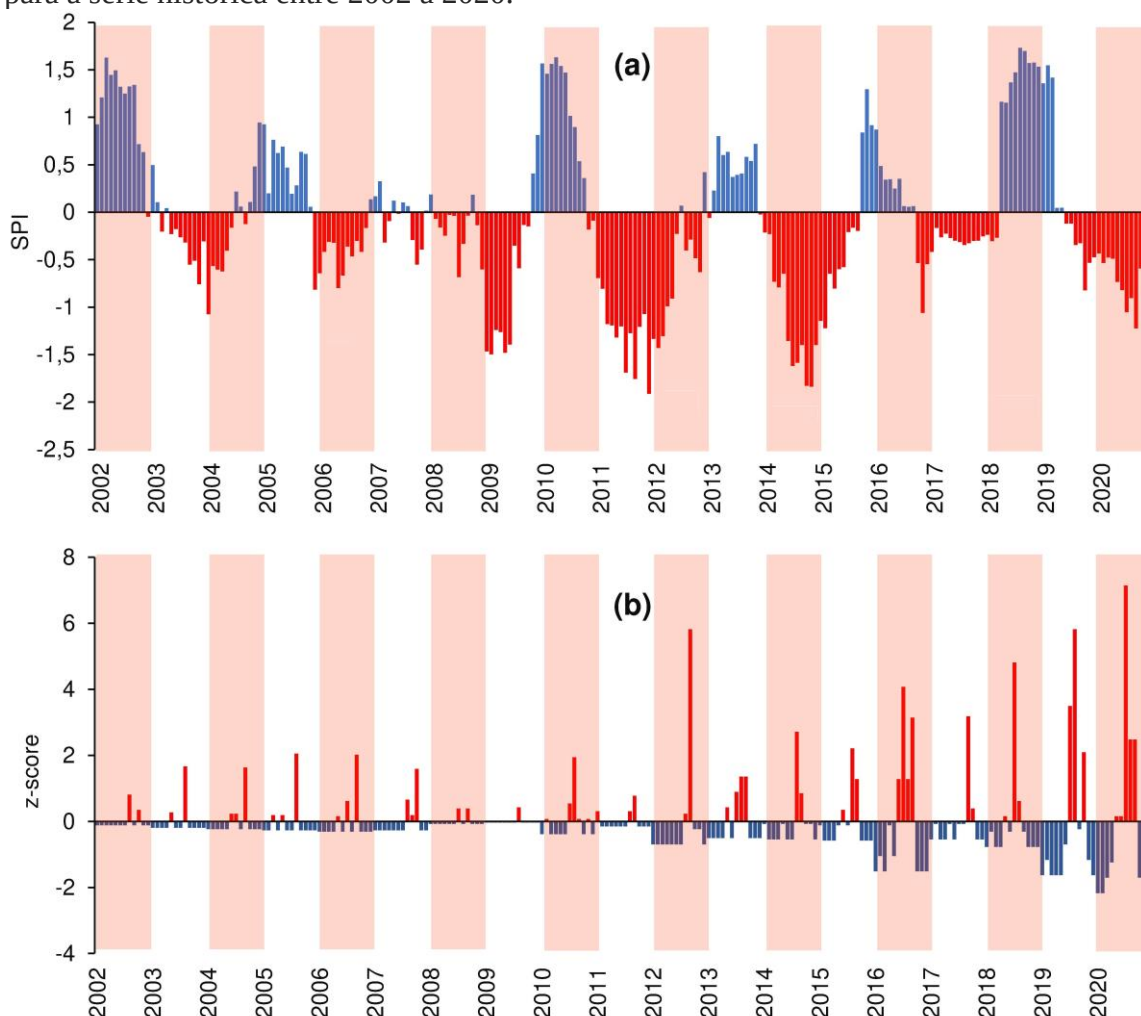


Observa-se que a partir de 2012 o número de ocorrência de focos de incêndios e queimadas na bacia hidrográfica aumentaram consideravelmente quando comparado aos valores observados nos anos anteriores (Figura 10a). Esse aumento pode ser explicado pelo fato que em 2011 foram lançados os satélites NPP-375 e NPP-750, e por possuírem melhor resolução espacial, possibilitou registrar focos de incêndios e queimadas em áreas menores do que os satélites utilizados até então, o que representou um salto na qualidade e na quantidade de registros de eventos (IEMBO; GALVANI, 2021). Dessa forma, o baixo número de incêndios e queimadas registrado nos anos anteriores a 2012 podem ter efetivamente ocorrido, bem como ter sido subdetectados em função da baixa resolução dos sensores utilizados até então.

Em relação a sazonalidade anual, observa-se que os incêndios ocorrem com maior frequência entre os meses que apresentam menor precipitação (maio, junho, julho, agosto e setembro), pois o material combustível torna-se menos úmido, assim como já relatado em estudos anteriores (LOPES et al. 2018; SOUSA, 2021). Além disso, os dados históricos mostram que, no período de ocorrência de incêndios, a frequência de ocorrência é mais intensa entre os meses de julho, agosto e setembro, e com menor frequência nos meses de maio, junho e outubro, meses que podem ser considerados como de transição (Figura 10b).

Com vista em avaliar a relação entre intensidade de ocorrência de incêndios e queimadas com a precipitação local, obteve-se o SPI (*Standardized Precipitation Index* – SPI) para a área de estudo com base nos dados de precipitação na bacia hidrográfica, apresentado na Figura 11 (a), e a Figura 11 (b) apresenta o número de ocorrências de incêndios e queimadas padronizado (z-score), ambos gerados com base em dados da série histórica entre 2002 a 2020.

Figura 11. Padrões temporais do SPI (a) e de focos de incêndio (b) padronizado (z-score) para a série histórica entre 2002 a 2020.



Com base no SPI (Figura 11a), observa-se que ao longo da série histórica ocorreram períodos que podem ser considerados como extremamente seco (2009, 2011 e 2014), e moderadamente seco (2003, 2006, 2018 e 2020), por outro lado, também foi observado períodos que podem ser considerados como extremamente chuvoso (2002, 2010 e 2018) moderadamente chuvoso (2005, 2013 e 2016), e o restante, classificados como quase normal, conforme classificação do SPI adaptada de McKee et al. (1993).

Avaliando o número de ocorrências de incêndios e queimadas padronizado (z-score) é possível identificar as anomalias dos eventos ocorridos na série histórica (ARAGÃO et al., 2018). Observa-se que entre 2002 a 2011 a distribuição z-score mostra a predominância de eventos amenos, e a partir de 2012 ocorre uma crescente nas anomalias, com maiores valores de z-score observados em 2020 ($7,14\sigma$), seguido de 2012 ($5,82\sigma$), 2019 ($5,82\sigma$), 2018 ($4,81\sigma$) e 2016 ($4,07\sigma$), mostrando que apesar de ocorrerem períodos considerado

extremamente chuvoso em 2018 e moderado em 2016, ocorreram anomalias acentuadas nesses períodos.

6.3. Descrição das variáveis de entrada

A Tabela 7 apresenta a estatística descritiva e geoestatística para os valores médios observados entre os meses de maio a setembro (período de menor taxa de precipitação e maior número de ocorrência de eventos na série histórica) para a precipitação, umidade do ar, temperatura próximo à superfície e umidade do solo regularizado por KO.

Tabela 7. Estatística descritiva e geoestatística para as variáveis regularizadas.

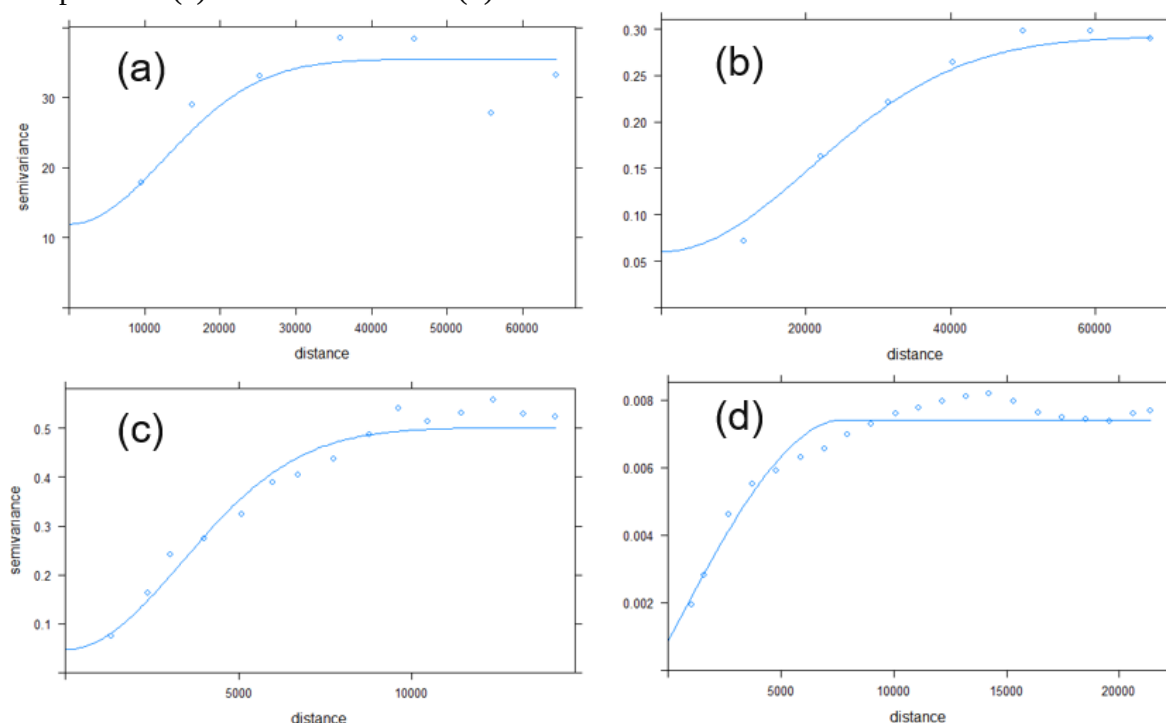
Estatística Descritiva				
Parâmetros	Pr (mm)	UA (kg.kg ⁻¹)	T (°C)	US (cm ³ .cm ⁻³)
Máximo	83,49	11,92	26,38	0,53
Mínimo	49,80	9,01	20,98	0,04
Média	63,68	9,94	23,73	0,24
Mediana	64,11	9,62	24,06	0,24
Desvio Padrão	7,15	0,79	1,17	0,08
Assimetria	0,12	1,16	-0,54	0,52
Curtose	-0,63	0,02	-0,75	0,49
CV	11,23	7,96	4,94	34,52
p-valor (KS)	0,906	2,07x10 ⁻⁶	2,24x10 ⁻⁶	0,0017
NT (ORQ)	-	1,02	1,15	1,15
Geoestatística				
Modelo	Gau.	Gau.	Gau.	Esf.
Efeito Pepita	11,93	0,0601	0,046	0,0008
Patamar	23,57	0,231	0,4545	0,0065
Alcance (m)	17.717,23	29.319,55	4.746,16	7.741,01
R ²	0,83	0,90	0,96	0,83
RMSE	4,06	0,17	0,22	0,03
GDE (%)	50,51	26,13	10,12	12,30

***Pr.:** Precipitação; **UA.:** Umidade do ar; **T.:** Temperatura; **US:** Umidade do solo; **CV:** Coeficiente de Variação; **KS:** Kolmogorov-Smirnov; **NT:** Normalization Transformation; **ORQ:** Ordered Quantile; **Esf.:** Esférico; **Gau.:** Gaussiano; **Exp.:** Exponencial; **RMSE.:** Root Mean Squared Error; **GDE.:** Grau de Dependência Espacial (GDE < 25% - Forte; 25 ≤ GDE < 75 - Moderado; GDE ≥ 0,75 - Fraco).

Observa-se que as variáveis umidade do ar, temperatura e umidade do solo não apresentaram distribuição normal, ao nível de 5% de significância pelo teste de Kolmogorov-Smirnov, nesse caso, foi realizado a transformação para normalização dos dados por meio do método *Ordered Quantile*. O histograma da distribuição normal para os dados de precipitação, e pós-transformação pelo método de *Ordered Quantile* para as demais variáveis são apresentados no Anexo C1.

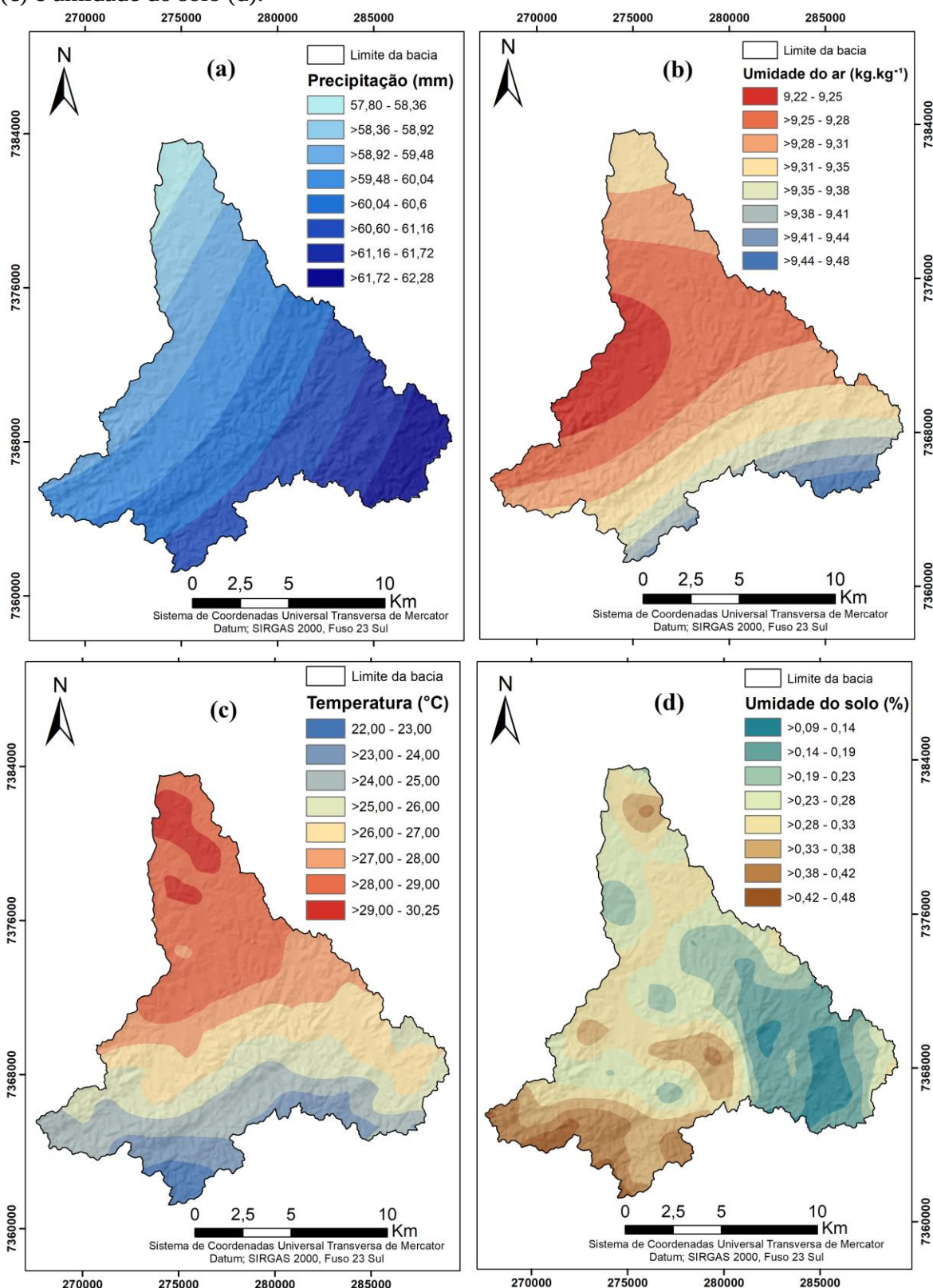
Em relação à análise geoestatística, os ajustes dos modelos teóricos foram avaliados por meio da avaliação do efeito pepita, patamar e alcance, seguido de análise dos melhores valores de R^2 e RMSE para cada modelo teórico com uso *Cross-Validation*. Os resultados mostraram que em ambos os atributos, o GDE foi considerado moderado a forte, sendo a precipitação, umidade do ar e temperatura apresentaram dependência espacial com melhores ajustes para o modelo gaussiano, enquanto a umidade do solo apresentou o melhor ajuste para o modelo esférico (Figura 12).

Figura 12. Semivariograma ajustado para a precipitação (a), umidade do ar (b), temperatura (c) e umidade do solo (d).



O semivariograma ajustado para a precipitação (Figura 12a), umidade do ar (Figura 12b), temperatura (Figura 12c) e umidade do solo (Figura 12d) apresentaram ajuste bem estruturado, conforme descrito por Yamamoto (2020). Os dados sobre a análise da *Cross-Validation* para a KO encontram-se no Anexo C2. Após ajuste dos modelos teóricos, foi realizada a Krigagem Ordinária, interpolação realizada com grid ajustado com dimensões de 30 m. Detalhes sobre a análise de *Cross-Validation* para cada variável encontra-se disponível no Anexo C2. O resultado da interpolação para a precipitação, umidade do ar, temperatura e umidade do solo são apresentados na Figura 13.

Figura 13. Mapas dos valores médios da precipitação (a), umidade do ar (b), temperatura (c) e umidade do solo (d).



A precipitação é um fator de suma importância, pois afetam diretamente na produtividade fotossintética de vegetação, influenciando no material combustível, na

umidade do solo e umidade do ar, impactando diretamente na suscetibilidade de ocorrência de incêndios e queimadas (VADREVU et al., 2006; VENKATESH et al. 2020). Na área de estudo, a precipitação média (Figura 13a) estimada entre os meses de maio a setembro de 2019 pelo sensor GPM variam entre 57,80 mm a 62,28 mm, sendo as maiores taxas observadas a montante da bacia, no sentido sudeste, e os menores valores foram observados a jusante da bacia, no sentido noroeste. Estudos anteriores mostraram que a precipitação estimada pelo sensor GPM apresenta boa similaridade com os valores observados por estação pluviométrica de superfície para o Brasil, e em especial, para a região sudeste, o que torna o GPM uma excelente fonte de dados que pode ser utilizado em diversos estudos, em especial, a nível de bacias hidrográficas devido a boa resolução espacial e temporal (ROZANTE et al., 2018; GADALHA et al., 2020).

A umidade do ar, por sua vez, altera de forma direta a inflamabilidade do material combustível, influenciando na suscetibilidade de incêndios e queimadas (POURGHASEMI et al., 2020; VENKATESH et al. 2020). Os dados médios entre os meses de maio a setembro de 2019 obtidos a partir de dados do FLDAS (Figura 13b) mostraram que existe baixa variabilidade espacial (variando entre 9,22 kg.kg⁻¹ a 9,48 kg.kg⁻¹), sendo que a região central da bacia possui os menores valores, locais onde se concentram as maiores extensões de atividade antrópica, enquanto as regiões que apresentaram maior umidade ocorrem a montante da bacia, principalmente na região sudeste, onde ocorrem as maiores taxas de precipitação, e onde se encontram os maiores maciços florestais (SOUSA, 2021).

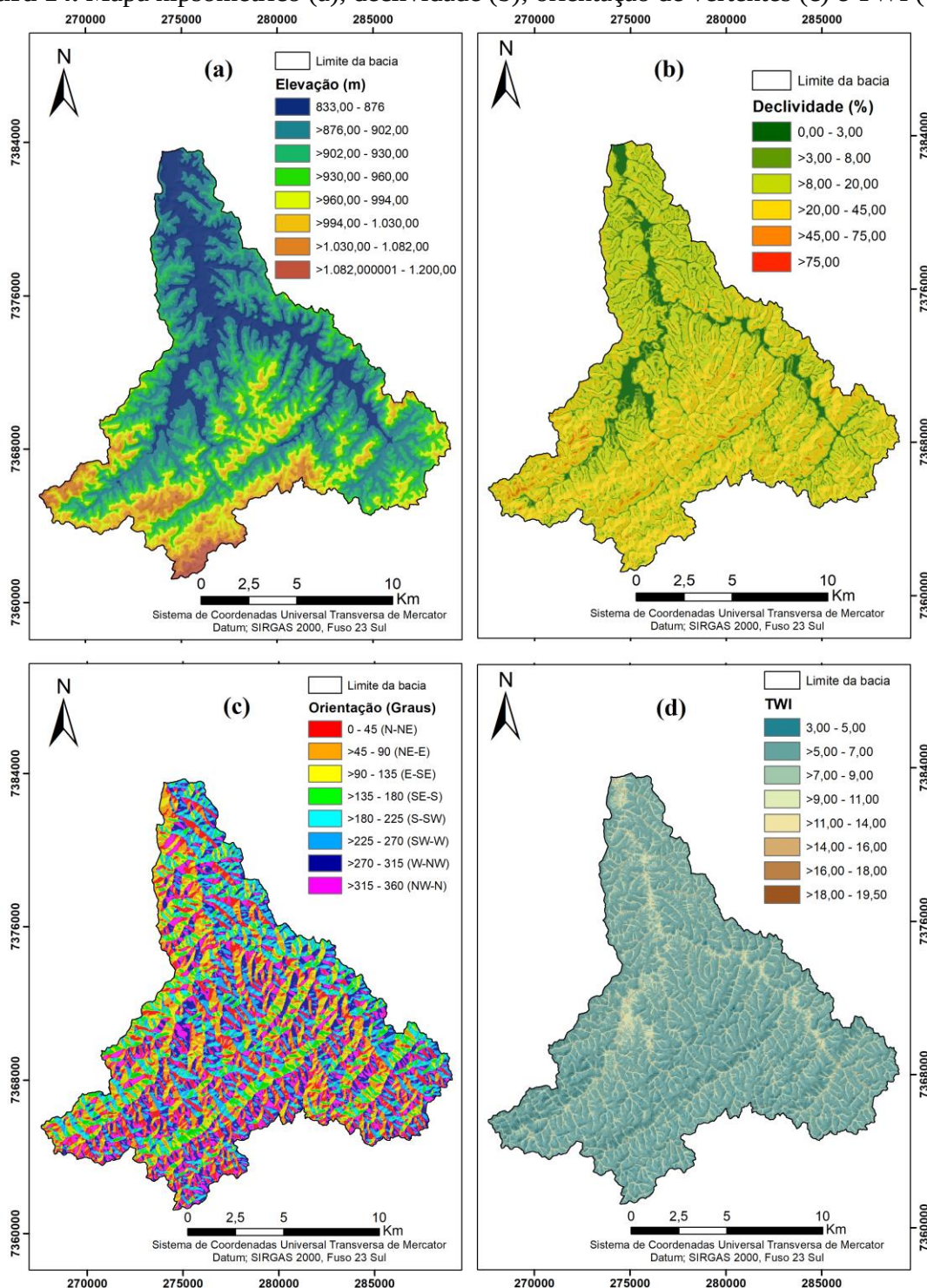
A temperatura, por sua vez, por afetar a umidade do ar e do solo, está diretamente ligada à combustão da vegetação, contribuindo de forma direta na suscetibilidade na ocorrência de incêndio e queimadas (POURGHASEMI et al., 2020). Na área de estudo, temperatura média observada entre os meses de maio a setembro de 2019 estimada pelo sensor MOD11 (Figura 13c) mostraram que as áreas próximas a foz apresentam valores variando entre 28°C a 30,25°C, e a montante é a região que apresentou as temperaturas mais amenas, variando entre 22°C a 23°C, respectivamente, indicando a influência da altitude do relevo na variabilidade da temperatura de superfície, conforme relatado por Padovanni et al. (2018).

Por regular a umidade da vegetação em longo prazo, a umidade do solo também é uma variável de suma importância (RAKHMATULINA et al., 2021). Os dados de umidade do solo (Figura 13d) estimados pelo sensor SMAP/Sentinel-1 entre os meses de maio a setembro de 2019 mostraram que a montante da bacia, na região sudeste, corresponde a área que apresenta os menores valores (0,09% a 0,23%), por outro lado, a sudoeste da bacia

encontram-se as áreas com maior umidade do solo, com valores variando entre 0,38% a 0,48%, respectivamente.

A Figura 14 apresenta o mapa hipsométrico (a), declividade (b), orientação de vertentes (c) e TWI (d), ambos obtidos com base em dados planialtimétricos provenientes de cartas topográficas do IGC (2010).

Figura 14. Mapa hipsométrico (a), declividade (b), orientação de vertentes (c) e TWI (d).



Segundo Chuvieco (2019), a elevação interfere diretamente na umidade, pois contribui com o acúmulo e/ou escoamento de água no solo. A Figura 14a apresenta a altimetria da bacia, observa-se que a elevação local varia entre 833 a 1.200 metros, sendo as áreas ao longo dos principais corpos hídricos e na região central os locais com menor elevação, enquanto as áreas mais elevadas se localizam a montante da bacia, principalmente na região sul e sudoeste.

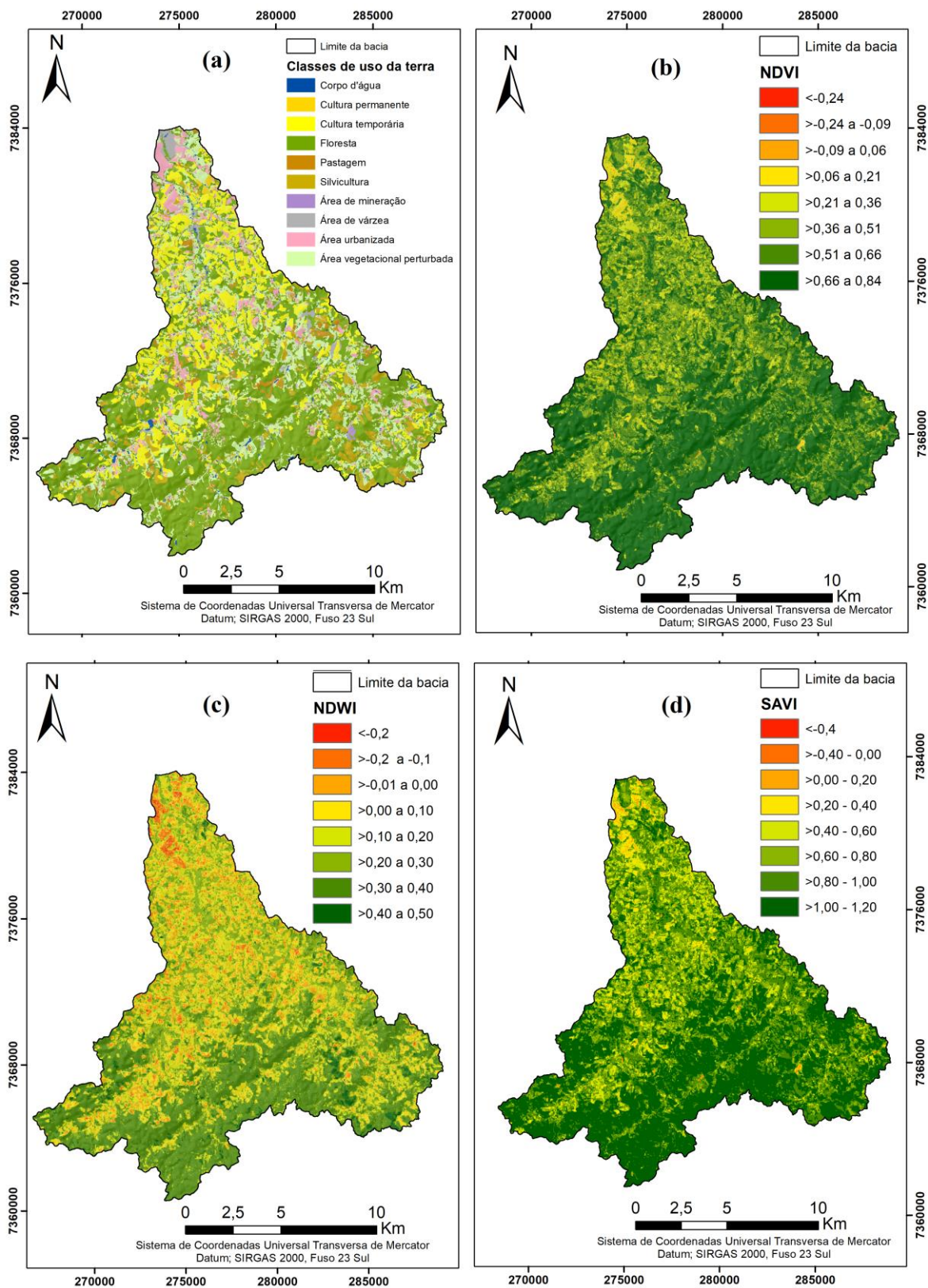
Considerando as classes de declividade estabelecidas pela EMBRAPA (2018), a bacia hidrográfica apresenta terreno predominantemente ondulado (46,50%) a fortemente ondulado (35,38%), seguido de suavemente ondulado (11,31%), plano (5,96%), montanhoso (0,82%) e escarpado (0,0003%). As regiões com baixas altitudes próximas aos corpos hídricos principais correspondem às áreas de terreno plano (declividade $< 3\%$), enquanto as áreas com terreno montanhoso e escarpado (declividade $> 45\%$) são representados por porções mais isoladas localizadas principalmente na região oeste e sudoeste da bacia hidrográfica (Figura 14b).

A orientação do terreno é um fator de suma importância para compreensão de incêndios e queimadas, pois influencia a quantidade de radiação solar incidente no terreno, e por consequência, na umidade do solo, temperatura do ar e velocidade do vento (EUGENIO et al., 2016; VENKATESH et al., 2020). Na região, as maiores incidências de radiação ao longo do ano se dão nos terrenos orientados nas direções N-NE e N-NW (MARTIN, 2021), sendo que na área de estudo, as encostas (Figura 14d) estão predominantemente orientadas na direção S-SW (17,79%), seguido de áreas no sentido W-NW (14,96%), E-NE (13,70%), N-NW (13,63%), N-NE (13,08), E-SE (12,33%), e com menor percentual, S-SE (9,90%) e W-SW (4,60%), respectivamente.

Em relação ao *Topographic Wetness Index*-TWI (Figura 14d), ele indica a estabilidade de umidade do solo em uma área, definida pela contribuição da encosta e da região a montante em influenciar sua habilidade/capacidade de reter água no solo (LEI et al., 2016; NÓBREGA et al., 2018). Na bacia hidrográfica, os valores de TWI variaram entre 3,00 a 19,50, sendo observado os menores valores com predominância em áreas com maior declividade, enquanto os valores de TWI mais elevados ocorrem em fundo de vale e áreas mais planas, locais que possuem maior potencial de retenção de água.

Na Figura 15a são apresentados as classes de cobertura e uso da terra para o ano de 2019, bem como os valores médios para o NDVI (Figura 15b), NDWI (Figura 15c) e SAVI (Figura 15d) observados entre os meses de maio a setembro para o mesmo ano.

Figura 15. Mapa de cobertura e uso da terra (a), NDVI (a), NDWI (b) e SAVI (c) para o ano de 2019.



Em relação à cobertura e uso da terra (Figura 15a), observou-se que a classe predominante refere-se aos remanescentes florestais (45,72%), seguido por área vegetacional perturbada (17%), cultura temporária (17%), área urbanizada (6,62%), silvicultura (6,32%), áreas de pastagem (2,17%), área de várzea (1,33%), corpos d'água (1,17%), agricultura perene (0,51%), e área de mineração (0,11%). Os remanescentes florestais são representados por pequenos, médio, grande e muito grandes fragmentos dispersos em toda bacia, totalizando cerca de 9.266,41 hectares, onde os fragmentos menores ocorrem principalmente próximo à jusante, e fragmentos florestais maiores e mais conectados ocorrem a montante, esses por sua vez predominam em áreas com maior declividade do terreno, o que não propicia o desenvolvimento de atividades agrícolas (SOUZA, 2021).

A segunda classe de uso da terra predominante refere-se a áreas vegetacionais perturbadas, ocupando cerca de 3.860,72 hectares. Sousa (2021) descreve que a maioria dessas áreas são resultantes do uso intensivo do solo por atividades agrícolas, sendo posteriormente abandonadas. As atividades agrícolas por sua vez, correspondem à terceira classe de uso e ocupação do solo predominante na bacia representada pelas culturas permanentes e temporárias, o que perfaz cerca de 3.548,15 hectares. Essas atividades são desenvolvidas principalmente por pequenos agricultores, cujas práticas desenvolvidas predominam o cultivo de hortaliças e legumes que são atividades comuns na região, conforme descrito por Silva et al. (2016) e Souza et al. (2018).

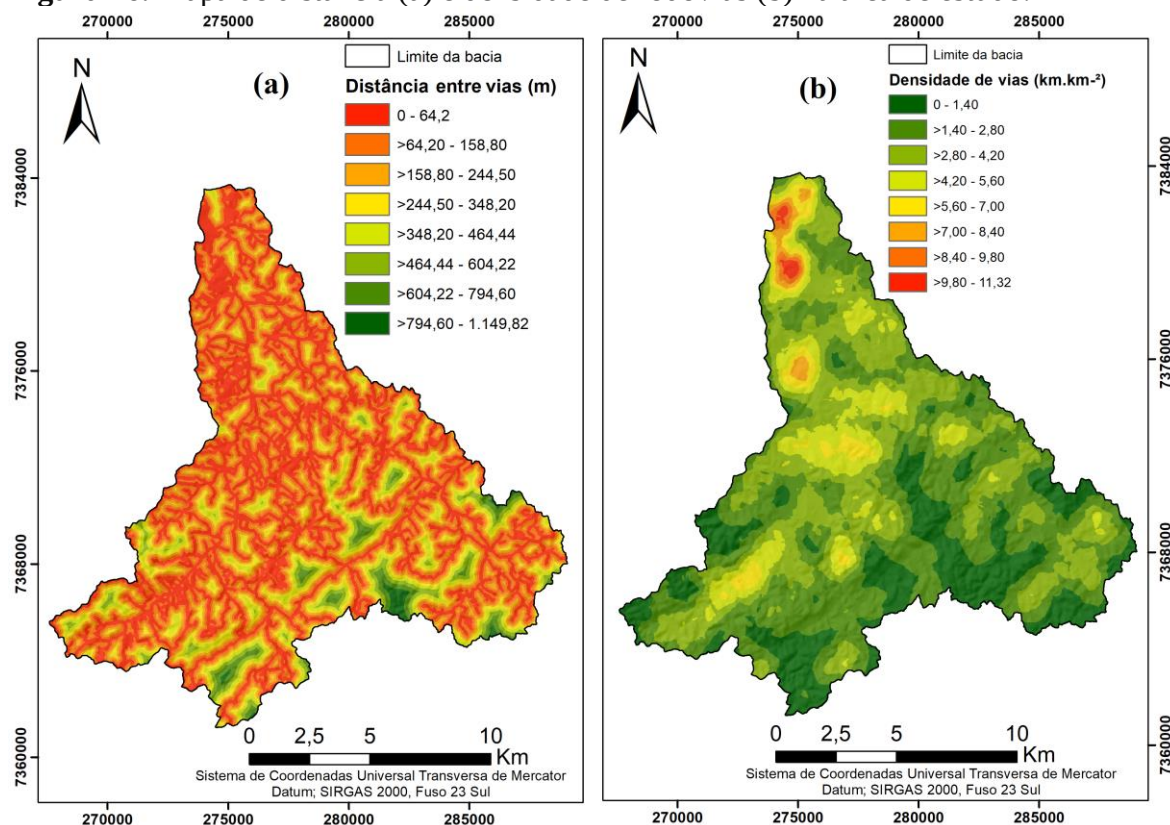
As áreas urbanizadas ocupam cerca de 1.341,29 hectares, sendo sua distribuição espacial predominante na região noroeste, onde se localiza a sede municipal de Ibiúna, e demais pequenos núcleos urbanos e rurais distribuídos em toda extensão da bacia hidrográfica. A prática de silvicultura representada pelo cultivo de *Eucalyptus* sp. e *Pinus* sp., cultivares comum na região (SOUZA et al., 2018; SOUZA, 2021), ocupam cerca de 1.279,91 hectares na bacia, essas atividades são distribuídas em toda extensão da área de estudo, porém, com predominância a montante da bacia hidrográfica.

Quanto aos índices de vegetação representados pelo NDVI (Figura 15b), NDWI (Figura 15c) e SAVI (Figura 15d), observa-se que em ambos, os maiores valores ocorrem em áreas ocupadas por fragmentos florestais e silvicultura, cuja produtividade fotossintética é elevada. As áreas que apresentaram valores moderados referem-se a locais ocupados por atividades agropecuárias e áreas vegetacional perturbada, enquanto os menores valores correspondem a vegetação morta e/ou alvos não vegetais tais com corpos d'água e área urbana (ROUSE et al., 1973; GAO, 1996; MKHABELA et al., 2011). Os índices de

vegetação são de suma importância, pois possibilitam avaliar o grau de produtividade fotossintética e/ou de ressecamento da vegetação (SILVA et al., 2016).

As vias de acesso comumente representadas por rodovias e estradas vicinais, trilhos de trem, dentre outros, também são variáveis influentes na ocorrência de incêndios e queimadas (ZUMBRUNNEN et al., 2011; TORRES et al., 2017). A Figura 16a apresenta o mapa de distância euclidiana entre as rodovias e estradas vicinais e a Figura 16b a densidade de rodovias e estradas vicinais respectivamente.

Figura 16. Mapa de distância (a) e densidade de rodovias (b) na área de estudo.



Quanto a distância de rodovias e estradas vicinais (Figura 16a), diferentemente aos locais com baixa densidade, como a região Amazônica, em áreas densamente povoadas a relação com incêndios e queimadas não é direta (ZUMBRUNNEN et al., 2011; TORRES et al., 2017; ARAGÃO et al., 2018). Na área de estudo, existem cerca de 611 km de rodovias e estradas vicinais. Como já esperado, as áreas com maior malha viária encontram-se no perímetro urbano da sede municipal de Ibiúna, bem como nos demais pequenos núcleos urbanos (SOUSA, 2021). Além disso, as rodovias e estradas vicinais no perímetro rural estão raramente localizadas em áreas mais declivosas, sendo possível notar, de modo geral, que as

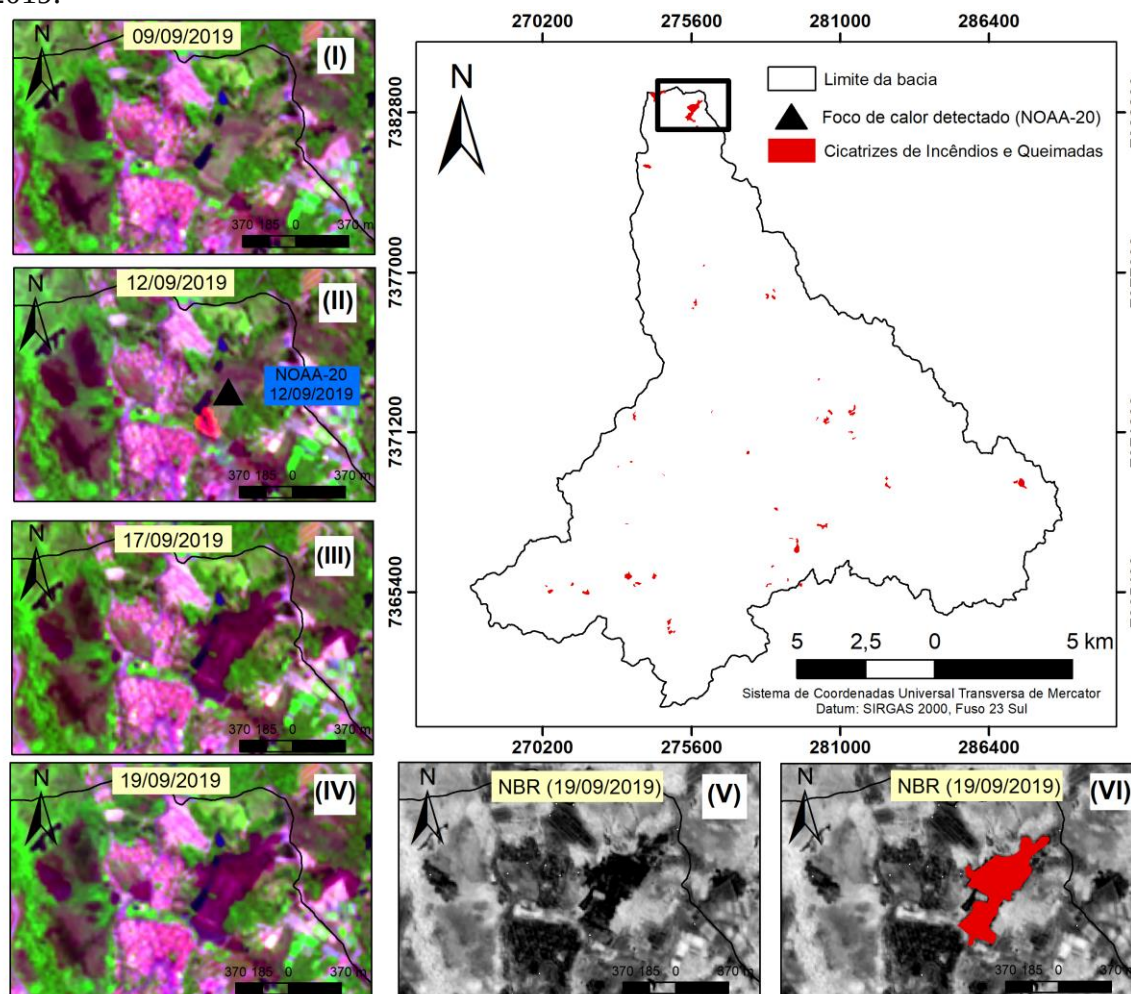
rodovias principais foram estabelecidas próximas aos fundos de vale, em áreas com elevação variando entre 800 a 900 metros.

Essa forma distribuição da malha viária faz com que a maior densidade de rodovias e estradas vicinais na área de estudo se dê na região noroeste da bacia (Figura 16b), local onde se encontra parte da área urbana do município de Ibiúna. Nas demais localidades da área rural, observa-se boa conectividade de vias na bacia, principalmente em decorrência dos pequenos núcleos urbanos distribuídos em toda extensão, entretanto, observa-se que as áreas com declividade menos acidentada são as que possuem maior densidade de vias, o que pode ser explicado em função de propiciar o desenvolvimento de práticas agrícolas, que por sua vez necessita de acessibilidade, facilitada em função de possuir terreno mais plano (SOUSA, 2021).

6.4. Análise exploratória dos dados

A Figura 17 apresenta as cicatrizes de incêndios e queimadas observadas em 2019 com base em imagens Sentinel-2, bem como exemplos de áreas pré e pós-evento de incêndios e queimadas com base em composição colorida falsa cor (RGB - B12/B8A/B04) obtida na base do EO Browser (EO Browser, 2021).

Figura 17. Exemplo de áreas identificadas como cicatrizes de incêndios e queimadas em 2019.

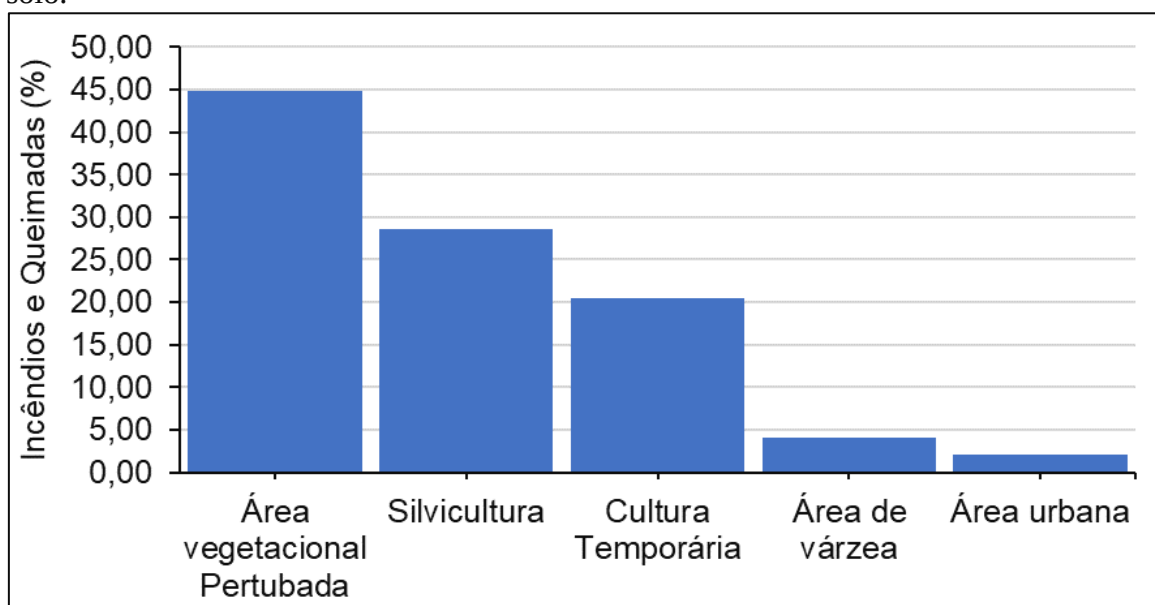


Considerando a sequência de cenas pré e pós evento apresentado na Figura 17, observa-se que na cena pré evento, no dia 09/09/2019 (Figura 17 - I) a área de ocorrência do foco de calor era ocupada por vegetação rasteira, possivelmente gramínea, e no dia 12/09/2019 foi detectado a anomalia de foco de calor pelo sensor NOAA-20 (Figura 17 - II), e pós evento, nos dias 17/09/2019 e 19/09/2019 (Figura 17 III e IV) é possível identificar a área de cicatriz do evento detectado pelo sensor NOAA-20. Com base em cenas de pós evento, obteve-se o NBR que possibilitou identificar a cicatriz de incêndio (Figura 17 - V) e, posteriormente, mapear essas áreas (Figura 17 - VI).

Com a identificação e mapeamento das cicatrizes de incêndios e queimadas na área de estudo, observou-se que a extensão de ocorrência desses eventos na bacia hidrográfica para o ano de 2019 foi de 95,48 hectares, com áreas queimadas variando entre 0,23 a 17,21, com predominância de áreas menores (média de 2,01 hectares), distribuídas em toda extensão da bacia hidrográfica.

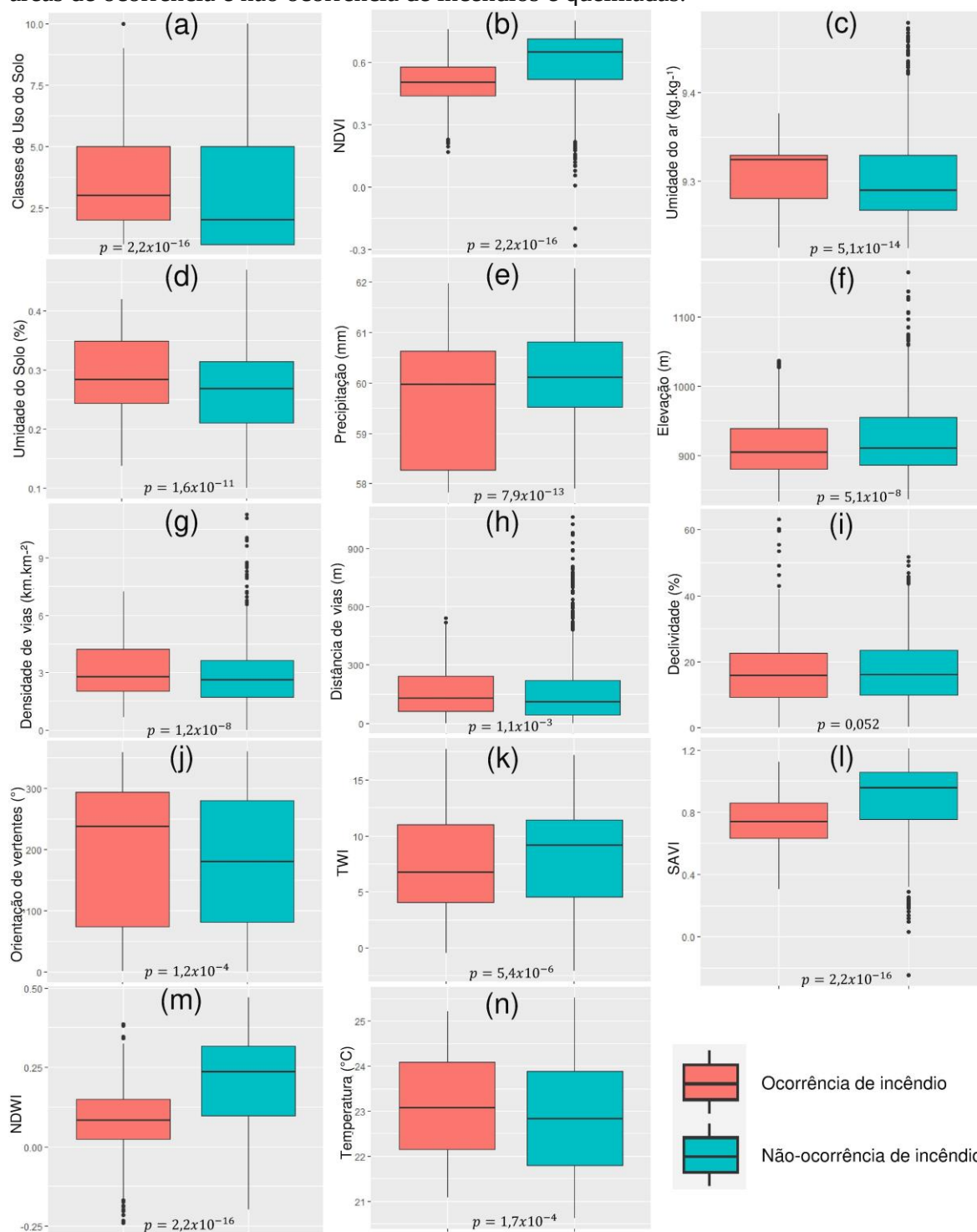
Pela análise de interseção entre os focos de incêndios e queimadas com as classes de uso e ocupação do solo foi possível identificar que esses eventos ocorrem com maior frequência em áreas ocupadas por área vegetacional perturbada (44,90%), ou seja, em áreas abandonadas, seguido com menor percentual em áreas ocupadas por silvicultura (28,57%), cultura temporária (20,41%), áreas de várzea (4,08%) e área urbana (2,04%), respectivamente, conforme apresentado na Figura 18.

Figura 18. Interseção entre focos de incêndios e queimadas e classes de uso e ocupação do solo.



A Figura 19 apresenta gráficos boxplots e resultado da estatística do teste de Wilcoxon ($\alpha=0,05$) para verificar a existência de diferença na distribuição dos valores para cada atributo em áreas de ocorrência e de não ocorrência de incêndios e queimadas. Vale ressaltar que os valores dos aspectos climáticos, bem como do NDVI, NDWI e SAVI correspondem a valores médios observados entre os meses de maio a setembro de 2019, período de maior ocorrência dos eventos de incêndios e queimadas.

Figura 19. Gráficos boxplot e valores do teste Wilconxon para os atributos avaliados em áreas de ocorrência e não ocorrência de incêndios e queimadas.



**p*-valor para o teste de Wilcoxon a 0,05 de significância: H0 = as médias entre as amostras são iguais, H1= Existe diferença entre as médias das amostras; *p*-valor < 0,05, rejeita-se H0.

Com base nos gráficos de boxplot, é possível identificar os atributos potenciais para limiarização de áreas com ocorrência e não ocorrência de incêndios e queimadas, tais como o NDVI (Figura 19b), umidade do solo (Figura 19d), precipitação (Figura 19e), SAVI (Figura 19l) e NDWI (Figura 19m), pois tais atributos não apresentam sobreposição, o que

facilita a separação entre as classes (QUEVEDO et al., 2020). Por outro lado, os demais atributos apresentaram medianas semelhantes, tais como observado para a elevação (Figura 19f), densidade de rodovias (Figura 19g), distância entre as rodovias (Figura 19h), e declividade (Figura 19i).

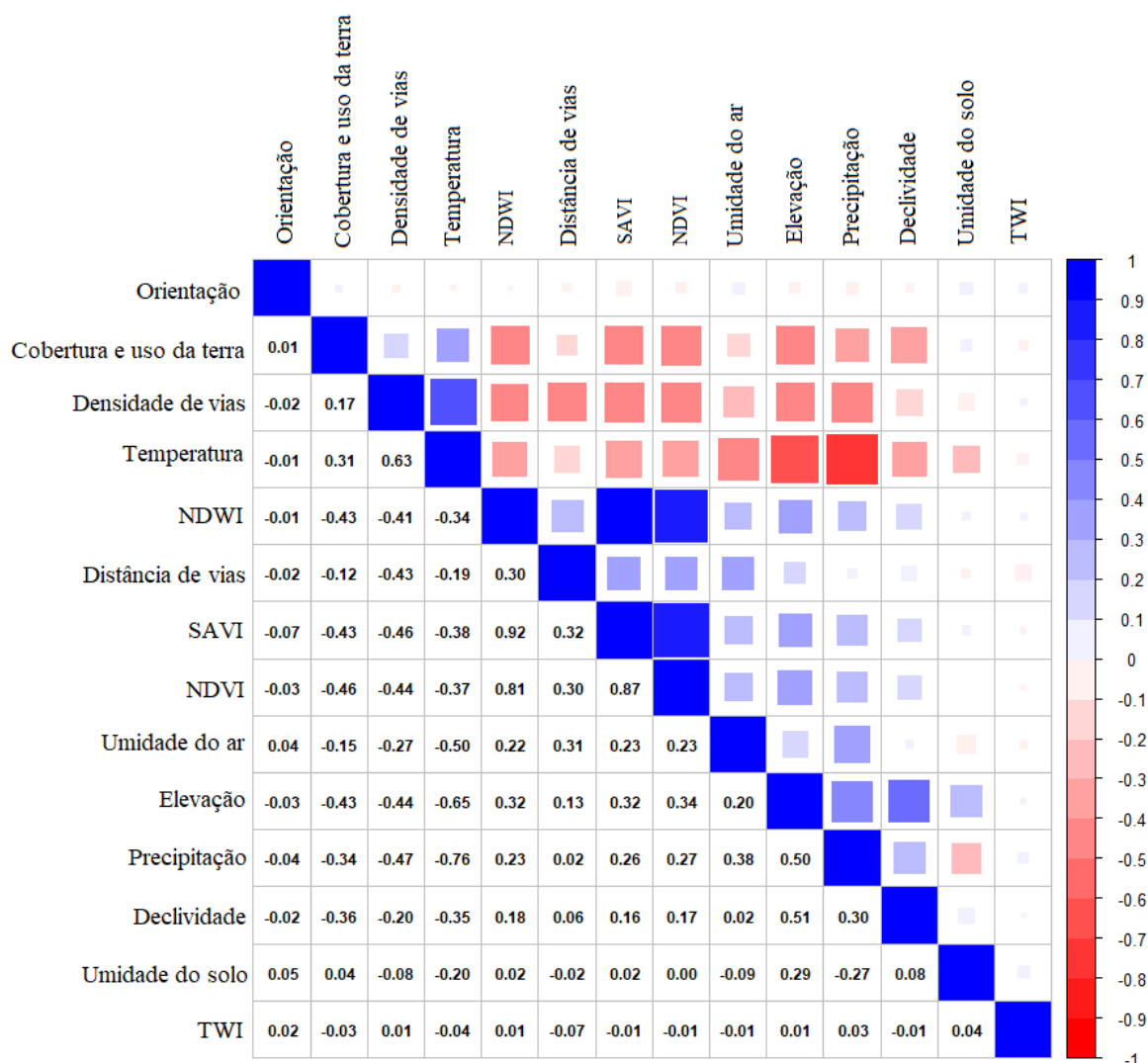
Com exceção para a declividade, todos os atributos apresentaram diferença significativa entre as medianas pelo teste de Wilcoxon a 5% de significância (demais detalhes sobre o teste Wilcoxon encontra-se no Anexo C3), implicando que em termos significativos, os valores observados em áreas de ocorrência de incêndio queimadas diferem quando comparado aos valores observados em áreas de não ocorrência desses eventos. Em relação à declividade, o teste de Wilcoxon mostrou que não existe diferença significativa ($p=0,052$), conforme mostrado no gráfico boxplot (Figura 19i), onde os valores do primeiro, segundo e terceiro quartil praticamente se sobrepõem, diferindo apenas nas ocorrências de outliers (pontos localizados acima do valor máximo).

Entretanto, o teste de Wilcoxon apenas indica se existe ou não diferença entre os dados em áreas de queimada e não queimada, e não necessariamente, se esses atributos possuem potencial para contribuição na classificação dos eventos quando implementados em um modelo de classificação. Dessa forma, a seleção de fatores com contribuição significativa é de suma importância, pois avalia a contribuição dos atributos de forma individual e associada, possibilitando remover as variáveis redundantes, altamente correlacionadas e irrelevantes, além de identificar os fatores condicionantes a ocorrência dos eventos, o que proporciona um melhor desempenho e eficiência do modelo de aprendizagem de máquina (KUHN; JOHNSON, 2013).

6.5. Importância das variáveis

A Figura 20 apresenta a correlação de Pearson entre as variáveis independentes com potencial de previsão dos eventos de incêndios e queimadas. Em suma, os valores do coeficiente de correlação próximos de ± 1 representam uma associação positiva/negativa quase perfeita, enquanto os valores próximos de 0 representam pouca ou nenhuma associação (ASUERO et al., 2006).

Figura 20. Correlação de Pearson entre as variáveis potenciais de previsão de incêndios e queimadas.



Com base na Figura 20, é possível identificar as variáveis que possuem alta correlação entre si, tais como observado entre SAVI e NDWI (0,92), NDVI e SAVI (0,87), NDVI e NDWI (0,81), e entre precipitação e temperatura (0,76). Dentre essas variáveis, apenas aquelas com maior contribuição serão selecionadas para compor o modelo. Dessa forma, a importância individual das 14 variáveis foi avaliada considerando o potencial de explicação da ocorrência de incêndios e queimadas pelo algoritmo Boruta.

Para identificação da Importância Relativa (IR), a variável dependente considerada foram os pontos identificados pelo mapeamento das cicatrizes de incêndios e queimadas mapeadas para o ano de 2019 (como verdadeiro positivo), e como verdadeiro negativo, uma distribuição randômica de pontos de não ocorrência de incêndios.

Com a aplicação do algoritmo Boruta, observou-se que todas as variáveis apresentaram contribuição significativa (a estatística individual das 14 variáveis originais correspondentes a essa análise encontra-se no Anexo C4), nesse caso, as variáveis que apresentaram alta correlação e baixo IR foram removidas, tais como o NDWI e SAVI, pois apresentaram IR inferior ao NDVI. Dessa forma, o NDWI e SAVI foram consideradas variáveis redundantes por possuírem equivalência funcional em decorrência de sua relação direta com a fotossíntese (VERBESSELT et al., 2006). Da mesma forma, observou-se que a temperatura apresentou um IR inferior ao da precipitação, nesse caso, optou-se pela remoção da variável temperatura do conjunto de dados. Em seguida, a importância individual do conjunto de dados restante foi verificada pelo algoritmo Boruta, sendo os resultados apresentados na Tabela 8.

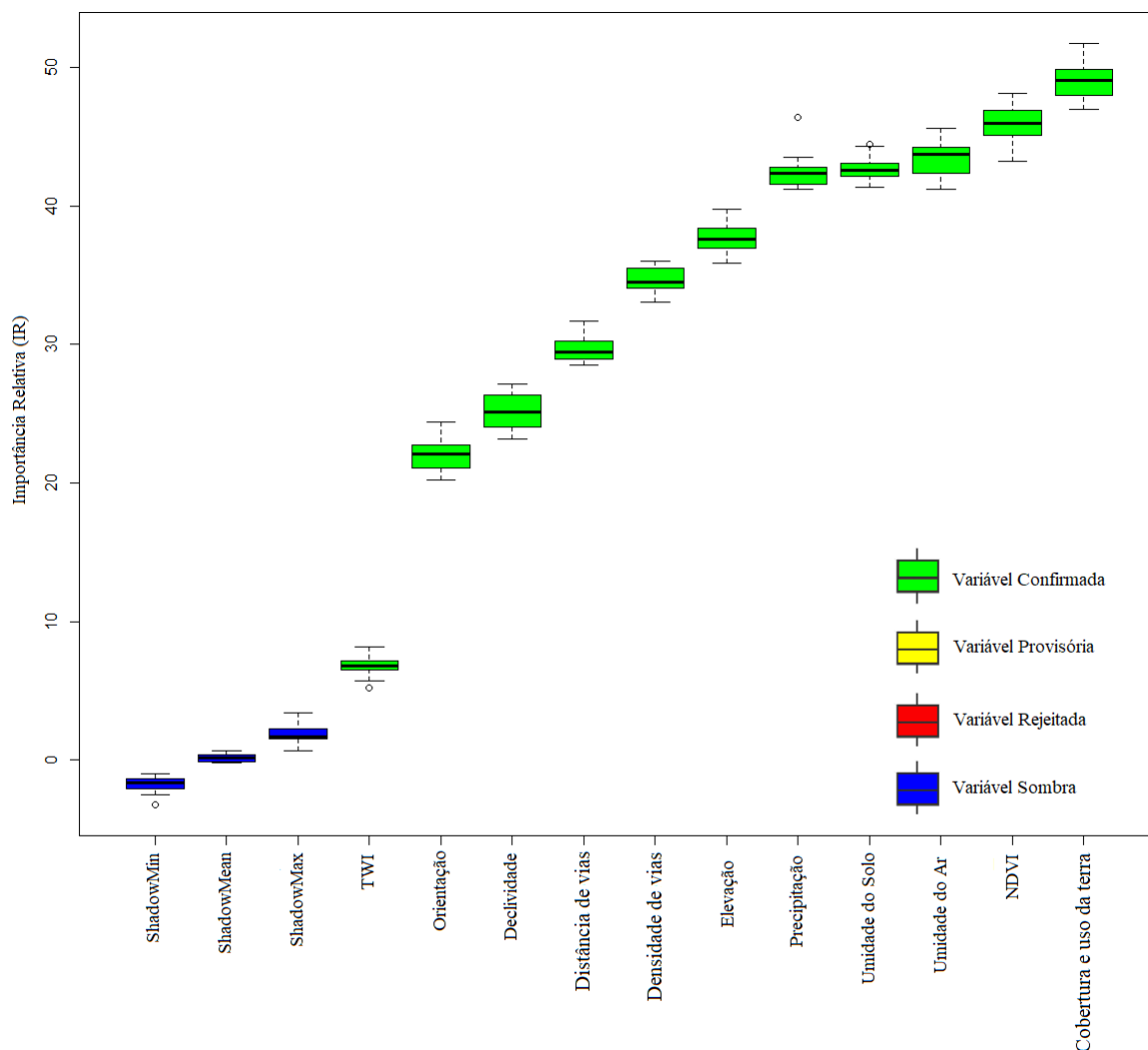
Tabela 8. Resultado do algoritmo Boruta para selecionar os fatores mais importantes com potencial de explicação dos eventos de incêndios e queimadas.

Variáveis	IRMean	IRMedian	IRMin	IRMax	Decisão
Cobertura e uso da terra	48,97	49,06	46,96	51,75	Confirmado
NDVI	45,87	45,98	43,23	48,12	Confirmado
Umidade do ar	43,34	43,69	41,17	45,57	Confirmado
Umidade do solo	42,69	42,57	41,33	44,44	Confirmado
Precipitação	42,54	42,33	41,18	46,41	Confirmado
Elevação	37,77	37,60	35,86	39,76	Confirmado
Densidade de vias	34,62	34,48	33,07	35,98	Confirmado
Distância entre vias	29,69	29,47	28,49	31,67	Confirmado
Declividade	25,14	25,11	23,15	27,15	Confirmado
Orientação	21,95	22,08	20,22	24,39	Confirmado
TWI	6,80	6,81	5,20	8,18	Confirmado

IRMean: Importância Relativa Média; **IRMedian:** Importância Relativa Mediana;
IRMin: Importância Relativa Mínima; **IRMax:** Importância Relativa Máxima;

A Figura 21 apresenta um gráfico *boxplot* da importância relativa das variáveis potenciais para explicação dos eventos de incêndios e queimadas na área de estudo identificado pelo algoritmo Boruta, onde os *boxplots* na cor azul corresponde aos z-scores mínimos, médios e máximos de uma variável sombra (*ShadowMin*, *ShadowMax*, *ShadowMean*), enquanto os *boxplots* nas cores verde indicam pontuações z-score das variáveis confirmadas, já as variáveis provisórias e rejeitadas seriam representadas por *boxplots* nas cores amarelo e vermelho, respectivamente (GHOLAMI et al., 2021), variáveis não observadas nesse estudo.

Figura 21. Importância relativa das variáveis potenciais para explicação dos eventos pelo método Boruta.



Como apresentado na Tabela 8 e Figura 21, todas as variáveis foram confirmadas como importante para explicação dos eventos de incêndios e queimadas. Seguindo o ranking de importância relativa, observa-se que as variáveis relacionadas às características da vegetação, tais como a cobertura e uso da Terra (IR = 48,97) e NDVI (IR = 45,87%) apresentaram as maiores contribuições, seguido pelos fatores climáticos representados pelas variáveis umidade do ar (43,34%), umidade do solo (IR = 42,69%) e precipitação (IR = 42,54%), e com menor contribuição os fatores relacionados a topografia, tais como elevação (IR = 37,77%), declividade (IR = 25,14%) e orientação (IR = 21,95%) e os fatores relacionados aos aspectos antrópicos, tais como a densidade de rodovias (IR = 34,62%) e proximidade de rodovias (IR = 29,69%).

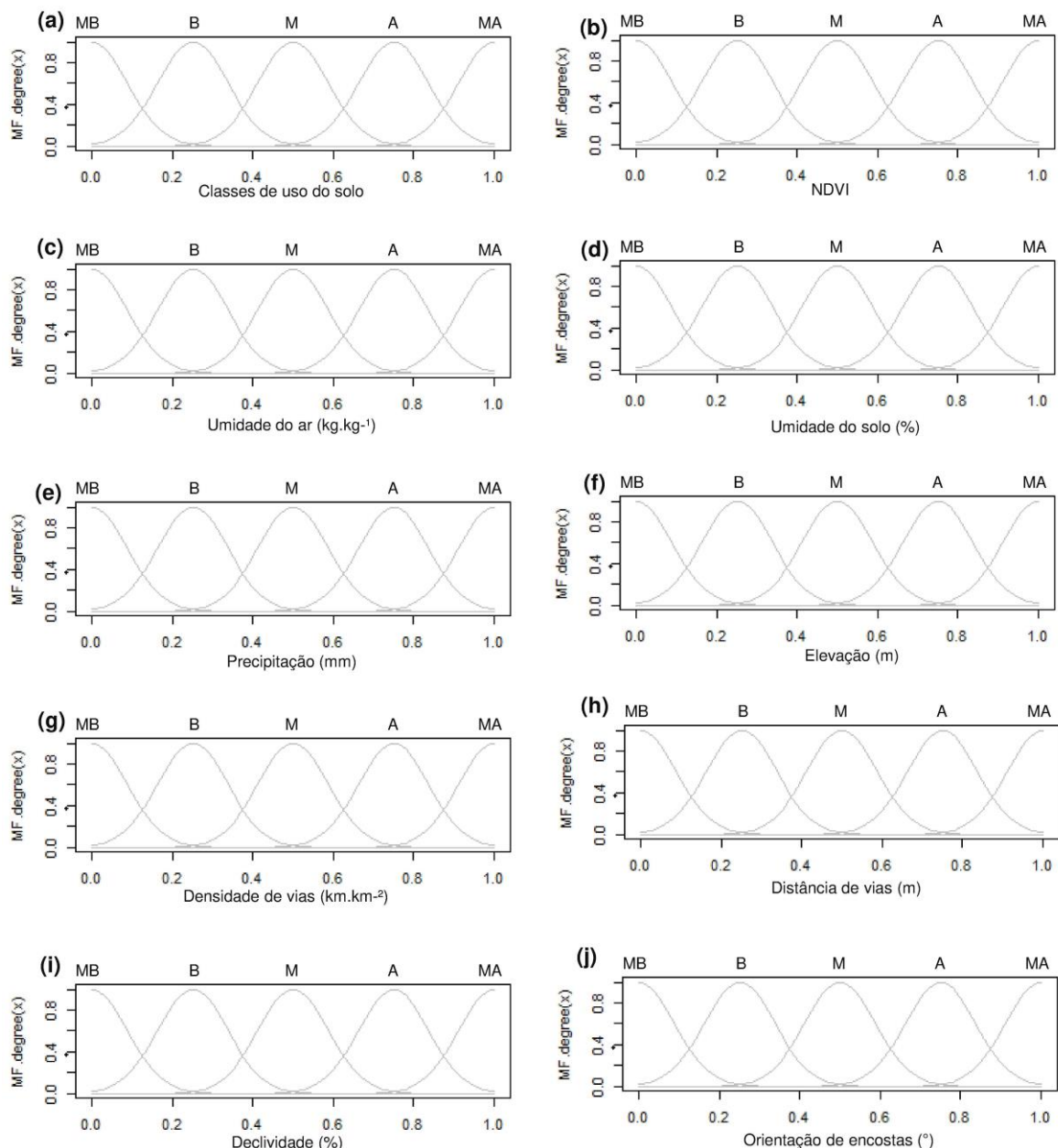
Por outro lado, o TWI apresentou o menor valor (IR = 6,80), se aproximando do desempenho observado para a variável sombra (*ShadowMax*) mais importante, ou seja, a IR

observado para o TWI se aproxima da casualidade no que se refere ao potencial de explicar os eventos de incêndios e queimadas, dessa forma, optou-se por não utilizar a variável para composição do modelo de mapeamento, pois apesar de ser confirmada como importante pelo método Boruta, o baixo IR da variável quando adicionada ao modelo pode representar maior complexidade e menor desempenho no modelo de predição (KUHN; JOHNSON, 2013). Dessa forma, as variáveis consideradas para compor o modelo de mapeamento foram as que apresentaram IR médio superior a 20,00, compreendendo as dez primeiras variáveis selecionadas pelo algoritmo Boruta (Tabela 8).

6.6. Previsão da Suscetibilidade de Incêndios e Queimadas (SQI)

Após a identificação e remoção das variáveis redundantes, e identificação das variáveis mais importantes, foi implementado o Sistema de Inferência Fuzzy Híbrido (HFIS). Foram testadas várias funções de associação fuzzy (triangular, trapezoidal, gaussiana e sigmoidal), sendo observado melhor desempenho o modelo com a implementação de funções de pertinência gaussiana (o resultado dos testes das diferentes funções de associações encontram-se no Anexo C5) com particionamento em cinco valores linguísticos (Muito Baixo, Baixo, Moderado, Alto, Muito Alto), sendo as variáveis normalizadas em intervalos de 0 a 1. A Figura 22 apresenta os valores linguísticos para cada variável utilizada neste estudo.

Figura 22. Valores linguísticos adotado para as variáveis de entrada.



(a) Uso e ocupação do solo; **(b)** NDVI; **(c)** umidade do ar; **(d)** umidade do solo; **(e)** precipitação; **(f)** elevação; **(g)** densidade de vias; **(h)** distância de vias; **(i)** declividade; **(j)** orientação de encostas. **MB:** Muito Baixo; **B:** Baixo; **M:** Moderado; **A:** Alto; **MA:** Muito Alto.

De modo geral, em um sistema fuzzy de partição do tipo grade, o número de regras ($n = a^p$) é dado pela combinação de valores linguísticos (a), ou seja, a quantidade de regras pode aumentar de forma exponencial em função do número de variáveis de entrada (p) (BRESSANE et al., 2020). Nesse estudo, o método Botura identificou 10 variáveis com relevância para composição do modelo de SIQ, a saber, a cobertura e uso da Terra, NDVI, umidade do ar, umidade do solo, precipitação, elevação, densidade de rodovias, distâncias

de rodovias, declividade e orientação, sendo cada variável de entrada foi normalizada em intervalo de 0 a 1, particionado em cinco valores linguísticos, o que resultaria em um sistema com 9.765.625 regras.

Entretanto, o método HFIS utilizados nesse estudo considera o grau de certeza fuzzy (CF) obtido por aprendizagem de máquina de forma automática, o que possibilita a construção de uma base de regras compacta, sendo que o grau de certeza identifica as regras que apresentam maior compatibilidade (mais próxima de 1), enquanto as regras redundantes e/ou com menor compatibilidade são excluídas do sistema (RIZA et al., 2015). Dessa forma, no presente trabalho, o uso de graus de certeza fuzzy (CF) possibilitou a construção de uma base de regras compacta, resultando em apenas 799 regras de tomada de decisão capazes de identificar as áreas com suscetibilidade a ocorrência de incêndio e queimadas, tais como:

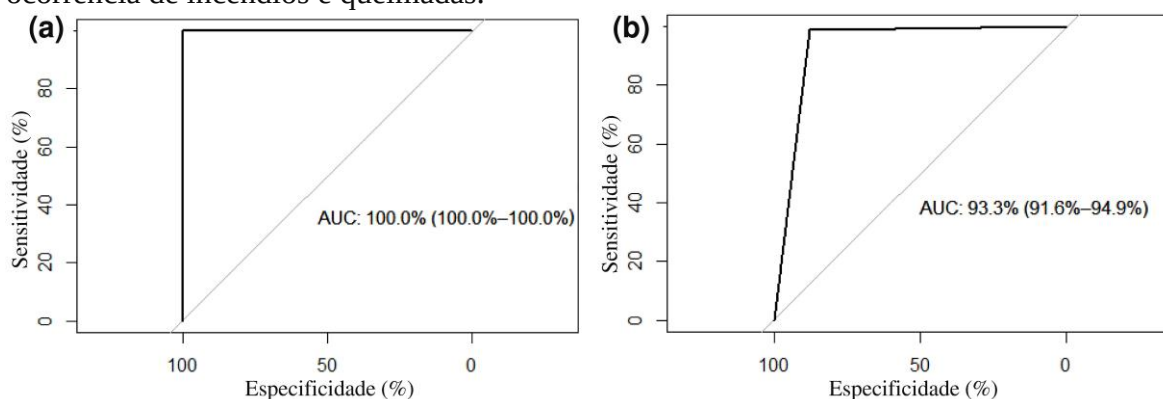
R1: *IF v1 is small and v2 is very.small and v3 is small and v4 is large and v5 is very.small and v6 is very.small and v7 is small and v8 is small and v9 is medium and v10 is small THEN Ocorrência Focos is 0; CF 0,658818507.*

R410: *IF v1 is medium and v2 is small and v3 is very.small and v4 is large and v5 is small and v6 is small and v7 is medium and v8 is large and v9 is small and v10 is medium THEN Ocorrência Focos is 1; CF 0,341181493.*

R799: *IF v1 is very.large and v2 is small and v3 is medium and v4 is large and v5 is very.small and v6 is small and v7 is medium and v8 is very.large and v9 is small and v10 is small THEN Ocorrência Focos is 1; CF 0,341181493.*

Onde os parâmetros v1 a v10 correspondem às variáveis independentes utilizadas nesse estudo e CF a certeza fuzzy para cada regra otimizada (demais regras obtidas pelo HFIS encontram-se no Anexo C6). Utilizando o método HFIS e considerando as variáveis identificadas pelo algoritmo Boruta, foi realizada a classificação dos eventos de incêndios e queimadas levando em consideração o conjunto de dados de treinamento (70%) e teste (30%) balanceado. A Figura 23 apresenta a curva ROC (*Receiver Operating Characteristic*) e AUC (*Area Under the Curve*) obtido pelo classificador HFIS com base em dados de treinamento (Figura 23a) e teste (Figura 23b) para classificação de eventos de incêndios e queimadas em 2019.

Figura 23. Curva ROC e AUC para os dados de treinamento (a) e teste (b) para previsão de ocorrência de incêndios e queimadas.



Observa-se que os resultados obtidos com base nos dados de treinamento apresentaram desempenho esperado, ou seja, as observações foram perfeitamente separáveis (Figura 23), conforme observado nos valores das métricas de desempenho (especificidade, sensibilidade, acurácia e índice Kappa), pois os dados utilizados para o treinamento do HFIS foram os mesmos utilizados para previsão, dessa forma, a curva ROC é representada por uma linha horizontal e uma vertical, e consequentemente, apresentando um AUC de 100% (Figura 23a). Pourghasemi et al. (2020) descrevem que a curva ROC construída com base em dados de treinamento não é um indicador adequado para avaliar o potencial de previsão do modelo, entretanto, os autores ressaltam que esses resultados podem ajudar a determinar o quão bem os mapas de suscetibilidade construídos pelos modelos combinam com os locais de incêndios e queimadas ocorridos anteriormente.

Por outro lado, considerando dados independentes (Figura 23b) foi possível avaliar o potencial de replicação do método HFIS, pois foi realizado com base em um conjunto de dados independentes não utilizados no ajuste do modelo. O resultado obtido mostrou que o modelo HFIS apresentou boa precisão na predição (Acurácia de 0,93 e índice Kappa de 0,86), pois a curva ROC se aproxima das linhas horizontais e verticais, apresentando um valor de AUC de 93,3%, se distanciando de predições aleatórias (quando AUC é igual a 0,5), o que significa que o classificador HFIS apresentou desempenho excelente tanto para os dados de treinamento quanto para os dados de teste.

A Tabela 9 apresenta as métricas de desempenho obtidas para os dados de treinamento e teste no ano de 2019, os dados sobre a matriz de erro encontram-se no Anexo C7.

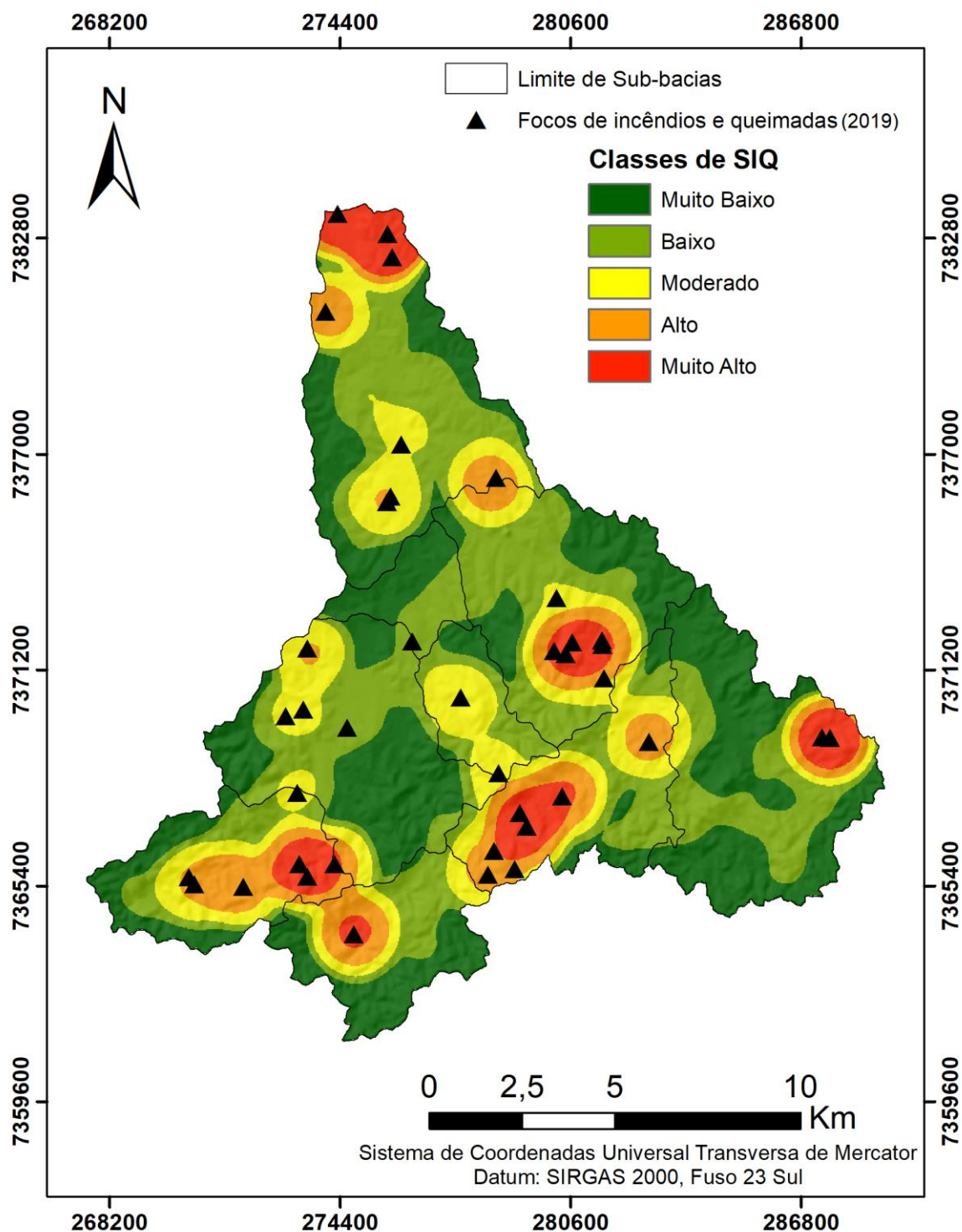
Tabela 9. Métricas de desempenho para os dados de treinamento e teste.

Métricas	Treinamento	Teste
Sensitividade	1,0	0,878
Especificidade	1,0	0,9873
Valor Preditivo Positivo	1,0	0,9863
Valor Preditivo Negativo	1,0	0,8858
Prevalência	0,5269	0,506
Taxa de detecção	0,5269	0,4483
Prevalência de detecção	0,5269	0,4545
Acurácia balanceada	1,0	0,9327
Acurácia	1,0	0,9315
Índice Kappa	1,0	0,8633

Utilizando diferentes técnicas de aprendizagem de máquina, Pourghasemi et al. (2020) observaram que o método de Análise Discriminante de Mistura (Mixture Discriminant Analysis - MDA) e árvore de regressão impulsionada (Boosted Regression Tree - BRT) apresentaram valores de AUC entre 82,5% a 88,90% com dados de treinamento e teste. Li et al. (2020) observaram que o método de Redes Neurais Artificiais (RNA) e Máquina de Vetores de Suporte (SVM) apresentaram valores de AUC variando entre 89,89% a 92,16%. Com uso de métodos conjuntos, Eskandari et al. (2021) observaram que o método GAM-MARS-SVM (*Generalized Additive Model - Multivariate Adaptive Regression Spline - Support Vector Machine*) apresentou AUC de 83,00%, valor superior aos modelos individuais utilizados pelos autores, da mesma forma que Mohajane et al. (2021), que observaram valor de AUC de 98,90% para o método RF-FR (*Random Forest-Frequency Ratio*). Em ambos os casos, os modelos com melhores ajustes (em geral com AUC superior a 80%) foram considerados satisfatórios e adequados para prever a suscetibilidade de incêndios e queimadas na área que foram propostos (POURGHASEMI et al., 2020; LI et al., 2020; ESKANDARI et al., 2021; MOHAJANE et al., 2021).

Com base no modelo ajustado, foi realizado a previsão dos eventos, em seguida, os pontos de ocorrência foram espacializados pelo método de densidade de Kernel considerando o número de ocorrências por km² (NASCIMENTO et al., 2017), posteriormente, foi realizado uma classificação considerando o nível de densidade observada por km², sendo discretizadas nas classes Muito Baixo, Baixo, Moderado, Alto e Muito Alto, representando o mapa de suscetibilidade de ocorrência de incêndios e queimadas, conforme descrito por Pourghasemi et al. (2020). A Figura 24 apresenta a distribuição espacial da Suscetibilidade de Incêndios e Queimadas-SIQ pelo método HFIS.

Figura 24. Mapa de SIQ obtido pelo método HFIS para o ano de 2019.



A Tabela 10 apresenta os valores de área (em hectares e percentual) para cada classe de Suscetibilidade de Incêndios e Queimadas - SQI na bacia hidrográfica, bem como o número de ocorrência dos eventos observados em 2019.

Tabela 10. Área e número de focos de incêndios e queimadas por classe de Suscetibilidade de Incêndios e Queimadas-SIQ em 2019.

Classes de SIQ	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	6.824,97	33,68	0
Baixo	6.536,09	32,25	2
Moderado	3.497,36	17,26	7
Alto	2.051,05	10,12	13
Muito Alto	1.356,22	6,69	17
Total	20.265,68	100	39

Observa-se que as áreas identificadas como Alto e Muito Alto SIQ prevista pelo método HFIS para o ano de 2019 ocorreram predominantemente a montante da bacia, principalmente na região sudoeste e sul, com demais pequenos núcleos também ocorrendo na região leste e região central, bem como no extremo norte (Figura 24). Por outro lado, as áreas classificadas como Muito Baixo SIQ encontram-se em locais com baixa densidade de rodovias, principalmente em áreas próximo ao divisor topográfico da bacia hidrográfica, região com declividade acentuada e normalmente ocupadas por vegetação florestal devido ao difícil acesso local. Já as áreas classificadas como Baixo SIQ ocorrem em toda extensão da bacia, distribuída predominantemente no entorno das principais rodovias, locais com terreno menos acidentado.

Em relação ao total de área ocupado por cada classe de SIQ, observa-se que o percentual de áreas classificadas entre Muito Baixo a Muito Alto SIQ é decrescente, enquanto o número de ocorrência é inverso, ou seja, à medida que a SIQ aumenta, o número de incêndios e queimadas também crescem (Tabela 10), relação que indica a adequabilidade do método proposto, conforme destacado por Pourghasemi et al. (2020) e Venkatesh et al. (2020).

Além do mais, considerando a classificação proposta por Venkatesh et al. (2020), a previsão da SIQ para 2019 apresenta uma acurácia de 76,92%, o que refere-se ao número de ocorrências de incêndios e queimadas em áreas classificadas como Alto e Muito Alto SIQ. Ressalta-se que o percentual ocupado por essas classes perfaz apenas 16,81% da área de estudo, entretanto, é justamente nessas áreas onde cerca de 76,92% dos eventos ocorreram. Por outro lado, apesar do modelo HFIS ter indicado uma predominância de áreas classificadas como Muito Baixo e Baixo SIQ (65,93%), nessas áreas foram identificados apenas 2 eventos, resultados que confirmam a eficiência do método proposto.

A Figura 25 apresenta a distribuição espacial da SIQ para as sub-bacias do rio Sorocabaçu e a Tabela 11 apresenta a área e o número de ocorrências de focos de incêndios e queimadas.

Figura 25. SIQ obtido pelo método HFIS para as sub-bacias do rio Sorocabaçu.

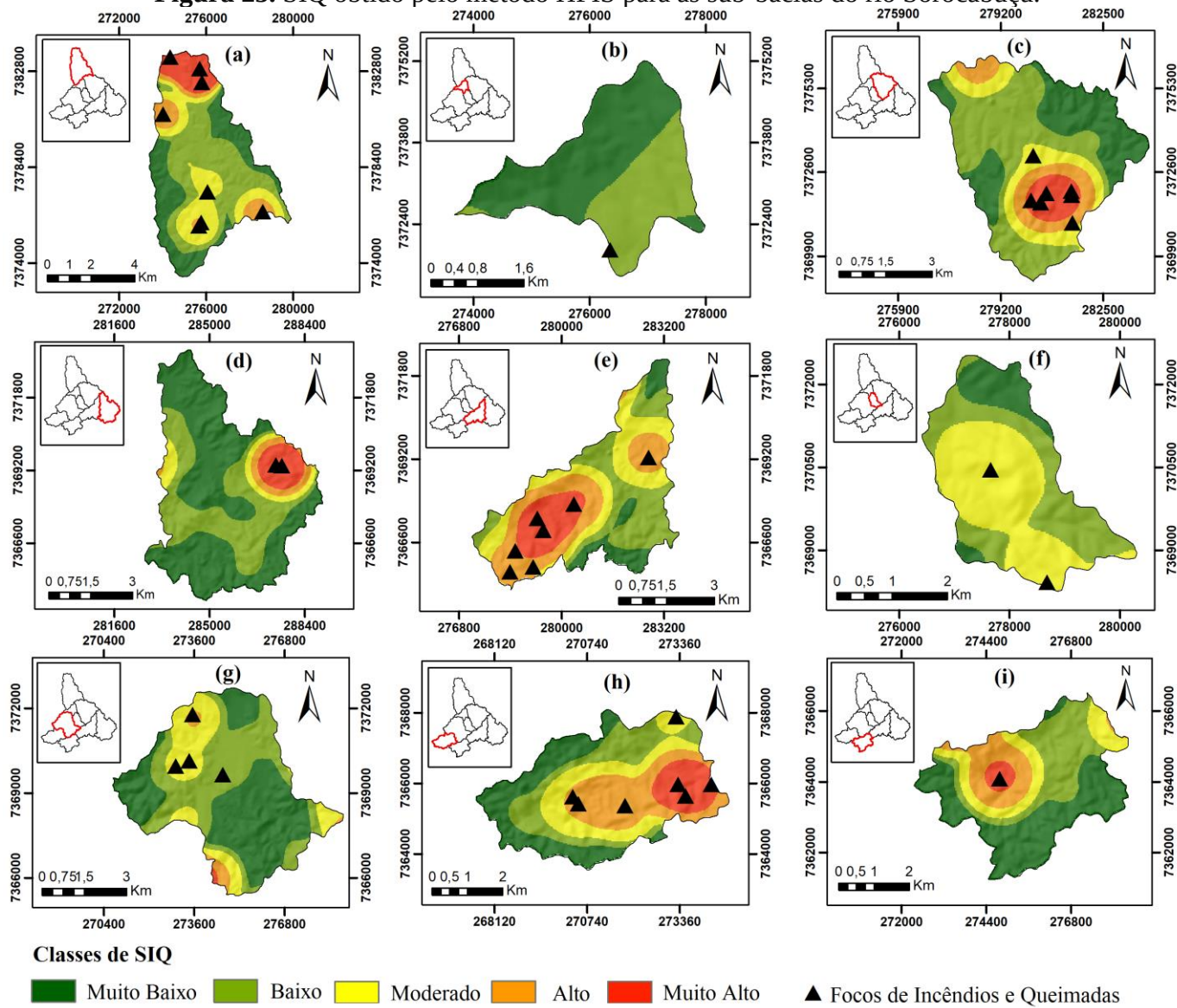


Tabela 11. Área e número de focos de incêndios e queimadas por classe de SIQ nas sub-bacias.

Sub-bacia 1 (a)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	946,93	25,40	0
Baixo	1.342,74	36,01	0
Moderado	796,11	21,35	1
Alto	307,31	8,24	4
Muito Alto	335,36	8,99	3
Sub-bacia 2 (b)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	404,92	61,36	0
Baixo	254,97	38,64	1
Sub-bacia 3 (c)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	805,30	27,78	0
Baixo	1124,61	38,79	0
Moderado	422,06	14,56	1
Alto	333,66	11,51	1
Muito Alto	213,55	7,37	5
Sub-bacia 4 (d)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	1784,83	54,85	0
Baixo	960,95	29,53	0
Moderado	169,65	5,21	0
Alto	123,91	3,81	0
Muito Alto	214,42	6,59	2
Sub-bacia 5 (e)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	215,89	10,02	0
Baixo	604,20	28,04	0
Moderado	501,07	23,26	0
Alto	526,12	24,42	4
Muito Alto	307,23	14,26	3
Sub-bacia 6 (f)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	117,88	14,01	0
Baixo	301,02	35,78	0
Moderado	422,47	50,21	2
Sub-bacia 7 (g)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	1372,63	44,45	0
Baixo	1090,74	35,32	1
Moderado	544,52	17,63	2
Alto	71,28	2,31	1
Muito Alto	9,01	0,29	0
Sub-bacia 8 (h)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	457,69	24,07	0
Baixo	404,73	21,29	0
Moderado	382,48	20,12	1
Alto	441,44	23,22	3
Muito Alto	214,76	11,30	3
Sub-bacia 9 (i)	Área (ha)	Área (%)	Nº de Focos
Muito Baixo	677,63	40,14	0
Baixo	446,37	26,44	0
Moderado	262,00	15,52	0
Alto	242,94	14,39	0
Muito Alto	59,37	3,52	1

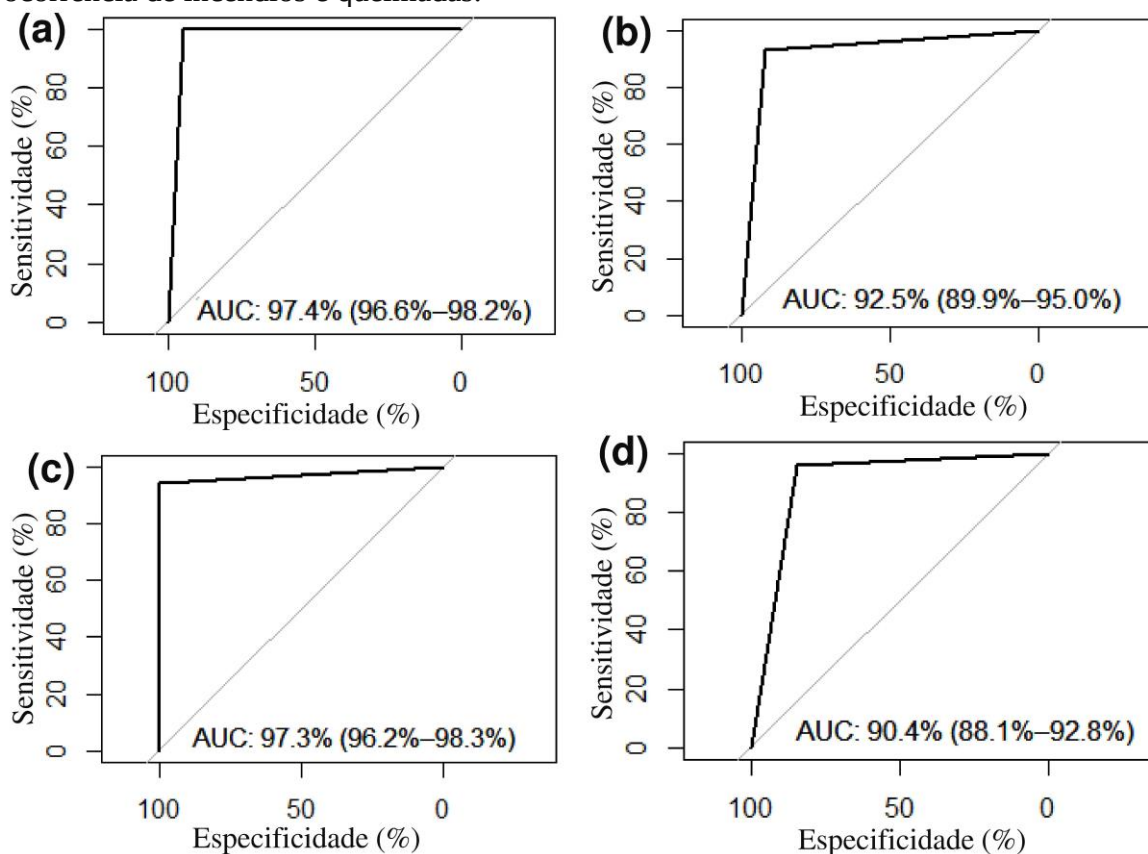
Conforme apresentado na Figura 25 e Tabela 11, observa-se que a sub-bacia 2 apresentou apenas as classes Muito Baixo e Baixo SIQ, sendo observado apenas uma ocorrência de incêndios e queimadas em áreas classificadas como Baixo SIQ. Em seguida, a sub-bacia 6, apresentando as classes Muito Baixo, Baixo e Moderado SIQ, contabilizando dois eventos, ambos em áreas classificadas com Moderado SIQ. Essas sub-bacias constituem as unidades de gerenciamento com a menor suscetibilidade de ocorrência de incêndios e queimadas observado em 2019. Por outro lado, a sub-bacia 5 (Figura 25e) 8 (Figura 25h) correspondem às unidades de gerenciamento mais críticas, sendo mais de 30% da área ocupadas por classe Alto e Muito Alto SIQ em ambas as sub-bacias. Nessas unidades ocorreram sete eventos de incêndios e queimadas, sendo que na sub-bacia 5 foram três ocorrências em áreas classificadas com Muito Alto SIQ e quatro ocorrências em áreas classificadas com Alto SIQ, enquanto na sub-bacia 8, também foram observadas sete ocorrências, entretanto, três delas foram identificadas em áreas classificadas como Muito Alto e Alto SIQ e uma ocorrência em área classificada como Moderado SIQ.

Com vista a verificar o potencial do modelo HFIS em períodos de ocorrências de eventos extremos, o modelo também foi implementado para o ano considerado moderadamente chuvoso (2018) e moderadamente seco (2020). Os dados médios observados entre os meses de maio a setembro utilizados como entrada (precipitação, umidade do ar, umidade do solo, NDVI), bem como a cobertura e uso da Terra retificada para os dois períodos com base em imagens do Sentinel-2 com resolução espacial de 30 m encontram-se no Anexos C8 a C10, os demais fatores (elevação, declividade, orientação, distância e densidade de rodovias) foram considerados estáticos, ou seja, na escala de tempo analisada as mudanças foram tão ínfimas ao ponto de serem considerados os valores mapeados no ano de 2019. A área de cicatrizes de incêndios e queimadas mapeada, bem como a distribuição espacial do conjunto de dados balanceado (verdadeiro positivo e verdadeiro negativo) para os dois períodos encontram-se no Anexo 11 e 12.

A Tabela 12 apresenta as métricas de desempenho observado para os eventos extremos, e a Figura 26 apresenta a curva ROC e os valores de AUC obtido pelo classificador HFIS com base em dados de treinamento (Figura 26a) e teste (Figura 26b) para classificação de eventos de incêndios e queimadas, dados sobre a matriz de erro encontra-se no Anexo C13.

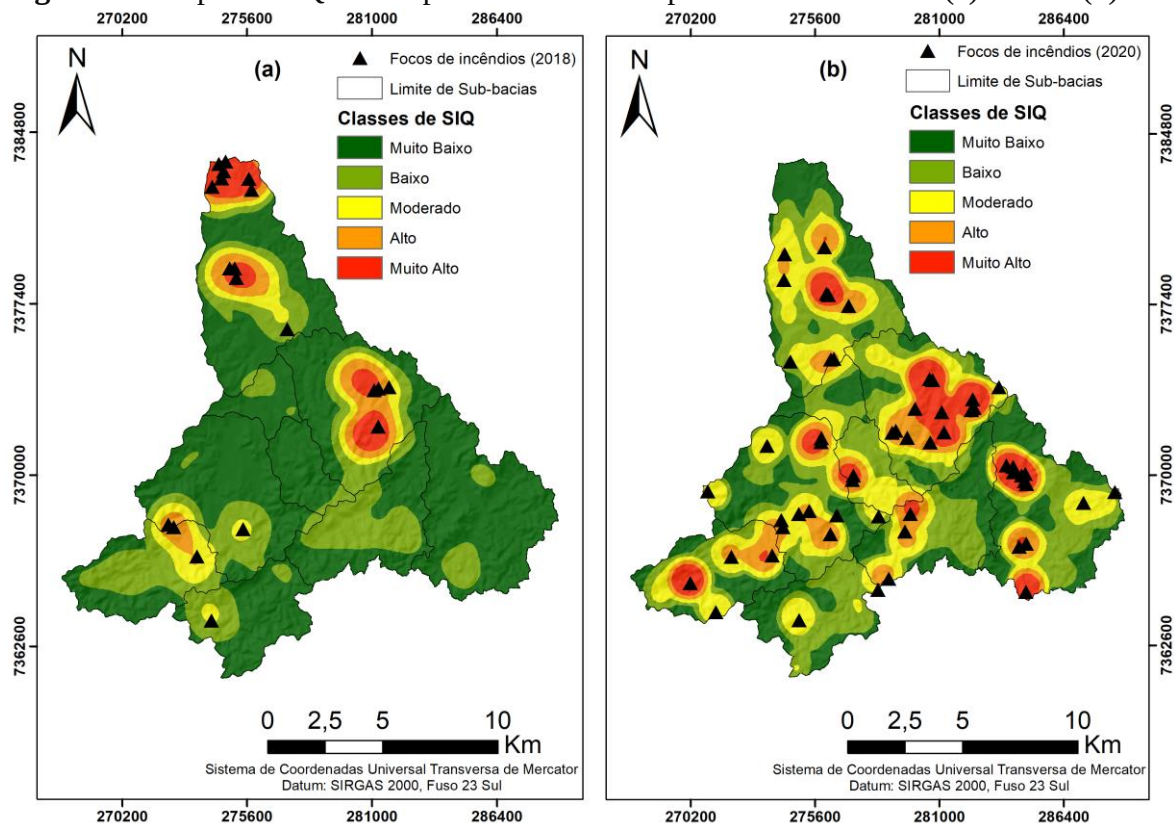
Tabela 12. Métricas de desempenho para os dados de treinamento e teste.

Métricas	Treinamento		Teste	
	2018	2020	2018	2020
Sensitividade	1,00	0,948	0,921	0,8471
Especificidade	0,945	1,00	0,928	0,9617
Valor Preditivo Positivo	0,952	1,00	0,934	0,9638
Valor Preditivo Negativo	1,00	0,9492	0,914	0,8395
Prevalência	0,524	0,5097	0,525	0,5461
Taxa de detecção	0,524	0,4834	0,484	0,4626
Prevalência de detecção	0,551	0,4843	0,51	0,4800
Acurácia balanceada	0,972	0,9743	0,924	0,9044
Acurácia	0,973	0,9738	0,924	0,8991
Índice Kappa	0,947	0,9475	0,848	0,799

Figura 26. Curva ROC e AUC para os dados de treinamento e teste para previsão de ocorrência de incêndios e queimadas.

*(a) treinamento (b) teste para o ano de 2018; (c) treinamento (d) teste para o ano de 2020

Observa-se que os valores de AUC para dados de treinamento e teste em ambos os períodos foram superiores a 90%, resultados que podem ser considerados satisfatórios (POURGHASEMI et al., 2020; MOHAJANE et al., 2021). A Figura 27 apresenta a distribuição espacial do mapeamento da SIQ pelo método HFIS para o período de evento moderadamente chuvoso (2018) e moderadamente seco (2020), respectivamente.

Figura 27. Mapa de SIQ obtido pelo método HFIS para os anos de 2018 (a) e 2020 (b).

A Tabela 13 apresenta a área em hectare e seus respectivos percentuais para cada classe de SIQ obtidos para o período de evento moderadamente chuvoso (2018) e moderadamente seco (2020), respectivamente.

Tabela 13. Área (hectares e percentual) e número de focos por classe de Suscetibilidade de Incêndios e Queimadas-SIQ para os anos de 2018 e 2020.

Classes de SIQ	Área (ha)		Área (%)		Nº de Focos	
	2018	2020	2018	2020	2018	2020
Muito Baixo	12.289,09	5.329,41	60,64	26,30	0	0
Baixo	4.991,10	6.291,22	24,63	31,05	1	3
Moderado	1.236,78	4.664,16	6,10	23,02	2	8
Alto	967,18	2.362,25	4,77	11,66	5	19
Muito Alto	781,20	1.616,61	3,85	7,98	12	26
Total	20.265,36	20.263,67	100,00	100,00	20	56

A previsão da SIQ pelo método HFIS para o período moderadamente chuvoso (2018) e moderadamente seco (2020) mostra que, no ano de anomalia positiva (Figura 27a) as áreas previstas com SIQ Alto e Muito Alto foram menores, já no período de anomalia negativa (Figura 27b) as classes de SIQ Alto e Muito Alto foram mais expressivas, isso pode ser explicado em virtude da ocorrência de menores índices de precipitação no ano de 2020 quando comparado aos observados em 2018. Os menores índices de precipitação observado

em 2020 influenciaram de forma direta na SIQ na área de estudo, uma vez que com o menor volume precipitado, maior será o grau de ressecamento da vegetação, e por consequência, aumenta a suscetibilidade de ocorrência de incêndios e queimadas (POURGHASEMI et al., 2020; VENKATESH et al. 2020; RAKHMATULINA et al., 2021).

Em relação ao total de área ocupado por cada classe de SIQ (Tabela 13), observa-se que para o período moderadamente chuvoso (2018), o percentual de áreas classificadas com SIQ Muito Baixo a Muito Alto é decrescente, enquanto o número de ocorrência de focos de incêndios e queimadas é crescente entre essas classes, ou seja, à medida que a SIQ aumenta, o número de incêndios e queimadas também cresce. No período moderadamente seco (2020), observou-se uma predominância de áreas com SIQ Baixo, seguido de Muito Baixo, Moderado, Alto e Muito Alto, sendo observado os maiores números de ocorrências de incêndios nas classes de maior SIQ, assim como observado em 2018.

No ano moderadamente chuvoso o método implementado atingiu um percentual de acerto de 85% enquanto no ano seco esse percentual foi de 80,35%, valores consideravelmente superiores ao sugerido por White et al. (2016), que estabelece um percentual de acerto igual ou superior a 60% para que um modelo de suscetibilidade seja válido. Vale ressaltar que no ano considerado moderadamente chuvoso o volume médio de precipitação observado entre os meses de maio a setembro foi de 59,92 mm, enquanto no ano considerado moderadamente seco, a precipitação média observada no mesmo período foi de 34,44 mm, o que reflete consideravelmente na inflamabilidade do material combustível, não por acaso que no ano moderadamente seco as ocorrências de incêndios e queimadas foram 180% superior aos eventos observados no ano moderadamente chuvoso.

Entre 2018, 2019 e 2020 as áreas abandonadas (área vegetacional perturbada) foi a classe de cobertura e uso da Terra mais atingida pelos eventos de incêndios e queimadas (57,14% em 2018, 44,89% em 2019, e 36,36% em 2020) na bacia hidrográfica, e em menor proporção áreas ocupadas por silvicultura, pastagem e cultura temporária. As áreas ocupadas por silvicultura, pastagem e cultura temporária podem passar por um processo de manejo que muitas vezes é utilizado o fogo para a limpeza do terreno, caracterizando-se como queima controlada (ICMBio, 2010; PMAPA, 2012; SOUSA, 2021), entretanto, estes podem servir como fonte de ignição para o início de incêndios em áreas abandonadas ou com demais usos na bacia hidrográfica.

Em suma, os resultados obtidos com a implementação do método HFIS foram superiores quando comparado aos resultados observados em trabalhos recentes desenvolvidos ao nível de bacia hidrográfica que utilizaram métodos de aprendizagem de

máquina tradicional (POURGHASEMI et al., 2020; LI et al., 2020; ESKANDARI et al., 2021; MOHAJANE et al., 2021). Como a ocorrência de incêndios e queimadas são eventos dinâmicos no espaço e tempo, e dependentes de vários fatores (APONTE et al., 2016; BUI et al., 2017; CHUVIECO et al., 2019), o modelo proposto pode ser utilizado de forma contínua para previsão da SIQ na bacia hidrográfica, visto que os dados utilizados são disponibilizados quase em tempo real, o que possibilita prever as áreas com Alto e Muito Alto SIQ com antecedência a ocorrência desses eventos, possibilitando indicar as áreas prioritárias onde os esforços devem ser concentrados.

Dessa forma, os resultados obtidos neste estudo podem auxiliar os gestores públicos a implementar medidas mais efetivas de combate e prevenção em áreas prioritárias, tais como com campanhas educativas, construção de aceiros, intensificação de fiscalização, definição de estratégia e alocação de brigadistas, para evitar e minimizar a ocorrência desses eventos nos locais identificados como prioritários. A implementação de ações efetivas pode minimizar a fragmentação de habitats, degradação do solo e diminuir a emissão de gases de efeito estufa, contribuindo com os Objetivos de Desenvolvimento Sustentável (ODS) proposto pela Organização das Nações Unidas (ONU, 2015).

Vale ressaltar que, apesar dos bons resultados obtidos com a utilização de métodos de aprendizagem de máquina para previsão de incêndios e queimadas nos últimos anos, os resultados ainda têm sido subutilizados para sua gestão (BUI et al., 2019; TANG et al., 2020; TONINI et al., 2020). Para Jain et al. (2020), um desafio importante é permitir a transição desses resultados para a operação de gestão desses eventos, sendo frequentemente agravado pela falta de recursos e/ou financiamento, bem como pelas diferentes prioridades e culturas institucionais de pesquisadores e órgãos responsáveis pela gestão. Uma possível solução seria a formação de grupos de trabalho dedicados a viabilizar essa transição, o que possibilitaria a integração e colaboração entre a comunidade de pesquisa e os agentes responsáveis pela gestão desses eventos (JAIN et al., 2020; POURGHASEMI et al., 2020).

7. CONCLUSÕES E RECOMENDAÇÕES

O desenvolvimento deste estudo possibilitou implementar uma metodologia para a previsão de áreas com suscetibilidade de incêndios e queimadas a partir de dados de sensoriamento remoto com uso do método HFIS. A metodologia adotada pode ser replicada em outras áreas, visto que os dados utilizados são disponibilizados de forma gratuita quase em tempo real, e os resultados obtidos podem auxiliar os gestores na tomada de medidas

eficazes, possibilitando prevenir e antecipar as ações de combate de incêndios e queimadas, minimizando as perdas em caso de ocorrência desses eventos.

A análise das variáveis potencialmente explicativas de incêndios e queimadas realizada com uso do algoritmo Boruta indicou que das 14 variáveis utilizadas, apenas 10 foram utilizadas para compor o modelo de previsão, pois três variáveis foram removidas por serem consideradas redundantes (NDWI, SAVI e temperatura) e uma por possuir potencial de explicação próximo a casualidade (TWI). Quanto a importância das variáveis restantes, o algoritmo indicou que os fatores relacionados a características da vegetação, tais como o uso e ocupação do solo e NDVI foram as mais importantes, seguido pelos fatores climáticos (umidade do ar, umidade do solo e precipitação), antrópicos (distância e densidade de rodovias) e topográficos (elevação, declividade e orientação), respectivamente.

A implementação do método HFIS para prever a suscetibilidade de incêndios e queimadas demonstrou que a função de associação fuzzy gaussiana com particionamento em cinco valores linguísticos (Muito Baixo, Baixo, Moderado, Alto, Muito Alto) apresentou os melhores resultados. O método implementado considerando o grau de certeza fuzzy (CF) obtido por aprendizagem de máquina possibilitou a construção de uma base de regras compacta (799 regras) capazes de identificar as áreas com suscetibilidade a ocorrência de incêndio e queimadas.

Os resultados obtidos mostraram que o modelo HFIS apresentou boa precisão na predição para o ano de 2019 (Acurácia de 0,93 e índice Kappa de 0,86), apresentando um valor de AUC de 93,3%, o que significa que o classificador HFIS apresentou desempenho excelente quando avaliado considerando dados independentes, o que representa boa capacidade de replicação do modelo proposto. Os resultados mostraram valores de AUC superiores a 90% para os períodos climáticos extremos (0,97-0,94 de Acurácia-Kappa para 2018 e 0,89-0,79 para a Acurácia-Kappa em 2020), resultados que podem ser considerados satisfatórios.

A espacialização da SIQ previsto com base no modelo HFIS ajustado mostraram que em 2019, as classes com menor suscetibilidade (Muito Baixo e Baixo) são predominantes (65,93%) na bacia hidrográfica, enquanto as áreas com maior suscetibilidade (Muito Alto e Alto) são minoritárias (16,18%), e apesar disso, é neste último onde ocorrem cerca de 76,69% dos eventos de incêndios e queimadas, enquanto nas classes de menor suscetibilidade concentram apenas 5,12% dos eventos, resultados que confirmam a eficiência do método proposto e sua capacidade em modelar problemas não lineares complexos.

A previsão da SIQ mostrou que em 2019, as áreas identificadas como Alto e Muito Alto SIQ pelo método HFIS ocorreram predominantemente a montante da bacia, principalmente na região sudoeste e sul, com demais pequenos núcleos também ocorrendo na região leste e região central, bem como no extremo norte. Já as áreas identificadas como Muito Baixo SIQ encontram-se em locais com baixa densidade de rodovias, principalmente em áreas próximo ao divisor topográfico da bacia hidrográfica, região com declividade acentuada e normalmente ocupadas por vegetação florestal devido ao difícil acesso local.

Em suma, este estudo demonstrou que é possível prever a suscetibilidade de incêndios e queimadas com uso do método HFIS. O modelo implementado possibilitou elaborar mapas de suscetibilidade, indicando os locais com maior predisposição de ocorrência do evento em função de fatores climáticos, antrópicos, topográficos e características da vegetação. Os resultados obtidos podem auxiliar os gestores públicos no planejamento para a mitigação dos eventos, principalmente em períodos de seca quando o número de ocorrências são maiores, visto que possibilita direcionar medidas de gerenciamento em áreas que apresentaram Alta e Muito Alta SIQ, tais como a construção de aceiros preventivos, construção de estradas de acesso em pontos estratégicos e a reorganização de práticas de manejo do solo, delineamento de áreas prioritárias para a intensificação de fiscalização, alocação de equipes de brigadistas, bem como a organização de ações no combate de incêndios.

Como recomendações e sugestões para estudos futuros:

- I) Considerando que os eventos de incêndios e queimadas ocorrem em função de características intrínsecas de cada região, a importância das variáveis deve ser avaliada em cada caso, não sendo considerado o nível de importância observado neste estudo.
- II) A disponibilidade de dados sobre os fatores climáticos com melhor resolução espacial é de suma importância, uma vez que esses fatores são preponderantes para ocorrência desses eventos, dessa forma, a compreensão das variações em escala microclimática desses fatores podem melhorar a acurácia do modelo de previsão da SIQ.
- III) Em relação ao particionamento do conjunto de dados, demais métodos de amostragem aqui não implementados em função do baixo poder computacional podem ser testados, com vista a minimizar o tamanho reduzido do conjunto de dados de treinamento e teste.
- IV) Em estudos futuros, sugere-se utilizar diferentes métodos de amostragem para os casos de verdadeiro negativo para identificar a distribuição mais eficiente, possibilitando melhorar o desempenho do modelo implementado, e consequentemente, a eficiência da classificação das áreas com suscetibilidade de incêndios e queimadas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAM, N.J.; HENLEY, B.J.; SEN GUPTA, A. Connections of climate change and variability to large and extreme forest fires in southeast Australia. **Commun Earth Environ**, v. 2, n. 8. 2021. <https://doi.org/10.1038/s43247-020-00065-8>
- ADAB, H.; KANNIAH, K. D.; SOLAIMANI, K. Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. **Natural Hazards**, v. 65, p. 1723–1743. 2013. <https://doi.org/10.1007/s11069-012-0450-8>
- ADIRI, Z.; LHISSOU, R.; HARTI, A. E.; JELLOULI, A.; MOHCINE, C. Recent advances in the use of public domain satellite imagery for mineral exploration: A review of Landsat-8 and Sentinel-2 applications. **Ore Geology Reviews**. v. 117, 103332. 2020. <https://doi.org/10.1016/j.oregeorev.2020.103332>
- AMIRI, M.; POURGHASEMI, H. R.; GHANBARINAN, G. A.; AFZALI, S. F. Assessment of the importance of gully erosion effective factors using Boruta algorithm and its spatial modeling and mapping using three machine learning algorithms. **Geoderma**, v. 340, p. 55-69. 2019. <https://doi.org/10.1016/j.geoderma.2018.12.042>
- AMORIM, A. T.; SOUSA, J. A. P.; LOURENÇO, R. W. Indicador dos Estágios de Sucessão de Fragmentos Florestais do Bioma Mata Atlântica. **Revista Brasileira de Cartografia**, v. 71, n. 3, p. 756 -780, 2019. <https://doi.org/10.14393/rbcv71n3-48546>
- ANDRADE, C. F.; DELGADO, R. C.; BARBOSA, M. L. F.; TEODORO, P. E.; JUNIOR, C. A. S.; WANDERLEY, H. S.; SILVA, G. F. C. Fire regime in Southern Brazil driven by atmospheric variation and vegetation cover. **Agricultural and Forest Meteorology**, v. 295, 2020, <https://doi.org/10.1016/j.agrformet.2020.108194>
- ANDRADE, S. C.; FERREIRA, A. F. Mapeamento Geoecológico da susceptibilidade a ocorrência de incêndios no Parque Estadual da Serra da Concórdia – Valença RJ. **Revista Teccen**. v. 12, n. 2, p. 45-58. 2019.
- APONTE, C.; GROOT, W. J.; WOTTON, B. M. Forest fires and climate change: causes, consequences and management options. **International Journal of Wildland Fire**, v. 25, n. 8, 2016. https://doi.org/10.1071/WFv25n8_FO
- AQUILUÉ, N.; FORTIN, M-J.; MESSIER, C.; BROTONS, L. The Potential of Agricultural Conversion to Shape Forest Fire Regimes in Mediterranean Landscapes. **Ecosystems**. v. 23. p. 34-51. 2020. <https://doi.org/10.1007/s10021-019-00385-7>
- ARAGÃO, L.; ANDERSON, L.O.; FONSECA, M.G.; ROSAN, T.M.; VEDOVATO, L. B.; WAGNER, F.H.; SILVA, C. V. J.; SILVA JUNIOR, C. H. L.; ARAI, E.; AGUIAR, A. P.; BARLOW, J.; BERENGUER, E.; DEETER, M. N.; DOMINGUES, L. G.; GATTI, L.; GLOOR, M.; MALHI, Y. MARENGO, J. A.; MILLER, J. B.; PHILLIPS, O. L.; SAATCHI, S. 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions. **Nature Communications**, p. 1–12. 2018. <https://doi.org/10.1038/s41467-017-02771-y>

ASUERO, A. G.; SAYAGO, A.; GONZÁLES, A. G. The correlation Coefficient: An Overview. **Critical Reviews in Analytical Chemistry**, v. 36. p. 41–59. 2006. <https://doi.org/10.1080/10408340500526766>

BADIA, A.; PALLARES-BARBERA, M.; VALLDEPERAS, N.; GISBERT, M. Wildfires in the wildland-urban interface in Catalonia: Vulnerability analysis based on land use and land cover change. **Science of The Total Environment**, v. 673, p. 184-196. 2019. <https://doi.org/10.1016/j.scitotenv.2019.04.012>

BARMPOUTIS, P.; PAPAIOANNOU, P.; DIMITROPOULOS, K.; GRAMMALIDIS, N. A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing. **Sensors**, v. 20, n 22. 2020. <https://doi.org/10.3390/s20226442>

BEGUERÍA, S.; SERRANO, S. M. V. 2017. SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index. R package version 1.7. <https://CRAN.R-project.org/package=SPEI>

BERLINCK, C. N.; BATISTA, E. K. L. Good fire, bad fire: It depends on who burns. **Flora**, v. 268, 2020. <https://doi.org/10.1016/j.flora.2020.151610>

BIVAND, R.; KEITT, T.; ROWLINGSON, B. 2020. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-16. <https://CRAN.R-project.org/package=rgdal>

BONAZOUNTAS, M.; KALLIDROMITOU, D.; KASSOMENOS, P. A.; PASSAS, N. Forest Fire Risk Analysis. **Human and Ecological Risk Assessment: An International Journal**, v. 11, n. 3. 2005. p. 617-626. <https://doi.org/10.1080/10807030590949717>

BRESSANE, A.; SILVA, P. M.; FIORE, F. A.; CARRA, T. A.; EWBANK, H.; DE-CARLI, B. P.; MOTA, M. T. Fuzzy-based computational intelligence to support screening decision in environmental impact assessment: A complementary tool for a case-by-case project appraisal. **Environmental Impact Assessment Review**, v. 85. 2020. <https://doi.org/10.1016/j.eiar.2020.106446>

BUI, D. T.; BUI, Q. T.; NGUYEN, Q. P.; PRADHAN, B.; H. NAMPAK, H.; TRINH, P. T. A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. **Agricultural and Forest Meteorology**, v. 233, p. 32-44, 2017. <https://doi.org/10.1016/j.agrformet.2016.11.002>

BUI, D. T.; HOANG, N-D.; SAMUI, P. Spatial pattern analysis and prediction of forest fire using new machine learning approach of Multivariate Adaptive Regression Splines and Differential Flower Pollination optimization: A case study at Lao Cai province (Viet Nam). **Journal of Environmental Management**, v. 237, p. 476-487. 2019. <https://doi.org/10.1016/j.jenvman.2019.01.108>

CAMARGO, L. S.; SILVA, R. W.; AMARAL, S. S.; SILVA, A. P.; FERRELI, A. P.; SILVA, M. P. D. Mapeamento de Áreas Suscetíveis a Incêndios Florestais do Município

de Petrópolis – RJ. **Anuário do Instituto de Geociências – UFRJ**. v. 42, n. 1, p. 630-641. 2019. http://dx.doi.org/10.11137/2019_1_630_641

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in central Iowa soils. **Soil Science Society of America Journal**, v. 58, n. 1, p. 1501–1511. <https://doi.org/10.2136/sssaj1994.03615995005800050033x>

CANDANEDO, L. M.; FELDHEIM, V.; DERAMAIX, D. Data driven prediction models of energy use of appliances in a low-energy house. **Energy and Buildings**, v. 140, p. 81-97. 2017. <https://doi.org/10.1016/j.enbuild.2017.01.083>

CAO, J.; ZHANG, Z.; DU, J.; ZHANG, L.; SONG, Y.; SUN, G. Multi-geohazards susceptibility mapping based on machine learning—a case study in Jiuzhaigou, China. **Natural Hazards**, v. 102, p. 851-871, 2020. <https://doi.org/10.1007/s11069-020-03927-8>

CARMO, M.; MOREIRA, F.; CASIMIRO, P.; VAZ, P. Land use and topography influences on wildfire occurrence in northern Portugal. **Landscape and Urban Planning**. v. 100. n. 1-2. 2011. p. 169-176. <https://doi.org/10.1016/j.landurbplan.2010.11.017>

CAÚLA, R. H.; OLIVEIRA-JÚNIOR, J. F. LYRA, G. B.; DELGADO, R. C.; FILHO, P. F. L. Overview off fire foci causes and localitions in Brazil based on meteorological satellite data from 1998 to 2011. **Enviromental Earth Science**. v. 74. p. 1497-1508. 2015. <https://doi.org/10.1007/s12665-015-4142-z>

CHEN, C.; GARIBALDI, J.; RAZAK, T. FuzzyR: Fuzzy Logic Toolkit for R. R package version 2.3.1. 2021. <https://CRAN.R-project.org/package=FuzzyR>

CHEN, F.; NIU, S.; TONG, X.; ZHAO, J.; SUN, Y.; HE, T. The Impact of Precipitation Regimes on Forest Fires in Yunnan Province, Southwest China. **The Scientific World Journal**, v. 2014. p. 1-9. 2014. <https://doi.org/10.1155/2014/326782>

CHUVIECO, E.; MOUILLOT, F.; WERF, G. R. V.; MIGUEL, J. S.; TANASE, M.; KOUTSIAS, N.; GARCÍA, M.; YEBRA, M.; PADILLA, M.; GITAS, L.; HEIL, A.; HAWBAKER, T. J.; GIGLIO, L. Historical background and current developments for mapping burned area from satellite Earth observation. **Remote Sensing of Environment**, v. 225, p. 45-64. 2019. <https://doi.org/10.1016/j.rse.2019.02.013>

CHUVIECO, E.; AGUADO, I.; SALAS, J.; GARCÍA, M.; YEBRA, M.; OLIVA, P. Satellite Remote Sensing Contributions to Wildland Fire Science and Management. **Current Forestry Reports**. v. 6, p. 81-96. 2020. <https://doi.org/10.1007/s40725-020-00116-5>

CIABURRO, G.; VENKATESWARAN, B. Neural Networks with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles. Packt>. BIRMINGHAM – MUMBAI, 2017. 314p.

CIIAGRO. Centro Integrado de Informações Agrometeorologias. Disponível em: <http://www.ciiagro.sp.gov.br/#>, Acesso em: 16/09/2021.

CORRÊA, C. V. S. A Geotecnologia no monitoramento, alerta e prevenção de desastres – os desastres naturais e tecnológicos no cenário brasileiro e estudos de caso do emprego de geotecnologias como subsídios a sua mitigação. **Revista Ciência Geográfica**. p. 390-417. 2020.

COSTAFREDA-AUMEDES, S.; COMAS, C.; VEJA-GARCIA, C. Human-caused fire occurrence modelling in perspective: a review. **International Journal of Wildland Fire**. v. 26. n. 12. p. 938-998, 2017. <https://doi.org/10.1071/WF17026>

DAS, N. N.; ENTEKHABI, D.; DUNBAR, R. S.; CHAUBELL, M. J.; COLLIANDER, A.; YUEH, S.; JAGDHUBER, T.; CHEN, F.; CROW, W.; O'NEILL, P. E.; WALKER, J. P.; BERG, A.; BOSCH, D. D.; CALDWELL, T.; COSH, M. H.; COLINS, C. H.; LOPEZ-BEAZA, E.; THIBEAULT, M. The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. **Remote Sensing of Environment**, v. 233, 2019. <https://doi.org/10.1016/j.rse.2019.111380>

DAS, N., ENTEKHABI, D.; DUNBAR, R. S.; KIM, S.; YUEH, S.; COLLIANDER, A.; O'NEILL, P. E.; JACKSON, T.; JAGDHUBER, T.; CHEN, F.; CROW, W. T.; WALKER, J.; BERG, A.; BOSCH, D.; CALDWELL, T.; COSH, M. SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Soil Moisture, Version 2. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018. <https://doi.org/10.5067/KE1CSVXMI95Y>

DIAS, G. F. Queimadas e incêndios florestais: cenários e desafios: subsídios para a educação ambiental. – Brasília: MMA, IBAMA, 2008. 32p

DONG, X.; LI, F.; HARRISON, S. P.; CHEN, Y.; KUG, J. S. Climate influence on the 2019 fires in Amazonia. **Science of The Total Environment**, v. 794, 2021. <https://doi.org/10.1016/j.scitotenv.2021.148718>

DRISCOLL, D. A.; ARMENTERAS, D.; BENNETT, A. F.; BROTONS, L. How fire interacts with habitat loss and fragmentation. **Ecosystem, Frontiers in Ecology and Evolution**, v. 9, 2021. <https://doi.org/10.3389/fevo.2021.647557>

EK, M.B.; MITCHELL, K.E.; LIN, Y.; ROGERS, E.; GRUNMANN, P.; KOREN, V.; GAYNO, G.; TARPLEY, J.D. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. **Journal of Geophysical Research Atmospheres**, v. 108, n. D22. 2003. <https://doi.org/10.1029/2002JD003296>

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Sistema brasileiro de classificação de solos. Brasília, Serviço de Produção de Informação, 2018. 355p.

ENTEKHABI, D.; DAS, N.; NJOKU, E.; YUEH, S.; JOHNSON, J.; SHI, J. Soil Moisture Active Passive (SMAP) - Algorithm Theoretical Basis Document L2 & L3 Radar/Radiometer Soil Moisture (Active/Passive) Data Products. California Institute of Technology, p. 89, 2014.

EO Browser - Earth Observation Browser. Sentinelhub. European Space Agency (ESA). Disponível em: <https://www.sentinel-hub.com/explore/eobrowser/>, acessado em: 10/05/2021.

ESA – European Space Agency. User Guides - Sentinel-2 MSI Introducion. Disponível em: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi>, Acessado em: 01/06/2021.

ESKANDARI, S.; POURGHASEMI, H. R.; TIEFENBACJER, J. P. Fire-susceptibility mapping in the natural areas of Iran using new and ensemble data-mining models.

Environmental Science and Pollution Research, v. 28, p. 47395–47406, 2021.

<https://doi.org/10.1007/s11356-021-13881-y>

ESRI - ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. **ArcGis 10.5**. 2016.

ESRI - Environmental Systems Research Institute. How line density work. Disponível em: <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-line-density-works.htm>, Acessado em: 24/03/2021.

EUGENIO, F. C.; SANTOS, A. R.; FIELDLER, N. C.; RIBEIRO, G. A.; SILVA, A. G.; Santos, Á. B.; Paneto, G. G.; Schettino, V. R. Applying GIS to develop a model for forest fire risk: A case study in Espírito Santo, Brazil. **Journal of Environmental Management**, v. 173, p. 65-71. 2016. <https://doi.org/10.1016/j.jenvman.2016.02.021>

FAO - Food and Agriculture Organization of the United Nations. Forest fire and the law: A guide for nationl drafters based on the Fire Management Voluntary Guidelines. **United Nations**, Rome, 2009. 175p.

FERREIRA, M. V. M. Avaliação da relação entre incêndios e precipitação nas unidades de conservação Parque Estadual do Biribiri e Parque Estadual do Rio Preto. **Caderno Prudentino de Geografia**, n. 39, v. 2 p. 118-140, 2017.

FINNEY, M. A. The challenge of quantitative risk analysis for wildland fire. **Forest Ecology and Management**, v. 211, n. 1–2, p. 97-108. 2005. <https://doi.org/10.1016/j.foreco.2005.02.010>

FLANNIGAN, M.; CANTIN, A. S.; GROOT, W. J.; WOTTON, M.; NEWBERY, A.; GOWMAN, L. M. Global wildland fire season severity in the 21st century. **Forest Ecology and Management**, v. 294, p. 54-61. 2013. <https://doi.org/10.1016/j.foreco.2012.10.022>

FONSECA, M. G.; ALVES, L. M.; AGUIAR, A. P.; ARAI, E.; ANDERSON, L. O.; ROSAN, T. M.; SHIMABUKURO, Y. E.; ARAGÃO, O, L. E. C. Effects of climate and land-use change scenarios on fire probability during the 21st century in the Brazilian Amazon. **Global Change Biology**, 2019. <https://doi.org/10.1111/gcb.14709>

GADALHA, A. N.; COELHO, V. H. R.; XAVIER, A. C.; BARBOSA, L. R.; MELO, D. C. D.; XUAN, Y.; HUFFMAN, G. J.; PETERSEN, W. A.; ALMEIDA, C. N. Grid box-level evaluation of IMERG over Brazil at various space and time scales. **Atmospheric Research**, v. 218, p. 231-244. 2019. <https://doi.org/10.1016/j.atmosres.2018.12.001>

GAO, B.C. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, v. 58, p. 257-266, 1996. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)

GHOLAMI, H.; MOHAMMADIFAR, A.; GOLZARI, S.; KASKAOUTIS, D. G.; COLLINS, A. L. Using the Boruta algorithm and deep learning models for mapping land susceptibility to atmospheric dust emissions in Iran. **Aeolian Research**, v. 50, 2021. <https://doi.org/10.1016/j.aeolia.2021.100682>

GOMES, L.; MIRANDA, H. S.; SILVÉRIO, D. V.; BUSTAMANTE, M. M. C. Effects and behaviour of experimental fires in grasslands, savannas, and forests of the Brazilian Cerrado. **Forest Ecology and Management**, v. 458, 2020. <https://doi.org/10.1016/j.foreco.2019.117804>

GOOGLE LLC. Google Earth Pro. Versão 7.3.3. 2020. Disponível em: <https://www.google.com.br/earth/download/gep/agree.html>, Acessado em: 24/03/2020.

GRAINGER, M. J.; ARAMYAN, L.; PIRAS, S.; QUESTED, T. E.; RIGHI, S.; SETTI, M.; VITTUARI, M.; STEWART, G. B. Model selection and averaging in the assessment of the drivers of household food waste to reduce the probability of false positives. **Plos One**, 2018. <https://doi.org/10.1371/journal.pone.0192075>

GRALEWICZ, N. J.; NELSON, T. A.; WULDER, M. A. Factors influencing national scale wildfire susceptibility in Canada. **Forest Ecology and Management**. v. 265, n. 1. 2012. p. 20-29. <https://doi.org/10.1016/j.foreco.2011.10.031>

GUEDES, B. J.; MASSI, K. G.; ERVS, C.; NIELSEN-PINCUS, M. Vulnerability of small forest patches to fire in the Paraíba do Sul River Valley, southeast Brazil: Implications for restoration of the Atlantic Forest biome. **Forest Ecology and Management**, v. 465, 2020. <https://doi.org/10.1016/j.foreco.2020.118095>

HANES, C.; WOTTON, M.; WOOLFORD, D. G.; MARTELL, D.; FLANNIGAN, M. Preceding Fall Drought Conditions and Overwinter Precipitation Effects on Spring Wildland Fire Activity in Canada. **Fire**, v. 3, n. 2, 2020. <https://doi.org/10.3390/fire3020024>

HARDY, C. C. Wildland fire hazard and risk: Problems, definitions, and context. **Forest Ecology and Management**, v. 211, n. 1-2, p. 73-82, 2005. <https://doi.org/10.1016/j.foreco.2005.01.029>

HIJMANS, R. J. raster: Geographic Data Analysis and Modeling. R package version 3.3-13. 2020. <https://CRAN.R-project.org/package=raster>

HOSSEINI, M.; POWELL, M.; COLLINS, J.; CALLAHAN-FLINTOFT, C.; JONES, W.; BOWMAN, H.; WYBLE, B. I tried a bunch of things: The dangers of unexpected overfitting in classification of brain data. **Neuroscience & Biobehavioral Reviews**, v. 119, p. 456-467. 2020. <https://doi.org/10.1016/j.neubiorev.2020.09.036>

- HUETE, A.R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v.25, p.295-309, 1988. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- HUFFMAN, G. J.; BOLVIN, D. T.; BRAITHWAITE, D.; HSU, K.; JOYCE, R.; KIDD, C.; NELKIN, E. J.; SOOSHIAN, S.; TAN, J.; XIE, P. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). **National Aeronautics and Space Administration**, 39p. 2020.
- HUFFMAN, G. J.; STOCKER, E. F.; BOLVIN, D. T.; NELKIN, E.J.; TAN, J. GPM IMERG Final Precipitation L3 1 month 0.1 degree x 0.1 degree V06, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), 2019. Accessed: [Data Access Date], <https://doi.org/10.5067/GPM/IMERG/3B-MONTH/06>
- IBGE – Instituto Brasileiro de Geografia e Estatística. Manual Técnico de Uso da Terra. 3ª Edição. Rio de Janeiro. 2013. 171p.
- IEMBO, J. L. V.; GALVANI, E. Vulnerabilidade a incêndios florestais: reflexes acerca do Parque Estadual do Juquery e seu entorno. **Revista Brasileira de Climatologia**, v. 28. p. 74-100. 2021.
- IGC - INSTITUTO GEOGRÁFICO E CARTOGRÁFICO. Ortofoto Leste de SP. 2010. Disponível em: <https://datageo.ambiente.sp.gov.br/>, Acessado em: 01/03/2021.
- IGC - INSTITUTO GEOGRÁFICO E CARTOGRÁFICO. **Carta topográfica**. Serviço Gráfico do IGC, 1979. Escala 1:10.000.
- INPE – Instituto Nacional de Pesquisas Espaciais. Preguntas Frequentes/Queimadas. 2021. Disponível em: <https://queimadas.dgi.inpe.br/queimadas/portal/informacoes/perguntas-frequentes#p40>, Acessado em: 14/07/2021.
- IPCC - Intergovernmental Panel on Climate Change. Impacts of 1.5°C Global Warming on Natural and Human Systems. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Waterfield (eds.)). In Press, 2018. 138p. Disponível em: <https://www.ipcc.ch/sr15/chapter/chapter-3/>
- ISDR -International Strategy for Disaster Reduction. 2009 UNISDR Terminology on Disaster Risk Reduction. **United Nations**, Geneva, 2009. 35p.
- ISHIBUCHI, H.; NAKASHIMA, T. Effect of rule weights in fuzzy rule-based classification systems. **IEEE Transactions on Fuzzy Systems**, v. 9, n. 4, p. 506-515, 2001. <https://doi.org/10.1109/91.940964>
- JAIN, P.; COOGAN, S. C. P.; SUBRAMANIAN, S. G.; CROWLEY, M.; TAYLOR, S.; FLANNIGAN, M. D. A review of machine learning applications in wildfire science and management. **Environmental Reviews**, v. 28, n. 4. 2020. <https://doi.org/10.1139/er-2020-0019>

JAISWAL, R. K.; MUKHERJEE, S.; RAJU, K. D.; SAXENA, R. Forest fire risk zone mapping from satellite imagery and GIS. **International Journal of Applied Earth Observation and Geoinformation**. v. 4, n. 1, 2002. p. 1-10.
[https://doi.org/10.1016/S0303-2434\(02\)00006-5](https://doi.org/10.1016/S0303-2434(02)00006-5)

JESUS, J. B.; ROSA, C. N.; BARRETO, I. D. C.; FERNANDES, M. M. Análise da incidência temporal, espacial e de tendência de fogo nos biomas e unidades de conservação do Brasil. **Ciência Florestal**. v. 30 n. 1. p. 176-191, 2020.
<https://doi.org/10.5902/1980509837696>

JIMÉNEZ, F.; MARTÍNEZ, C.; MARZANO, E.; PALMA, J. T.; SÁNCHEZ, G.; SCIAVICCO, G. Multiobjective Evolutionary Feature Selection for Fuzzy Classification. **IEEE Transactions on Fuzzy Systems**, v. 27, n. 5. p. 1085-1099. 2019.
<https://doi.org/10.1109/TFUZZ.2019.2892363>

JÚNIOR, L. A. S.; DELGADO, R. C.; PEREIRA, M. G.; TEODORO, P. E.; HUNIOR, C. A. S. Fire dynamics in extreme climatic events in western amazon. **Environmental Development**, v. 32, 2019. <https://doi.org/10.1016/j.envdev.2019.06.005>

KEY, C. H.; BENSON, N. C. Landscape assessment: Sampling and analysis methods USDA Forest Service General Technical Report RMRS-GTR-164-CD, 2006.

KNORR, W., KAMINSKI, T., ARNETH, A., AND WEBER, U. Impact of human population density on fire frequency at the global scale, **Biogeosciences**, v. 11, p. 1085–1102, 2014. <https://doi.org/10.5194/bg-11-1085-2014>

KOHAVI, R. AND JOHN, G. H. Wrappers for feature selection. *Artificial Intelligence*, v. 97, n. 1-2, p. 273–324, 1997. [https://doi.org/10.1016/S0004-3702\(97\)00043-X](https://doi.org/10.1016/S0004-3702(97)00043-X)

KUHN, M. caret: Classification and Regression Training. R package version 6.0-86. 2020.
<https://CRAN.R-project.org/package=caret>

KUHN, M.; JOHNSON, K. Applied Predictive Modeling, Springer, New York, 2013. Ed. 1. 615p. <https://doi.org/10.1007/978-1-4614-6849-3>

KURSA, M. B.; RUDNICKI, W. R. Feature Selection with the Boruta Package. **Journal of Statistical Software**, v. 36, n. 11, p. 1-13. 2010. <http://www.jstatsoft.org/v36/i11/>

LAWRENCE, M. G. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. **Bulletin of the American Meteorological Society**. v. 86, n. 2, p. 225-234. 2005.
<https://doi.org/10.1175/BAMS-86-2-225>

LE, T. T.; URBANOWICZ, R. J.; MOORE, J. H.; MCKINNEY, B. A. Statistical Inference Relief (STIR) feature selection. **Bioinformatics**, v. 35, n. 8, p. 1358–1365, 2019.
<https://doi.org/10.1093/bioinformatics/bty788>

LEI, S.; CHEN, H.; BIAN, Z.; LIU, Z. Evaluation of integrating topographic wetness index with backscattering coefficient of TerraSAR-X image for soil moisture estimation in

a mountainous region. **Ecological Indicators**, v. 61, p. 624-633. 2016.
<https://doi.org/10.1016/j.ecolind.2015.10.013>

LEUENBERGER, M.; PARENTE, J.; TONINI, M.; PEREIRA, M. G.; KANEVSKI, M. Wildfire susceptibility mapping: Deterministic vs. stochastic approaches. **Environmental Modelling & Software**, v. 101, p. 194-203. 2018.
<https://doi.org/10.1016/j.envsoft.2017.12.019>

LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. Classification evaluation. **Nature Methods**, v. 13, p. 603-604, 2016. <https://doi.org/10.1038/nmeth.3945>

LI, Y.; FENG, Z.; CHEN, S.; ZHAO, Z.; WANG, F. Application of the Artificial Neural Network and Support Vector Machines in Forest Fire Prediction in the Guangxi Autonomous Region, China. **Discrete Dynamics in Nature and Society**. v. 2020. p. 1-14. 2020. <https://doi.org/10.1155/2020/5612650>

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

LIU, H.; COCEA, M. Semi-random partitioning of data into training and test sets in granular computing context. *Granular Computing*. V. 2, p. 357-386. 2017.
<https://doi.org/10.1007/s41066-017-0049-2>

LOPES, E. R. N.; SILVA, A. P. P.; PERUCHI, J. F.; LOURENÇO, R. W. Zoneamento de risco de incêndio e queimadas no município de Sorocaba – São Paulo. **Revista do Departamento de Geografia**. v. 36. p. 117-129. 2018.
<https://doi.org/10.11606/rdg.v36i0.148048>

LOURENÇO, R. W.; SILVA, D. C. C.; MARTINS, A. C. G.; SALES, J. C. A.; ROVEDA, S. R. M.; ROVEDA, J. A. F. Use of fuzzy systems in the elaboration of an anthropic pressure indicator to evaluate the remaining forest fragments. **Environmental Earth Sciences**, v. 74, p. 2481-2488, 2015. <https://doi.org/10.1007/s12665-015-4253-6>

MACHADO, J. A.; LOPES, A. M. Analysis of Forest Fires by means of Pseudo Phase Plane and Multidimensional Scaling Methods. **Mathematical Problems in Engineering**. v. 2014. p. 1-8. <http://dx.doi.org/10.1155/2014/575872>

MACHADO, N. G.; SILVA, F. C. P.; BIUDES, M. S. Efeito das condições meteorológicas sobre o risco de incêndio e o número de queimadas urbanas e focos de calor em Cuiabá-MT, Brasil. **Ciência e Natura**, v. 36 n. 3. p. 459-469, 2014.

MAMDANI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis whit a fuzzy logic controller. **International Journal of Man-Machine Studies**. v. 7, n. 1, p. 1-13. 1975.
[https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2)

MARLON, J. R.; BARTLEIN, P. J.; DANIAU, A. L.; HARRISON, S. P.; MAEZUMI, S. Y.; POWER, M. J.; TINNER, W.; VANNIÉRE, B. Global biomass burning: a synthesis and review of Holocene paleofire records and their controls. **Quaternary Science Reviews**, v. 65, p. 5-25. 2013. <https://doi.org/10.1016/j.quascirev.2012.11.029>

MARTIN, N. A. Avaliação do potencial de suprimento energético a partir de energia fotovoltaica e resíduos sólidos urbanos, para municípios do Estado de São Paulo. Tese (Doutorado em Ciências Ambientais) – Universidade Estadual Paulista, Sorocaba, 2021. 96p.

MARTINS, A. P.; ROSA, R. Procedimentos Metodológicos para Estimativa de Variáveis Relacionadas ao Balanço de Energia no Sistema Terra-Atmosfera – Uma Revisão de Literatura. **Caderno Prudentino de Geografia**, n. 41, v. 1, p. 100-119, 2019.

MCKEE, T.B.; DOESKEN, N. J.; KLEIST, J. The relationships of drought frequency and duration to time scales. In: CONFERENCE ON APPLIED CLIMATOLOGY, 8, Proceedings... Boston: **American Meteorological Society**, 1993, p. 233-236.

MCNALLY, A. NASA/GSFC/HSL, FLDAS Noah Land Surface Model L4 Global Monthly 0.1 x 0.1 degree (MERRA-2 and CHIRPS), Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), 2018. Accessed: [16/04/2021], <https://doi.org/10.5067/5NHC22T9375G>

MENEZES, G. S.; CAZETTA, E.; DODONOV, P. Vegetation structure across fire edges in a Neotropical rain forest. **Forest Ecology and Management**, v. 453, 2019. <https://doi.org/10.1016/j.foreco.2019.117587>

MICHAEL, Y.; HELMAN, D.; GLICKMAN, O.; GABAY, D.; BRENNER, S.; LENSKY, I. M. Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series. **Science of The Total Environment**, v. 764, 2021. <https://doi.org/10.1016/j.scitotenv.2020.142844>

MILLER, J. D.; THODE, A. E. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). **Remote Sensing of Environment**, v. 109, n. 1, p. 66-80. 2007. <https://doi.org/10.1016/j.rse.2006.12.006>

MKHABELA, M. S.; BULLOCK, P.; RAJ, S.; WANG, S.; YANG, Y. 2011. Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. **Agricultural and Forest Meteorology**, v. 151, n. 3, p. 385-393. <https://doi.org/10.1016/j.agrformet.2010.11.012>

MOAYEDI, H. MEHRABI, M.; BUI, D. T.; PRADHAN, B.; FOONG, L. K. Fuzzy-metaheuristic ensembles for spatial assessment of forest fire susceptibility. **Journal of Environmental Management**, v. 260. 2020. <https://doi.org/10.1016/j.jenvman.2019.109867>

MOHAJANE, M.; COSTACHE, R.; KARIMI, F.; PHAM, Q. B.; ESSAHLAOUI, A.; NGUYEN, H.; LANEVE, G.; OUDIJA, F. Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean área. **Ecological Indicators**, v. 129, 2021. <https://doi.org/10.1016/j.ecolind.2021.107869>

MOTA, P. H. S.; ROCHA, S. J. S. S.; CASTRO, N. L. M.; FRANÇA, L. C. J.; SCHETTINI, B. L. S.; VILLANOVA, P. H.; SANTOS, H. T.; SANTOS, A. R. Forest fire hazard zoning in Mato Grosso State, Brazil. **Land Use Policy**. v. 88, p. 1-6. 2019. <https://doi.org/10.1016/j.landusepol.2019.104206>

NAPOLI, M. D.; CAROTENUTO, F.; CEVASCO, A.; CONFUORTO, P.; MARTIRE, D. D.; FIRPO, M.; PEPE, G.; RASO, E.; CALACATERRA, D. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. **Landslides**, v. 17, p. 1897–1914, 2020. <https://doi.org/10.1007/s10346-020-01392-9>

NASA - National Aeronautics and Space Administration. DAAC - Land Processes Distributed Active Archive Center LP. SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Soil Moisture V003. Disponível em: <https://search.earthdata.nasa.gov/search>, Acessado em: 14/04/2021.

NASCIMENTO, O. S.; SOUZA, I. R.; NASCIMENTO, W. P. S.; NASCIMENTO, J. S. Análise espaço-temporal dos focos de calor no município de Parnaíba Piauí, entre 2008 a 2015. **Revista Brasileira de Gestão Ambiental**. v.11, n.1, p.193 -199. 2017.

NJOKU, E. G.; ENTEKHABI, D. Passive microwave remote sensing of soil moisture, **Journal of Hydrology**, v. 184, p. 101-129, 1996.

NÓBREGA, L. O.; LAZZARINI, G. J.; VIOLA, M. R.; BATISTA, A. C.; CARVALHO, E. V. GIONGO, M. Forest Fire susceptibility index for assessing the history of fire occurrences in the indigenous land of Krahôlândia, Brazil. **Advances in Forestry Science**, v. 5, n. 2. p. 325-332, 2018.

NUNES, J. R. S.; BEUTLING, A.; KOPROSKI, L. P.; MELO, L. A. N.; BIONDI, D.; BATISTA, A. C. Relação entre a qualidade da paisagem e o risco de incêndios florestais. **FLORESTA**, v. 38, n. 1, p. 145- 154, 2008.

ONU - United Nations. General Assembly, Transforming Our World: the 2030 Agenda for Sustainable Development. 2015. 35p. Disponível em: <http://www.refworld.org/docid/57b6e3e44.html>, Acessado em: 30 de setembro de 2021.

OSM Foundation. 2021. OpenStreetMap Foundation. United Kingdom. Disponível em: <https://www.openstreetmap.org/copyright>

PADOVANNI, N. G.; LOPES, E. R.; SOUZA, J. C.; MORAIS, M. C. M.; LOURENÇO, R. W. Análise Espacial da Temperatura e Albedo de Superfície na Bacia Hidrográfica do rio Una – Ibiúna/São Paulo, a partir de imagens MODIS. **Revista Brasileira de Geografia Física**, v.11, n.5, p. 1832-1845. 2018.

PARENTE, J.; PEREIRA, M. G. 2016. Structural fire risk: the case of Portugal. **Science of the Total Environment**. v. 573, n. 15. 2016. p. 883-893. <https://doi.org/10.1016/j.scitotenv.2016.08.164>

PEBESMA, E. J. Multivariable geostatistics in S: the gstat package. **Computers & Geosciences**, v. 30, p. 683-691. 2004.

PMCRMAI – Plano Municipal de Conservação e Recuperação da Mata Atlântica de Ibiúna. Disponível em: <https://ibiuna.sp.gov.br/download/meio-ambiente-plano-municipal-de-conservacao-e-recuperacao-da-mata-atlantica/>, Acesso em: 17/09/2021.

PMAPA – Plano de Manejo de Área de Proteção Ambiental (APA) Itupararanga. Fundação Florestal. Disponível em: <https://www.infraestruturameioambiente.sp.gov.br/fundacaoflorestal/planos-de-manejo>, Acessado em: 17/09/2021.

POURGHASEMI, H. R.; GAYEN, A.; LASAPONARA, R.; TIEFENBACHER, J. P. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. **Environmental Research**. v. 184, p. 1-12. 2020. <https://doi.org/10.1016/j.envres.2020.109321>

POURGHASEMI, H. R.; KARIMINEJAD, N.; AMIRI, M.; EDALAT, M.; ZARAFSHAR, M.; BLASCHKE, T.; CERDA, A. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. **Scientific reports**. v. 10, 2020. <https://doi.org/10.1038/s41598-020-60191-3>

PRASAD, P.; LOVESON, V. J.; DAS, B.; KOTHA, M. Novel ensemble machine learning models in flood susceptibility mapping. **Geocarto International**. 2021. <https://doi.org/10.1080/10106049.2021.1892209>

QUEVEDO, R. P.; OLIVEIRA, G. G.; GUASSELLI, L. A. Mapeamento de Suscetibilidade a Movimentos de Massa a partir de Redes Neurais Artificiais. **Anuário do Instituto de Geociências – UFRJ**. v. 43, n.2, p. 128-138, 2020.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. <https://www.R-project.org/>

RAKHMATULINA, E.; STEPHENS, S.; THOMPSON, S. Soil moisture influences on Sierra Nevada dead fuel moisture content and fire risks. **Forest Ecology and Management**, v. 496, 2021. <https://doi.org/10.1016/j.foreco.2021.119379>

RIBEIRO, L. SOARES, R. V.; BEPLER, M. Mapeamento do risco de incêndios florestais no município de Novo Mundo, Mato Grosso, Brasil. **CERNE**, v. 18, n. 1. p. 117-12, 2012. <http://dx.doi.org/10.1590/S0104-77602012000100014>

RIBEIRO, T. M.; MENDONÇA, B. A. F.; OLIVEIRA-JÚNIOR, J. F.; FERNANDES-FILHO, E. I. Fire foci assessment in the Western Amazon (2000–2015). **Environment, Development and Sustainability**, 2020. <https://doi.org/10.1007/s10668-020-00632-1>

RIZA, L. S.; BERGMEIR, C.; HERRERA, F.; BEN'ITEZ, J. M. frbs: Fuzzy Rule-Based Systems for Classification and Regression in R. **Journal of Statistical Software**, v. 65, n. 6, p. 1-30. 2015.

ROMÁN-FLORES, H.; CHALCO-CANO, Y.; FIGUEROA-GARCÍA, J. C. A note on defuzzification of type-2 fuzzy intervals. **Fuzzy Sets and Systems**. 2020. <https://doi.org/10.1016/j.fss.2019.06.011>

ROUSE, J. JR.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Remote Sensing Center Texas A&M University College Station, Texas. 93p. Disponível em: < <https://core.ac.uk/download/pdf/42887948.pdf>>; acesso em 20/01/2021.

ROZANTE, J. R.; VILA, D. A.; CHIQUETTO, J. B.; FERNANDES, A. D.; ALVIM, D. S. Evaluation of TRMM/GPM Blended Daily Products over Brazil. **Remote Sensing**, v. 10, n. 6, 2018. <https://doi.org/10.3390/rs10060882>

SACHDEVA, S.; BHATIA, T.; VERMA, A. K. GIS-based evolutionary optimized Gradient Boosted Decision Trees for forest fire susceptibility mapping. **Natural Hazards**, v. 92, p. 1399-1418, 2018. <https://doi.org/10.1007/s11069-018-3256-5>

SALES, G. M.; PEREIRA, J. L. G.; THALÊS, M. C.; POCCARD-CHAPUISI, R.; ALMEIDA, A. S. Emprego dos focos de calor na avaliação das áreas queimadas e em incêndios florestais em Paragominas, Pará, Brasil. **Boletim do Museu Paraense Emílio Goeldi. Ciências Naturais**, v. 14, n. 1, p. 55-77, 2019.

SCHMIDT, D. A.; TAYLOR, A. H.; SKINNER, C. N. The influence of fuels treatment and landscape arrangement on simulated fire behavior, Southern Cascade range, California. **Forest Ecology and Management**, v. 255 p. 3170-3184, 2008. <https://doi.org/10.1016/j.foreco.2008.01.023>

SEADE – Fundação Sistema Estadual de Análise de Dados. Portal de estatísticas do Estado de São Paulo. Disponível em: <https://www.seade.gov.br/>, Acessado em: 15/09/2021.

SEIBERT, J.; MCGLYNN, B. A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models. **Water Resources Research**, v. 3. 2007. <https://doi.org/10.1029/2006WR005128>

SILVA, D. C. C.; ALBUQUERQUE FILHO, J. L.; SALES, J. C. A.; LOURENÇO, R. W. Uso de indicadores morfométricos como ferramentas para avaliação de bacias hidrográficas. **Revista Brasileira de Geografia Física**, v.9, n.2, p.221-217, 2016.

SILVA, D. C. C.; FILHO, J. L. A.; OLIVEIRA, R. A.; LOURENÇO, R. W. Aplicação do NDWI para avaliar a disponibilidade de água em bacias hidrográficas. **Revista da Universidade Vale do Rio Verde**, v. 15, n1, p. 774-783. 2017.

SILVA, D. C. C.; FILHO, J. L. A.; SALES, J. C. A.; LOURENÇO, R. W. Uso de indicadores morfométricos como ferramentas para avaliação de bacias hidrográficas. **Revista Brasileira de Geografia Física**. v. 9. n. 2. p. 627-642. 2016. <https://doi.org/10.26848/rbgf.v9.2.p627-642>

SILVA, V. V.; CARDOZO, F. S.; PEREIRA, G.; ZANIN, G. D.; COSTA, J. C.; ROCHA, L. C.; FIGUEIREDO, M. A.; RUFFINO, P. R. Análise da suscetibilidade da vegetação a incêndios florestais no estado de Minas Gerais. **Geografias**. v. 16, n 1. P. 125-139. 2018.

SILVEIRA, A. H. M.; QUEIROZ, A. F. S.; SILVA, B. C. O.; SILVA, F. M.; JUNIOR, N. P. C. Proposta metodológica para risco de incêndios florestal: estudo de caso na Zona de Proteção Ambiental (ZPA-1) em Natal/RN. **Revista Brasileira de Geografia Física**. v. 6, n. 5, p. 1174-1192. 2013.

SOUSA, J. A. P. Elaboração de um índice de vulnerabilidade Ambiental dos fragmentos florestais da Mata Atlântica. Tese (Doutorado em Ciências Ambientais) – Universidade Estadual Paulista, Sorocaba, 2021. 162p.

SOUZA, J. C.; LOPES, E. R. N.; SOUSA, J. A. P.; LOURENÇO, R. W. Avaliação dos Aspectos Físico-Químicos dos Solos em Diferentes Coberturas Vegetais. **Brazilian Journal of Environmental Sciences**, n.49, p.123-139, 2018. <https://doi.org/10.5327/Z2176-947820180366>

SUHET - Sentinel User Handbook and Exploitation Tools. Sentinel-2 User Handbook. ESA Standard Document. n. 1. 64p. 2013. Disponível em: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook, Acessado em: 06/05/2021.

SZPAKOWSKI, D. M.; JENSEN, J. LR. A Review of the Applications of Remote Sensing in Fire Ecology. **Remote Sensing**. v. 11. n. 22. 2019. <https://doi.org/10.3390/rs11222638>

TANG, X.; MACHIMURA, T.; LI, J.; LIU, W.; HONG, H. A novel optimized repeatedly random undersampling for selecting negative samples: A case study in an SVM-based forest fire susceptibility assessment. **Journal of Environmental Management**, v. 271, 2020. <https://doi.org/10.1016/j.jenvman.2020.111014>

TERMEH, S. V. R.; KORNEJADY, A.; POURGHASEMI, H. R.; KEESSTRA, S. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. **Science of The Total Environment**, v. 615, p. 438-451. 2018. <https://doi.org/10.1016/j.scitotenv.2017.09.262>

TONINI, M.; D'ANDREA, M.; BIONDI, G.; SPOSTI, S. D. TRUCCHIA, A.; FIOCUCCI, P. A Machine Learning-Based Approach for Wildfire Susceptibility Mapping. The Case Study of the Liguria Region in Italy. **Geosciences**, v. 10, n. 3. 2020. <https://doi.org/10.3390/geosciences10030105>

TORRES, F. T. P.; ROQUE, M. P. B.; LIMA, G. S.; MARTINS, S. V.; FARIA, A. L. Mapeamento do risco de incêndios florestais utilizando técnicas de geoprocessamento. **Floresta e Ambiente**, v. 24, p. 1-10. 2017.

TUCKER, G. E.; LANCASTER, S. T.; GASPARINI, N. M.; BRAS, R. L.; RYBARCZYK, S. M. An object-oriented framework for distributed hydrologic and geomorphic modeling using triangulated irregular networks. **Computers & Geosciences**, v. 27, n. 8, p. 959-973. 2001. [https://doi.org/10.1016/S0098-3004\(00\)00134-5](https://doi.org/10.1016/S0098-3004(00)00134-5)

URBANOWICZ, R. J.; MEEKER, M.; CAVA, W. L.; OLSON, R. S.; MOORE, J. H. Relief-based feature selection: Introduction and review. **Journal of Biomedical Informatics**, v. 85, p. 189-203. 2018. <https://doi.org/10.1016/j.jbi.2018.07.014>

VADREVU, K. P.; EATURU, A.; BADARINATH, K. V. S. 2006. Spatial distribution of forest fires and controlling factors in Andhra Pradesh, India using spot satellite datasets. **Environmental Monitoring and Assessment**, v. 123, n. 1–3, p. 75-96. <https://doi.org/10.1007/s10661-005-9122-4>

VENKATESH, K.; PREETHI, K.; RAMESH, H. 2020. Evaluating the effects of forest fire on water balance using fire susceptibility maps. **Ecological Indicators**, v. 110, p. 1-14. <https://doi.org/10.1016/j.ecolind.2019.105856>

VERBESSELT, J.; SOMERS, B.; AARDET, J. V.; JONCKHEERE, I.; COPPIN, P. Monitoring herbaceous biomass and water content with SPOT VEGETATION time-series to improve fire risk assessment in savanna ecosystems. **Remote Sensing of Environment**, v. 101, n. 3, p. 399-414. 2006. <https://doi.org/10.1016/j.rse.2006.01.005>

WAN, Z.; HOOK, S.; HULLEY, G. MOD11A1 MODIS / Terra Temperatura / Emissividade da superfície terrestre Diária L3 Global 1km SIN Grid V006. 2015, distribuído pela NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MOD11A1.006>

WHITE, L. A. S. WHITE, B. L.; RIBEIRO, G. T. Modelagem espacial do risco de incêndio florestal para o Município de Inhambupe, BA. **Pesquisa Florestal Brasileira**, v. 36, n. 85, 2016. <https://doi.org/10.4336/2016.pfb.36.85.850>

YAMAMOTO, J. K. Estatística, análise e Interpolação de Dados Geoespaciais. São Paulo, GEO KRIGAGEM. 2020. 308p.

ZAI – Zoneamento Ambiental do Ibiúna. Mapa do Zoneamento do Ibiúna. Disponível em: <https://sapl.ibiuna.sp.leg.br/materia/12540/documentoacessorio>, Acesso em: 20/09/2021.

ZHENG, Z.; HUANG, W.; LI, S.; ZENG, Y. Forest fire spread simulating model using cellular automaton with extreme learning machine. **Ecological Modelling**, v. 348, p. 33-43, 2017. <https://doi.org/10.1016/j.ecolmodel.2016.12.022>

ZUMBRUNNEN, T.; PAZZATTI, G. B.; MENÉNDEZ, P.; BUGMANN, H.; BURGI, M.; CONEDERA, M. Weather and human impacts on forest fires: 100 years of fire history in two climatic regions of Switzerland. **Forest Ecology and Management**, v. 261, n. 12, p. 2188-2199. 2011. <https://doi.org/10.1016/j.foreco.2010.10.009>

ANEXOS

A. Scripts utilizados no Google Earth Engine

O código pode ser colado diretamente no prompt de comando do GEE, seguida da configuração do intervalo temporal desejado ('AAAA-DD-MM'). Para exportação dos arquivos no formato GEOTIFF, é necessário criar uma geometria de interesse, definir o nome do arquivo (NAME), nome da pasta para hospedagem no Google Drive ('NAME_FOLDER'), número de pixels (NN), escala (NN), Sistema de Referência de Coordenadas - CRS ('EPSG:NNNN') e região (NNNNN).

A1. Script utilizado para exportação de dados mensais GPM no Google Earth Engine

```
var dataset = ee.ImageCollection('NASA/GPM_L3/IMERG_MONTHLY_V06')
    .filterDate('AAAA-DD-MM', 'AAAA-DD-MM');

// Select the max precipitation and mask out low precipitation values.
var precipitation = dataset.select('precipitation').max();
var mask = precipitation.gt(0.25);
var precipitation = precipitation.updateMask(mask);

var palette = [
    '000096', '0064ff', '00b4ff', '33db80', '9beb4a',
    'ffeb00', 'ffb300', 'ff6400', 'eb1e00', 'af0000'
];

var precipitationVis = {min: 0.0, max: 1.5, palette: palette};
Map.addLayer(precipitation, precipitationVis, 'Precipitation');
Map.setCenter(-76, 33, 3);

//-- Export
Export.image.toDrive({
    image: dataset.select('precipitation').mean(),
    description: 'NAME',
    folder: 'NAME_FOLDER',
    maxPixels: NN,
```

```

scale: NN,
crs: 'EPSG:NNNN',
region: geometry
});

```

A2. Script utilizado para exportação de dados mensais do MOD11 com uso do Google Earth Engine

```

var dataset = ee.ImageCollection('MODIS/006/MOD11A1')
    .filter(ee.Filter.date('AAAA-DD-MM', 'AAAA-DD-MM'));
var MOD = dataset.select('LST_Day_1km');
var LTS = MOD.map(function(image){
    var x = image.expression(
        '(N * 0.02) - T0',
        {
            N: image.select('LST_Day_1km'),
            T0: 273.15,
        }).float();
    return x.rename('LTS');
});
var ltsVis = {
    min: NNN,
    max: NNN,
    palette: [
        '040274', '040281', '0502a3', '0502b8', '0502ce', '0502e6',
        '0602ff', '235cb1', '307ef3', '269db1', '30c8e2', '32d3ef',
        '3be285', '3ff38f', '86e26f', '3ae237', 'b5e22e', 'd6e21f',
        'fff705', 'ffd611', 'ffb613', 'ff8b13', 'ff6e08', 'ff500d',
        'ff0000', 'de0101', 'c21301', 'a71001', '911003'
    ],
};
Map.setCenter(6.746, 46.529, 2);
Map.addLayer(

```

```

LTS, ItsVis,
'Land Surface Temperature');

//Export to Google Drive
Export.image.toDrive({image: LTS.first(),
folder: 'NAME_FOLDER',
description: 'NAME',
maxPixels: NN,
scale: NN,
crs: 'EPSG:NNNN',
region: geometry
});

```

A3. Script utilizado para exportação de dados mensais do FLDAS com uso do Google Earth Engine

```

var dataset=ee.ImageCollection('NASA/FLDAS/NOAH01/C/GL/M/V001')
    .filter(ee.Filter.date('NNNN-NN-NN', 'NNNN-NN-NN'));
var RH = dataset.select('Qair_f_tavg');
var rhVis = {
  min: NNNN,
  max: NNNN,
  palette: ['a50026', 'd73027', 'f46d43', 'fdae61', 'fee090', 'e0f3f8', 'abd9e9', '74add1',
'4575b4', '313695'],
};
Map.setCenter(30.0, 30.0, 2);
Map.addLayer(RH, rhVis, 'Specific humidity');

//-- Export
Export.image.toDrive({
  image: dataset.select('Qair_f_tavg').mean(),
  folder: 'NAME_FOLDER',
  description: 'NAME',

```

```

maxPixels: NN,
scale: NN,
crs: 'EPSG:NNNN',
region: geometry
});

```

A4. Script utilizado para exportação de dados mensais de NDVI com uso do Google Earth Engine

```

var s2 = ee.ImageCollection("COPERNICUS/S2");
var filtered = s2.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', NN))
  .filter(ee.Filter.date('NNNN-NN-NN', 'NNNN-NN-NN'))
  .filter(ee.Filter.bounds(geometry))
var image = filtered.median();
// 'NIR' (B8) and 'RED' (B4)
var ndvi = image.normalizedDifference(['B8', 'B4']).rename(['ndvi']);
var ndviVis = {min:0, max:1, palette: ['white', 'green']}
Map.addLayer(ndvi.clip(geometry), ndviVis, 'ndvi')

//Export to Google Drive
Export.image.toDrive({image: ndvi,
folder: 'NAME_FOLDER',
description: 'NN',
scale: NN,
maxPixels: NN,
region: geometry
});

```

A5. Script utilizado para exportação de dados mensais de NDWI com uso do Google Earth Engine

```

var s2 = ee.ImageCollection("COPERNICUS/S2");
var filtered = s2.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', NN))
  .filter(ee.Filter.date('NNNN-NN-NN', 'NNNN-NN-NN'))
  .filter(ee.Filter.bounds(geometry))

var image = filtered.median();

```

```
// Calculate Modified Normalized Difference Water Index (NDWI)
// 'GREEN' (B3) and 'SWIR1' (B11)
var ndwi = image.normalizedDifference(['B8', 'B11']).rename(['ndwi']);

var rgbVis = {min: 0.0, max: 3000, bands: ['B4', 'B3', 'B2']};
var ndviVis = {min:0, max:1, palette: ['white', 'green']}
var ndwiVis = {min:0, max:0.5, palette: ['white', 'blue']}

Map.addLayer(image.clip(geometry), rgbVis, 'Image');
Map.addLayer(ndwi.clip(geometry), ndwiVis, 'ndwi')

//Export to Google Drive
Export.image.toDrive({image: ndwi,
folder: 'NAME_FOLDER',
description: 'NN',
scale: NN,
maxPixels: NN,
region: geometry
});
```

A6. Script utilizado para exportação de dados mensais de SAVI com uso do Google Earth Engine

```
var s2 = ee.ImageCollection("COPERNICUS/S2");
var filtered = s2.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 30))
  .filter(ee.Filter.date('NNNN-NN-NN', 'NNNN-NN-NN'))
  .filter(ee.Filter.bounds(geometry))
var image = filtered.median();
// Calculate Soil-adjusted Vegetation Index (SAVI)
//  $1.5 * ((\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED} + 0.5))$ 

// For more complex indices, you can use the expression() function
var savi = image.expression(
  '1.5 * ((NIR - RED) / (NIR + RED + 0.5))', {
    'NIR': image.select('B8'),
    'RED': image.select('B4'),
  }).rename('savi');
var rgbVis = {min: 0.0, max: 3000, bands: ['B4', 'B3', 'B2']};
```

```
var ndviVis = {min:0, max:1, palette: ['white', 'green']}
var ndwiVis = {min:0, max:0.5, palette: ['white', 'blue']}
Map.addLayer(image.clip(geometry), rgbVis, 'Image');
Map.addLayer(savi.clip(geometry), ndviVis, 'savi')
```

```
//Export to Google Drive
Export.image.toDrive({image: savi,
folder: 'NAME_FOLDER',
description: 'NN',
scale: NN,
maxPixels: NN,
region: geometry
});
```

B. Script utilizado para análise de dados no software R

O código pode ser colado diretamente no prompt de comando do RStudio, em seguida, deve ser verificado os pacotes a ser instalados, indicar o caminho do diretório de arquivos (NN), e o nome do arquivo (DATA) e a variável utilizada (NN).

B1. Script utilizado para reamostragem de dados por Krigagem Ordinária

Bibliotecas utilizadas

```
library(readxl)
library(gstat)
library(rstudioapi)
library(sp)
library(raster)
library(hydroGOF)
library(sf)
library(rgdal)
library(spData)
library(mapview)
library(fBasics)
library(bestNormalize)
```

Limpar o ambiente de trabalho

```
rm(list=ls())
```

Define o diretório

```
PATH <- "NN"
setwd (PATH)
```

Inserir dado em GeoTIFF

```
DATA <- raster("NN.tif")
```

#Verifica o SRC

```
crs(DATA)
```

#Faz a reprojeção do SRC

```
projection(DATA) <- CRS("+proj=longlat +datum=WGS84")
DATA <- projectRaster(FLDAS, crs="+proj=utm +zone=23 +south +ellps=GRS80
+units=m +no_defs")#SIRGAS 2000/23S
crs(DATA)
```

#Conversão de raster para pontos

```
FLDAS = rasterToPoints(DATA, spatial = TRUE, bylayer=TRUE)
```

#Plot

```
plot(DATA)
```

#Estatística Descritiva

```
max(DATA$med2020)# maximo
min(DATA$med2020)# minimo
mean(DATA$med2020) #média
median(DATA$med2020, na.rm = FALSE)# mediana
sd(DATA$med2020)# Desvio padrão
skewness(DATA$med2020) # Assimetria
kurtosis(DATA$med2020) # curtose
cv(DATA$med2020) # coeficiente de variação
ks.test(DATA$med2020,"pnorm",mean(DATA$med2020),sd(DATA$med2020))#
Kolmogorov - Smirnov# p-value > 0,05 é normal
```

Transformação normal de dados, caso não seja

```
DATA2 <- bestNormalize(DATA$NN)
DATA2
hist(DATA2$NN)
DATA_2$NN <- DATA2$NN
```

#Separar dado a ser avaliado

```
DATA <- DATA2$NN
```

#Visualizar os dados espaciais

```
sp::bubble(DATA, "NN")
DATA <- DATA$NN
g = gstat(id="NN", formula = NN ~ 1, data = DATA)
```

Máxima distância entre os pontos

```
print(max(dist(DATA@coords))/2)
```

Mínima distância entre os pontos

```
print(min(dist(DATA@coords)))
```

#Semivariograma de pontos

```
var_could = gstat::variogram(g, cloud=T)
plot(var_could)
```

#Mapa variografico / alterar os valores de cutoff e width conforme distância máxima e mínima

```
var_map = gstat::variogram(g, cutoff=70843.56, width=9150, map=T)
plot(var_map)
var_exp = gstat::variogram(g, cutoff= 70843.56, width=9150)
plot(var_exp)
g=gstat(id="umidadea", formula= umidade~x+y, data=FLDAS)
var_exp = gstat::variogram(g, cutoff= 70843.56, width=9150)
plot(var_exp)
```

Etapas do ajuste do semivariograma

```
# 1° subtrair o fim pelo início da semivariância (contribuição)
# 2° definir o modelo (Sph, Exp, Gau)
# 3° identificar o ponto de estabilização da distância (alcance)
# 4° identificar o início da variância (efeito pepita)
```

#Semivariograma ajustado para cada modelo teórico

```
fit.sph = fit.variogram(var_exp, vgm(NN, "Sph",NN, NN))
plot(var_exp, fit.sph)
fit.sph
```

#Modelo de semivariograma

```
fit.exp = fit.variogram(var_exp, vgm(NN, "Exp",NN, NN))
plot(var_exp, fit.exp)
fit.exp
```

#Modelo de semivariograma

```
fit.gau = fit.variogram(var_exp, vgm(NN, "Gau",NN, NN))
plot(var_exp, fit.gau)
fit.gau
```

Validação cruzado dos modelos - LOOCV**#Validacao cruzada modelo Sph**

```
xvalid.sph = kriges.cv(NN~x+y, locations = DATA, model=fit.sph)
plot(xvalid.sph$var1.pred ~ umidade, cex=1.2, lwd=2)
```

```
abline(0,1, col="lightgrey", lwd=2)
lm_sph=lm(xvalid.sph$var1.pred ~ NN)
abline(lm_sph, col="red", lwd=2)
r2_sph = summary(lm_sph)$r.squared
rmse_sph= hydroGOF::rmse(xvalid.sph$var1.pred, NN)
```

#validacao cruzada modelo Exp

```
xvalid.exp = kriges.cv(NN~x+y, locations = DATA, model=fit.exp)
plot(xvalid.exp$var1.pred ~ NN, cex=1.2, lwd=2)
```

```
abline(0,1, col="lightgrey", lwd=2)
lm_exp=lm(xvalid.exp$var1.pred ~ NN)
abline(lm_exp, col="red", lwd=2)
r2_exp = summary(lm_exp)$r.squared
rmse_exp= hydroGOF::rmse(xvalid.exp$var1.pred, NN)
```

#validacao cruzada modelo Gau

```
xvalid.gau = krige.cv(NN~x+y, locations = DATA, model=fit.gau)
plot(xvalid.gau$var1.pred ~ NN, cex=1.2, lwd=2)
abline(0,1, col="lightgrey", lwd=2)
lm_gau=lm(xvalid.gau$var1.pred ~ NN)
abline(lm_gau, col="red", lwd=2)
r2_gau = summary(lm_gau)$r.squared
rmse_gau= hydroGOF::rmse(xvalid.gau$var1.pred, NN)
```

#Visualização dos resultados da validação cruzada

```
df.r2 = data.frame(r2_exp, r2_gau, r2_sph)
df.rmse= data.frame(rmse_exp, rmse_gau, rmse_sph)
temp = data.frame(cbind(t(df.r2), t(df.rmse)))
colnames(temp)=c('R2', 'RMSE')
rnames = gsub('R2_',",", rownames(temp))
rownames(temp)= rnames
print(temp)
```

Obtenção de contorno para interpolação de dados

```
poly <- shapefile("NN.shp")
```

#Verificar o SRC do Shp

```
crs(poly)<-('+init=epsg:31983')
plot(poly)
crs(poly)
```

#Rasterizar contorno, definir resolução do pixel

```
r = raster(poly, res=30) #resolução espacial
rp = rasterize(poly, r, 0)
grid <- as(rp, "SpatialPixelsDataFrame")
plot(grid)
dados <- as(DATA, "SpatialPixelsDataFrame")
crs(dados)
```

#Atribuindo um CRS para a base de dados

```
DATA3 <- DATA2
crs(DATA3) <- NA
crs(DATA3)
crs(DATA3) <- CRS("NNNN")
crs(DATA3)
```

Krigagem Ordinária com base no modelo ajustado

```
KO <- krige(DATA2$NN ~ 1, dados, grid, model = fit.exp)
plot(KO)
```

```
#Converter para raster
mapaRaster <-raster(KO)
```

```
#Exportar rastrear
writeRaster(mapaRaster, filename = "NNN.tif",
            format='GTiff',
            overwrite=F)
```

B2. Script utilizado para seleção de variáveis pelo método Boruta

```
# Bibliotecas utilizadas
```

```
library(readxl)
library(Boruta)
library(mlbench)
library(caret)
library(randomForest)
library(corrplot)
```

```
# Limpar o ambiente de trabalho
rm(list=ls())
```

```
# Define o diretório
```

```
PATH <- "NN"
setwd(PATH)
```

```
# Inserir dados
```

```
dados <- read_excel("NN.xlsx")
data <- dados
```

```
# Definir classe alvo como fator
```

```
data$NN = as.factor(data$NN)
```

```
# Algoritmo Boruta
```

```
Boruta.class <- Boruta(NN ~ ., data = data, doTrace = 2, ntree = 500)
Boruta.class
```

```
#Plot e resultados estatísticos
```

```
plot(Boruta.class)
TentativeRoughFix(Boruta.class)
getConfirmedFormula(Boruta.class)
attStats(Boruta.class)
```

B3. Script utilizado para o mapeamento da suscetibilidade de incêndios e queimadas

```
# Bibliotecas utilizadas
```

```

library(readxl)
library(frbs)
library(ROSE)
library(caret)

# Limpar o ambiente de trabalho
rm(list=ls())

# Define o diretório
PATH <- "NN"
setwd(PATH)

# Inserir dados
DATA <- read_excel("NN.xlsx")
data <- DATA
View(data)

# Escalonamento
data[,1:15]=scale(data[,1:15])
str(data)

# Definir classe alvo como fator
data$NN <- as.factor(data$NN)

# Verificando os dados
table(data$Focos)

# Verificar distribuição de classes
prop.table(table(data$Focos))

# Verificar distribuição de classes
table(data_balanced$NN)

# Implementando Sistema de Inferência Fuzzy Híbrido (HFIS)
# Separando dados de treinamento de teste
dataShuffled <- data_balanced[sample(nrow(data_balanced)), ]
tra.data <- dataShuffled[N : NN, ]
tst.data <- dataShuffled[N : nrow(dataShuffled), N : N]
real.data <- matrix(dataShuffled[N : nrow(dataShuffled), N], ncol = N)

## Definir intervalo de dados de entrada do IFIS
range.data.input <- apply(data_balanced[, -ncol(data_balanced)], 2, range)

## Defina o método e seus parâmetros
method.type <- "FRBCS.W" # FRBCS.CHI, FRBCS.W,
control <- list(num.labels = NN, type.mf = "GAUSSIAN", type.tnorm = "MIN",
               #TRIANGLE; #TRAPEZOID; GAUSSIAN; SIGMOID; BELL
               type.snorm = "MAX", type.implication.func = "ZADEH")

# Etapa de aprendizagem

```

```
object.cls <- frbs.learn(tra.data, range.data.input, method.type, control)
```

Etapa de previsão

```
res.test <- predict(object.cls, tst.data)
```

Exibir o modelo IFIS

```
summary(object.cls)
```

```
summary(real.data)
```

Plot as funções de associação

```
plotMF(object.cls)
```

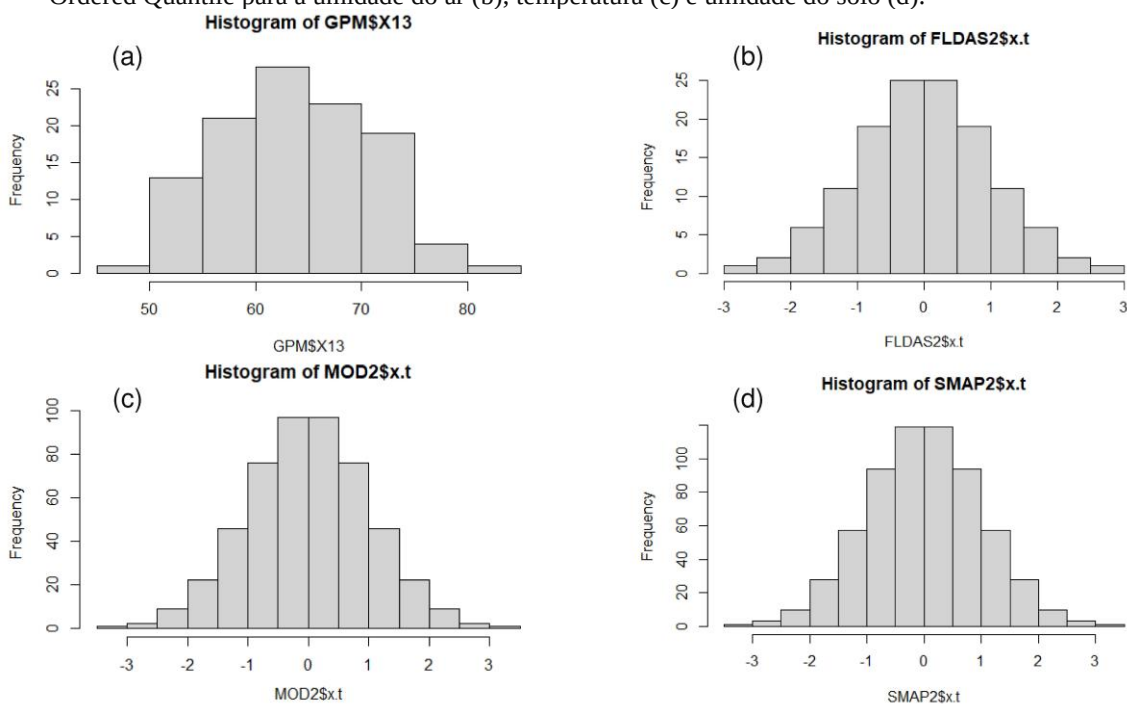
Matrix de confusão

```
confusionMatrix(factor(res.test, levels=1:2), factor(real.data, levels=1:2))
```

C. Resultados complementares da análise de dados

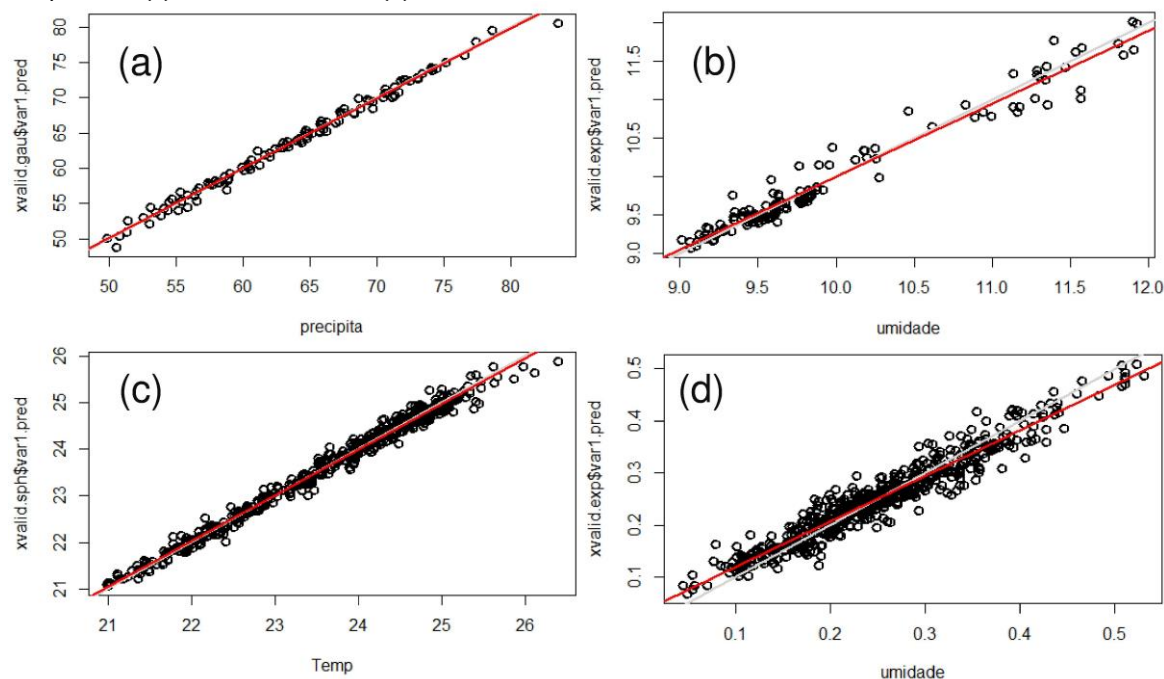
C1. Resultado da análise de normalidade dos dados

Histograma da distribuição normal para os dados de precipitação (a), e pós-transformação pelo método de Ordered Quantile para a umidade do ar (b), temperatura (c) e umidade do solo (d).



C2. Resultado da validação cruzada da krigagem

Resultado da validação cruzada da krigagem ordinária para os dados de precipitação (a), umidade do ar (b), temperatura (c) e umidade do solo (d).



C3. Resultado da estatística do teste de Wilcoxon

Resultado da estatística do teste de Wilcoxon com nível de significância de 0,05

Atributos	Estatística de Wilcoxon	p-valor
NDWI	308537	$2,2 \times 10^{-16}$
Precipitação	520415	$7,94 \times 10^{-13}$
Declividade	600783	0,05261
Temperatura	688429	0,0001707
Umidade do ar	746398	$5,199 \times 10^{-14}$
NDVI	330968	$2,2 \times 10^{-16}$
Distância de rodovias	680692	0,001114
Elevação	546742	$5,01 \times 10^{-8}$
Uso do solo	816266	$2,2 \times 10^{-16}$
Orientação de vertentes	689559	0,000127
Umidade do solo	734241	$1,621 \times 10^{-11}$
TWI	560744	$5,44 \times 10^{-6}$
Densidade de rodovias	718344	$1,17 \times 10^{-8}$
SAVI	331149	$2,2 \times 10^{-16}$

Valores de p para o teste de Wilcoxon: #H0 = as médias entre as amostras são iguais, #H1= Existe diferença entre as médias das amostras; $p\text{-valor} < 0,05$, rejeita-se H0.

C4. Resultado do algoritmo Boruta para as 14 variáveis originais

Resultado do algoritmo Boruta para as 14 variáveis com potencial de explicação dos eventos de incêndios e queimadas.

Variáveis	MeanImp p	MedianImp p	MinImp	MaxImp	NormHits	Decisão
Precipitação	38,94	38,65	36,57	40,49	38,94	Confirmado
Umidade do Ar	38,59	38,24	37,66	40,17	38,59	Confirmado
Umidade do Solo	37,98	37,73	36,69	39,42	37,98	Confirmado
Uso do solo	36,97	36,71	35,01	39,89	36,97	Confirmado
Elevação	34,54	34,00	32,56	37,70	34,54	Confirmado
Temperatura	34,11	33,75	32,70	36,55	34,11	Confirmado
NDVI	33,63	33,60	32,24	35,89	33,63	Confirmado
Densidade de vias	30,09	29,78	28,67	31,61	30,09	Confirmado
NDWI	29,33	28,86	27,88	31,01	29,33	Confirmado
SAVI	26,87	26,56	24,95	28,51	26,87	Confirmado
Distância entre vias	25,91	26,13	24,30	27,03	25,91	Confirmado
Declividade	22,54	22,75	20,74	23,47	22,54	Confirmado
Orientação	18,48	18,43	17,90	19,06	18,48	Confirmado
TWI	6,70	7,08	4,84	7,46	6,70	Confirmado

C5. Resultado de testes entre as diferentes funções de associação fuzzy

Métricas do teste entre as diferentes funções de associação fuzzy testadas

Métricas	Triangular	Trapezoidal	Gaussiano	Sigmoidal
Sensitividade	0,2442	0,4884	0,878	1,00
Especificidade	1,00	0,9970	0,9873	0,00
Valor Preditivo Positivo	1,00	0,9941	0,9863	0,5089
Valor Preditivo Negativo	0,5608	0,6529	0,8858	0,000
Prevalência	0,5089	0,5089	0,506	0,5089
Taxa de detecção	0,1243	0,2485	0,4483	0,5089
Prevalência de detecção	0,1243	0,2500	0,4545	1,00
Acurácia balanceada	0,6221	0,7427	0,9327	0,500
Acurácia	0,6154	0,7382	0,9315	0,5089
Índice Kappa	0,2409	0,4809	0,8633	0,00

C6. Exemplo da base de regras do Sistema Fuzzy Híbrido-HFIS

R1: IF Precipitação is small and Declividade is very.small and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is very.small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R2: IF Precipitação is small and Declividade is very.small and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is very.small and Uso do solo is small and Orientação de vertentes is large and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R3: IF Precipitação is very.large and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is very.large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R4: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is medium and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is large and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 1; CF 0,341181493

R5: IF Precipitação is very.large and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is small and Elevação is medium and Uso do solo is medium and Orientação de vertentes is large and Umidade do solo is medium and Densidade de rodovias is very.small THEN Ocorrência Focos is 1; CF 0,341181493

R6: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is very.large and Distância de rodovias is very.small and Elevação is medium and Uso do solo is very.small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R7: IF Precipitação is medium and Declividade is medium and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is medium and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R8: IF Precipitação is small and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R9: IF Precipitação is medium and Declividade is medium and Umidade do ar is small and NDVI is very.large and Distância de rodovias is medium and Elevação is medium and Uso do solo is very.small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is very.small THEN Ocorrência Focos is 1; CF 0,341181493

R10: IF Precipitação is small and Declividade is small and Umidade do ar is very.small and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R400: IF Precipitação is very.small and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is small and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 0; CF 0,658818507

R401: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is very.large and Umidade do solo is large and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R402: IF Precipitação is small and Declividade is very.small and Umidade do ar is very.small and NDVI is very.large and Distância de rodovias is very.small and Elevação is very.small and Uso do solo is medium and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R403: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R404: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is medium and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R405: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is medium and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is very.large and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 1; CF 0,341181493

R406: IF Precipitação is medium and Declividade is medium and Umidade do ar is small and NDVI is very.large and Distância de rodovias is small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is medium and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R407: IF Precipitação is large and Declividade is medium and Umidade do ar is large and NDVI is very.large and Distância de rodovias is very.small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is large and Umidade do solo is very.small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R408: IF Precipitação is very.small and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is small and Orientação de vertentes is medium and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R409: IF Precipitação is very.large and Declividade is medium and Umidade do ar is large and NDVI is large and Distância de rodovias is very.small and Elevação is medium and Uso do solo is small and Orientação de vertentes is large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R410: IF Precipitação is medium and Declividade is small and Umidade do ar is very.small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is medium and Orientação de vertentes is large and Umidade do solo is small and Densidade de rodovias is medium THEN Ocorrência Focos is 1; CF 0,341181493

R700: IF Precipitação is very.small and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 0; CF 0,658818507

R701: IF Precipitação is small and Declividade is small and Umidade do ar is small and NDVI is medium and Distância de rodovias is very.small and Elevação is small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 0; CF 0,658818507

R702: IF Precipitação is large and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is medium and Uso do solo is very.small and Orientação de vertentes is large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R703: IF Precipitação is small and Declividade is very.small and Umidade do ar is very.small and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 1; CF 0,341181493

R704: IF Precipitação is very.large and Declividade is medium and Umidade do ar is large and NDVI is large and Distância de rodovias is small and Elevação is medium and Uso do solo is medium and Orientação de vertentes is very.large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R705: IF Precipitação is medium and Declividade is medium and Umidade do ar is medium and NDVI is large and Distância de rodovias is medium and Elevação is small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is large and Densidade de rodovias is very.small THEN Ocorrência Focos is 0; CF 0,658818507

R706: IF Precipitação is large and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is very.large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R707: IF Precipitação is large and Declividade is medium and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is small and Orientação de vertentes is very.small and Umidade do solo is very.small and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R708: IF Precipitação is large and Declividade is large and Umidade do ar is medium and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is very.large and Orientação de vertentes is small and Umidade do solo is small and Densidade de rodovias is very.small THEN Ocorrência Focos is 1; CF 0,341181493

R709: IF Precipitação is small and Declividade is medium and Umidade do ar is small and NDVI is medium and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is large and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R710: IF Precipitação is medium and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is very.large and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is very.small THEN Ocorrência Focos is 1; CF 0,341181493

R790: IF Precipitação is very.large and Declividade is small and Umidade do ar is large and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is small and Orientação de vertentes is very.small and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R791: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is medium and Uso do solo is very.small and Orientação de vertentes is very.small and Umidade do solo is large and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R792: IF Precipitação is small and Declividade is medium and Umidade do ar is very.small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is medium and Orientação de vertentes is medium and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R793: IF Precipitação is small and Declividade is small and Umidade do ar is small and NDVI is medium and Distância de rodovias is very.small and Elevação is very.small and Uso do solo is medium and Orientação de vertentes is small and Umidade do solo is medium and Densidade de rodovias is medium THEN Ocorrência Focos is 1; CF 0,341181493

R794: IF Precipitação is medium and Declividade is small and Umidade do ar is very.small and NDVI is very.large and Distância de rodovias is small and Elevação is small and Uso do solo is large and Orientação de vertentes is large and Umidade do solo is medium and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R795: IF Precipitação is medium and Declividade is small and Umidade do ar is small and NDVI is large and Distância de rodovias is small and Elevação is small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 0; CF 0,658818507

R796: IF Precipitação is large and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is small and Orientação de vertentes is small and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R797: IF Precipitação is very.large and Declividade is very.small and Umidade do ar is very.large and NDVI is very.large and Distância de rodovias is small and Elevação is medium and Uso do solo is very.small and Orientação de vertentes is very.small and Umidade do solo is small and Densidade de rodovias is very.small THEN Ocorrência Focos is 1; CF 0,341181493

R798: IF Precipitação is large and Declividade is medium and Umidade do ar is small and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is very.small and Orientação de vertentes is medium and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

R799: IF Precipitação is very.large and Declividade is small and Umidade do ar is medium and NDVI is large and Distância de rodovias is very.small and Elevação is small and Uso do solo is medium and Orientação de vertentes is very.large and Umidade do solo is small and Densidade de rodovias is small THEN Ocorrência Focos is 1; CF 0,341181493

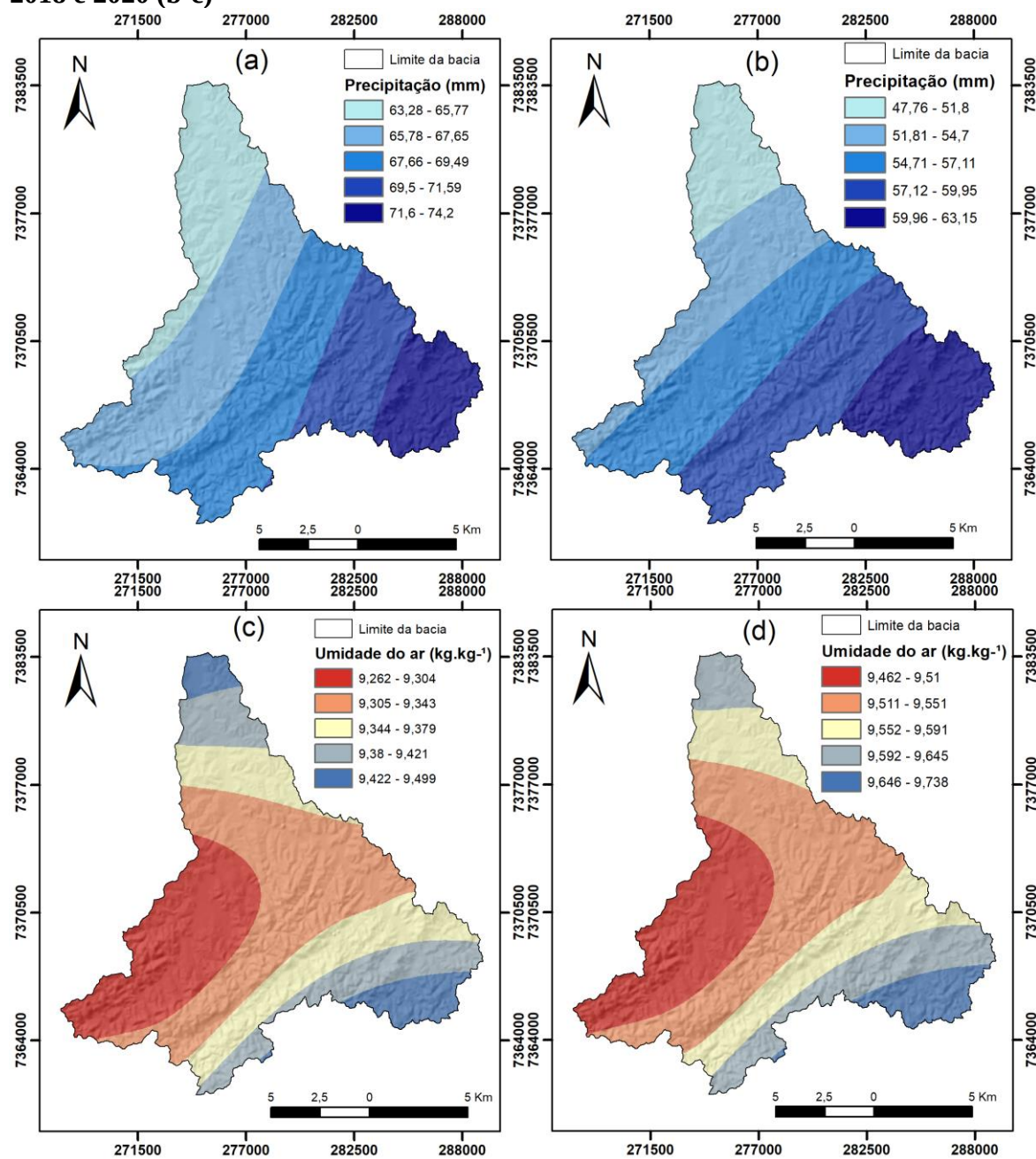
C7. Matriz de erro para os dados de treinamento e teste para o ano de 2019

Matriz de erro para os dados de treinamento e teste para o ano de 2019.

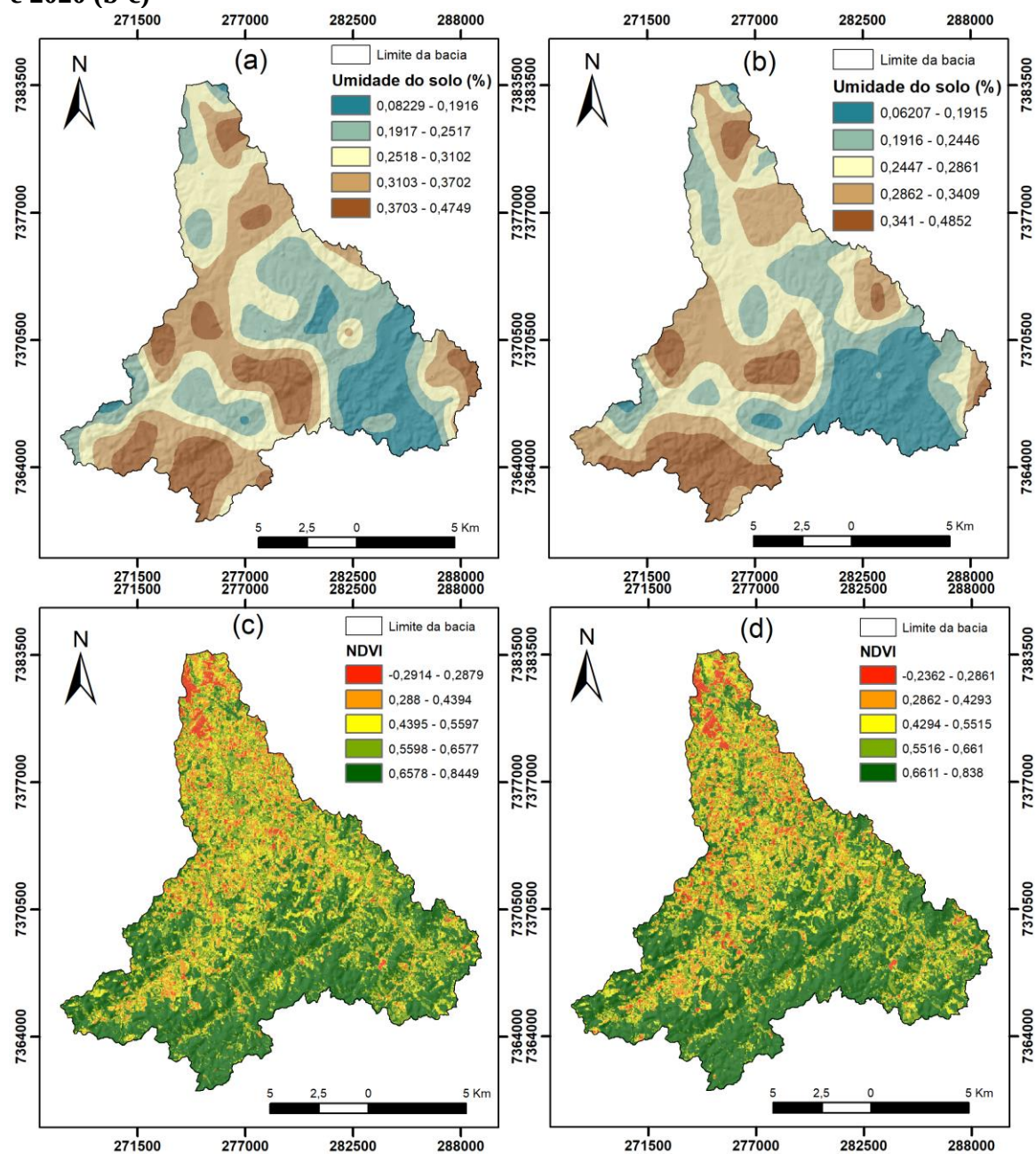
		Classe esperada	
Dados de Treinamento (70%)		Não-Fogo	Fogo
Classe Prevista	Não-Fogo	746	0
	Fogo	0	829
		Classe esperada	
Dados de Teste (30%)		Não-Fogo	Fogo
Classe Prevista	Não-Fogo	328	32
	Fogo	4	312

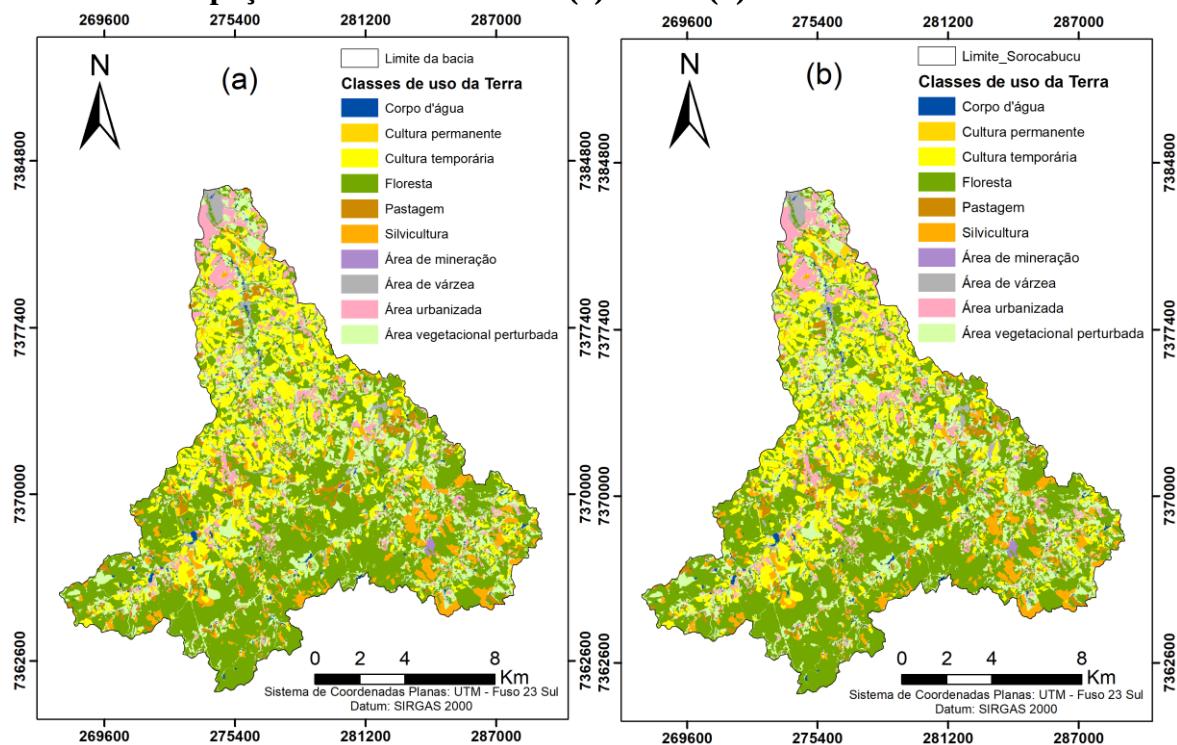
*Conjunto de dados de treinamento (1.575 observações) e teste (676 observações)

C8. Distribuição espacial da temperatura entre 2018 e 2020 (a-b) e umidade do ar em 2018 e 2020 (b-c)



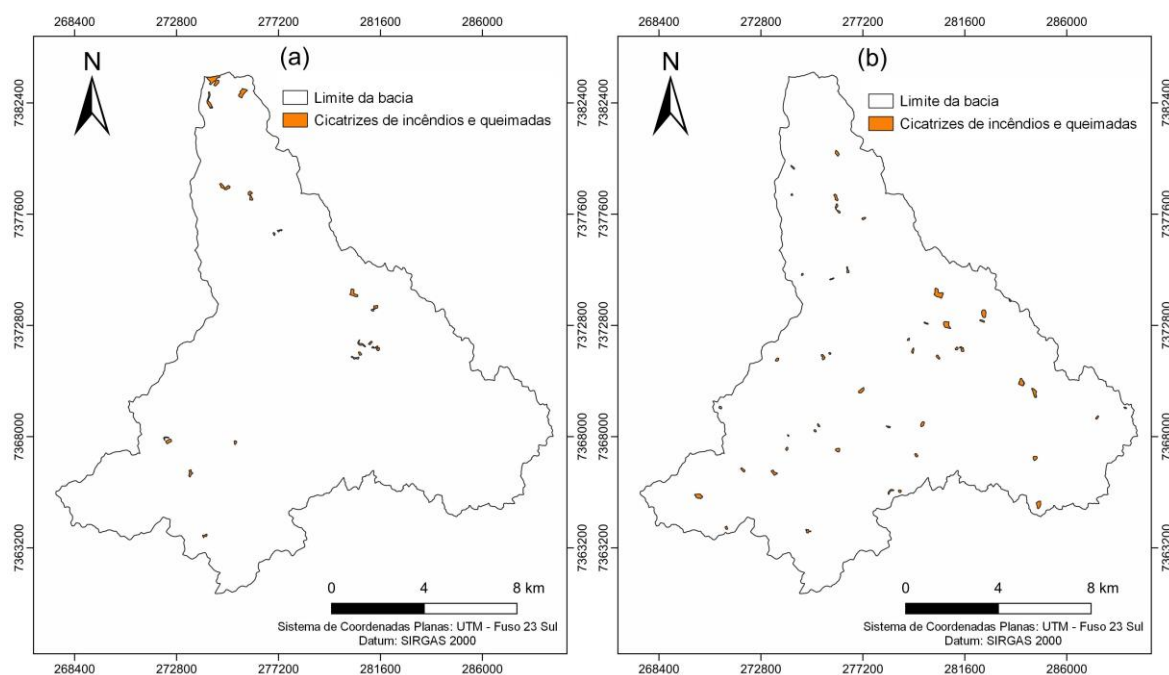
C9. Distribuição espacial da umidade do solo entre 2018 e 2020 (a-b) e NDVI em 2018 e 2020 (b-c)



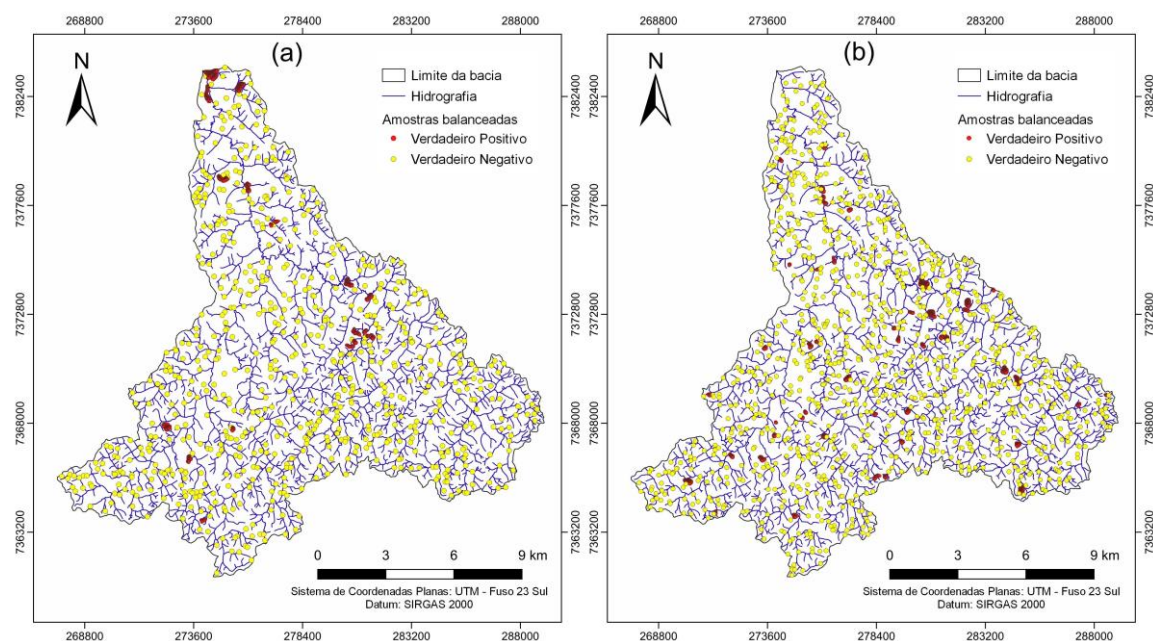
C10. Uso e ocupação da terra entre 2018 (a) e 2020 (b).

Classes de uso da terra	Área (ha)	Área (%)	Área (ha)	Área (%)
Silvicultura	1.299,65	6,41	1.188,31	5,86
Área vegetacional perturbada	3.831,50	18,90	3.865,44	19,06
Área urbanizada	1.349,39	6,66	1.392,90	6,87
Corpo d'água	236,19	1,16	236,19	1,16
Cultura temporária	3.398,16	16,76	3.467,10	17,10
Pastagem	544,06	2,68	525,08	2,59
Área de várzea	265,96	1,31	265,96	1,31
Cultura permanente	81,82	0,40	81,82	0,40
Área de mineração	22,98	0,11	22,98	0,11
Floresta	9.246,15	45,60	9.230,06	45,52
Área urbanizada	0,0359	0,0002	0,0359	0,0002

C11. Mapa de cicatrizes de incêndios e queimadas para 2018 (a) e 2020 (b)



C12. Mapa de pontos balanceados para 2018 (a) e 2020 (b)



C13. Matriz de erro para o período moderadamente chuvoso e moderadamente seco

Matriz de erro para os dados de treinamento e teste no período moderadamente chuvoso.

Classe Esperada				Classes Esperada			
Dados de Treinamento		Não-Fogo	Fogo	Dados de Teste		Não-Fogo	Fogo
Classe Prevista	Não-Fogo	502	25	Classe Prevista	Não-Fogo	199	14
	Fogo	0	430		Fogo	17	181

*Conjunto de dados de treinamento (957 observações) e teste (411 observações)

Matriz de erro para os dados de treinamento e teste no período moderadamente seco.

Classe Esperada				Classes Esperada			
Dados de Treinamento		Não-Fogo	Fogo	Dados de Teste		Não-Fogo	Fogo
Classe Prevista	Não-Fogo	695	38	Classe Prevista	Não-Fogo	264	48
	Fogo	0	683		Fogo	12	283

*Conjunto de dados de treinamento (1.416 observações) e teste (607 observações)