



Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia - FAENG



ISABELA DE OLIVEIRA GALLINDO

**MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS EM
ÁREA URBANA UTILIZANDO APRENDIZAGEM DE
MÁQUINA: ABORDAGEM BASEADA EM RELATOS DA
IMPrensa E FATORES NATURAIS E ANTRÓPICOS**

P G R N

Campo Grande (MS)

2024

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
FACULDADE DE ENGENHARIAS, ARQUITETURA E URBANISMO E GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS NATURAIS

ISABELA DE OLIVEIRA GALLINDO

**MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS EM
ÁREA URBANA UTILIZANDO APRENDIZAGEM DE
MÁQUINA: ABORDAGEM BASEADA EM RELATOS DA
IMPrensa E FATORES NATURAIS E ANTRÓPICOS**

Dissertação submetida ao Programa de Pós-graduação em Tecnologias Ambientais da Universidade Federal do Mato Grosso do Sul como requisito parcial a obtenção do título de Mestre em Recursos Naturais. Estudo realizado dentro da linha de pesquisa de Análise Integrada e Geotecnologias.

Orientador: Prof. Dr. Jamil Alexandre Ayach Anache
Coorientador: Profa. Dra. Eliane Guaraldo

Aprovada em: 29/04/2024

Banca Examinadora:

Prof. Dr. Jamil Alexandre Ayach Anache
Orientador PGRN - UFMS

Profa. Dra. Eliane Guaraldo
UFMS

Prof. Dr. Cristiano Capellani Quaresma
Universidade Nove de Julho

Prof. Dr. José Macato Junior
UFMS

Campo Grande (MS)
2024

DEDICATÓRIA

Dedico este trabalho, primeiramente, à Deus e aos meus mentores que me guiam e protegem. Também dedico aos meus pais, pois sem eles eu nada seria e nada me tornaria.

EPÍGRAFE

“Esse mundo que aí está foi feito por nós, portanto, pode ser por nós reinventado.”

Mário Sergio Cortella

AGRADECIMENTOS

Primeiramente agradeço à Deus a oportunidade de estar realizando este grande passo em direção ao meu sonho profissional. Sei que tudo aconteceu da forma como Ele quis e que me prestigiou como nenhum outro.

Meus pais são meus grandes apoiadores, meus amores e meus amigos. Nada seria sem eles e não seria como sou. Agradeço de corpo e alma por me ajudarem e me acolherem nessa caminhada chamada vida, por me ensinarem que a educação é a chama da alma e que devemos seguir nossos sonhos. A caminhada não foi e nem será fácil, mas sei que alcançarei tudo aquilo que busco, pois sei que vocês estarão sempre ao meu lado.

Agradeço também meu orientador, mentor e amigo. Obrigada, Jamil Anache, por me enriquecer de conhecimento e inspirar meus objetivos. Obrigada por acreditar no meu potencial e apostar no meu sucesso, pelos seus valiosos conselhos e sugestões. Obrigada por transformar minha ideia inicial neste trabalho do qual tenho tanto orgulho. Acredito que você foi uma das ações de Deus em meu caminho e serei eternamente grata por tudo o que vivemos neste período.

Além dos meus pais, meus apoios foram meus amigos. Amigos estes que sempre estiveram ao meu lado, me ouvindo, auxiliando e guiando durante todo o processo da pós-graduação. Obrigada Andreliz Souza, por entrar no meu caminho desde o primeiro dia e por cuidar de mim durante todo esse tempo, por ser minha dupla e por estar sempre comigo. Agradeço à Ana Munaro por toda parceria desde a primeira indagação sobre entrar nesta pós-graduação. E quero agradecer, também, meus amigos de longa data e imenso amor que, mesmo a centenas de quilômetros de distância, se fazem presentes todos os dias: Rebeca Lemos e Cauê de Lima.

Não menos importante, quero agradecer ao PGRN, Programa de Pós-Graduação em Recursos Naturais (PPGRN), da Faculdade de Engenharias, Arquitetura e Urbanismo (FAENG) da Fundação Universidade Federal de Mato Grosso do Sul (UFMS), que me possibilitou todas essas experiências, todo o aprendizado e conhecimento. Obrigada por permitir que eu me aproxime do meu sonho. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (Capes) - Código de Financiamento 001.

SUMÁRIO

RESUMO GERAL	11
GENERAL ABSTRACT.....	11
INTRODUÇÃO GERAL	12
OBJETIVOS.....	15
OBJETIVO GERAL.....	15
OBJETIVOS ESPECÍFICOS	15
PRIMEIRO CAPÍTULO: MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS BASEADO EM APRENDIZAGEM DE MÁQUINA: REVISÃO SISTEMÁTICA DA LITERATURA.....	16
1. INTRODUÇÃO.....	17
2. MATERIAL E MÉTODO.....	17
3. RESULTADOS E DISCUSSÃO	23
3.1 ARTIGOS DO PORTFÓLIO	23
3.2 EVOLUÇÃO DAS PUBLICAÇÕES SOBRE O TEMA.....	23
3.3 AUTORES, FILIAÇÕES E PAÍSES DE ORIGEM	24
3.4 ANÁLISE DAS PALAVRAS-CHAVE	26
3.5 ANÁLISE DOS PERIÓDICOS	27
3.6 ANÁLISE DOS PROCEDIMENTOS METODOLÓGICOS	28
3.7 LACUNAS E OPORTUNIDADES DE PESQUISA	30
3.7.1 Lacunas em Mato Grosso do Sul.....	31
4. CONCLUSÕES	33
5. REFERÊNCIAS.....	34
SEGUNDO CAPÍTULO: MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS BASEADO EM APRENDIZAGEM DE MÁQUINA A PARTIR DE RELATOS DA IMPRENSA	38
1. INTRODUÇÃO.....	39
2. MATERIAL E MÉTODO.....	40
2.1 ÁREA DE ESTUDO	40

2.2	COLETA DE OCORRÊNCIAS CITADAS	44
2.3	FATORES DE ENTRADA.....	49
2.4	APRENDIZAGEM DE MÁQUINA.....	51
3.	RESULTADOS E DISCUSSÃO	54
3.1	OCORRÊNCIAS CITADAS.....	55
3.2	FATORES DE ENTRADA.....	59
3.3	APRENDIZAGEM DE MÁQUINA.....	63
3.4	COMPARAÇÕES ENTRE PONTOS CITADOS NA IMPRENSA E OCORRÊNCIAS SIMULADAS	68
3.5	DIRETRIZES URBANÍSTICAS GERAIS	71
4.	CONCLUSÕES	73
5.	REFERÊNCIAS.....	74
	CONCLUSÕES GERAIS	76
	APÊNDICE 1: MATERIAL SUPLEMENTAR	78
	APÊNDICE 2: CÓDIGO PARA DEFINIR AS COORDENADAS DE CADA PUNTO MENCIONADO NOS RELATOS DE IMPRENSA	86
	APÊNDICE 3: CÓDIGO PARA APLICAR OS MÉTODOS DE MACHINE LEARNING.....	88

LISTA DE FIGURAS

Figura 1: Inundação e alagamento.....	12
Figura 2: Processo de seleção do portfólio.....	19
Figura 3: Questões da etapa de qualificação.....	20
Figura 4: Fluxograma do processo de seleção dos artigos.....	21
Figura 5: Evolução das publicações sobre o tema.....	24
Figura 6: Palavras-chave de maior destaque no portfólio bibliográfico.....	27
Figura 7: Sistema Integrado de Informação sobre Desastres.....	31
Figura 8: Atlas Digital de Desastres no Brasil.....	32
Figura 9: Mapa de pontos críticos monitorados pela Defesa Civil.....	33
Figura 10: Localização do município de Campo Grande-MS; (a) Contexto Brasil; (b) Contexto Mato Grosso do Sul; (c) Contexto Perímetro Municipal de Campo Grande..	41
Figura 11: Mapas e cartas da área de estudo; (a) Carta Geotécnica; (b) Carta de Drenagem; (c) Zonas Ambientais do Plano Diretor.....	43
Figura 12: Rotatória da Avenida Presidente Ernesto Geisel com a Rachid Neder.....	44
Figura 13: Rua São Roque.....	44
Figura 14: Fluxograma geral de processos.....	45
Figura 15: Busca nos veículos de imprensa.....	46
Figura 16: Fluxograma de coleta de dados nos veículos de imprensa.....	48
Figura 17: Mapa de localização dos pontos citados com a grade.....	49
Figura 18: Mapa de localização dos pontos usados para treinamento e validação.....	53
Figura 19: Fluxograma síntese do processo Python.....	54
Figura 20: Tipificação das ocorrências citadas.....	55
Figura 21: Mapa das ocorrências citadas e tipificadas no município.....	56
Figura 22: Mapa do grid de ocorrências citadas.....	59

Figura 23: Coleção dos Fatores de entrada; (a) Carta Geotécnica; (b) Carta de Drenagem; (c) Declividade; (d) MDE; (e) Direção de drenagem; (f) Proximidade de corpos hídricos; (g) Acumulação de fluxos hídricos; (h) Índice de potência de fluxo; (i) Índice de Umidade Topográfica; (j) NDVI; (k) Uso e ocupação do solo; (l) Solos; (m) Vegetação; (n) Distância vertical para a drenagem mais próxima.....	62
Figura 24: Resultados das rodadas do método <i>Gradient Boosting Classifier</i> (GBC)	65
Figura 25: Mapa síntese das ocorrências simuladas.....	66
Figura 26: Comparação entre ocorrências citadas e simuladas.....	68
Figura 27: Comparação entre ocorrências citadas e simuladas nas RU.....	69
Figura 28: Comparação entre ocorrências citadas e simulados nos bairros.....	70
Figura 29: Comparação entre ocorrências simuladas (GBC) e Suscetibilidade de inundação pela RIGeo (2022).....	71
Figura 30: Visão integrada do planejamento dos aspectos da água no ambiente (TUCCI e BERTONI, 2003).....	72

LISTA DE TABELAS

Tabela 1: Objetivos da metanálise dos artigos	20
Tabela 2: Autores relevantes do portfólio	24
Tabela 3: Países de maior destaque em publicações	25
Tabela 4: Características dos periódicos presentes no portfólio.....	28
Tabela 5: Métodos de Aprendizagem de Máquina e fatores de entrada.....	29
Tabela 6: Tempo de retorno.....	47
Tabela 7: Características dos fatores de entrada.....	50
Tabela 8: Extração dos fatores de entrada	51
Tabela 9: Regiões Urbanas mais citadas	56
Tabela 10: Bairros mais citados.....	57
Tabela 11: Pontos mais citados	58
Tabela 12: Ranking dos Fatores de entrada segundo o SVM.....	63
Tabela 13: Métodos na codificação em Python.....	63
Tabela 14: Total de ocorrências/ano simuladas nas RU	67
Tabela 15: Total de ocorrências/ano simuladas nos bairros	67
Tabela 16: Artigos do Portfólio	79

RESUMO GERAL

GALLINDO, I. O. (2024). Mapeamento de inundações e alagamentos em área urbana utilizando Aprendizagem de Máquina: abordagem baseada em relatos da imprensa e fatores naturais e antrópicos. 2024. 95 páginas. Dissertação - Programa de Pós-graduação em Recursos Naturais. Universidade Federal de Mato Grosso do Sul, Brasil.

Inundações e alagamentos são considerados os eventos adversos extremos mais frequentes do planeta e, em áreas urbanas são potencializados pelas mudanças ambientais locais e globais, trazendo grandes impactos econômicos e prejudicando a vida de milhões de pessoas. Esta dissertação tem por objetivo investigar, no primeiro capítulo, o estado da arte de inundações e alagamentos, dentro de um portfólio final com 43 artigos que destaca as técnicas e as metodologias mais disseminadas para promover o mapeamento de áreas inundáveis e alagáveis em cidades, obtendo indicativos metodológicos, lacunas e oportunidades de pesquisa e esclarecendo a discussão atual sobre a temática no meio acadêmico. A segunda parte deste trabalho disserta sobre a utilização de metodologias de *Machine Learning* para mapear zonas com potenciais de inundação e alagamento no perímetro urbano. Aplicando o método *Gradient Boosting Classifier* (GBC), verificamos a possibilidade de inundação e alagamento com 433 ocorrências/ano na sede municipal de Campo Grande (MS), mediante uso dos pontos noticiados pelos veículos de imprensa como verdade de campo e fatores de entrada que caracterizam a área de estudo em aspectos naturais e antrópicos.

Palavras-chave: Mapa de Ocorrências, Área urbana, Eventos Extremos, Sociohidrologia, Machine Learning, Mudanças Ambientais.

GENERAL ABSTRACT

Floods and inundations are considered the most frequent extreme adverse events on the planet, and in urban areas, they are exacerbated by local and global environmental changes, bringing significant economic impacts and harming the lives of millions of people. This dissertation aims to investigate, in the first chapter, the state of the art of floods and inundations, within a final portfolio of 43 articles that highlights the most disseminated techniques and methodologies to promote the mapping of flood-prone and inundation-prone areas in cities, identifying methodological indicators, research gaps, and opportunities, and clarifying the current discussion on the topic in the academic field. The second part of this work discusses the use of Machine Learning methodologies to map potential flood and inundation zones in the urban perimeter. By applying the Gradient Boosting Classifier (GBC) method, we identified the possibility of floods and inundations with 433 occurrences per year in the municipal area of Campo Grande (MS), using locations reported by the press as ground truth and input factors that characterize the study area in natural and anthropogenic aspects.

Keywords: Occurrences Map, Urban area, Extreme Events, Sociohydrology, Machine Learning, Environment Changes.

INTRODUÇÃO GERAL

As mudanças climáticas acentuam os fenômenos climáticos drásticos presentes em todo o globo (LEE et al. 2017; CHOUBIN et al. 2019). Os eventos extremos acelerados são ocorrências danosas advindas da somatória dos processos naturais do planeta Terra com ações antrópicas nocivas ao meio ambiente (TIEN BUI et al., 2019). Um dos maiores e mais frequentes eventos extremos conhecidos, responsáveis por perda de vidas, materiais e financeiras é o das inundações (TEHRANY et al., 2014; TELLA, BALOGUN, 2020).

A inundação ocorre quando uma área que está fora dos limites padrões de um curso d'água submerge, acontecendo de forma gradual, normalmente em zonas planas e geralmente resultado de chuvas intensas. Já o processo de alagamento é resultado de uma extrapolação na capacidade do atual sistema de drenagem urbana e se deve à somatória de águas pluviais que escoam pelas ruas e calçadas (CEMADEN, 2021). A Figura 1 ilustra a diferença entre inundações e alagamento.

Figura 1: Inundação e alagamento



Fonte: GONÇALVES, BAPTISTA e RIBEIRO (2016)

O cenário atual do crescimento urbano acelerado está intimamente atrelado ao aumento da população nas últimas décadas, contribuindo para a ampliação dos percentuais de pessoas que residem em cidades. Este aumento promove consequentemente, um rápido crescimento na expansão urbana, trazendo riscos de danos irreversíveis à agricultura, ao tecido urbano e ao meio ambiente (CHOUBIN et al., 2019; RAHMATI et al., 2019). As áreas urbanas podem sofrer inundações e alagamentos devido a precipitações intensas; estes problemas geralmente atestam um serviço de drenagem urbana sobrecarregado e deficiente, amplificado pela falta de manutenção. A urbanização espalhada, sem considerar os sistemas drenantes, geralmente amplifica o alcance desses fenômenos atingindo maiores extensões de áreas urbanizadas e

provocando graves eventos de inundações e alagamentos (DARABI et al., 2019; HOSSAIN e MENG, 2020).

De acordo com o Diagnóstico de Capacidades e Necessidades Municipais em Proteção e Defesa Civil de 2021, o número de desastres relacionados ao clima e ações antrópicas aumentaram consideravelmente nos últimos 20 anos em todo o Brasil. A Agência Nacional de Águas (ANA) alega que houve um aumento de 409% nas declarações municipais de emergência acerca das secas no país e, no mesmo período, os deslizamentos aumentaram, aproximadamente, 22 vezes. Estes acontecimentos representam gastos cada vez maiores para reparo de danos e mitigação, visto que entre 1995 e 2019, o Banco Mundial apurou que cerca R\$ 333,36 bilhões foram gastos em desastres relativos aos danos materiais e à prejuízos somente no Brasil. Nos setores de infraestrutura e habitações foram gastos, mensalmente, valores superiores a R\$ 1 bilhão (BRASIL, 2021b). Deste montante, o Centro-Oeste representa 6,1%, aproximadamente R\$20,4 bilhões (BRASIL, 2021c).

A fim de evitar devastações urbanas causadas por inundações e alagamentos em grande escala, é necessária a elaboração de medidas de prevenção e proteção. Essas medidas podem incluir sistemas de alerta antecipado de inundação e implementação de estações de bombeamento d'água, por exemplo (LEE et al., 2017). Entretanto, para promover esse tipo de mitigação e adaptação, é vital conhecer de maneira aprofundada as áreas de maior vulnerabilidade a esse desastre natural (TEHRANY et al. 2014).

As consequências de inundações e alagamentos que acometem o tecido urbano podem ser mitigadas e adaptadas através de planejamento e do conhecimento aprofundado da área utilizando os mapas de risco. O Estatuto da Cidade, Lei Federal n° 10.257 de 2001 e alterado pela Lei Federal n°12.608 de 2012, por exemplo, regulamenta os Art. 182 e 183 da Constituição Federal de 1988 e discorre sobre instrumentos de política urbana. Em seu Art. 42 menciona as exigências mínimas para produção do Plano Diretor em áreas com suscetibilidade de deslizamentos de grande impacto, inundações bruscas ou processos geológicos ou hidrológicos (BRASIL, 2001).

Decisões de planejamento urbano devem partir do princípio de conhecimentos prévios e estudos atualizados, principalmente ao alterar ou implementar leis e instrumentos urbanísticos que afetam o uso e cobertura do solo, infraestrutura urbana, sistemas de drenagem e permeabilidade e, também se referenciar de projetos que

tenham por objetivo controlar esses eventos extremos e alertar a população (SARHADI, SOLTANI e MODARRES, 2012; RAHMATI et al., 2019).

O trabalho está organizado por capítulos que seguem a ordem metodológica realizada para a produção do mapeamento de suscetibilidade de inundações e alagamentos em território urbano através de relatos de imprensa. O primeiro capítulo apresenta o portfólio base do estado da arte da área e compreensão das metodologias mais utilizadas atualmente. O segundo capítulo aborda o levantamento dos dados base para a elaboração do mapeamento e explica a metodologia utilizada para o processo de coleta, análise dos fatores de entrada e a produção do mapeamento de suscetibilidade com o uso do *Machine Learning*.

OBJETIVOS

Objetivo geral

Verificar a aplicabilidade da utilização do Aprendizado de Máquina (*Machine Learning*), treinado a partir dos relatos da imprensa, para o mapeamento de áreas com potencial de inundações e alagamentos em área urbana de Campo Grande (MS) almejando contribuir com diretrizes de políticas públicas voltadas ao planejamento urbano.

Objetivos Específicos

- Investigar a metodologia *Machine Learning*, por meio de uma Revisão Sistemática da Literatura, para identificar na produção científica quais as técnicas e fatores físicos e antrópicos são mais utilizados no tema;
- Levantar os relatos de imprensa em um recorte temporal de 10 anos para identificar os pontos e áreas com problemas de inundações e alagamentos em área urbana;
- Treinar e validar técnicas de Aprendizado de Máquina relacionando pontos de inundação e alagamento com fatores físicos e antrópicos da área de estudo;
- Mapear regiões não amostradas da área de estudo utilizando técnica validada de aprendizado de máquina, com fatores físicos naturais e antrópicos; e
- Identificar a importância dos fatores físicos e antrópicos da área de estudo na ocorrência de inundações e alagamentos de modo a nortear diretrizes urbanísticas de mitigação e adaptação.

PRIMEIRO CAPÍTULO: MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS BASEADO EM APRENDIZAGEM DE MÁQUINA: REVISÃO SISTEMÁTICA DA LITERATURA

Resumo: Derivado de deficiências no planejamento e gestão urbana e acréscimo de efeitos das mudanças climáticas, a inundação é um grave problema que impacta a qualidade de vida de milhões de pessoas que habitam zonas urbanas. A fim de entender mais a fundo a problemática e promover o mapeamento de suscetibilidade, é necessário compreender conceitos e metodologias capazes de contribuir com o tema. Neste estudo investigamos o estado da arte de inundações e alagamentos, as técnicas e as metodologias mais disseminadas para promover o mapeamento de áreas inundáveis e alagáveis em cidades. Por meio de uma Revisão Sistemática da Literatura, o portfólio elaborado possui 43 artigos, dentro de um recorte temporal de 9 anos (2012 a 2021), que abrange metodologias de mapeamento, possíveis causas e condições urbanas para a ocorrência da problemática. As informações obtidas indicam caminhos metodológicos, tais como métodos de *Machine Learning* comumente usadas para estudos da temática, lacunas e oportunidades de pesquisa que evidenciam a discussão do tema no meio acadêmico devido ao significativo aumento de artigos publicados e o aprimoramento das métricas e das metodologias para produção de mapeamentos de suscetibilidade de inundação e alagamento em área urbana.

Palavras-chave: Parsifal, Bibliometria, Mapeamento de Suscetibilidade, Inundação Urbana, Aprendizagem de Máquina.

Abstract: Due to deficiencies in urban planning and management, combined with the effects of climate change, flooding is a serious problem that impacts the quality of life of millions of people living in urban areas. To better understand this issue and promote susceptibility mapping, it is necessary to comprehend concepts and methodologies capable of contributing to the topic. In this study, we investigate the state of the art of floods and inundations, the most disseminated techniques, and methodologies to promote the mapping of flood-prone and inundation-prone areas in cities. Through a Systematic Literature Review, the portfolio comprises 43 articles within a 9-year time frame (2012 to 2021), covering mapping methodologies, possible causes, and urban conditions for the occurrence of the issue. The obtained information indicates methodological pathways, such as Machine Learning methods commonly used for studies on this topic, research gaps, and opportunities that highlight the academic discussion due to the significant increase in published articles and the improvement of metrics and methodologies for producing flood and inundation susceptibility maps in urban areas.

Keywords: Parsifal, Bibliometrics, Susceptibility Mapping, Urban Flooding, Machine Learning.

1. INTRODUÇÃO

Intensificados pelas mudanças climáticas, os eventos extremos se fazem recorrentes em todo o globo (LEE et al. 2017; CHOUBIN et al. 2019). As inundações provocam perdas humanas, materiais e financeiras (TEHRANY et al., 2014; TELLA, BALOGUN, 2020). O aumento expressivo da população em zonas urbanas deve ser abordado associado com a capacidade de drenagem, o uso do solo e um limite aceitável de impermeabilização de superfícies (TELLA, BALOGUN, 2020; KARYMBALIS et al., 2021).

Reconhecer e identificar as problemáticas existentes no perímetro urbano é uma das maneiras que promovem o fim dos danos causados por inundações e alagamentos em grande escala. É necessária a elaboração do mapeamento de suscetibilidade, fundamental para esse tipo de investigação, e pode ser executado através de diferentes métodos, como os baseados em estatística: Razão de Frequência (FR) e Logística de Regressão ou modelos que utilizam técnicas de Aprendizagem de Máquina (*Machine Learning*) como Floresta Aleatória (RF), Redes Neurais Artificiais (RNA) e Máquina de Vetores de Suporte (SVM) (CHOUBIN et al., 2019). Mediante o uso deste mapeamento é possível aprofundar-se na metodologia *Machine Learning* para prever quais pontos podem ocorrer inundações e alagamentos e, quando usados em conjunto, aumentam a qualidade de ajuste e a precisão da previsão.

O presente capítulo tem como objetivo revisar a literatura de referência acerca do tema de inundações em áreas urbanas e evidenciar a utilização de método de Aprendizagem de Máquina para o mapeamento de inundações e alagamentos em tecido urbano, ampliando e aglutinando conhecimentos prévios.

2. MATERIAL E MÉTODO

Conforme as diretrizes propostas por De Jesus-Lopes et al. (2022), esta pesquisa, como um todo, se concentra na área das Ciências Ambientais, de natureza interdisciplinar. Configura-se como uma pesquisa com abordagem de método misto. Caracteriza-se, também, como uma pesquisa experimental por método de modelagem e simulação. Trata-se de uma pesquisa exploratória com dados secundários. Os procedimentos metodológicos incluem estudos de caso e os dados são analisados utilizando a técnica de análise mista.

Foi elaborada uma revisão sistemática de literatura por meio do programa Parsifal versão 2.1.1 (2021) que visa identificar, compreender e interpretar outros estudos a fim de responder questionamentos levantados de acordo com o objetivo do pesquisador (KEELE, 2007). Este programa permite criar uma *string* de busca seguindo um processo de identificação do tema proposto (Figura 2). Para tal, o programa adota uma sequência de palavras-chave nomeada PICOC (PETTICREW; ROBERTS, 2008) com os seguintes parâmetros: População (*Population*), Intervenção (*Intervention*), Comparação (*Comparison*), Resultados (*Outcome*) e Contexto (*Context*). As palavras-chave definidas foram: *Population* – *Inundat**, *Flood**; *Intervention* – *River*, *Stream*, *Watersh**, *Hydrology*; *Comparison* – *Urban**, *City*, *Cities*; *Outcome* – *Compilation Data*, *Monitor**, *GIS*. Neste caso, não foi empregado termos referentes ao Contexto visto que não queríamos definir qual a abrangência (acadêmico ou não) dos artigos.

O uso das simbologias de truncamento (*) foi empregado para que mais palavras relacionadas fossem atraídas à pesquisa, visto que é gerado o restante da palavra automaticamente, já o termo booleano “OR” foi implementado para aumentar a abrangência e o termo “AND” foi utilizando visando relacionar as palavras-chave empregadas. Dessa forma a *string* de busca foi estruturada: ("*Flood**" OR "*Inundat**") AND ("*Hydrology*" OR "*River*" OR "*Stream*" OR "*Watersh**") AND ("*Cities*" OR "*City*" OR "*Urban**") AND ("*GIS*" OR "*Compilation Data*" OR "*Monitor**").

As bases de dados indexadas utilizadas *Scopus*, *Science Direct* e *Web of Science*, foram escolhidas pela sua ampla gama de títulos, qualidade dos artigos publicados e disponibilidade dos artigos de forma livre (BAKKALBASI et al., 2006; MARTÍN-MARTÍN et al., 2018). Além dessa seleção via *string* outros artigos foram selecionados através da técnica conhecida por *snowball*, que tem por intuito identificar os trabalhos relevantes dentro das referências dos artigos evidenciados (KEELE; CHARTERS, 2007; JALALI; WOHLIN, 2012).

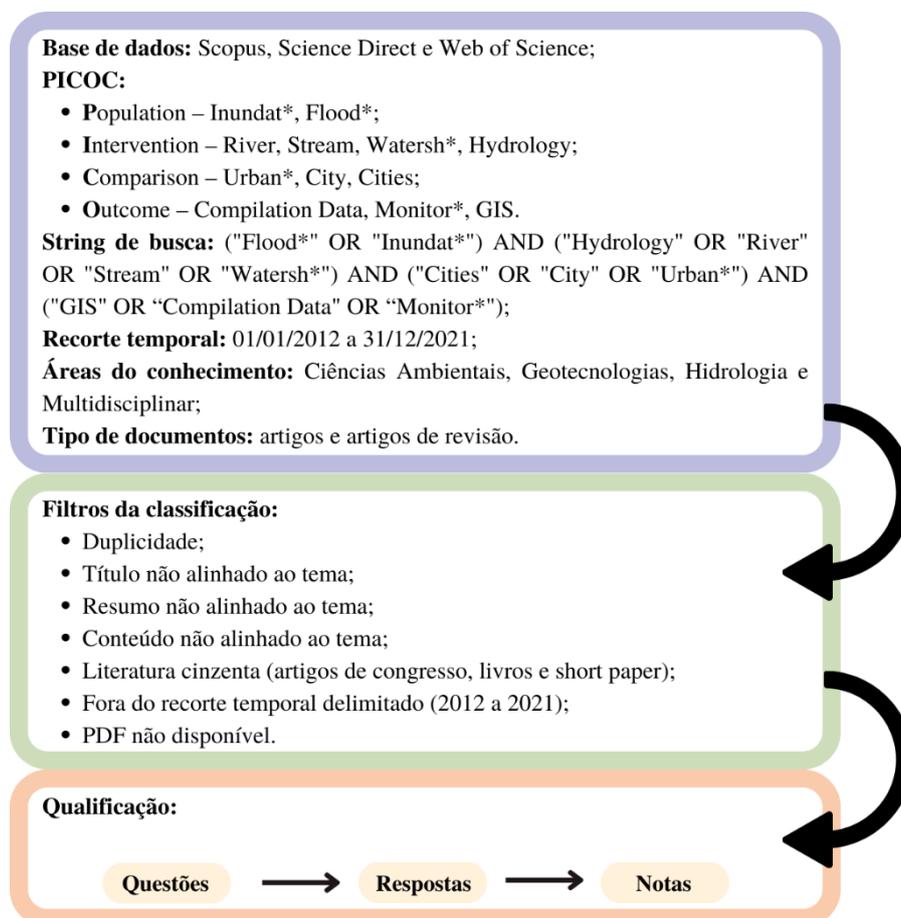
Esse montante passou por classificações de compatibilidade com o estudo proposto para definir os caminhos e interesses para com a pesquisa (SNYDER, 2019) e, mediante isso, alguns foram rejeitados, dentro do ambiente Parsifal (2021), segundo os critérios de duplicidade, título não alinhado ao tema, resumo não alinhado ao tema, conteúdo do artigo não alinhado ao tema, não alinhado as áreas do conhecimento preestabelecidas (Ciências Ambientais, Geotecnologias, Hidrologia e Multidisciplinar),

literatura cinzenta (artigos de congresso, livros e *short paper*), fora do recorte temporal delimitado (2012 a 2021) e PDF não disponível.

A escolha do recorte temporal deu-se pela pouca quantidade de artigos voltados ao Machine Learning anterior ao ano de 2012, período em que o número de artigos que voltavam seus esforços para a hidrologia com o uso da metodologia Aprendizado de Máquina é praticamente inexistente, não acrescentando muito à pesquisa.

Após a exclusão dos artigos e artigos de revisão que não foram aprovados pela classificação, o montante restante deste material seguiu para o processo de qualificação. Essa qualificação consiste em elaborar um questionário com perguntas que tem por objetivo encontrar, dentro do aglomerado de artigos classificados previamente, respostas para a pesquisa em questão (BANDARA et al., 2015).

Figura 2: Processo de seleção do portfólio

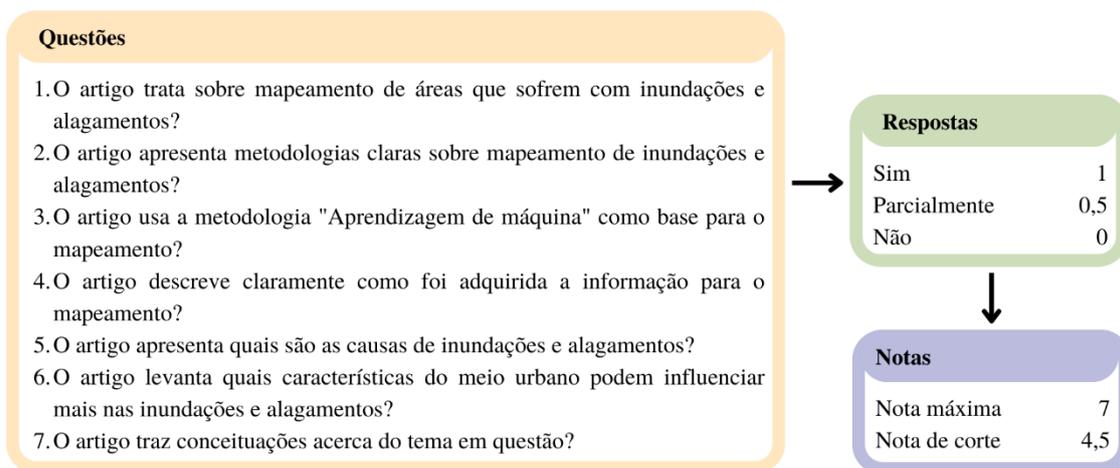


Fonte: Elaborado pela autora

Foram elaboradas 7 questões para o aglomerado de artigos, conforme a Figura 3. O critério para a elaboração de cada uma das perguntas levou em consideração os pontos cruciais da pesquisa: mapeamento de inundações, metodologias necessárias

para o mapeamento, metodologias acerca do Aprendizagem de Máquina e as condições urbanas para inundações e alagamentos. Desta forma, os artigos foram lidos e selecionados de acordo com a resposta a cada uma das questões, indicando sua relevância para esta pesquisa. Para cada uma das respostas “sim”, “parcialmente” e “não” foram atribuídas pontuações. A nota máxima total que cada artigo poderia receber seria 7 pontos, com uma nota de corte estabelecida em 4,5 pontos (metade do total acrescido de um ponto), permitindo, assim, um portfólio para extração de dados minuciosos como conceituações e práticas metodológicas.

Figura 3: Questões da etapa de qualificação



Fonte: Elaborada pela autora

Após concluída a etapa de qualificação dos artigos, o montante resultante gerou o portfólio final. A partir desse portfólio foi possível extrair todos os metadados de cada um dos artigos (ano de publicação; nome do periódico; país da instituição a que o primeiro autor está vinculado; palavras-chave; número de citações em cada base indexada; fatores de impacto dos periódicos no JCR e SJR; referências *snowball*) para uma metanálise. Posteriormente a leitura de cada um dos artigos que compõem o portfólio, os seguintes itens foram evidenciados (Tabela 1):

Tabela 1: Objetivos da metanálise dos artigos

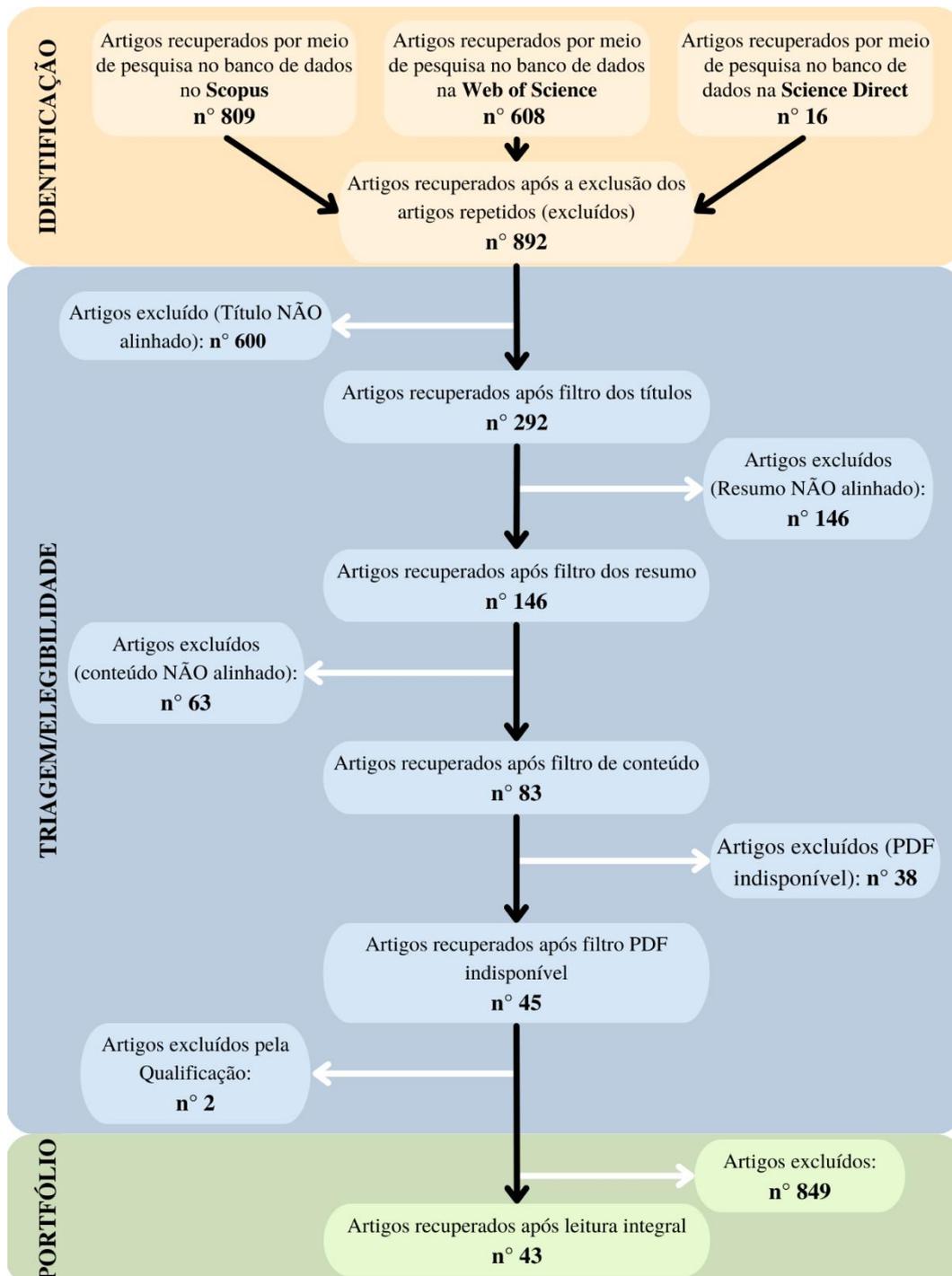
Áreas de interesse	Objetivos
Conceito de inundações, mapeamento de risco e <i>Machine Learning</i>	Compreender qual é o conceito de inundação e como ele é abordado pela literatura escolhida
Métodos de pesquisa	Almeja assimilar quais são os métodos empregados, seus componentes e como aplicá-los
Resultados da pesquisa	Visa evidenciar os resultados obtidos nas pesquisas realizadas
Lacunas ou oportunidades de pesquisa	Compreender e apontar quais são as lacunas de conhecimento ou limitações que existem nos artigos

Áreas de interesse	Objetivos
	evidenciados a fim de propor novos estudos e recomendações para outros pesquisadores

Fonte: Elaborado pela autora

A Figura 4 ilustra o fluxograma do processo final da seleção dos artigos para o portfólio e como os filtros de triagem foram aplicados em cada um dos processos.

Figura 4: Fluxograma do processo de seleção dos artigos



Fonte: Elaborado pela autora

Foram colhidos 809 artigos da base *Scopus*, 608 artigos da base *Web of Science* e 16 artigos da base *Science Direct*, totalizando 1.433 publicações. Após a exclusão dos artigos duplicados, o conjunto resultou em 892 artigos. A leitura de todos os títulos foi feita e foram rejeitados aqueles que não atenderam ao tema proposto, resultando em 292 artigos. Posteriormente, os resumos destes 292 artigos foram lidos e metade desse montante não foi aprovado, resultando em 146 artigos que, após lidos e avaliados pela pertinência do conteúdo de forma geral, reduziram o conjunto para 83 artigos, devido a exclusão de 63 artigos. Por fim, os artigos foram baixados em formato PDF e os que não tiveram acesso disponível foram rejeitados (38 artigos), chegando, assim, ao montante final de 45 publicações. Por fim, os 45 textos foram relidos integralmente, aplicando a etapa de qualificação e obtendo o portfólio ideal com 43 artigos para extração das primeiras informações desta revisão sistemática.

Finalizado o portfólio, uma parte importante subsequente é a elaboração do fichamento de cada um dos artigos. O fichamento permitiu organizar, classificar e evidenciar os autores envolvidos na publicação, o ano de sua publicação, o nome do periódico, o país de origem dos autores e do periódico, as citações das bases indexadas, a qualificação feita via Parsifal (2021) e, por fim, indicar sucintamente quais são os objetivos e o método utilizado pelos autores.

Desta forma, outras análises podem ser realizadas, tal como a Lei do Elitismo de Price que estabelece que o número de autores considerados “elite”, dentro de um campo acadêmico, compreende a raiz quadrada do total de autores do portfólio evidenciado e que, a metade da quantidade de artigos selecionados é considerado o critério para saber se essa elite é tida como produtiva (ARAÚJO, 2007).

Outra análise importante são as palavras-chave. Nesta revisão sistemática de literatura, a nuvem de palavras (Figura 6) foi a forma escolhida de estudo pelo fato de expor as palavras usadas de forma simples, didática e direta. No centro da nuvem, estão destacadas as palavras mais recorrentes. A localização, a espessura e o tamanho da fonte que cada palavra possui dentro da nuvem revelam a sua importância. Bueno, Banuls e Gallego (2021) explicam que a espessura é compatível com a força de relações que a palavra detém dentro da pesquisa, o tamanho de sua fonte caracteriza a frequência no qual ela é empregada dentro dos textos e a localização unifica esses parâmetros.

3. RESULTADOS E DISCUSSÃO

Este item apresenta os resultados da revisão sistemática de literatura, incluindo os metadados evidenciados (ano de publicação, país de análise, título de periódicos, citações por autor e as principais teorias e tópicos de pesquisa desses artigos).

3.1 Artigos do portfólio

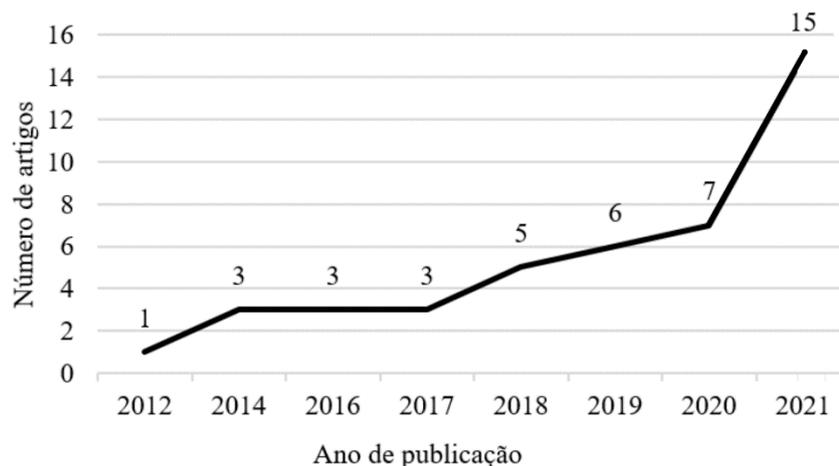
Na Tabela 15 (Apêndice 1), observam-se os 43 artigos do portfólio com as pontuações obtidas de acordo com os critérios de qualidade estabelecidos. Podemos observar que, em todos os artigos, o objetivo é a elaboração de um mapa de risco ou de suscetibilidade de áreas com problemas de inundações e alagamentos, devidos todos os danos e prejuízos urbanos, ambientais e socioeconômicos causados (SARHADI, SOLTANI e MODARRES, 2012).

É possível identificar, também, os métodos de Aprendizagem de Máquina utilizados em cada uma das publicações, apontando os caminhos metodológicos que os pesquisadores trilharam. No montante final do portfólio existem artigos bastante citados, porém esses não foram obrigatoriamente os melhores qualificados. A justificativa para essa ampla variação das notas finais dos artigos mais citados dá-se pelo fato de que as perguntas vitais para o estudo proposto não foram abordadas ou respondidas pelos artigos.

3.2 Evolução das publicações sobre o tema

A Figura 5 mostra a distribuição das publicações acerca do tema de inundações, dentro do portfólio construído previamente. Observa-se uma tendência crescente do número de pesquisas sobre inundações urbanas com o passar dos anos. Este padrão evidencia um grande interesse sobre o tema no meio acadêmico e indica a importância do assunto entre os problemas cotidianos da população global. De 2014 até 2017, o assunto aparece de forma bastante tímida (3 publicações), a partir de 2018 e até 2020 adquire mais relevância, com 5 a 7 publicações. O auge de publicações ocorre em 2021 (15 publicações), apontando a indiscutível importância do tema na atualidade.

Figura 5: Evolução das publicações sobre o tema



Fonte: Elaborado pela autora

3.3 Autores, filiações e países de origem

A seleção de autores de maior destaque no portfólio, por número de publicações, é apresentada na Tabela 2. É possível apontar a produção dos autores, contabilizando o montante de artigos pelo número de autores: 165 autores para 43 artigos, obtendo uma média de 3,84 autores por artigo. A Tabela 2 exhibe uma amostra do total de autores que atuaram nas publicações e evidencia os que detêm duas ou mais colaborações em artigos diversos.

Tabela 2: Autores relevantes do portfólio

Autor	Filiação do autor	País	Número de publicações
Biswajeet Pradhan	University of Technology	Austrália	6
Saro Lee	Korea Institute of Geoscience and Mineral Resources	Coreia do Sul	3
Ahmed M. Youssef	Sohag University	Egito	2
Ali Torabi Haghghia	University of Oulu	Finlândia	2
Bahram Choubin	University of Tehran	Irã	2
Costas Armenakis	York University	Canadá	2
Hamid Darabi	Esfarayen University of Technology	Irã	2
Hamid Reza Pourghasemi	Shiraz University	Irã	2
Jan Adamowski	McGill University	Canadá	2
Mahyat Shafapour Tehrany	Bogazici University	Turquia	2
Moung-Jin Lee	Korea Environment Institute	Coreia do Sul	2
Omid Rahmati	University of Tehran	Irã	2
Saleh A. Sefry	Saudi Geological Survey	Arábia Saudita	2
Soha A. Mohamed	Faculty of Computer Science & Artificial	Egito	2

Autor	Filiação do autor	País	Número de publicações
	Intelligence		

Fonte: Elaborado pela autora

Destacando os números extraídos no portfólio foi possível analisar a efetividade da produção do número total de autores e observar que cerca de 91,7% dos autores produziram ou participaram de somente um único artigo do portfólio. Caminhamos em direção à Lei do Elitismo de Price, visto que o número total de autores mais relevantes em produção, salientados no portfólio, é de 14 artigos, ou seja, acima do que prevê a lei neste caso (13 artigos). A produção média dos artigos feitos pela “elite” não se sobressai da média dos artigos evidenciados, pois foram contabilizados 14 artigos realizados pela “elite”, cerca de 31,8%, e não 22 artigos, como prevê a lei (ARAÚJO, 2007). A conclusão que podemos chegar é que há uma elite acadêmica que investiga o tema com profundidade, mas que ainda não detém publicações o suficiente para denominar-se como uma elite completa, segundo a Lei do Elitismo de Price.

A Tabela 3, apresentada a seguir, lista uma amostra dos países de origem dos pesquisadores do portfólio. Podemos observar que 10 países diferentes em todo globo tiveram autores que publicaram pelo menos 2 artigos sobre o tema inundações dentro do portfólio; destaca-se o continente asiático com 14 publicações, com ênfase para a China e o Irã, com 5 publicações cada. O continente americano vem logo após com um total de 8 publicações, o Canadá com 2 publicações e o Brasil e Estados Unidos da América com 2 artigos. O terceiro continente com mais autores é a Europa com 4 artigos publicados. Com isso podemos concluir que não há predomínio de conhecimento em somente um continente e que a produção deste tema interessa a uma grande quantidade de países no mundo e é objeto de publicações de qualidade.

Tabela 3: Países de maior destaque em publicações

País de origem do autor principal	Total
China	5
Irã	5
Egito	3
Canadá	3
Estados Unidos da América	2
Itália	2
Grécia	2
Malásia	2
Brasil	2
Arábia Saudita	2

Fonte: Elaborado pela autora

Pangali Sharma (2021) explica que, somente no continente asiático, as inundações repentinas provocaram a morte de mais de 22.000 pessoas e o prejuízo de mais de 34 bilhões de dólares em danos. O grande número de pesquisadores oriundos da Ásia explica-se pela necessidade de pesquisas visando medidas de prevenção de eventos climáticos extremos que precocemente ceifam vidas e para evitar prejuízos financeiros, corroborado pela World Meteorological Organization (2023) que atesta um aumento significativo de eventos extremos ligados à água no continente com mais de 50 milhões de pessoas afetadas em 2022. Os países desenvolvidos, como os Estados Unidos da América e o Canadá, também sofrem com eventos extremos, principalmente com as inundações e alagamentos e, de acordo com Hossain e Meng (2020), o problema se deve às extensas superfícies impermeáveis e ao fato de as inundações serem negligenciadas no planejamento urbano.

3.4 Análise das palavras-chave

As palavras-chave mais citadas são apresentadas na Figura 6, na qual se observa a ocorrência na busca inicial nas bases indexadas, o que de certa forma valida o processo de seleção do referencial.

Podemos destacar a palavra “inundação” como o termo com mais representatividade acerca do tema da pesquisa, vista a sua localização central, seu tamanho de fonte muito superior aos outros termos e sua espessura bastante acentuada. A importância da palavra-chave “inundação” já era esperada; os termos adjacentes que podem evidenciar novas discussões podem ser separados por grupos para compreendê-las melhor. Nessa nuvem de palavras há alguns termos que se referem ao “risco”, “vulnerabilidade” e “susceptibilidade”. Depreende-se também que é bastante abordada nas palavras-chave a questão da metodologia adotada, o que é constatado pelos termos: “sig”, “modelagem”, “modelo”, “mapa”, “ahp”, “neural”, “floresta”, “análise”, “multicritério”, “sensoriamento” e “remoto”, por exemplo. O método “ahp”, Processo de Hierarquia Analítica, caracteriza-se como um método multicritério de tomada de decisão a partir de medições relativas que têm por objetivo integrar alguns critérios e levar a escolhas justificadas de forma racional e sistemática (FAREGH; BENKHALED, 2021). Já o termo citado “neural” representa o método Rede Neural Artificial (RNA) que é uma abordagem de Aprendizagem de Máquina que se inspira no cérebro humano como base para elaborar algoritmos de relações e consegue identificar quais são as vulnerabilidades do local de acordo com os critérios (TAMIRU; DINKA, 2021). A palavra “floresta” remete ao

total de artigos selecionados. Desta forma, é observado que não há um predomínio do conhecimento do tema por parte dos periódicos, visto que os artigos estão distribuídos de forma mais uniforme já que 22 artigos estão em 22 periódicos diferentes, restando apenas 8 periódicos com 2 ou mais publicações.

O destaque fica para a revista científica *Natural Hazards*, periódico holandês de alto fator de impacto com 5 publicações e que tem por objetivo evidenciar estudos que abordem os aspectos dos eventos extremos acelerados, os riscos desses eventos, as previsões e quais são seus precursores. Além disso, explicita as interações entre esses problemas com a sociedade humana, tratando sobre governança, resposta as ocorrências e ações preventivas, de mitigação e adaptação. Outro periódico importante dentro do portfólio é a revista científica *Water*, de origem suíça, e que aborda a investigação dos setores hídricos tais como gestão, sistemas, monitoramento, proteção, planejamento e outros temas que envolvam água. Por fim, a revista científica *Sustainability*, também suíça, tem por prioridade a publicação sobre temas relacionados com a sustentabilidade ambiental, cultural, econômica e social, aprofundando temáticas voltadas ao desenvolvimento sustentável do planeta. Todos esses periódicos têm vertentes ligadas ao desenvolvimento sustentável e são relacionadas com a proteção do meio ambiente, por isso, são inevitavelmente multidisciplinares, pois contam com artigos com vários enfoques para uma mesma finalidade: o bem-estar desta e da próxima geração.

Tabela 4: Características dos periódicos presentes no portfólio

Periódico	Número de artigos	País de origem do periódico	F.I. JCR (WoS)	F.I. SJR (Scopus)
Natural Hazards	5	Holanda	3.102	0.76
Water	4	Suíça	3.103	0.72
Sustainability	3	Suíça	3.251	0.61
Environmental Earth Sciences	2	Alemanha	2.784	0.64
Geomatics, Natural Hazards and Risk	2	Reino Unido	3.528	0.9
Geosciences (Switzerland)	2	Suíça	0.69	0.61
Journal of Hydrology	2	Holanda	5.722	1.68
Remote Sensing	2	Suíça	4.848	1.29

Fonte: Elaborado pela autora

3.6 Análise dos procedimentos metodológicos

Neste item serão evidenciados os procedimentos metodológicos observados nas publicações do portfólio. Verificou-se o equilíbrio unânime da orientação

metodológica dos artigos: todos eles são estudos empíricos. A razão para este fato dá-se pela necessidade de evidenciar alguns conceitos no corpo do texto para que se possa justificar o uso da metodologia escolhida, já que as publicações almejam os métodos mais acurados mediante metodologias claras para replicação em diversas localidades. Dessa forma, é plausível afirmar que os artigos seguem uma ordenação lógica: introdução, conceitos, procedimentos metodológicos, dados e processos, resultados e discussões e, por fim, considerações finais.

A Tabela 5 aponta os métodos mais utilizados em todas as publicações selecionadas. Nos 43 estudos do portfólio, foram encontradas 16 metodologias diferentes para a promoção do mapeamento de suscetibilidade. Dos 16 métodos mencionados, foram descartados os que foram unicamente usados ou utilizados de forma experimental. Mantiveram-se os métodos de uso mais frequentes, chegando a um total de 8 métodos de Aprendizagem de Máquina mais empregados. Cada um dos artigos aplicou um processo metodológico que poderia contar com um ou mais métodos de *Machine Learning* somado com grupos de fatores de entrada definidos pelos autores, ou seja, alguns dados se mostraram imprescindíveis para o mapeamento de inundações e outros foram adicionados de forma a complementar os resultados almejando entender a complexidade da temática. Sobre os fatores de entrada utilizados no mapeamento, os 10 mais usados foram: elevação, declividade, distância até o rio, uso e cobertura do solo, litologia e geologia, precipitação, curvatura do solo, Índice de potência de fluxo (SPI), Índice de umidade topográfica (TWI) e drenagem.

Tabela 5: Métodos de Aprendizagem de Máquina e fatores de entrada

Fatores de Entrada	Métodos								Total
	SVM (1)	AHP (2)	MCE (3)	RNA (4)	RS + GIS (5)	LR (6)	RF (7)	FR (8)	
Elevação	2	6	8	3	5	4	3	2	33
Declividade	2	6	5	2	4	3	3	2	27
Distância até o rio	2	5	4	2	2	4	3	1	23
Uso/cobertura do solo (LULC)	2	5	6	1	2	2	3	2	23
Litologia e geologia	2	3	4	2	1	2	1	2	17
Precipitação		4	2	2	2	2	1	2	15
Curvatura do solo	2	1	-	2	1	3	2	-	11
Índice de potência de fluxo (SPI)	2	1	1	2	-	2	1	-	9
Índice de umidade topográfica (TWI)	2	1	-	1	-	1	1	1	7

Fatores de Entrada	Métodos								Total
	SVM (1)	AHP (2)	MCE (3)	RNA (4)	RS + GIS (5)	LR (6)	RF (7)	FR (8)	
Drenagem	1	2	1	-	-	-	-	2	6
Escoamento superficial	2	1	3	-	-	-	-	-	6
Dados históricos	-	-	2	1	1	-	1	-	5
Edificações	-	-	-	-	3	1	1	-	5
Sistema viário	-	-	-	-	3	-	1	-	5
Vegetação	-	-	-	1	1	-	2	-	4
Aspecto	2	-	-	-	-	-	1	-	3
Hidrologia do solo	1	1	1	-	-	-	-	-	3
Dados meteorológicos	-	-	-	-	1	-	1	-	2
Distância até o esgoto	-	-	1	-	1	-	-	-	2
Entrevista/ mídias digitais	-	1	1	-	-	-	-	-	2
Índice de vegetação por diferença normalizada (NDVI)	-	-	-	2	-	-	-	-	2
Profundidade da inundação	-	1	1	-	-	1	-	-	2
Profundidade do solo	1	-	1	-	-	-	-	-	2
Antropização	-	-	-	-	-	-	1	-	1
Duração da inundação	-	-	-	-	-	1	-	-	1
Fator de erosividade da chuva (fator R)	-	-	-	1	-	-	-	-	1
Temperatura/ clima	-	-	-	-	-	1	-	-	1

- (1) SVM: Máquina de Vetores de Suporte;
(2) AHP: Processo Hierárquico Analítico;
(3) MCE: Análise de Decisão Multicritério;
(4) RNA: Rede Neural Artificial;
(5) RS: Sensoriamento Remoto; GIS: Sistema de Informação Geográfica;
(6) LR: Regressão Logística;
(7) RF: Floresta Aleatória; e
(8) FR: Relação de Frequência.

Fonte: Elaborado pela autora

3.7 Lacunas e Oportunidades de Pesquisa

Neste item apresentamos as lacunas e as possíveis oportunidades de pesquisa encontradas no portfólio e que precisam ser trabalhadas para que o tema inundações avance.

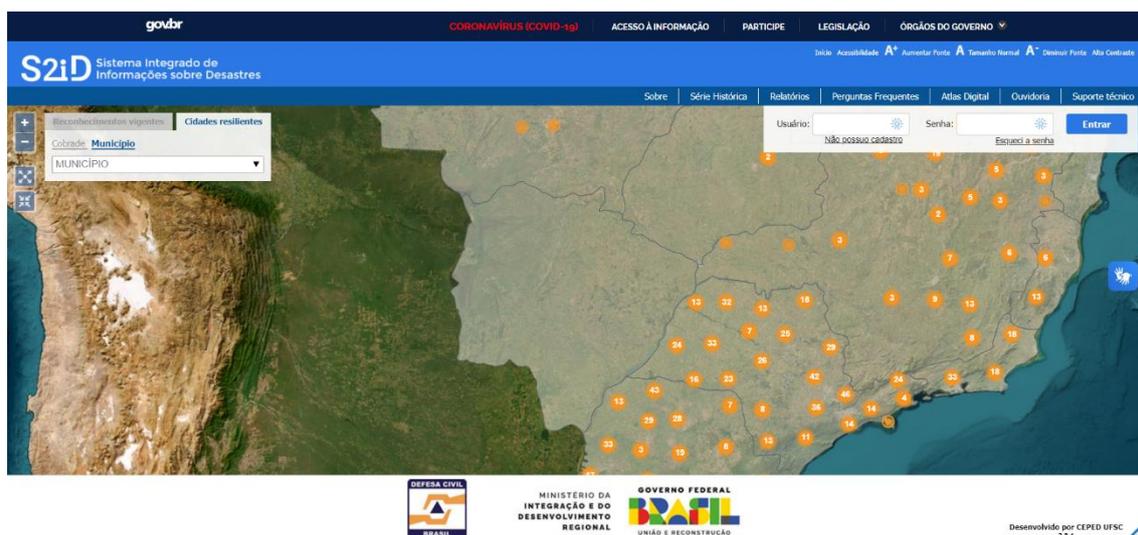
A grande lacuna observada pelos pesquisadores ao aplicarem as metodologias escolhidas dá-se pela precariedade dos dados levantados, visto que, em muitas pesquisas, se mostraram insuficientes. Chen et al. (2020) argumenta que uma grande limitação encontrada em sua pesquisa foi a disponibilidade de dados para

aplicação da metodologia Rede Neural, dessa forma, enfatiza a oportunidade de pesquisa para providenciar tais dados e ampliar a capacidade do método para prever inundações. Complementando, Waqas et al. (2021) explica que novos estudos sobre estimativa, previsão e mapeamento realizado com novos dados e outros modelos podem contribuir para o mapeamento de vulnerabilidade de inundações. Outra lacuna que cabe investigação mais aprofundada dá-se na questão da integração de GIS com novos modelos de Aprendizagem de Máquina, além de uma comparação entre os modelos de decisões de multicritério e Aprendizagem de Máquina para evidenciar quais são seus pontos fortes e suas fraquezas como método (TELLA E BALOGUN, 2020).

3.7.1 Lacunas em Mato Grosso do Sul

No *site* oficial do Governo Federal, Sistema Integrado de Informações sobre Desastres, o estado de Mato Grosso do Sul não possui nenhum município cadastrado na Iniciativa Construindo Cidades Resilientes 2030 (MCR2030), mostrada na Figura 7. Iniciativa liderada pelo Escritório das Nações Unidas para a Redução de Riscos (UNDRR), que tem por objetivo promover a resiliência local através de conhecimento e experiências, ações políticas e troca de aprendizagem entre cidades para reduzir riscos e possibilitar resiliência urbana, promovendo cidades inclusivas e seguras e, ainda, contribuindo para com os Objetivos de Desenvolvimento Sustentável (BRASIL, 2023a).

Figura 7: Sistema Integrado de Informação sobre Desastres

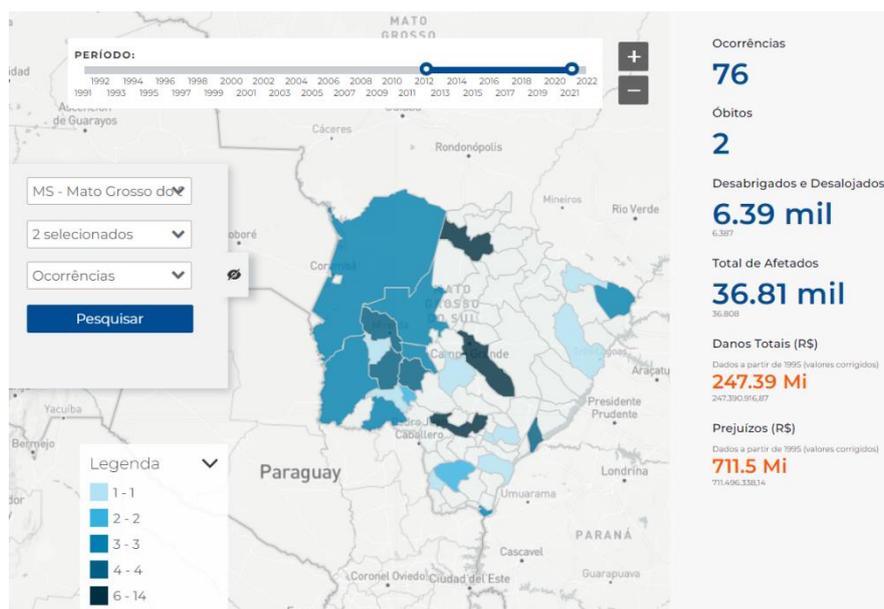


Fonte: Brasil (2023b)

Outra lacuna encontrada advém de uma cooperação entre o Banco Mundial e a Universidade de Santa Catarina que proporcionaram a aglutinação de ocorrências de desastres por meio do Atlas Digital de Desastres no Brasil; essa ferramenta aponta

ocorrências no território brasileiro cronologicamente. Dessa forma, ao observar o estado de Mato Grosso do Sul, na Figura 8, verificamos que nos anos 2012 a 2021 houve 76 ocorrências de inundações e alagamentos em todo o estado, com um montante de aproximadamente 37 mil pessoas afetadas. O município de Campo Grande, de acordo com o Atlas, teve um total de 6 ocorrências durante os 9 anos selecionados (ATLAS DIGITAL, 2023).

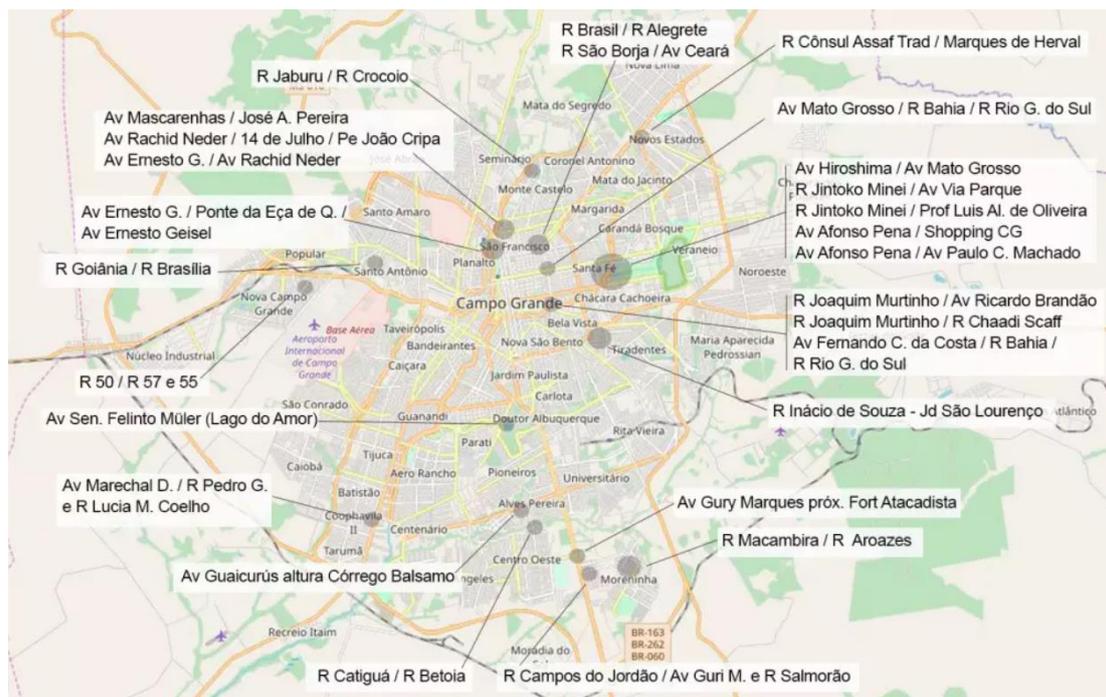
Figura 8: Atlas Digital de Desastres no Brasil



Fonte: Atlas Digital (2023)

A Proteção e Defesa Civil de Campo Grande disponibiliza um documento de Setorização de áreas em alto e muito alto risco a movimentos de massa, enchentes e inundações realizado em julho de 2019 pelo Repositório Institucional de Geociências. No ano seguinte, 2020, uma matéria realizada pelo veículo de imprensa Campo Grande News divulgou um mapa com pontos críticos (Figura 9) que podem inundar ou alagar segundo a Defesa Civil Municipal (CAMPO GRANDE NEWS, 2020).

Figura 9: Mapa de pontos críticos monitorados pela Defesa Civil



Fonte: Campo Grande News (2020)

Confirmando os apontamentos feito pelos autores do portfólio, há uma grande lacuna na atualização de dados sobre inundações e alagamentos. Os dados encontram-se desatualizados na cidade de Campo Grande (MS) e confirmam a necessidade de novas investigações e pesquisas sobre o tema.

4. CONCLUSÕES

Este capítulo discutiu a Revisão Sistemática da Literatura em destaque sobre o tema, ou seja, artigos e artigos de revisão dentro do recorte temporal de 2012 a 2021.

A conceituação sobre um dos eventos extremos mais frequentes do globo, explicados com o auxílio do portfólio levantado, trouxe contribuições sobre a importância e urgência do tema, informou sobre as possíveis causas mais comuns que promovem as inundações em meio urbano e salientou os imensos prejuízos financeiros e socioambientais que decorrem de eventos climáticos extremos. Além disso, os resultados apontam as metodologias Máquina de Vetores de Suporte (SVM), Processo Hierárquico Analítico (AHP), Análise de Decisão Multicritério (MCE), Rede Neural Artificial (RNA), Regressão Logística (LR), Floresta Aleatória (RF), RS (Remote Sensing e GIS) e Relação de Frequência (FR) como as mais usadas para promover o mapeamento dessas áreas conflituosas, principalmente zonas que detém irregularidades no uso do solo e retirada da

cobertura vegetal e, por fim, quais são os requisitos para um mapeamento de qualidade, ou seja, quais os fatores de entrada são fundamentais para estudar as inundações urbanas.

5. Referências

- ARAÚJO, C. A. A. **Bibliometria: evolução histórica e questões atuais**. Em *Questão*, v. 12, n. 1, 2007
- ARMENAKIS, Costas et al. **Flood risk assessment in urban areas based on spatial analytics and social factors**. *Geosciences*, v. 7, n. 4, p. 123, 2017.
- ATLAS DIGITAL. **Mapa Interativo**. Disponível em: <http://atlasdigital.mdr.gov.br/paginas/mapa-interativo.xhtml>. Acesso em: 3 de outubro de 2023.
- BAKKALBASI, N. et al. **Three options for citation tracking: Google Scholar, Scopus and Web of Science**. *Biomedical digital libraries*, v. 3, n. 1, p. 1-8, 2006.
- BANDARA, W. et al. **Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support**. *Communications of the Association for Information systems*, v. 37, n. 1, p. 8, 2015.
- BRASIL. 2023a. Ministério do Desenvolvimento Regional. **Proteção e Defesa Civil. Cidades Resilientes**. Disponível em: <https://www.gov.br/mdr/pt-br/assuntos/protecao-e-defesa-civil/cidades-resilientes>. Acesso em: 3 de outubro de 2023.
- BRASIL. 2023b. S2iD - **Sistema de Informações Integradas de Defesa Social**. Disponível em: <https://s2id.mi.gov.br/paginas/index.xhtml#>. Acesso em: 3 de outubro de 2023.
- BUENO, S.; BANULS, V. A.; GALLEGO, M. D. **Is urban resilience a phenomenon on the rise? A systematic literature review for the years 2019 and 2020 using textometry**. *International Journal of Disaster Risk Reduction*, v. 66, p. 14, Dec 2021.
- CAMPO GRANDE NEWS. **Mapeamento indica 33 pontos críticos de alagamento em Campo Grande**. Disponível em: <https://www.campograndenews.com.br/cidades/capital/mapeamento-indica-33-pontos-criticos-de-alagamento-em-campo-grande>. Acesso em: 3 de out. 2023.
- CHEN, Junfei et al. **A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: a case study of the Yangtze River Delta, China**. *International journal of environmental research and public health*, v. 17, n. 1, p. 49, 2020.
- CHOUBIN, Bahram et al. **An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines**. *Science of the Total Environment*, v. 651, p. 2087-2096, 2019.

- DARABI, Hamid et al. **Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques.** Journal of hydrology, v. 569, p. 142-154, 2019.
- DE JESUS-LOPES, J. C.; MACIEL, W. R. E.; CASAGRANDA, Y. G. **Check-List Dos Elementos Constituintes Dos Delineamentos Das Pesquisas Científicas.** Desafio Online, v. 10, n. 1, 2022.
- FAREGH, Wail; BENKHALED, Abdelkader. **GIS-based multicriteria approach for flood risk assessment in Sigus city, east Algeria.** Arabian Journal of Geosciences, v. 14, n. 12, p. 1-9, 2021.
- FRANCI, Francesca et al. **Satellite remote sensing and GIS-based multi-criteria analysis for flood hazard mapping.** Natural Hazards, v. 83, n. 1, p. 31-51, 2016.
- HOSSAIN, Mohammad Khalid; MENG, Qingmin. **A fine-scale spatial analytics of the assessment and mapping of buildings and population at different risk levels of urban flood.** Land use policy, v. 99, p. 104829, 2020.
- HUSSAIN, Muhammad et al. **GIS-Based Multi-Criteria Approach for Flood Vulnerability Assessment and Mapping in District Shangla: Khyber Pakhtunkhwa, Pakistan.** Sustainability, v. 13, n. 6, p. 3126, 2021.
- JALALI, S.; WOHLIN, C. **Systematic literature studies: database searches vs. backward snowballing.** Proceedings of the 2012 ACM-IEEE international symposium on empirical software engineering and measurement, 2012. IEEE. p.29-38.
- KARYMBALIS, Efthimios et al. **Integration of GIS-Based Multicriteria Decision Analysis and Analytic Hierarchy Process for Flood-Hazard Assessment in the Megalo Rema River Catchment (East Attica, Greece).** Sustainability, v. 13, n. 18, p. 10232, 2021.
- KEELE, Staffs et al. **Guidelines for performing systematic literature reviews in software engineering.** 2007.
- LEE, Sunmin et al. **Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea.** Geomatics, Natural Hazards and Risk, v. 8, n. 2, p. 1185-1203, 2017.
- MARTÍN-MARTÍN, A. et al. **Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories.** Journal of Informetrics, v. 12, n. 4, p. 1160-1177, 2018.
- PANGALI SHARMA, Til Prasad et al. **A Geomorphic Approach for Identifying Flash Flood Potential Areas in the East Rapti River Basin of Nepal.** ISPRS International Journal of Geo-Information, v. 10, n. 4, p. 247, 2021.
- PARSIFAL. **Parsifal v.2.1.1: Perform Systematic Literature Reviews.** 2021. Disponível em: < <https://parsif.al/> >. Acesso em: 30 de mar. de 2022.

- PETTICREW, M.; ROBERTS, H. **Systematic reviews in the social sciences: A practical guide**. Oxford: John Wiley & Sons, 2008. 336 ISBN 1405150149.
- PROTEÇÃO E DEFESA CIVIL. **Planos de Contingência**. Campo Grande (MS). Disponível em: <https://protecaodefesacivil.campogrande.ms.gov.br/planos-de-contingencia/>. Acesso em: 3 de outubro de 2023.
- RAHMATI, Omid et al. **Urban flood hazard modeling using self-organizing map neural network**. *Water*, v. 11, n. 11, p. 2370, 2019.
- RIGEO. **Setorização de áreas em alto e muito alto risco a movimentos de massa, enchentes e inundações: Campo Grande, Mato Grosso do Sul**. Disponível em: <https://rigeo.sgb.gov.br/handle/doc/19545>. Acesso em: 3 de out. 2023.
- SARHADI, Ali; SOLTANI, Saeed; MODARRES, Reza. **Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis**. *Journal of Hydrology*, v. 458, p. 68-86, 2012.
- SKILODIMOU, Hariklia D.; BATHRELLOS, George D.; ALEXAKIS, Dimitrios E. **Flood hazard assessment mapping in burned and urban areas**. *Sustainability*, v. 13, n. 8, p. 4455, 2021.
- SNYDER, H. **Literature review as a research methodology: An overview and guidelines**. *Journal of business research*, v. 104, p. 333-339, 2019.
- TAMIRU, Habtamu; DINKA, Megersa O. **Artificial Intelligence in Geospatial Analysis for Flood Vulnerability Assessment: A Case of Dire Dawa Watershed, Awash Basin, Ethiopia**. *The Scientific World Journal*, v. 2021, 2021.
- TEHRANY, Mahyat Shafapour et al. **Flood susceptibility mapping using integrated bivariate and multivariate statistical models**. *Environmental earth sciences*, v. 72, n. 10, p. 4001-4015, 2014.
- TELLA, Abdulwaheed; BALOGUN, Abdul-Lateef. **Ensemble fuzzy MCDM for spatial assessment of flood susceptibility in Ibadan, Nigeria**. *Natural Hazards*, v. 104, n. 3, p. 2277-2306, 2020.
- TIEN BUI, Duie et al. **Flood spatial modeling in northern Iran using remote sensing and gis: A comparison between evidential belief functions and its ensemble with a multivariate logistic regression model**. *Remote Sensing*, v. 11, n. 13, p. 1589, 2019.
- WAQAS, Hassan et al. **Flash flood susceptibility assessment and zonation using an integrating analytic hierarchy process and frequency ratio model for the Chitral District, Khyber Pakhtunkhwa, Pakistan**. *Water*, v. 13, n. 12, p. 1650, 2021.
- WORLD METEOROLOGICAL ORGANIZATION (WMO). **Climate change impacts increase in Asia**. 2023. Disponível em: <https://wmo.int/news/media-centre/climate-change-impacts-increase-asia>. Acesso em: 11 de maio de 2024.

- YEGANEH, Nasim; SABRI, Soheil. **Flood vulnerability assessment in Iskandar Malaysia using multi-criteria evaluation and fuzzy logic**. Research Journal of Applied Sciences, Engineering and Technology, v. 8, n. 16, p. 1794-1806, 2014.
- YOUSSEF, Ahmed M.; PRADHAN, Biswajeet; SEFRY, Saleh A. **Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models**. Environmental Earth Sciences, v. 75, n. 1, p. 1-16, 2016.

SEGUNDO CAPÍTULO: MAPEAMENTO DE INUNDAÇÕES E ALAGAMENTOS BASEADO EM APRENDIZAGEM DE MÁQUINA A PARTIR DE RELATOS DA IMPRENSA

Resumo: As inundações em zonas urbanas acontecem de forma frequente e geram graves consequências para a sociedade, causando prejuízos socioeconômicos e ambientais. O enfrentamento dessa problemática sugere a produção de um mapeamento de suscetibilidade de inundações e alagamentos. A cidade de Campo Grande (MS) registrou 1.100 pontos de inundação e alagamento durante 9 (nove) anos amostrados, segundo a participação da população por meio de relatos de imprensa. Este trabalho demonstra a utilização de *Machine Learning* para mapear zonas potenciais de inundação e alagamento em perímetro urbano. Utilizando o método *Gradient Boosting Classifier* (GBC) verificamos que é possível aplicar *Machine Learning* para o mapeamento de suscetibilidade mediante o uso de localizações noticiadas pelos veículos de imprensa e fatores de entrada que caracterizam a área de estudo, estimando-se o número de ocorrências por ano de eventos de inundação e alagamentos em todo o perímetro urbano. O número total de ocorrências (eventos) simulados pelo GBC foi de 433 ocorrências/ano na área urbana de Campo Grande (MS) e se mostrou uma ferramenta com potencial de utilização. As descobertas feitas servem como ferramenta para políticas públicas e de planejamento urbano para mitigação e adaptação das áreas problemáticas.

Palavras-chave: Sociohidrologia, Eventos Extremos, Inundação Urbana, *Gradient Boosting Classifier*

Abstract: Floods in urban areas occur frequently and have severe consequences for society, causing socio-economic and environmental damage. Addressing this issue suggests producing a susceptibility mapping of floods and inundations. The city of Campo Grande (MS) recorded 1.100 flood and inundation points over nine sampled years, according to population participation through press reports. This work demonstrates the use of Machine Learning to map potential flood and inundation zones in urban areas. Using the Gradient Boosting Classifier (GBC) method, we found that it is possible to apply Machine Learning for susceptibility mapping by using locations reported by the press and input factors that characterize the study area, estimating the number of flood and inundation events per year across the entire urban perimeter. The total number of simulated occurrences (events) by the GBC was 433 occurrences per year in the urban area of Campo Grande (MS), and it proved to be a tool with potential use. The findings serve as a tool for public policy and urban planning for mitigation and adaptation in problematic areas.

Keywords: Sociohydrology, Extreme Events, Urban Flooding, Gradient Boosting Classifier

1. INTRODUÇÃO

Notícias são capazes de gerar recursos fundamentais para transmitir rapidamente as informações coletadas em campo (OKADA et al., 2021). O uso desse método de coleta de informações para pesquisas que abordam análises climáticas é bastante difundido e apontado como bem-sucedido pois é eficiente e conveniente (LIMA e AMORIM, 2015). Com os dados coletados e catalogados, advindos das mídias de notícias, é possível empregar o treinamento das informações em poucas horas (OKADA et al., 2021).

O uso de mídias sociais se apresenta como uma ferramenta de estudo de atividades humanas bastante promissoras devido seu uso difundido. Plataformas digitais e mídias sociais permitem que os pesquisadores estudem atividades humanas e naturais de forma bastante oportuna, vista a facilidade de colher de dados sobre crises, monitorar e reconhecer riscos naturais e coletar informações importantes para contribuir nos processos de planejamento urbano. Porém as redes sociais ainda detêm um caráter complexo para sua ampla utilização, principalmente por serem influenciadas pela desigualdade socioeconômica, pelas estruturas sociais e pela complexidade da topologia urbana (DE ANDRADE et al.; CAMARGO et al., 2021), às quais acrescentaríamos também os fatores político-ideológicos.

A metodologia chamada Aprendizagem de Máquina, *Machine Learning*, em inglês, é utilizada para descrever métodos de treinamentos organizados que possibilitam maior flexibilidade nas suposições de dados. Mediante o uso dessa metodologia, torna-se possível viabilizar previsões de inundações e alagamentos urbanos com mais rapidez e permitindo que esses dados sejam aplicados em avaliações e em mapas de suscetibilidade (RAHMATI et al., 2019; GABRIELS, WILLEMS e VAN ORSHOVEN, 2020). O mapeamento de suscetibilidade ou mapeamento de risco é um instrumento imprescindível para desenvolver um ordenamento e planejamento urbano eficientes, principalmente pelo fato de investigar a quantificação das zonas com potencial para inundações e alagamentos e a propagação deste risco, baseando na vulnerabilidade do local e seu entorno (FAREGH e BENKHALED, 2021). Com isso, é possível um mapeamento de vulnerabilidade física, socioeconômica e capacidade de enfrentamento do risco, através de estratégias traçadas para promoção de políticas públicas e planos de mitigação e gestão de riscos (HUSSAIN et al., 2021).

Somado a isso, o estudo se mostra alinhado com os Objetivos de Desenvolvimento Sustentável no Brasil da Organizações das Nações Unidas em pelo menos 4 objetivos: 3 – Saúde e bem-estar, 6 – Água potável e saneamento, 11 – Cidades e comunidades sustentáveis e, 13 – Ação contra a mudança global do clima.

O alinhamento se deve ao fato desta pesquisa, de caráter inédita, se debruçar nos estudos de acometimento do tecido urbano pelas inundações e alagamentos, elucidando diversas problemáticas por meio da percepção através do conhecimento social explicitada pela população local em relatos de imprensa, aproximando a comunidade da gestão municipal e visando a melhoria das políticas públicas vigentes a fim de melhorar a qualidade de vida dessa população.

O objetivo central deste capítulo é identificar, colher e analisar informações de zonas com problemas de inundação e alagamento na malha urbana do município de Campo Grande (MS), com a finalidade de reunir os casos denunciados pelos munícipes e noticiados pelos veículos da imprensa local Campo Grande News e G1 MS, dentro do recorte temporal de nove (9) anos (2012 a 2021). Posteriormente, efetuar o treinamento e a validação de nove (9) métodos de *Machine Learning* para viabilização da produção de um mapeamento de ocorrências anuais de inundações e alagamentos para toda a área urbana, observando a melhor acurácia a fim de mapear áreas já relatadas e não relatadas pela imprensa.

2. MATERIAL E MÉTODO

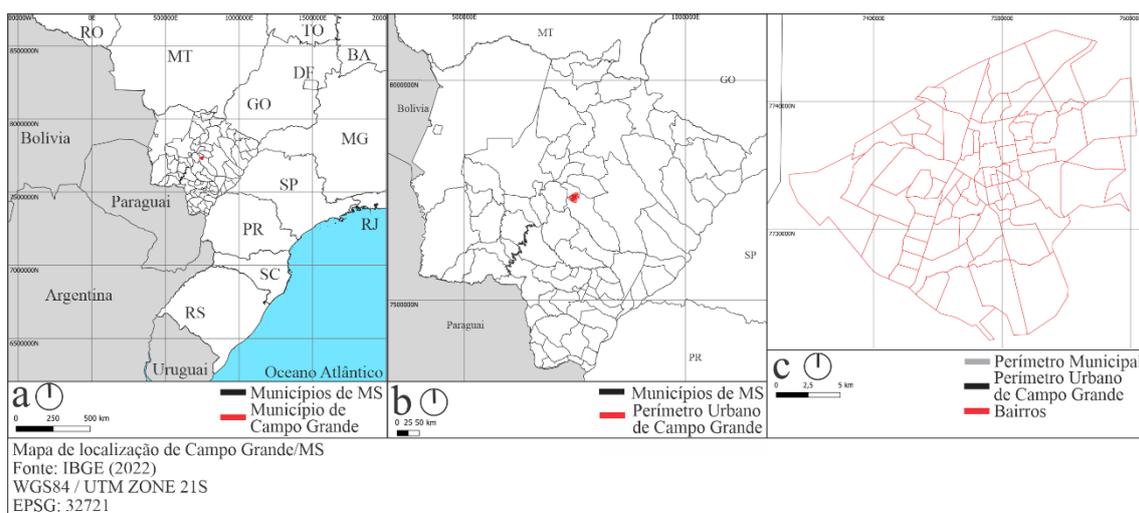
Neste item trataremos os processos de coleta de dados referentes à área de estudo, dos dados secundários utilizados e da realização da codificação em ambiente Python.

2.1 Área de estudo

A cidade escolhida para este estudo trata da capital do estado de Mato Grosso do Sul, Campo Grande (Figura 10b). Sua escolha se deu pela importância do município capital do estado de Mato Grosso do Sul, pelo histórico de inundações e alagamentos que prejudicam a qualidade de vida dos moradores dessa cidade e pela quantidade reduzida de estudos realizados sobre o tema neste local, sendo o último realizado em 2020.

O município conta com uma área total de 8.082,97 km² localizada no divisor de águas da Bacia do Paraná e da Bacia do Paraguai, estabelecida nas coordenadas geográficas 20°28'13 latitude sul e 54°37'25 longitude oeste. A cidade de Campo Grande encontra-se em duas zonas climáticas classificadas como Verão Quente e Savana (PEEL, FINLAYSON e MCMAHON, 2007), pertencente a região fitogeográfica do Cerrado (PLANURB, 2021).

Figura 10: Localização do município de Campo Grande-MS; (a) Contexto Brasil; (b) Contexto Mato Grosso do Sul; (c) Contexto Perímetro Municipal de Campo Grande



Fonte: Elaborado pela autora

A capital sul-mato-grossense situa-se sobre a Bacia Sedimentar do Paraná e possui solos classificados como Latossolos, predominantes no território, e Neossolos (EMBRAPA, 2018; PLANURB, 2021). Essa classificação de solos se faz importante para compreender a Carta Geotécnica do município (Figura 11a). A carta em questão é um instrumento de planejamento urbano que tem por objetivo reunir dados geológicos, hidrogeológicos, geomorfológicos, geotécnicos e pedológicos para embasar diretrizes condizentes com as condições físicas do município. O documento passou por uma revisão no ano de 2019 e seus produtos foram atualizados; desta forma podemos observar que o território urbano da cidade de Campo Grande (MS) possui 4 Unidades Homogêneas. A Unidade Homogênea I é a mais extensa de todo o município com predomínio de declividade plana (0 a 3%) na região oeste e declividade suave ondulada (3 a 6%) na região sudeste, apresentando um coeficiente de infiltração de 40 a 70 litros/m².dia, sendo considerada de vagarosa a média. Já a Unidade Homogênea II localiza-se mais a nordeste do território e é composta, majoritariamente, pela declividade plana (0 a 3%), seu coeficiente de infiltração de água no solo varia entre 60 a 120 litros/m².dia, considerado média a rápida. A Unidade Homogênea III encontra-se mais ao sul e na porção oeste da

cidade, possui predomínio de declividade plana (0 a 3%) e apresenta um coeficiente de infiltração médio, variando de 60 a 90 litros/m² dia. Por fim, a Unidade Homogênea IV possui, majoritariamente, declividade suave ondulado (3 a 6%) a ondulado (8 a 20%) e seu coeficiente de infiltração de água no solo varia de acordo com o solo predominante do local, visto que essa classe acompanha o leito dos corpos hídricos (CAMPO GRANDE, 2020).

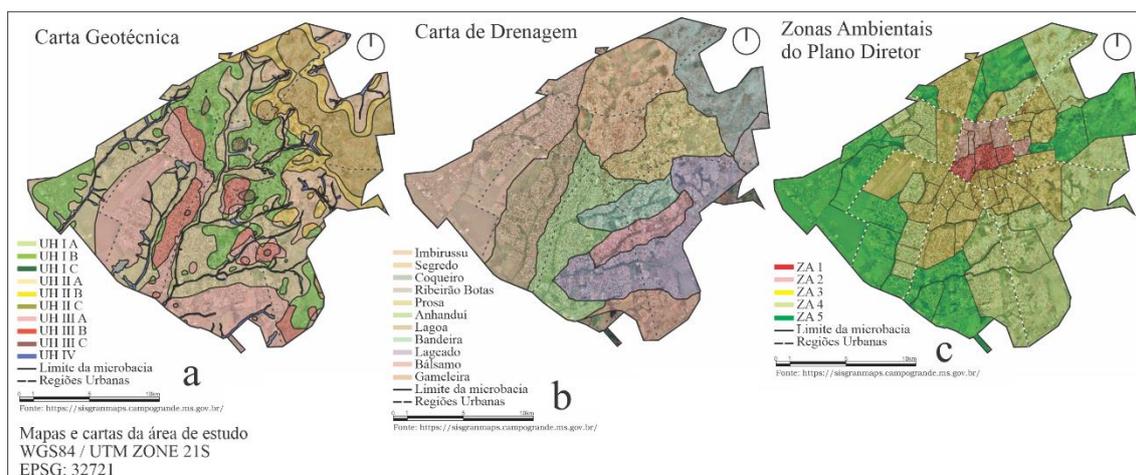
A capital sul-mato-grossense contém 33 cursos d'água que se dividem em 11 microbacias hidrográficas urbanas: Anhanduí, Bandeira, Bálsamo, Coqueiro, Gameleira, Imbirussu, Lageado, Lagoa, Ribeirão Botas, Prosa e, por fim, Segredo. Desses 33 cursos d'água há somente um (1) rio urbano que nasce do encontro dos córregos Prosa e Segredo e é o principal corpo hídrico do município: o Rio Anhanduí (PLANURB, 2021). A Carta de Drenagem municipal (Figura 11b) é um instrumento de planejamento urbano que evidencia os corpos hídricos pertencentes a cada microbacia urbana e visa indicar quais são as porcentagens de áreas permeáveis e impermeáveis em cada um dos seis (6) Graus de Criticidade e apontar problemas e serviços necessários, auxiliando na tomada de decisões. O Grau de Criticidade III, que ocorre na microbacia urbana Lageado, possui a maior área permeável do município com aproximadamente 96,28%. Já o Grau de Criticidade VI, na microbacia urbana Anhanduí, detém o menor percentual de permeabilidade com, aproximadamente, 70,77% (CAMPO GRANDE, 1997).

O Plano Diretor de Desenvolvimento Urbano Ambiental vigente, Lei Complementar n° 341 de 2018, é o instrumento básico de política urbana e objetiva ordenar o desenvolvimento social, ambiental, econômico e cultural do município (CAMPO GRANDE, 2018). O Plano Diretor define três (3) Macrozonas no perímetro urbano de Campo Grande, dividindo seu território por adensamento e densidade demográfica. Para ordenar o uso e ocupação do município, o Plano Diretor vigente agrupou bairros e elencou cinco (5) zonas urbanas. Outra forma de promover o zoneamento, porém priorizando a questão ambiental (Figura 11c) e a indicação mínima de permeabilidade, foi aglutinando bairros em outras cinco (5) zonas. De acordo com o Zoneamento Ambiental, a Zona Ambiental 1 (ZA1) localiza-se na centralidade da porção urbana do território municipal e estabelece uma taxa de permeabilidade mínima de 20%, já a Zona Ambiental 2 (ZA2) se encontra logo acima da ZA1 e prevê uma taxa de permeabilidade mínima de 25%, assim como a Zona Ambiental 3 (ZA3) que também define uma taxa de permeabilidade de 25% e situa-se ao redor das outras duas zonas. Já

a Zona Ambiental 4 (ZA4) e a Zona Ambiental 5 (ZA) detêm a taxa de permeabilidade mínima como 30% e se encontram nas áreas periféricas do perímetro urbano (CAMPO GRANDE, 2018).

No ano de 2018 o setor com maior atividade econômica no município foi o setor de serviços com 43,82%, seguido do setor do comércio com 36,86%. O comércio da cidade se faz vital não somente para seus habitantes urbanos, visto que no ano de 2020 obteve cerca de 34,82% do comércio de todo o estado (PLANURB, 2021).

Figura 11: Mapas e cartas da área de estudo; (a) Carta Geotécnica; (b) Carta de Drenagem; (c) Zonas Ambientais do Plano Diretor



Fonte: Elaborado da autora

Um exemplo claro da gravidade das inundações e alagamentos no perímetro urbano (Figura 10c) dá-se por uma reportagem disponibilizada pelo veículo Campo Grande News, no ano de 2015, segundo a qual o Córrego Segredo transbordou e invadiu uma rotatória muito importante e movimentada da cidade (Figura 12), a rotatória da Avenida Presidente Ernesto Geisel com a Avenida Rachid Neder (CAMPO GRANDE NEWS, 2015).

Figura 12: Rotatória da Avenida Presidente Ernesto Geisel com a Rachid Neder



Fonte: Campo Grande News (2015)

O alagamento na Rua São Roque (Figura 13), no bairro Taquarussu, é um exemplo mais recente do problema que também foi relatado por meio de denúncia no veículo Campo Grande News, evidenciando que a problemática é recorrente e ainda afeta os munícipes (CAMPO GRANDE NEWS, 2021).

Figura 13: Rua São Roque



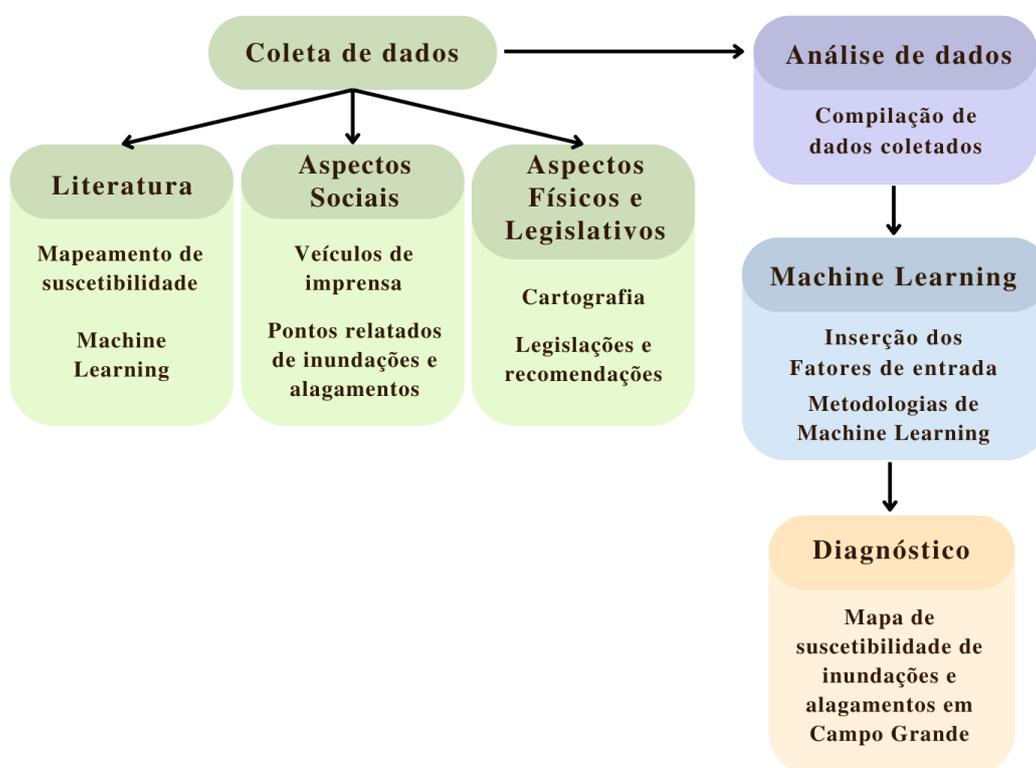
Fonte: Campo Grande News (2021)

2.2 Coleta de ocorrências citadas

A estrutura desse levantamento conta com quatro (4) fases: Coleta de dados, Análise de dados, *Machine Learning* e Diagnóstico (Figura 14). A primeira fase, a Coleta de Dados, foi dividida em três (3) outras partes: Literatura, Aspectos Sociais,

Aspectos Físicos e Legislativos. A coleta de dados de Literatura foi apresentada no capítulo 1 deste estudo e trabalha todo o levantamento e construção do portfólio base da pesquisa metodológicas do processo de mapeamento de suscetibilidade e *Machine Learning*. A segunda parte, a de Aspectos Sociais, será explicada a seguir, tendo em vista o processo de coleta das reportagens em mídias de imprensa e da identificação de cada elemento da matéria noticiada. A última parte do levantamento da Coleta de dados é a parte focada em conhecer a área de estudo nos âmbitos Físicos e Legislativos, observando aspectos físicos, socioambientais, leis vigentes e recomendações, por exemplo, e se encontra previamente citado na Área de Estudo.

Figura 14: Fluxograma geral de processos



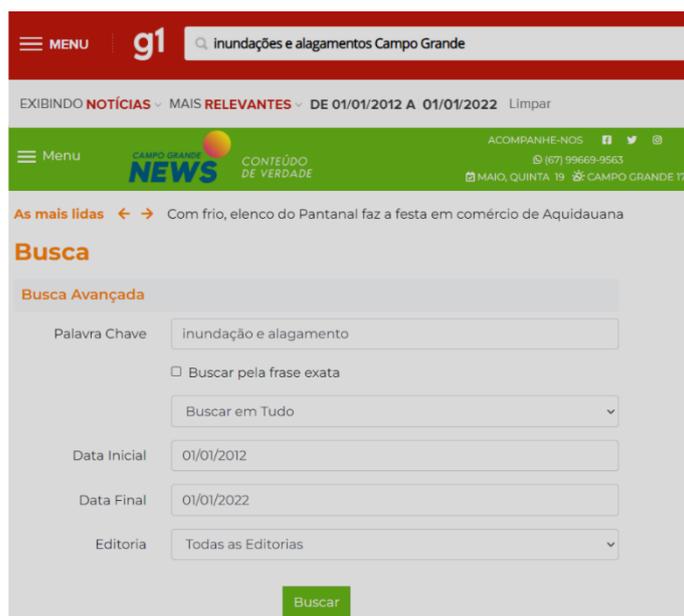
Fonte: Elaborado pela autora

A coleta de dados noticiados em veículos de imprensa decorreu da necessidade de coletar dados para aplicar o *Machine Learning* e para identificar a familiaridade da população com o problema tratado nesta pesquisa, fazendo com que haja um maior espaço amostral e uma conexão entre os residentes com o diagnóstico final. Portanto, evidenciando a necessidade da considerar a sua percepção para ações de melhoria da qualidade de vida dessas pessoas e uma cidade que avança em questões socioambientais (MUNARO, 2023).

Os mapeamentos participativos têm sido empregados cientificamente, de modo qualitativo, para conhecer a vivência da população residente. Tal modelo de mapeamento difundiu-se mais a partir de 1990 e tem sido utilizada por agências governamentais e organizações não governamentais na prevenção e enfrentamento de eventos adversos bem como para desenvolver e consolidar cidades resilientes (DA SILVA e DOS SANTOS, 2022).

Para identificar as áreas submetidas aos eventos extremos, foram escolhidos dois meios de comunicação digitais: o jornal digital de acesso aberto Campo Grande News (<https://www.campograndenews.com.br/>) e G1 MS (<https://g1.globo.com/ms/mato-grosso-do-sul/>) (Figura 15). A busca de dados nos veículos selecionados, deu-se pelas palavras-chave que definem o tema da pesquisa (inundação e alagamento), pelo recorte temporal da pesquisa (2012 a 2021) e, por fim, pelo município escolhido.

Figura 15: Busca nos veículos de imprensa



The image shows a screenshot of a search interface on the G1 website. At the top, there is a search bar with the text "inundações e alagamentos Campo Grande". Below the search bar, there are several filters and options: "Palavra Chave" (inundação e alagamento), "Data Inicial" (01/01/2012), "Data Final" (01/01/2022), and "Editoria" (Todas as Editorias). There is also a checkbox for "Buscar pela frase exata" and a dropdown menu for "Buscar em Tudo". A green "Buscar" button is located at the bottom of the search form.

Fonte: Elaborado pela autora

O recorte temporal foi definido a fim de obter um grande número de reportagens em um tempo hábil de investigação de acordo com o tempo de retorno de obras de microdrenagem, que variam entre 2 a 10 anos, como consta na Tabela 6, para avaliação da intervenção realizada (BRASIL, 2016), conseguindo, dessa forma, compreender como avança a participação do tema no cotidiano do cidadão e obter uma quantidade suficiente de informações de todo o perímetro urbano para um diagnóstico mais completo e fiel possível à opinião pública.

Tabela 6: Tempo de retorno

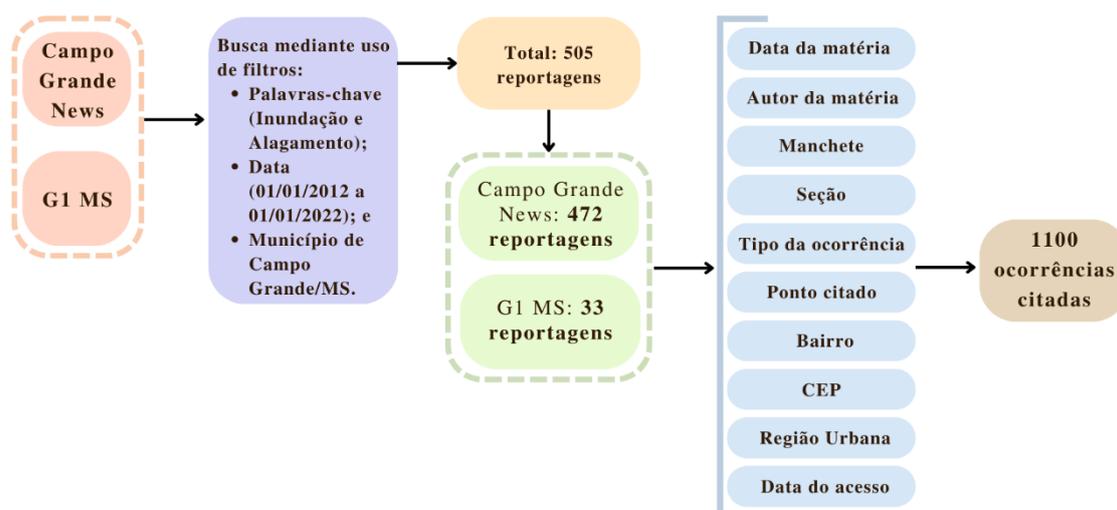
Tipo de obra	Tipo de ocupação	Período de retorno (anos)
Microdrenagem	Residencial	2
Microdrenagem	Comercial	5
Microdrenagem	Edifícios de serviços públicos	2 – 5
Microdrenagem	Aeroportos	5 – 10
Microdrenagem	Áreas comerciais e residenciais	50 – 100
Microdrenagem	Áreas de importância específica	500

Fonte: Brasil (2016) adaptado pela autora

Após finalizada a coleta de informações nos dois (2) veículos de imprensa, passou-se à sua catalogação. Cada uma das notícias classificadas dentro das linhas de busca foi lida e catalogada em um banco de dados elaborado que obteve 472 notícias do *site* Campo Grande News e 33 do *site* do G1 MS, totalizando 505 reportagens. Em posse dos *links* selecionados, foi iniciada uma tabulação aprofundada de cada uma das matérias emitidas pelos veículos de imprensa. Essa investigação, encontrada nos materiais suplementares, objetivou extrair o máximo de informações possível de cada uma das denúncias noticiadas, tais como data da matéria, autor da matéria, manchete, seção, tipo de ocorrência, ponto citado, bairro, Código de Endereçamento Postal (CEP), Região Urbana e data do acesso da matéria (GALLINDO et al., 2024) (Figura 16).

Cada uma das notícias coletadas e analisadas gerou o que chamamos até agora de “ocorrência”. A ocorrência, aqui mencionada, trata de uma localização apontada pelos munícipes, podendo ser uma via ou até mesmo um cruzamento de vias. Grande parte das notícias relata mais de um ponto com problemas, possibilitando um número final de 1100 ocorrências citadas com problemas de inundações e alagamentos dentro do perímetro urbano de Campo Grande.

Figura 16: Fluxograma de coleta de dados nos veículos de imprensa

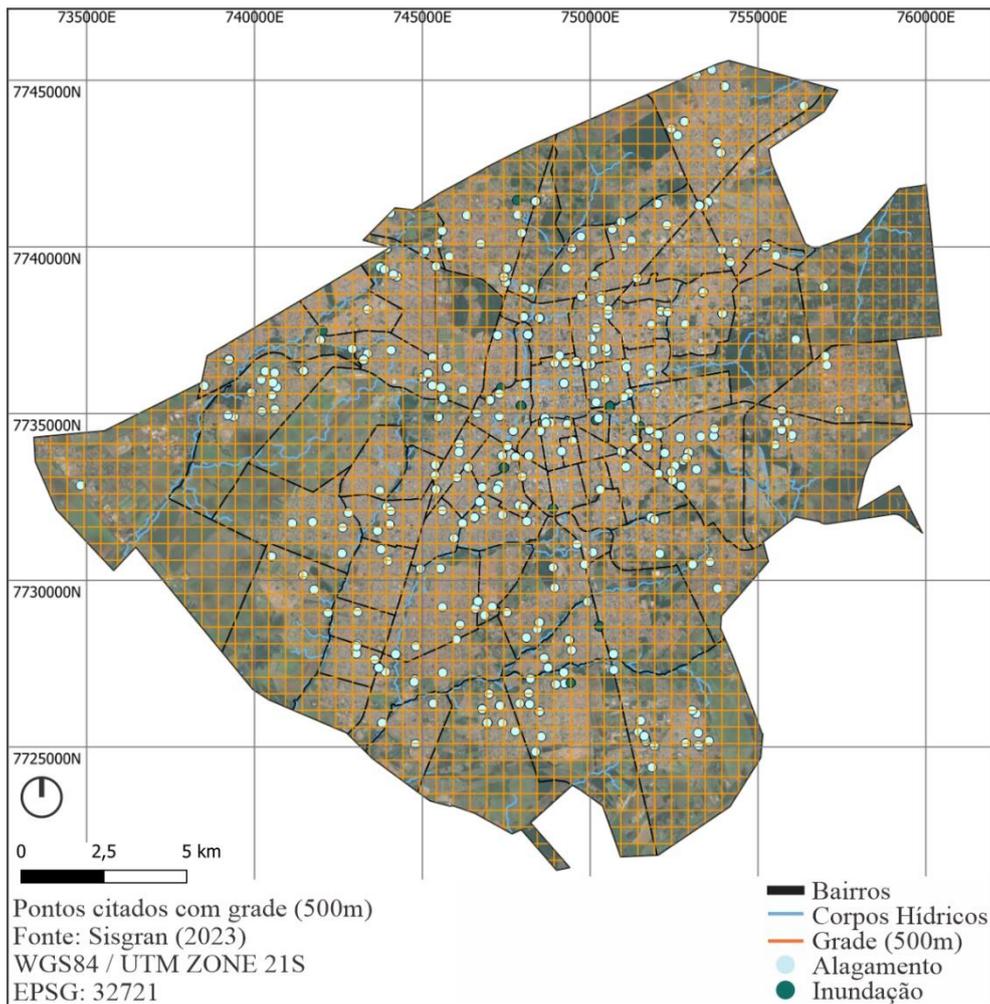


Fonte: Elaborado pela autora

Após listar as ocorrências citadas, identificamos as coordenadas geográficas de cada um desses pontos para corretamente em um mapa síntese de localização, capaz de informar visualmente, a extensão das zonas conflituosas dentro do perímetro urbano campo-grandense. Uma grade quadriculada de 500 metros foi proposta, buscando manter a forma mais simples para replicação e visando melhorar o nível de detalhamento das informações no perímetro urbano do município (*downscaling*). Essa grade foi testada de diferentes formas a fim de definir qual forma seria mais coerente e que agrupava os pontos colhidos de melhor forma. O primeiro formato testado foi o hexagonal de 1000 metros, que gerou resultados pouco detalhados. Já o hexagonal de 500 metros não foi escolhido devido aos ângulos formados que dificultavam a programação e, por fim, o hexagonal de 30 metros, que tornou difícil manipulação devido ao tamanho do arquivo trabalhado. Os mesmos testes foram feitos com o formato quadriculado (grade regular). que foi por fim o escolhido, adotando quadrícula de 500 metros.

O próximo passo foi mesclar os pontos de ocorrências citadas com a grade (Figura 17). Muitos deles, por terem sido citados mais de uma (1) vez estão contidos no mesmo CEP, ou seja, no mesmo ponto no mapa. Portanto, contando os pontos que estão localizados dentro de cada uma das células poligonais que compõem a grade para que as informações obtidas possam ser alocadas no centroide de cada quadrícula, a qual chamaremos de *grid*, permitindo uma trabalhabilidade facilitada das 1100 ocorrências citadas.

Figura 17: Mapa de localização dos pontos citados com a grade



Fonte: Elaborado pela autora

2.3 Fatores de entrada

Encerrada a coleta de pontos reportados (ocorrências citadas), partimos para a coleta dos fatores de entrada que caracterizam a área estudada, a fim de que essas informações foram compiladas para prosseguir com o estudo. O software livre QGIS, versão 3.16.0 com GRASS 7.8.4, foi utilizado para produzir esta etapa do estudo.

Os fatores de entrada são parte importante do estudo como um todo porque permitem compreender quais são as características físicas, socioeconômicas e legislativas de cada uma das ocorrências citadas. Fatores de entrada representativos possibilitam a elaboração da modelagem através de *Machine Learning*, visto que o Aprendizado de Máquina se “aprende” por meio dos fatores de entrada inseridos e, com isso, consegue obter uma maior flexibilidade nas suposições e nas avaliações de suscetibilidade de inundações e alagamentos (GABRIELS, WILLEMS, VAN ORSHOVEN, 2020).

A Revisão Sistemática da Literatura, realizada no capítulo 1, será base para compreender quais são os fatores de entrada mais utilizados pelos pesquisadores atualmente. A Tabela 5, apresentada previamente no capítulo 1, mescla tipos de Aprendizado de Máquina com os fatores de entrada evidenciados no portfólio, utilizando como sustento para escolher quais dados utilizar neste estudo, mas não limita as escolhas feitas para essa etapa da pesquisa. Os fatores de entrada mais aplicados entre os pesquisadores e que foram escolhidos são: Elevação, Declividade, Distância até o rio, Uso/cobertura do solo (LULC), Litologia e Geologia, Índice de Potência de Fluxo (SPI), Índice de Umidade Topográfica (TWI) e Direção de drenagem. Além destes, o Índice de Vegetação de Diferença Normalizada (NDVI), as Bacias Hidrográficas (Carta de Drenagem), a Distância vertical da drenagem mais próxima (HAND) e a Carta Geotécnica do município foram adicionadas aos fatores de entrada pela sua importância vital para a compreensão da área estudada. A Tabela 7 indica quais são as características, fontes e escalas ou resoluções espaciais de cada um dos fatores de entrada escolhidos.

Tabela 7: Características dos fatores de entrada

Fatores de entrada	Fonte	Escala ou Resolução Espacial	
Acumulação de fluxo hídrico	IBGE	Resolução Espacial	30m
Bacias Hidrográficas (Carta de Drenagem)	Prefeitura de Campo Grande	Escala	1:10.000
Carta Geotécnica	Prefeitura de Campo Grande	Escala	1:10.000
Declividade	IBGE	Resolução Espacial	30m
Direção de drenagem	IBGE	Resolução Espacial	30m
Elevação (MDE)	IBGE	Resolução Espacial	30m
Índice de Potência de Fluxo (SPI)	IBGE	Resolução Espacial	30m
Índice de Umidade Topográfica (TWI)	IBGE	Resolução Espacial	30m
Índice de Vegetação de Diferença Normalizada (NDVI)	IBGE	Resolução Espacial	30m
Proximidade de corpos hídricos	IBGE	Resolução Espacial	30m
Solos	IBGE	Escala	1:5.000.000
Uso e cobertura do solo (LULC)	Map Biomas	Resolução Espacial	30m
Vegetação	IBGE	Escala	1:250.000
Distância vertical da drenagem mais próxima (HAND)	IBGE	Resolução Espacial	30m

Fonte: Elaborado pela autora

O procedimento posterior foi o tratamento das informações, passo fundamental para retirar células sem dados e que não seriam utilizadas durante o processo, neste caso, células fora do perímetro urbano de Campo Grande. Todos os dados iniciais foram tratados e destacados na tabela de atributos da seguinte forma: mapas e cartas categóricas foram retratados como maioria e cada mapa ou carta de caráter numérico foram evidenciadas através de suas médias (Tabela 8). Esse processo é necessário para que possamos compreender a informação de cada ponto, trazendo o que se apresenta de forma mais fiel para com a realidade.

Tabela 8: Extração dos fatores de entrada

Fatores de entrada	Tipo de extração
Acumulação de fluxo hídrico	Média
Bacias Hidrográficas (Carta de Drenagem)	Média
Carta Geotécnica	Maioria
Declividade	Média
Direção de drenagem	Maioria
Elevação (MDE)	Média
Índice de Potência de Fluxo (SPI)	Média
Índice de Umidade Topográfica (TWI)	Média
Índice de Vegetação de Diferença Normalizada (NDVI)	Média
Proximidade de corpos hídricos	Média
Solos	Maioria
Uso e cobertura do solo (LULC)	Maioria
Vegetação	Maioria
Distância vertical da drenagem mais próxima (HAND)	Média

Fonte: Elaborado pela autora

2.4 Aprendizagem de Máquina

Todos estes dados foram compilados dentro do programa QGIS e possibilitaram a criação de uma tabela de atributos, que posteriormente foi inserida no ambiente Python, onde todos os fatores foram testados e manipulados para testar cenários (Apêndice 2). O Python é uma linguagem de programação livre e de código aberto que atua como umas das principais alternativas no desenvolvimento de sistemas, além de ser flexível e empregado em diversos fins, como automatizar tarefas e adicionar novas funções em programas, por exemplo (BORGES, 2014).

Para que os dados fossem selecionados de forma aleatória, a primeira parte do código pensada foi promover um número aleatório de simulações, evitando repetições de realizações promovidas pela inteligência artificial.

O passo seguinte é realizar o tratamento dos fatores inseridos, sendo assim, as células observadas com pelo menos uma (1) ocorrência de inundação ou alagamento

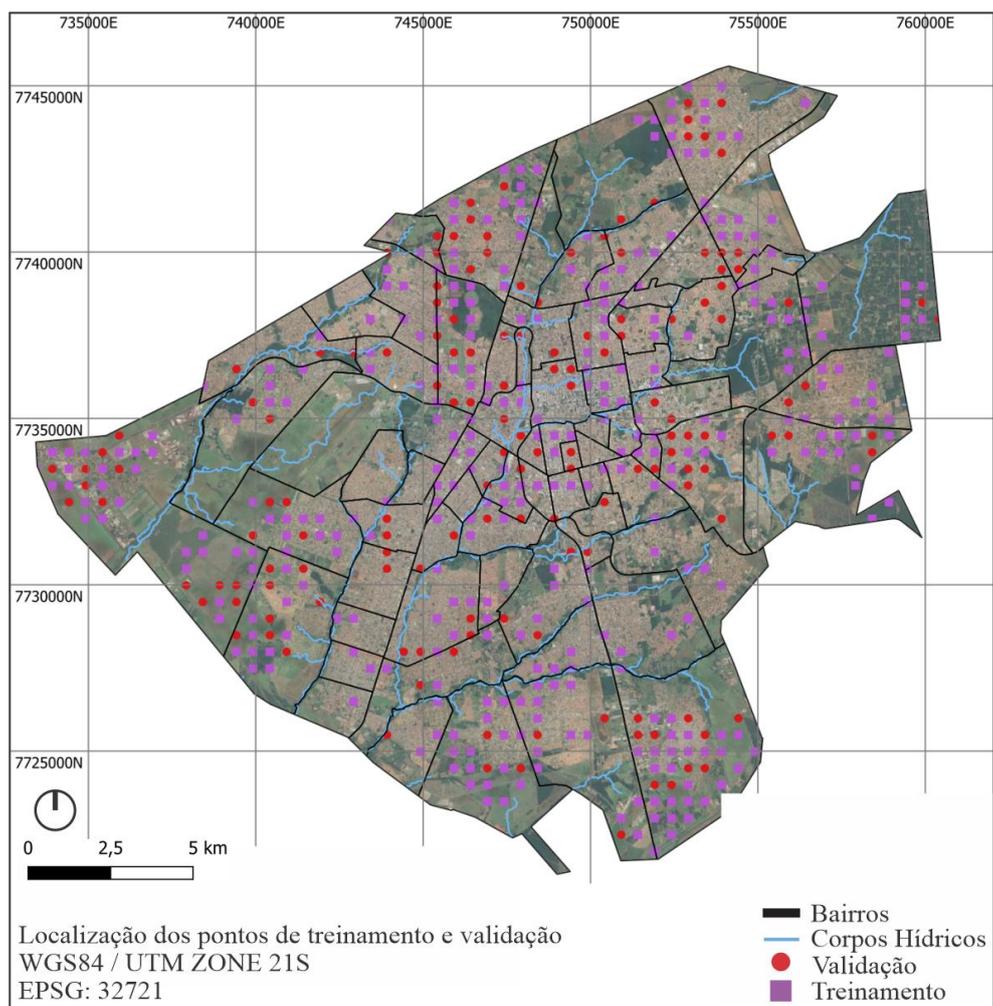
somam um total de 243. Entendendo que as células conhecidas partem de uma (1) ocorrência, se fez necessário adicionar pontos com nenhuma ocorrência para que os métodos tivessem parâmetros de comparação. Portanto, as células zero foram adicionadas sob a circunstância de serem pontos zeros de fato e não falsos zeros, sendo assim, foram selecionados, de forma aleatória, áreas elevadas dos quais não haveriam chances de sofrerem com inundações e alagamentos. Ou seja, sob a única regra de serem acima de um (1) metro da Proximidade de corpos hídricos, outras 243 células sem ocorrências foram adicionadas para a etapa de treinamento e validação do modelo.

Sanando a questão da aleatoriedade e finalizado o tratamento preliminar dos fatores de entrada, adicionamos os 14 fatores de entrada para caracterizar a área de estudo e seguimos para a separação dos *datasets* montados em *sub-datasets* para treinamento e validação. A separação desses fatores dá-se para que o modelo de *Machine Learning* tenha material hábil para treinamento.

A porcentagem escolhida para dados de treinamento e de validação do modelo acompanhou a literatura atual global, que indica um percentual de 70% dos dados coletados e conhecidos para treinamento do modelo e 30% para validação, entendendo se o modelo consegue produzir mapas de suscetibilidade com eficiência e de acordo com a realidade do local (YOUSSEF, PRADHAN, SEFRY, 2016).

A Figura 18 mostra a localização dos 486 pontos aleatoriamente separados em treinamento (70%) e validação (30%).

Figura 18: Mapa de localização dos pontos usados para treinamento e validação



Fonte: Elaborado pela autora

O próximo passo foi evidenciar quais foram os fatores de entrada que obtiveram um maior impacto no método de *Machine Learning* aplicado, intencionando compreender quais características do meio foram fundamentais para que houvesse alagamento ou inundação no dado local. Essa prática é factível mediante o uso do método SVM (*Support Vector Machine*), que tem como uma de suas funções ranquear os *features* inseridos utilizando-o para modular os dados de forma linear (“Kernel Linear”).

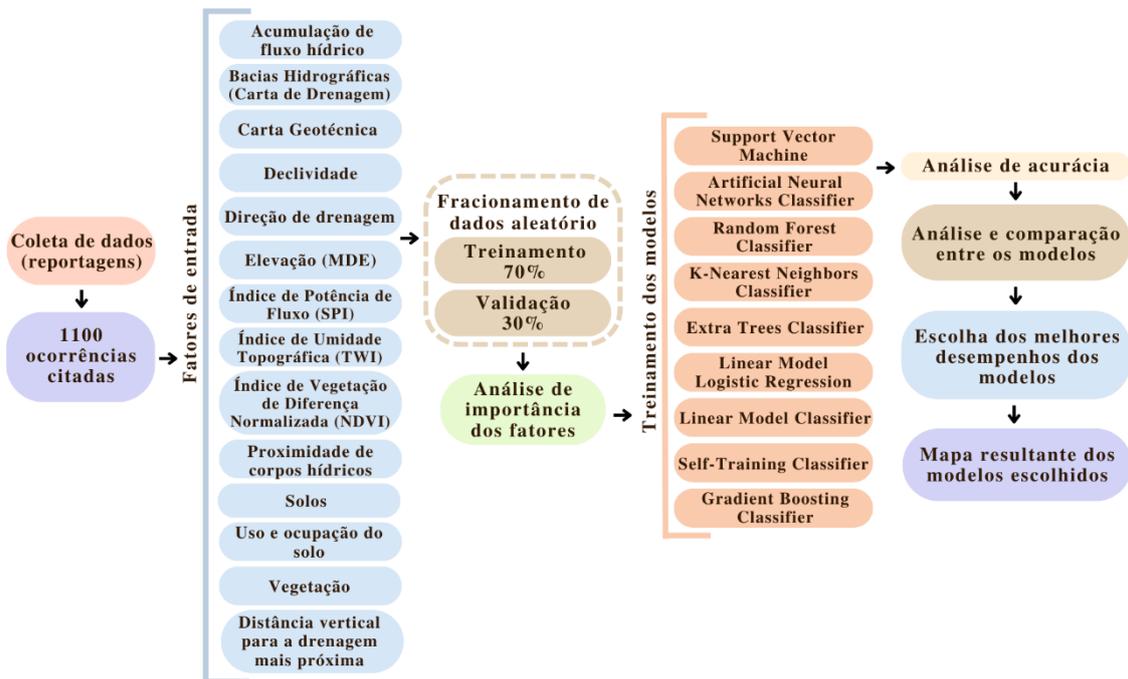
Posteriormente, foram incorporados nove (9) métodos de Aprendizado de Máquina, com base na Tabela 5 do capítulo 1. Entretanto, da mesma forma que na seleção dos fatores de entrada, a tabela não tem a intenção de restringir os métodos escolhidos no código implementado em ambiente Python: *Support Vector Machine*, *Artificial Neural Networks Classifier*, *Random Forest Classifier*, *K-Nearest Neighbors Classifier*, *Extra Trees Classifier*, *Linear Model Logistic Regression*, *Linear Model Classifier*, *Self-Training Classifier* e *Gradient Boosting Classifier* para checagem das acurácias de cada

um deles, almejando obter métodos suficientes para comparação e, por fim, escolher os métodos que obtiveram melhores métricas na comparação de acurácias excluindo os métodos que apresentarem *overfit*, que ocorre quando o modelo demonstra aproveitamento máximo na fase de treinamento, porém, na realidade, este modelo não compreendeu bem o conjunto ofertado para seu treinamento e lida de forma equivocada com o conjunto de validação (YING, 2019).

Para tal, foi preciso normalizar os dados inseridos e trabalhá-los dentro do montante de treinamento e validação randomicamente separada, ou seja, a averiguação das acurácias deu-se na predição dos valores de cada conjunto (treinamento, validação e total) de acordo com cada modelo para definir o método escolhido e o mapa síntese das ocorrências simuladas, após cinco (5) realizações do programa para coletar e promover a média destes, visto a aleatoriedade do código, fosse produzido.

O processo foi sintetizado na Figura 19, que mostra em fluxograma todas as etapas realizadas para que os melhores modelos fossem selecionados.

Figura 19: Fluxograma síntese do processo Python



Fonte: Elaborado pela autora

3. RESULTADOS E DISCUSSÃO

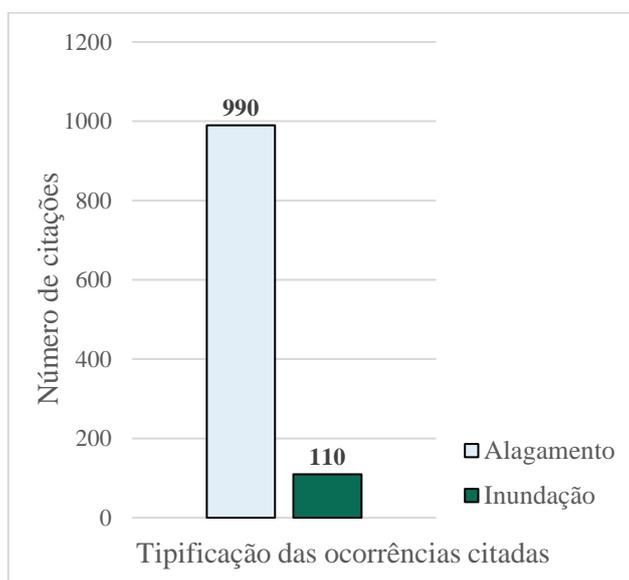
Neste item focaremos nos resultados e nas análises obtidas por meio do banco de dados gerado mediante averiguação das notícias levantadas e dos processos

posteriores em ambiente Python. Por meio dos indicativos da análise de dados, podemos aprofundar cada um dos fatores evidenciados.

3.1 Ocorrências citadas

Com o montante de pontos já identificados, precisamos tipificá-los entre inundação ou alagamento (Figura 20). Das 1100 ocorrências citadas nas reportagens, 990 foram classificados como pontos que sofreram com algum tipo de alagamento e o restante, 110 ocorrências foram apontadas como zonas inundáveis de acordo com os meios de comunicação.

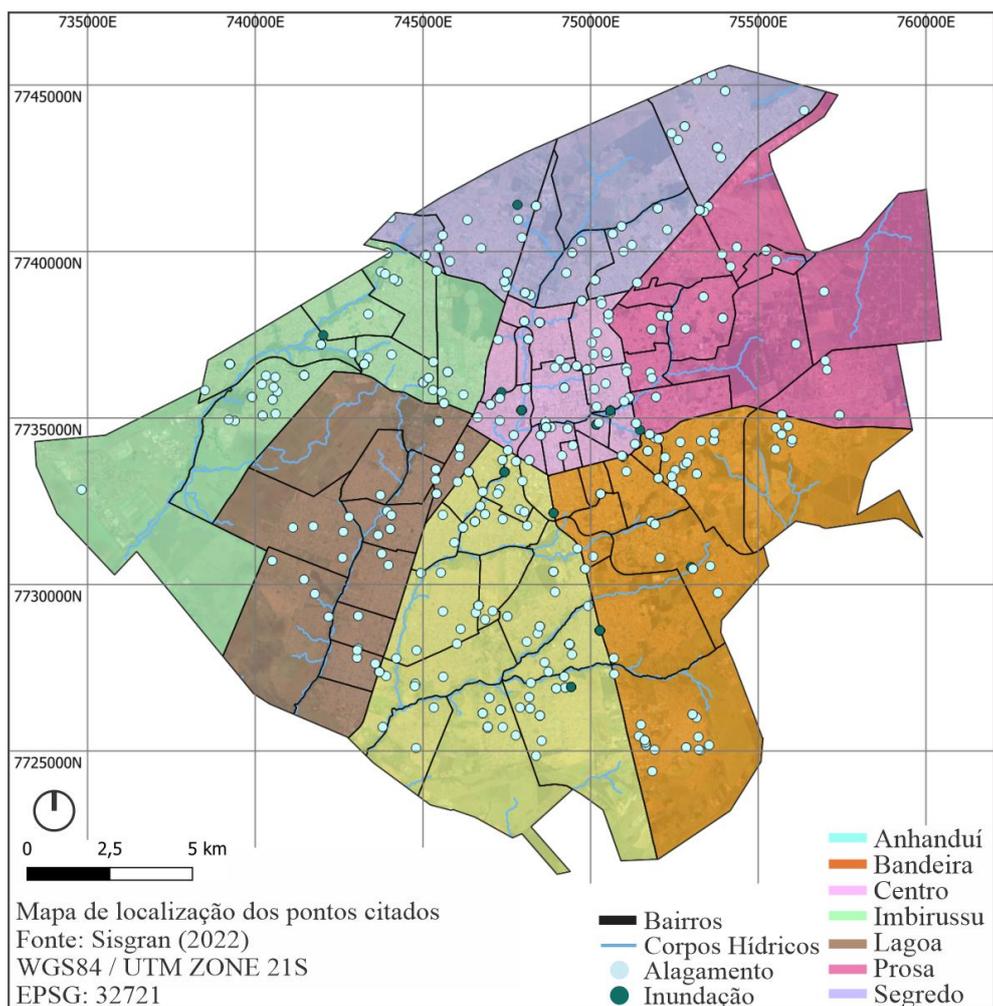
Figura 20: Tipificação das ocorrências citadas



Fonte: Elaborado pela autora

Outra análise importante é a alocação das 1100 ocorrências citadas nas reportagens averiguadas. O mapa das ocorrências citadas (Figura 21), expressa que há problemas de inundações e alagamentos em todas as Regiões Urbanas (RU) no município de Campo Grande. Essa informação é vital para credibilizar a análise dos dados, visto que não apresenta vícios elitistas (enfoque em áreas centrais e de maior concentração econômica) nos veículos de informação.

Figura 21: Mapa das ocorrências citadas e tipificadas no município



Fonte: Elaborado pela autora

A quantificação desses pontos (Tabela 9) aponta que a RU mais afetada é a Centro com 327 ocorrências citadas, a RU do Anhanduizinho com 227 ocorrências citadas são as RU mais sinalizadas. A RU do Centro é a área mais antiga do município de Campo Grande e vem passando por alterações em seu uso do solo desde a criação da cidade em 1875 quando era conhecida como Arraial dos Pereiras (PLANURB, 2021).

Tabela 9: Regiões Urbanas mais citadas

RU citadas	Total das RU citadas
Centro	327
Anhanduizinho	227
Bandeira	156
Imbirussu	135
Segredo	120
Prosa	96
Lagoa	39
Total	1100

Fonte: Elaborado pela autora

Foram ranqueados os 10 bairros que se mostraram mais frequentes quando tratamos de inundações e alagamentos a fim de obter uma amostragem factível dos bairros citados. O restante dos dados se encontra no material suplementar. Dos 74 bairros existentes no município de Campo Grande (MS) (PLANURB, 2021), o bairro São Francisco, localizado na RU do Centro, é o bairro que mais atestou pontos problemáticos (Tabela 10) com um total de 106 menções. Uma diferença de 32 citações do segundo bairro mais mencionado, o bairro Santo Antônio, na RU do Imbirussu.

Tabela 10: Bairros mais citados

Ranking	Bairros citados	RU	Total dos bairros citados
1°	São Francisco	Centro	106
2°	Santo Antônio	Imbirussu	74
3°	Centro	Centro	64
4°	Aero Rancho	Anhanduizinho	58
5°	Itanhangá Park	Centro	51
6°	Taquarussu	Anhanduizinho	43
7°	Nova Lima	Segredo	41
8°	Moreninha	Bandeira	37
9°	Santa Fé	Prosa	37
10°	Amambaí	Centro	36
Total			1100

Fonte: Elaborado pela autora

Assim como feito anteriormente com os bairros, as 10 vias mais citadas nos veículos de imprensa foram ranqueadas para melhor constatação da situação atual do município (Tabela 11). O restante dos dados está disponível no material suplementar. Dessa forma, a Avenida Presidente Ernesto Geisel foi a via mais apontada pelas denúncias de inundação e alagamento, contabilizando um total de 43 menções. A segunda via mais afetada, segundo o levantamento, refere-se à Avenida Cônsul Assaf Trad com 28 citações, juntamente da Rua Professor Luís Alexandre de Oliveira, conhecida popularmente como Via Park. Um destaque interessante a ser mencionado é que além da Avenida Presidente Ernesto Geisel ser apontada como a via mais acometida por inundações e alagamentos, o encontro desta via com a Avenida Rachid Neder também foi fortemente citado, com 21 menções. Sendo assim, a Avenida Presidente Ernesto Geisel é um forte foco para o estudo no município.

Tabela 11: Pontos mais citados

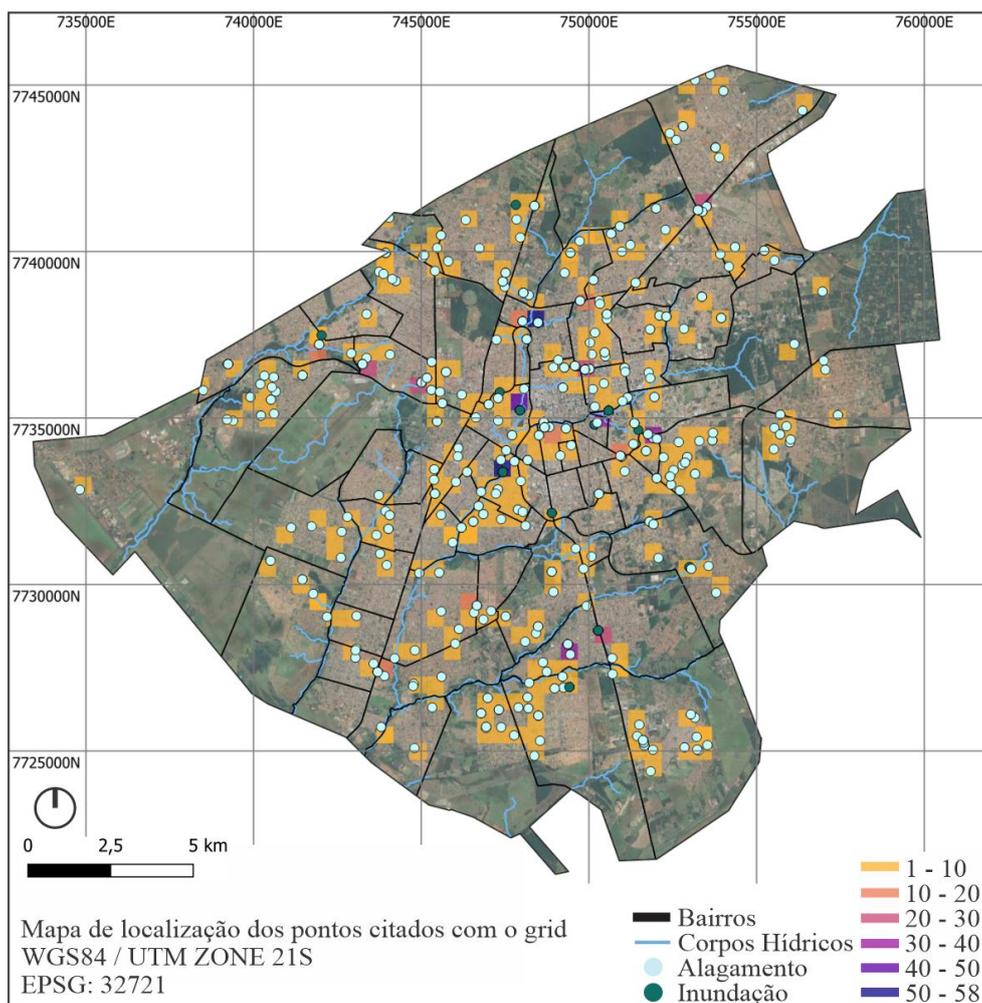
Ranking	Ocorrências citadas	Total das ocorrências citadas
1°	Avenida Presidente Ernesto Geisel	43
2°	Avenida Cônsul Assaf Trad	28
3°	Rua Professor Luís Alexandre de Oliveira	28
4°	Avenida Capibaribe	27
5°	Avenida Guaicurus	27
6°	Avenida Gury Marques	24
7°	Avenida Ricardo Brandão	22
8°	Rotatória da Avenida Presidente Ernesto Geisel e Avenida Rachid Neder	21
9°	Avenida Júlio de Castilho	18
10°	Avenida Afonso Pena	17
	Total	1100

Fonte: Elaborado pela autora

A manipulação dos dados é um processo importante e minucioso, visto que há uma grande quantidade de pontos e outros dados trabalhados neste estudo. Portanto, foi proposto um agrupamento dessas informações (*grid*).

Na Figura 22 podemos observar que a concentração de pontos mencionados pelos munícipes (64 pontos) coincide com a Avenida Presidente Ernesto Geisel, citada como uma via que sofre com muitas inundações e alagamentos. Outros quadrantes que merecem destaque no *grid* são os que englobam o entorno do cruzamento da Avenida Presidente Ernesto Geisel e Avenida Rachid Neder (54 pontos), o cruzamento da Avenida Presidente Ernesto Geisel e Avenida Afonso Pensa (50 pontos) e, por fim, cruzamento da Avenida Fernando Corrêa da Costa com Rua Joaquim Murtinho (39 pontos).

Figura 22: Mapa do grid de ocorrências citadas



Fonte: Elaborado pela autora

3.2 Fatores de entrada

Os fatores de entrada, como dito anteriormente, são informações vitais para sequência do estudo, e por essa razão, 14 fatores foram escolhidos a fim de compreender as características físicas da área de estudo (Figura 23).

O primeiro mapa trata da Carta Geotécnica municipal (Figura 23a) e o segundo mapa apresenta a Carta de Drenagem municipal (Figura 23b), ambas já foram explicadas previamente no item área de estudo. A Declividade, demonstrada pelo Figura 23c, governa o movimento e a penetração dos fluxos superficiais, visto que através dos ângulos da superfície do solo, obtemos sua forma e sua relação com a litologia e drenagem no local (WAQAS et al., 2021). O mapa de declividade do município indica, mediante classificação da Embrapa (2009), que o relevo de Campo Grande majoritariamente varia de Plano (0 a 3%) para Suave-ondulado (3 a 8%), na zona mais central da cidade há relevo classificado como ondulado (8 a 20%).

A Figura 23d trata do mapeamento do Modelo Digital de Elevação (MDE), que tem por objetivo representar as altitudes da superfície topográfica (IBGE, 2023). O município de Campo Grande, através do MDE, apresenta uma variação de altitude de 485 a 700, com as áreas mais elevadas ao norte da cidade. Já no mapa de Direção de drenagem (Figura 23e) podemos observar a direção do fluxo de drenagem do solo, compreendendo como as águas superficiais e pluviais se comportam.

A Figura 23f apresenta o mapa de proximidade de corpos hídricos, indicando a distância de todo território de estudo para com seus corpos hídricos, sendo 0 o próprio rio ou córrego e 2671,8 a maior metragem encontrada. O próximo mapa indicado aborda a acumulação de fluxo hídrico (Figura 23g) que demonstra a direção e as áreas de fluxo concentrado a fim de salientar os canais de transmissão de águas superficiais (NOVAIS, 2015). Dessa forma observamos que o acúmulo hídrico se dá na zona sul do município no Rio Anhanduí em sua porção final dentro do perímetro urbano campo-grandense.

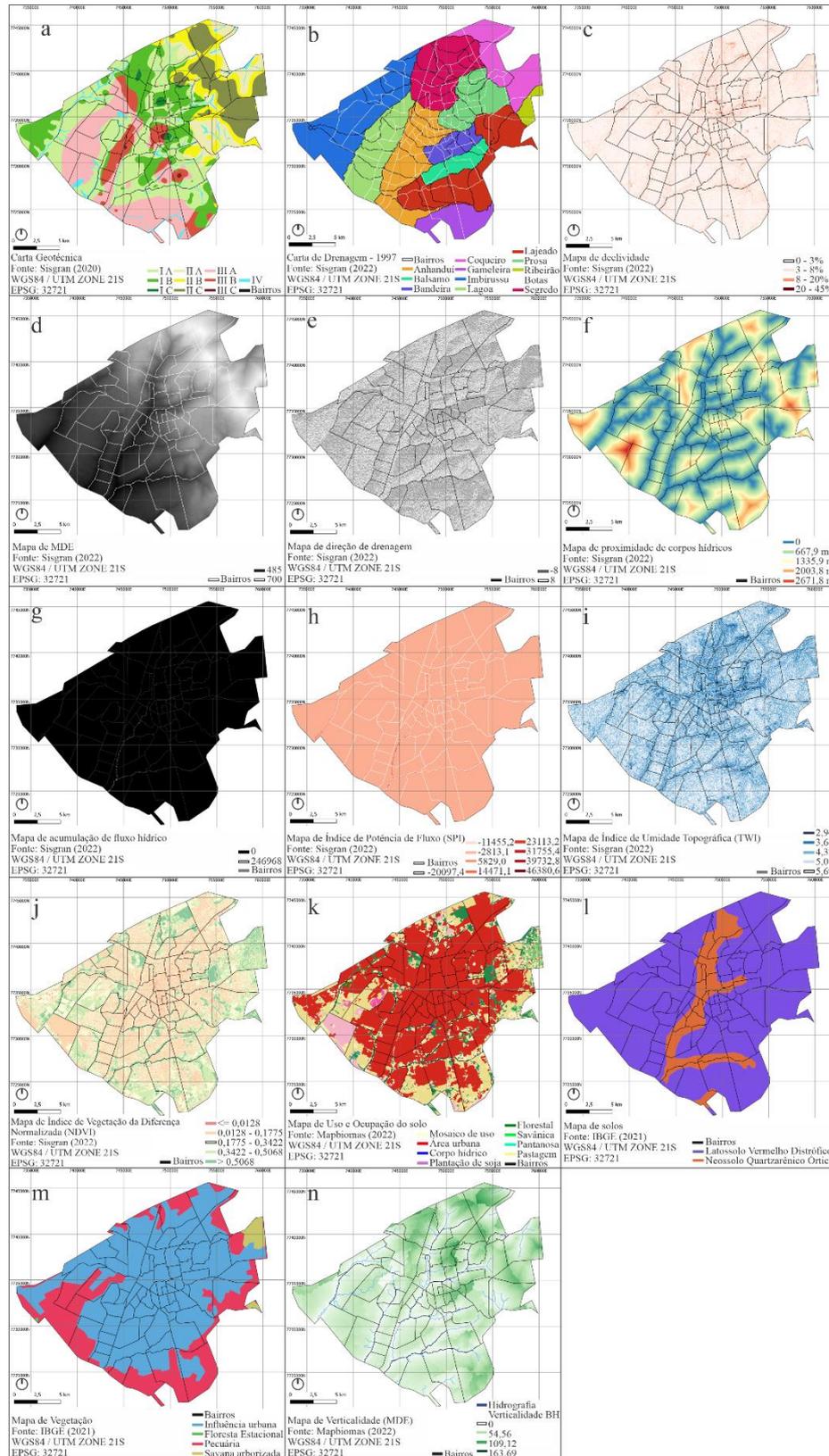
A Figura 23h expressa o Índice de Potência de Fluxo (SPI) da área de estudo, e tem por objetivo definir o transporte de sedimentos de um canal em referência a um ponto da superfície do terreno e sua magnitude em potencial aumenta quando há um aumento na inclinação do solo possibilitando o aumento do fluxo d'água (TELLA e BALOGUN, 2020). Portanto é observável que o Rio Anhanduí é o mais impactado. Já o Índice de Umidade Topográfica (TWI) evidenciada na Figura 23i, explica o acúmulo de fluxo (áreas de inundação) de acordo com a jusante aliada à declividade (TIEN BUI et al., 2019; TELLA e BALOGUN, 2020), assim é notável o acúmulo de umidade mais acentuado nas áreas próximas dos corpos hídricos no tecido urbano e pontos de maior umidade em lagos artificiais da cidade.

O Índice de Vegetação da Diferença Normalizada (NDVI) expressa a cobertura vegetal viva, esse índice varia entre -1 a +1 e seus valores são calculados através da refletância espectral nas bandas infravermelho próximo e vermelho, na Figura 23j observamos que a vegetação mais presente se encontra nos limites perimetrais, em zonas protegidas e ao redor dos corpos hídricos em sua maioria. Já no mapa que apresenta o uso e ocupação do solo (Figura 23k), podemos salientar a massiva ocorrência de área urbana maciça, de áreas de pastagem muito presentes margeando o perímetro urbano áreas de plantio de soja e áreas florestadas em zonas protegidas.

A Figura 23l aborda os solos do município e destaca predominantemente o Latossolo Vermelho distrófico que possui óxidos de ferro presentes, profundos e porosos, mas que apresenta baixa fertilidade; já o Neossolo Quartzarênico órtico, encontrado nos leitos dos corpos hídricos mais centrais do município possui material mineral ou orgânico pouco espesso com características que variam de rasos a maior profundidade e variando, também, na capacidade de permeabilidade (ZARONI e SANTOS, 2021).

O mapa de Vegetação (Figura 23m) demonstra a grande área de urbanizada juntamente com seus limites perimetrais voltados para a pecuária, já evidenciado pelo mapa de Uso e Ocupação. Por fim, o mapa de distância vertical da drenagem mais próxima (HAND) (Figura 23n), extraída através do Modelo Digital de Elevação (MDE), tem por intuito indicar uma dimensão em metros de verticalidade para cada área de acordo com o corpo hídrico mais próximo.

Figura 23: Coleção dos Fatores de entrada; (a) Carta Geotécnica; (b) Carta de Drenagem; (c) Declividade; (d) MDE; (e) Direção de drenagem; (f) Proximidade de corpos hídricos; (g) Acumulação de fluxos hídricos; (h) Índice de potência de fluxo; (i) Índice de Umidade Topográfica; (j) NDVI; (k) Uso e ocupação do solo; (l) Solos; (m) Vegetação; (n) Distância vertical para a drenagem mais próxima.



Fonte: Elaborado pela autora

3.3 Aprendizagem de Máquina

Dos 14 fatores de entrada observados, os seis (6) primeiros exercem mais influência para o resultado (Tabela 12), segundo o método SVM; eles refletem diretamente a tipologia do solo (Solo e Carta Geotécnica) somado com dados derivados da elevação da área (Direção da drenagem, Distância vertical da drenagem mais próxima, Elevação e Proximidade com corpos hídricos). A união desses fatores é preponderante para a discussão da impermeabilidade do solo e o seu fator de absorção, indicando que as áreas nas proximidades de corpos hídricos possam ter maior probabilidade de serem acometidas pelas inundações e alagamentos.

Tabela 12: Ranking dos Fatores de entrada segundo o SVM

Posição do ranking	Fatores de entrada inseridos
1°	Solo
2°	Proximidade de corpos hídricos
3°	Direção de Drenagem
4°	Carta Geotécnica
5°	Distância vertical da drenagem mais próxima (HAND)
6°	Elevação (MDE)
7°	Declividade
8°	Índice de Umidade Topográfica (TWI)
9°	Índice de Vegetação de Diferença Normalizada (NDVI)
10°	Índice de Potência de Fluxo (SPI)
11°	Bacia Hidrográfica (Carta de Drenagem)
12°	Acumulação de fluxo hídrico
13°	Vegetação
14°	Uso e ocupação do solo

Fonte: Elaborado pela autora

Um total de nove (9) métodos de *Machine Learning* foram inseridos em ambiente Python para aprenderem com os fatores de entrada inseridos, separados em treinamento e validação a fim de obter um comparativo entre suas acurácias (Tabela 13).

Tabela 13: Métodos na codificação em Python

Métodos	Acurácia		
	Treinamento	Validação	Final
Artificial Neural Networks Classifier (ANN)	65%	62%	65%
Extra Trees Classifier	100%	61%	87%
Gradient Boosting Classifier (GBC)	88%	52%	76%
K-Nearest Neighbors Classifier (KNN)	100%	42%	82%
Linear Model Classifier (LM)	66%	64%	65%
Linear Model Logistic Regression (LMLR)	66%	63%	65%

Métodos	Acurácia		
	Treinamento	Validação	Final
Random Forest Classifier (RF)	100%	63%	88%
Self-Training Classifier (STC)	66%	64%	66%
Support Vector Machine (SVM)	66%	63%	65%

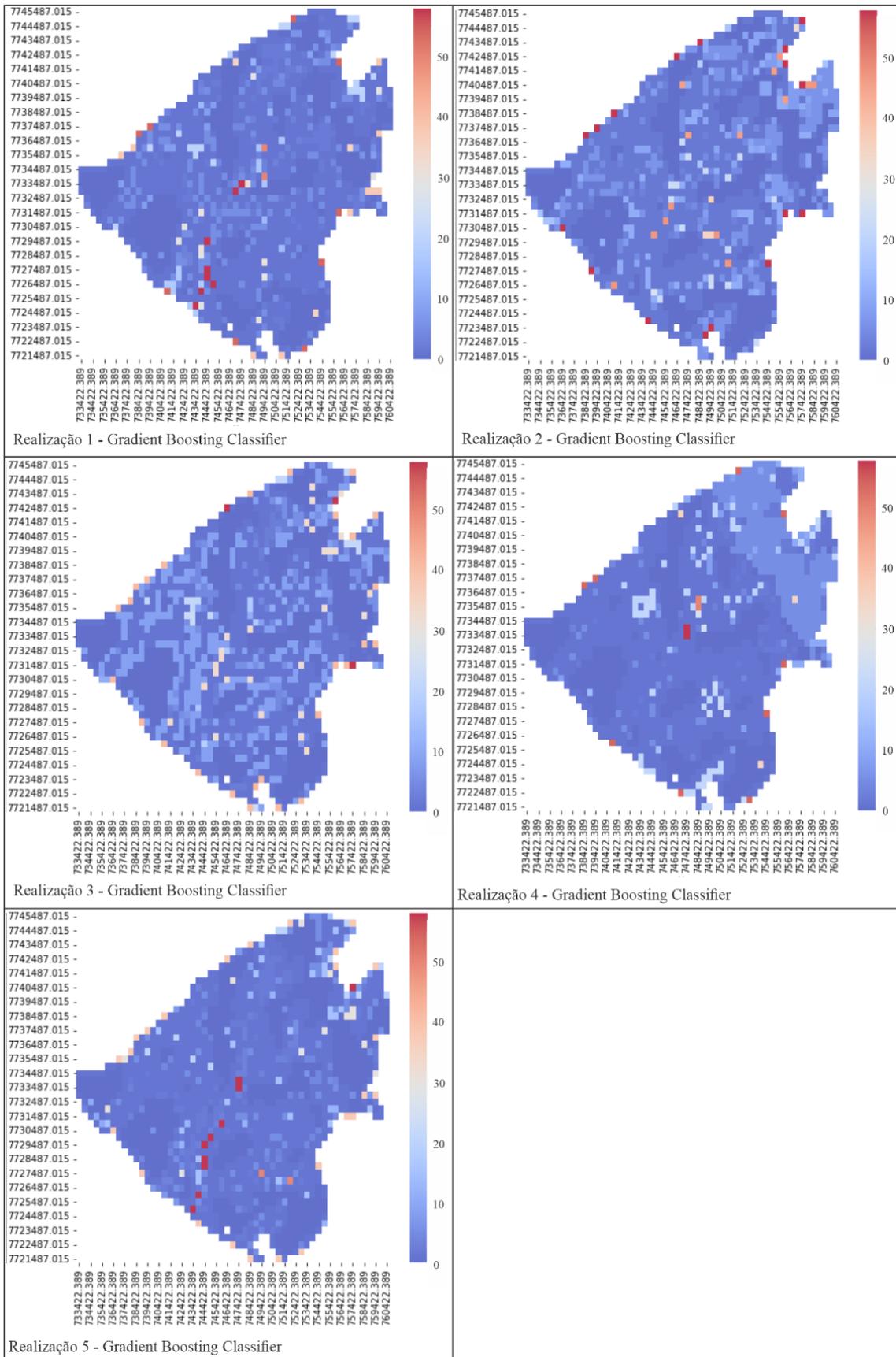
Fonte: Elaborada pela autora

Os modelos *Random Forest Classifier*, *K-Nearest Neighbors Classifier* e *Extra Trees Classifier* foram excluídos dos processos subsequentes por conta do “*overfitting*” em sua fase de treinamento. O restante dos métodos foi avaliado de acordo com a porcentagem de treinamento, validação e porcentagem final do processo. Deste modo, o método *Gradient Boosting Classifier* foi escolhido para dar prosseguimento ao estudo visto seu alto desempenho na fase de treinamento (88%) e seus resultados também satisfatórios na fase de validação (52%) e final (76%) e pela exclusão dos outros modelos devido suas respostas insatisfatórias e incoerentes com a realidade dos fatos, portanto, mesmo que as métricas finais não foram as mais altas dos modelos testados, a escolha se deu pelo respeito à ordem de grandeza e distribuição das ocorrências citadas pelos veículos de comunicação já registradas (Tabela 13).

O método *Gradient Boosting* tem sido empregado em modelos de localização, escala e formas, para estruturação de previsão de classes (LUSA et al., 2017). Para isso, diferente de outros métodos, como *Random Forest*, que utilizam a média simples dos modelos do conjunto, o *boosting* implementa novos modelos ao conjunto e identifica aqueles que se mostraram mais “fracos” e a cada novo modelo proposto, o erro do modelo “fraco” é treinado em relação ao erro de todo o conjunto estruturado promovendo uma estimativa mais precisa para variável de resposta, dessa forma o método consegue implementar novos conjuntos correlacionados com gradientes de perda ou erro, sendo da escolha do pesquisador qual perda ele poderá implementar na função. Portanto, essa flexibilidade do método permite sua personalização para qualquer tarefa baseada em dados (NATEKIN e KNOLL, 2013; LUSA et al., 2017), para maiores informações acessar o *site* Scikit-Learn (<https://scikit-learn.org/stable/>).

O método escolhido, GBC, passou por uma sequência de 5 realizações na programação devido a aleatoriedade do código formulado. Sendo assim, a cada realização um resultado diferente foi gerado (Figura 24). É possível observar que não há grande disparidade entre as realizações, mas ainda assim, há pequenas diferenças entre elas.

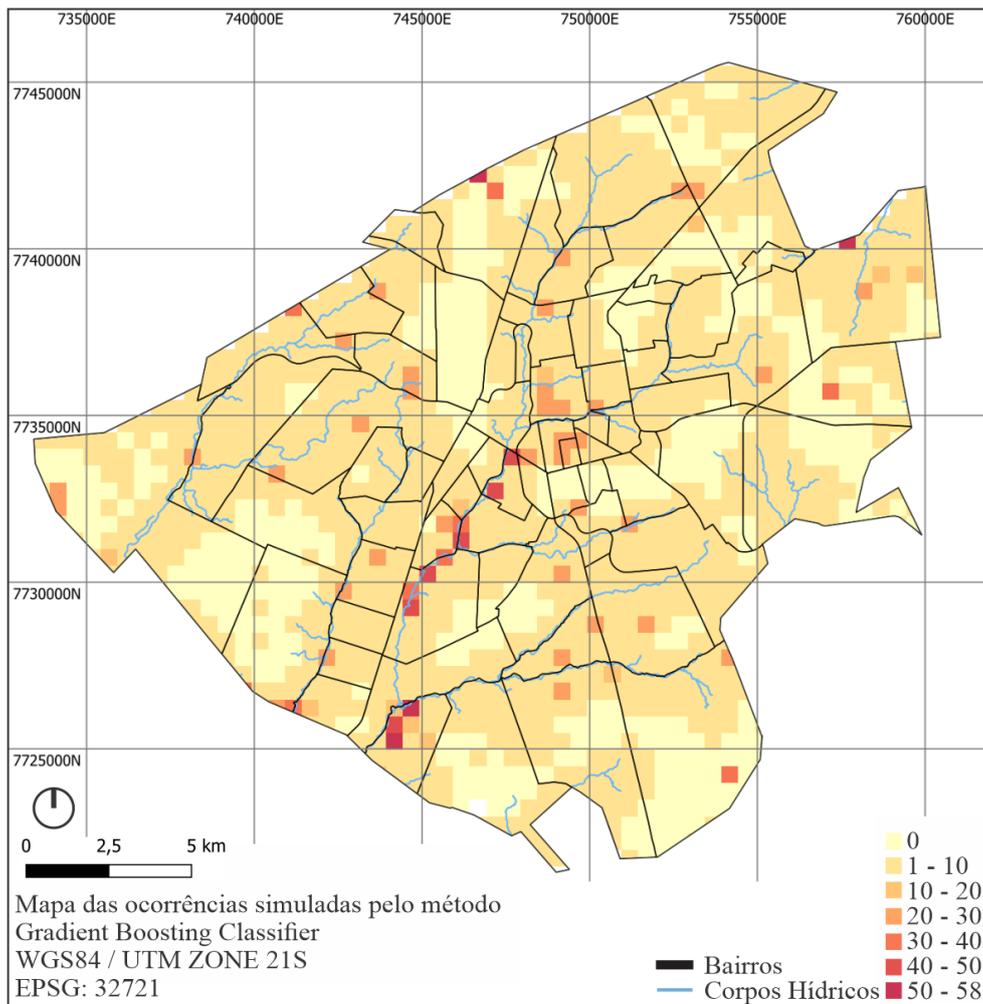
Figura 24: Resultados das rodadas do método *Gradient Boosting Classifier* (GBC)



Fonte: Elaborado pela autora

Os resultados obtidos nas cinco (5) realizações aleatórias foram tabuladas e uma média foi proposta para a elaboração de um produto. A média final proporcionou as ocorrências simuladas (Figura 25), além de sua localização de acordo com a proposição da inteligência artificial, a quantidade de possíveis ocorrências que cada um desses pontos poderá vir a experienciar, partindo de um (1) até o máximo sugerido de 54 ocorrências.

Figura 25: Mapa síntese das ocorrências simuladas



Fonte: Elaborado pela autora

Há um total de 1562 células com ao menos uma ocorrência simulada pelo método *Gradient Boosting Classifier* distribuídos em todas as RU do município de Campo Grande (Tabela 14), que somam 3897 ocorrências simuladas. Tais ocorrências foram divididas pelo total de anos do recorte temporal, ou seja, nove (9) anos, obtendo 433 ocorrências/ano. Deste montante, 113 ocorrências/ano encontram-se na RU Anhanduizinho, seguida pela RU Prosa com 65 ocorrências/ano apontadas pelo GBC. A RU do Centro conta com 41 ocorrências/ano simuladas.

Tabela 14: Total de ocorrências/ano simuladas nas RU

RU	Total de ocorrências/ano simuladas nas RU
Anhanduizinho	113
Prosa	65
Lagoa	57
Segredo	55
Imbirussu	52
Bandeira	50
Centro	49
Total por ano	433

Fonte: Elaborado pela autora

Aprofundando os dados, na Tabela 15 foram ranqueados os 10 bairros que obtiveram mais pontos com possíveis inundações e alagamentos. O restante dos dados se encontra no material suplementar no Apêndice 1. Dos 74 bairros do município de Campo Grande, 73 deles foram assinalados com possíveis pontos de inundação e alagamento, exceto o bairro TV Morena, sendo o único que não obteve ocorrência simulada. O bairro com o maior número de ocorrências foi o Núcleo Industrial, na RU Imbirussu, com 228 possíveis ocorrências, seguido pelo Guanandi, pertencente à RU Anhanduizinho com 194 projeções.

Tabela 15: Total de ocorrências/ano simuladas nos bairros

Ranking	Bairros	RU	Total de ocorrências/ano simuladas nos bairros
1°	Núcleo Industrial	Imbirussu	25
2°	Guanandi	Anhanduizinho	22
3°	Nova Lima	Segredo	20
4°	Caiobá	Lagoa	19
5°	Chácara dos Poderes	Prosa	18
6°	Centenário	Anhanduizinho	17
7°	Moreninha	Bandeira	15
8°	Noroeste	Prosa	14
9°	Lageado	Anhanduizinho	14
10°	Taquarussu	Anhanduizinho	13
Total por ano			433

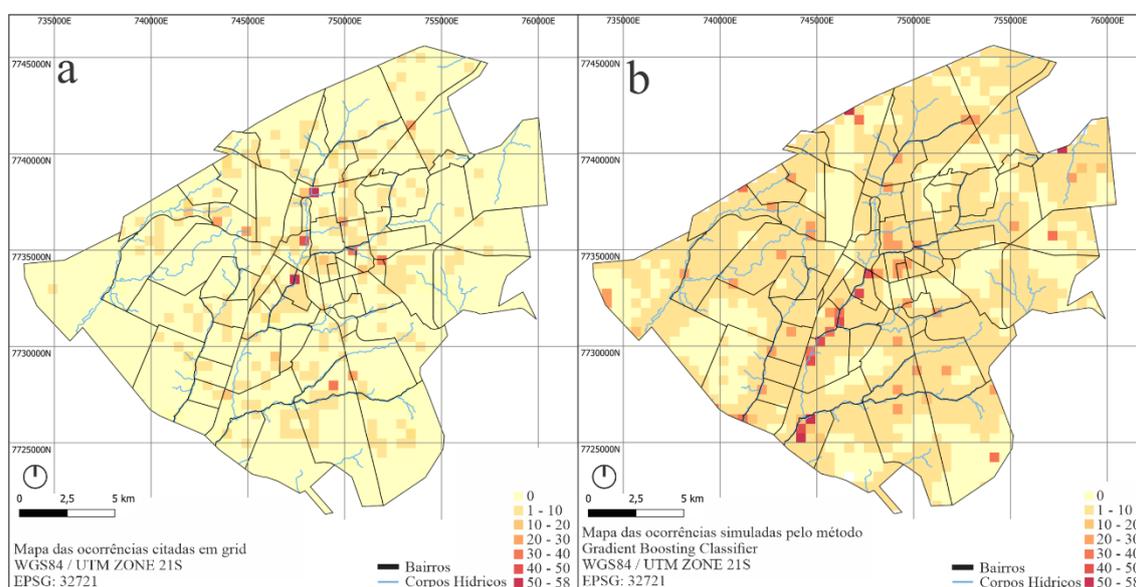
Fonte: Elaborado pela autora

3.4 Comparações entre pontos citados na imprensa e ocorrências simuladas

Decorrente de todo material evidenciado e averiguado, iremos comparar os produtos fornecidos pelas ocorrências citadas mediante relatos dos munícipes via veículos de imprensa com os produtos das ocorrências simuladas obtidas através do processo realizado pelo método *Gradient Boosting Classifier* em ambiente Python. Para estas análises, as cores foram padronizadas visando a melhor compreensão e comparação entre os resultados obtidos.

A primeira comparação aborda as ocorrências citadas e simuladas (Figura 26). Podemos observar que a Figura 26a dispõe de mais citações zeradas do que a Figura 26b, além das ocorrências mais citadas (Figura 26a), se encontrarem mais ao centro da sede municipal de Campo Grande. A Figura 26b detém mais área com possibilidades baixas de inundação e alagamento (1 a 10), indicando zonas com zero probabilidade de ocorrências em áreas entre corpos hídricos, ou seja, em zonas divisores de água. Podemos observar, também, que na Figura 26b as ocorrências simuladas com mais ocorrências estão localizadas mais ao sul do município, na Região Urbana Anhanduizinho.

Figura 26: Comparação entre ocorrências citadas e simuladas

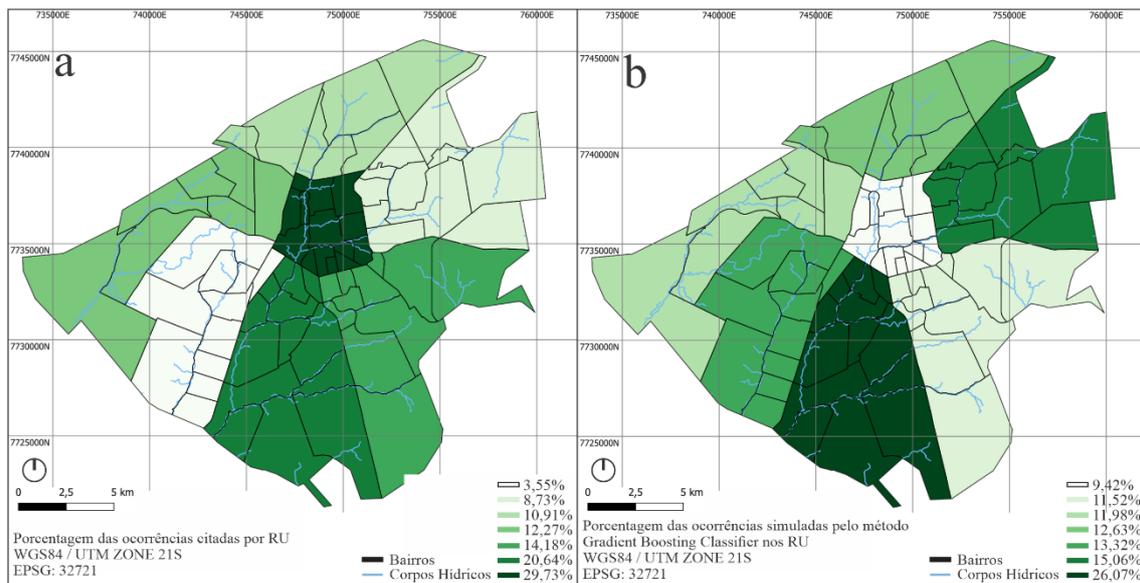


Fonte: Elaborado pela autora

Almejando compreender estatisticamente as diferenças entre os resultados das ocorrências citadas para as ocorrências simuladas no tecido urbano, foram determinados os percentuais totais em suas devidas proporções.

Na Figura 27a, observamos que a RU Centro foi a região mais citada pelos veículos de imprensa com 29,73% do total de pontos observados (1100 pontos), seguido pela RU Anhanduizinho com 20,64% dos pontos coletados. A RU menos citada nas reportagens é a Lagoa, com 3,55% das citações. Ao nos voltarmos para a Figura 27b, observamos que a RU com mais ocorrências simuladas é a Anhanduizinho com 26,07%, a segunda região que mais aparece é a RU Prosa com 15,06% e a que menos detém ocorrências simuladas é a RU Centro com 9,42%.

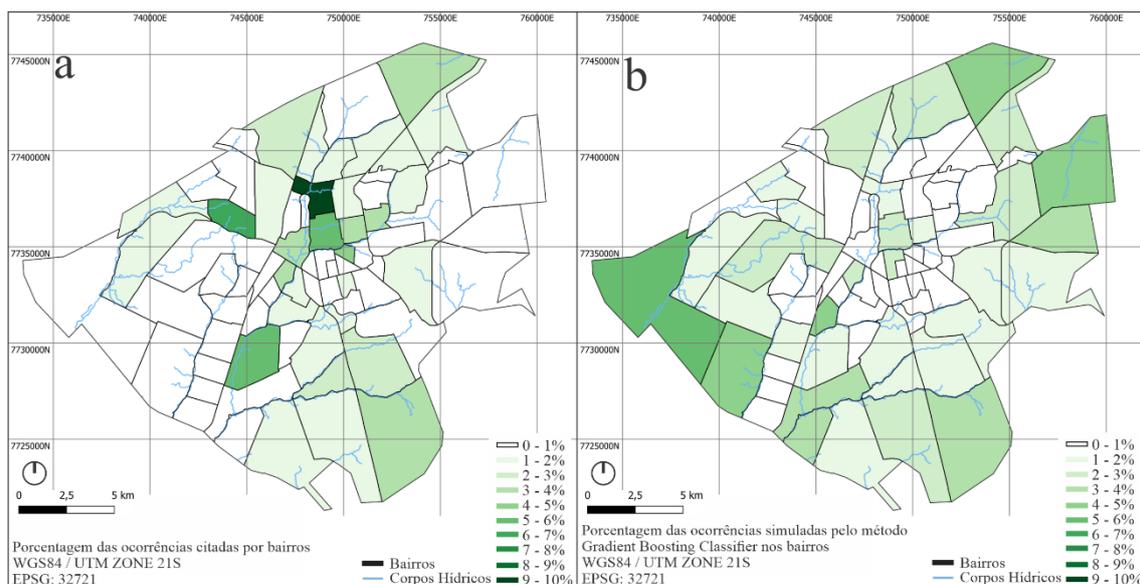
Figura 27: Comparação entre ocorrências citadas e simuladas nas RU



Fonte: Elaborado pela autora

Examinando com mais profundidade, ao se tratar dos bairros, no âmbito das ocorrências citadas, podemos observar pela Figura 28a, que o bairro mais mencionado pelos moradores é o bairro São Francisco com 9,72%, seguido pelo bairro Santo Antônio com 74% citações. Notamos que, de acordo com a Figura 28b, o bairro com mais ocorrências simuladas, com 5,85%, foi o bairro Núcleo Industrial, seguido pelo Guanandi com 4,98%. Já o bairro com menos ocorrências simuladas foi o bairro TV Morena, sendo este o único bairro que não obteve nenhuma ocorrência simulada.

Figura 28: Comparação entre ocorrências citadas e simulados nos bairros

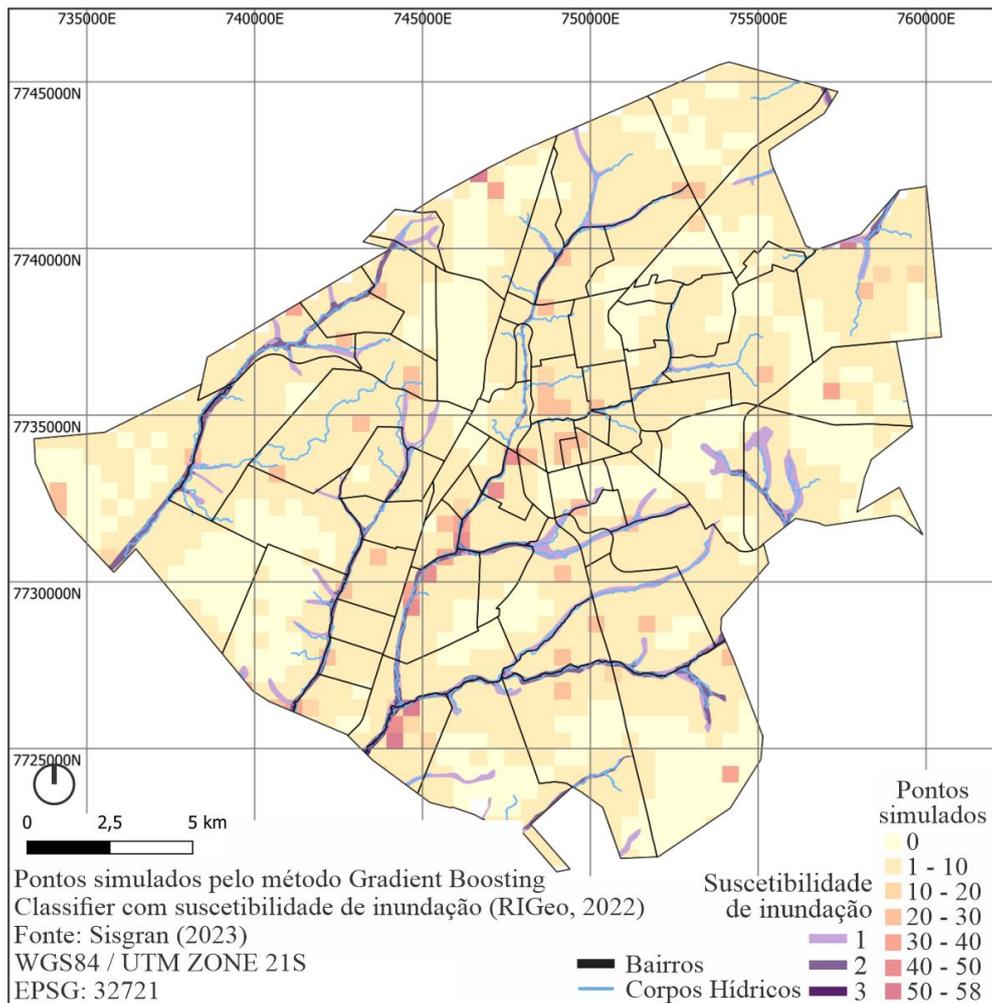


Fonte: Elaborado pela autora

O *site* Repositório Institucional de Geociências (RIGeo), que tem por objetivo permitir o acesso às produções intelectuais e preservar a produção científica com o passar dos anos, detém salvo um estudo realizado para averiguar a suscetibilidade a movimentos gravitacionais de massa e inundação no município de Campo Grande (MS). Este trabalho classifica quais trechos dos corpos hídricos do município são suscetíveis à inundação em três (3) níveis: nível 1 – baixa suscetibilidade, nível 2 – média suscetibilidade e, por fim, nível 3 – alta suscetibilidade.

Na Figura 29 podemos comparar o estudo de suscetibilidade a inundações da RIGeo com o mapa das ocorrências simuladas pelo método GBC, resultado desta pesquisa, e assim comprovar a veracidade e fidedignidade dos dados aqui produzidos. Dado pelo fato que as áreas assinaladas como < 1, são os divisores de água, condiz com quase nula possibilidade de inundação e alagamento em tais áreas. Ocorrências simuladas coincidem com as classificações de suscetibilidade dos corpos hídricos, principalmente quando observamos o Rio Anhanduí. A similaridade ainda acontece na localização das ocorrências simuladas de maiores ocorrências, mais ao sul do perímetro urbano da cidade, nos bairros Centenário e Lageado.

Figura 29: Comparação entre ocorrências simuladas (GBC) e Suscetibilidade de inundação pela RIGeo (2022)



Fonte: Elaborado pela autora

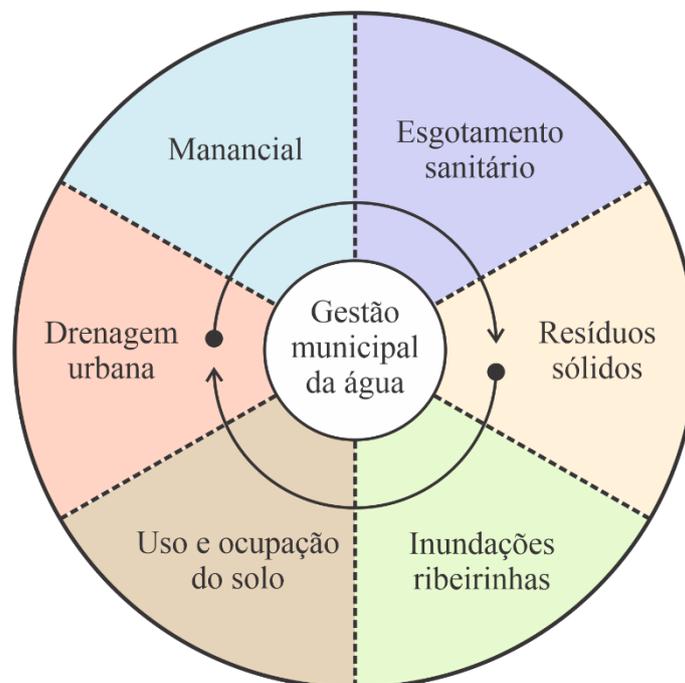
3.5 Diretrizes urbanísticas gerais

Como mostrado previamente, inundações e alagamentos ocasionam problemas em diferentes esferas, mas ao tratarmos do meio urbano, esses danos podem acometer desde a infraestrutura de um município até danos à vida dos moradores (YOUSSEF, PRADHAN e SEFRY, 2016). Esses problemas podem decorrer de muitas mazelas urbanas: planejamento inadequado, problemas de escoamento, impermeabilização do solo, sobrecarga no sistema de drenagem, obstrução de fluxos e muitos outros (HOSSAIN e MENG, 2020; FAREGH e BENKHALED, 2021).

Tendo em vista a necessidade de avanço nas questões de enfrentamento do problema em questão, é preciso gerir o município de forma integrada, adotando um planejamento de caráter preventivo e indutivo. O planejamento hídrico integrado com o desenvolvimento urbano é capaz de abranger várias frentes (Figura 30), mas priorizando

o uso e ocupação do solo, visto que o desenvolvimento urbano é caracterizado pelo planejamento do desenvolvimento, transporte, abastecimento de água e saneamento, drenagem e controle de inundações, resíduos sólidos e controle ambiental (TUCCI e BERTONI, 2003)

Figura 30: Visão integrada do planejamento dos aspectos da água no ambiente (TUCCI e BERTONI, 2003)



Fonte: Tucci e Bertoni (2003) adaptado pela autora

O município de Campo Grande conta com o Plano Diretor de Desenvolvimento Urbano Ambiental (PDDUA), Lei Complementar n° 341 de 2018. Este plano caracteriza-se como o instrumento básico para promoção da política de desenvolvimento urbano de acordo com a Constituição Federal de 1988. Dentro do mesmo, encontra-se, pertencente aos Instrumentos de Gestão Ambiental, o Plano Diretor de Drenagem. Este plano data de 2008, necessitando urgentemente de revisão e conta com medidas de controle não-estruturais que tratam das normas e legislações municipais.

Mais recentemente, no ano de 2022, Lei n° 6.914 é estabelecida. Em suma, esta legislação implementa a aplicação do Índice de Relevância Ambiental, que juntamente com outras estratégias municipais, como a Taxa de Permeabilidade, prevê parâmetros urbanísticos para o uso e ocupação do solo, controlando a drenagem local com a cobertura vegetal existente e futura (CAMPO GRANDE, 2022). Medidas como essa são essenciais para o enfrentamento das ocorrências de inundação e alagamento dentro do perímetro urbano, principalmente se somadas as medias estruturais, que modificam o

meio de forma mitigatória e adaptativa, ou seja, investindo em infraestrutura verde e melhorando os sistemas de drenagem.

4. CONCLUSÕES

Este trabalho discutiu e produziu o mapeamento de suscetibilidade de inundações e alagamentos no perímetro urbano de Campo Grande (MS) de acordo com dados físicos e dados coletados em veículos de imprensa dentro do recorte temporal de 2012 a 2021 utilizando a metodologia deste trabalho de forma inédita.

O processo de produzir mapas de suscetibilidade de inundações e alagamentos em perímetro urbano depende de várias fases e etapas para gerar um produto de qualidade. O mapeamento realizado aqui respondeu a algumas perguntas previamente realizadas e averigou a paridade dos dados simulados com a realidade do município.

Ademais, podemos observar e assimilar a importância da coleta de dados físicos que caracterizem o local estudado tanto quanto a coleta de dados secundários mediante relatos de imprensa, visto que através deles foi possível analisar quais fatores da área de estudo detém maior impacto perante a ocorrência ou não de inundações e alagamentos e como estes se correlacionam com fatores físicos e antrópicos, indicando problemáticas urbanas, tal como mostrado neste estudo.

Esta pesquisa evidenciou a gravidade da impermeabilização do solo e a necessidade de repensar a cobertura vegetal urbana, intencionando um ganho de permeabilidade e a consequente retenção do escoamento superficial de acordo com o ranqueamento realizado utilizando o método SVM de *Machine Learning*. Este método indicou o solo, a proximidade com corpos hídricos, a direção da drenagem e a Carta Geotécnica como os fatores mais preponderantes para a ocorrência de inundações e alagamentos. Em síntese, podemos afirmar que os métodos aplicados se mostram eficazes para a análise apresentada e que é possível conseguir dados secundários sobre o tema através de relatos de imprensa, ou seja, dados que agregam a vivência do munícipe e que enriquecem o estudo para que sirva de suporte fomento para atualizações e novas regulamentações.

Os resultados obtidos contribuem ativamente para o aprimoramento das normas e legislações municipais. Portanto, o uso ativo de Aprendizagem de máquina para a fiscalização, previsão e mapeamento de inundações e alagamentos em perímetro urbano

se mostra exequível e deve ser incentivado futuramente para que se torne ferramenta habitual nos processos de planejamento urbano.

5. Referências

BORGES, Luiz Eduardo. **Python para desenvolvedores: aborda Python 3.3**. Novatec Editora, p.14-16 2014.

BRASIL. **Lei nº 10.257, de 10 de julho de 2001**. Dispõe sobre o Estatuto da Cidade. Diário Oficial da União, Brasília, DF, 11 jul. 2001. Seção 1, p. 1.

CAMPO GRANDE NEWS. 2015. **Córrego Segredo transborda e pessoas ficam presas em veículos**. Disponível em: <https://www.campograndenews.com.br/cidades/capital/corrego-segredo-transborda-e-pessoas-ficam-presas-em-veiculos>. Acesso em: 15 nov. 2023

CAMPO GRANDE NEWS. 2021. **Transtornos para moradores: alagamentos danificaram até Kombi durante chuva**. Disponível em: <https://www.campograndenews.com.br/cidades/capital/transtornos-para-moradores-alagamentos-danificaram-ate-kombi-durante-chuva>. Acesso em: 15 nov. 2023

CAMPO GRANDE. Prefeitura Municipal de Campo Grande. Secretaria Municipal do Meio Ambiente e Desenvolvimento Urbano (SEMADUR). **Índice de Relevância Ambiental (IA)**. Disponível em: <https://www.campogrande.ms.gov.br/semadur/indice-de-relevancia-ambiental-ia/>. Acesso em: 1 dez. 2023

DA SILVA, Ananda Ramos; DOS SANTOS, Vânia Maria Nunes. **O papel da participação social na Redução de Riscos de Desastres no Brasil**. Labor e Engenharia, v. 16, p. e022012-e022012, 2022.

DE ANDRADE, Sidgley Camargo et al. **The effect of intra-urban mobility flows on the spatial heterogeneity of social media activity: investigating the response to rainfall events**. International Journal of Geographical Information Science, p. 1-26, 2021.

DULHARE, Uma N.; AHMAD, Khaleel; AHMAD, Khairol Amali Bin (Ed.). **Machine learning and big data: concepts, algorithms, tools and applications**. John Wiley & sons, 2020.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos (Rio de Janeiro, RJ). **Reunião Técnica de Levantamento de Solos**. Rio de Janeiro, 2009. (EMBRAPA-SNLCS. Miscelânea, 1).

FRANCI, Francesca et al. **Satellite remote sensing and GIS-based multi-criteria analysis for flood hazard mapping**. Natural Hazards, v. 83, n. 1, p. 31-51, 2016.

- GABRIELS, Karen; WILLEMS, Patrick; VAN ORSHOVEN, Jos. **A data-driven analysis, and its limitations, of the spatial flood archive of Flanders, Belgium to assess the impact of soil sealing on flood volume and extent.** PloS one, v. 15, n. 10, p. e0239583, 2020.
- GALLINDO, I. et al. (2024). **Mapeamento de inundações e alagamentos baseado em aprendizagem de máquina a partir de relatos da imprensa e fatores naturais e antrópicos.** Zenodo. <https://doi.org/10.5281/zenodo.10818968>.
- HUSSAIN, Muhammad et al. **GIS-Based Multi-Criteria Approach for Flood Vulnerability Assessment and Mapping in District Shangla: Khyber Pakhtunkhwa, Pakistan.** Sustainability, v. 13, n. 6, p. 3126, 2021.
- IBGE, 2023. **MDE - Modelo Digital de Elevação.** Disponível em: <https://www.ibge.gov.br/geociencias/modelos-digitais-de-superficie/modelos-digitais-de-superficie/10856-mde-modelo-digital-de-elevacao.html?=&t=o-que-e>. Acesso em: 6 jan. 2023.
- LIMA, Altieris Porfírio; AMORIM, Margarete Cristiane de Costa Trindade. **Análise de episódios de alagamentos e inundações urbanas na cidade de São Carlos a partir de notícias de jornal.** Revista Brasileira de Climatologia, v. 15, 2015.
- LUSA, Lara et al. **Gradient boosting for high-dimensional prediction of rare events.** Computational Statistics & Data Analysis, v. 113, p. 19-37, 2017.
- MUNARO, Ana Carolyne John. **Mapeamento da percepção social e aspectos físicos dos alagamentos e inundações no contexto do microplanejamento urbano.** Dissertação (Mestrado em Recursos Naturais) - Programa de Pós-Graduação em Recursos Naturais, Universidade Federal de Mato Grosso do Sul, Campo Grande (MS), 2023.
- NATEKIN, Alexey; KNOLL, Alois. Gradient boosting machines, a tutorial. **Frontiers in neurorobotics**, v. 7, p. 21, 2013.
- NOVAIS, Marcos Paulo – Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR, João Pessoa-PB, Brasil, INPE, 2015. **Análise Espacial de Bacias Hidrográficas a partir de SIG: um estudo da Bacia Hidrográfica do Itapicuru – Bahia.** Disponível em: <http://www.dsr.inpe.br/sbsr2015/files/p0541.pdf>. Acesso em: 15 Fev. 2023.
- OKADA, Genki et al. **The potential role of news media to construct a machine learning based damage mapping framework.** Remote Sensing, v. 13, n. 7, p. 1401, 2021.
- ONU, 2023. **Objetivos de Desenvolvimento Sustentável no Brasil.** Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 28 abril 2023.
- RANKING PESQUISA. **Pesquisa: Saiba quais são os principais sites de notícias e outros meios de comunicação mais utilizados em Campo Grande.** Disponível em: <https://rankingpesquisa.com.br/noticias/pesquisa-saiba-quais-sao-os-principais-sites-de-noticias-e-outros-meios-de-comunicacao-mais-utilizados-em-campo-grande/>. Acesso em: 24 nov. 2023

- TAMIRU, Habtamu; DINKA, Megersa O. **Artificial Intelligence in Geospatial Analysis for Flood Vulnerability Assessment: A Case of Dire Dawa Watershed, Awash Basin, Ethiopia**. The Scientific World Journal, v. 2021, 2021.
- TEHRANY, Mahyat Shafapour et al. **Flood susceptibility assessment using GIS-based support vector machine model with different kernel types**. Catena, v. 125, p. 91-101, 2015.
- TELLA, Abdulwaheed; BALOGUN, Abdul-Lateef. **Ensemble fuzzy MCDM for spatial assessment of flood susceptibility in Ibadan, Nigeria**. Natural Hazards, v. 104, n. 3, p. 2277-2306, 2020.
- TIEN BUI, Duie et al. **Flood spatial modeling in northern Iran using remote sensing and gis: A comparison between evidential belief functions and its ensemble with a multivariate logistic regression model**. Remote Sensing, v. 11, n. 13, p. 1589, 2019.
- WAQAS, Hassan et al. **Flash flood susceptibility assessment and zonation using an integrating analytic hierarchy process and frequency ratio model for the Chitral District, Khyber Pakhtunkhwa, Pakistan**. Water, v. 13, n. 12, p. 1650, 2021.
- YING, Xue. **An overview of overfitting and its solutions**. In: Journal of physics: Conference series. IOP Publishing, 2019. p. 02, 2022.
- YOUSSEF, Ahmed M.; PRADHAN, Biswajeet; SEFRY, Saleh A. **Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models**. Environmental Earth Sciences, v. 75, n. 1, p. 1-16, 2016.
- ZARONI, Maria José; SANTOS, Humberto – Embrapa solos 2021. **Solos tropicais**. Disponível em: <https://www.embrapa.br/agencia-de-informacao-tecnologica/tematicas/solos-tropicais/sibcs/chave-do-sibcs/latossolos/latossolos-vermelhos>. Acesso em: 6 jan. 2023.

CONCLUSÕES GERAIS

Esta dissertação teve por objetivo averiguar a possibilidade de produzir mapeamentos de ocorrência de inundações e alagamentos dentro do perímetro urbano, utilizando metodologias apoiadas em Inteligência Artificial de *Machine Learning*.

O primeiro capítulo deste trabalho evidenciou a fundamental importância da elaboração de mapeamentos de suscetibilidade para áreas urbanas acometidas por inundações e alagamentos. Indicou a concentração da pesquisa científica nos caminhos metodológicos viáveis, sugerindo abordagens e fatores físicos necessários para a produção desse mapeamento.

O segundo capítulo atestou a aplicabilidade da utilização de dados secundários (obtidos mediante veículos de imprensa) somado à fatores físicos para a elaboração de mapeamentos simulados de áreas suscetíveis à inundação e alagamento mediante uso de Inteligência Artificial (*Machine Learning* com o método *Gradient Boosting Classifier*) de forma coerente com a realidade da área.

O processo percorrido destaca a importância da temática, possíveis causas e prejuízos de tais eventos, propiciou o aprofundamento nas metodologias atuais possíveis e os requisitos fundamentais para a promoção de mapeamentos que sejam capazes de nortear futuros planejamentos, mitigações e adaptações em áreas acometidas com inundações e alagamentos urbanos realizados pelo Poder Público.

Com isso, conclui-se que o mapeamento da ocorrência de inundações e alagamentos, mediante uso de dados preditores (características ambientais, naturais e antrópicas) somado à coleta de dados secundários, obtidos mediante reportagens em veículos de imprensa, se mostrou como uma ferramenta emergente para detectar possíveis ocorrências de inundações e alagamentos em áreas previamente não amostradas pela mídia em áreas urbanas.

APÊNDICE 1: Material Suplementar

Tabela 16: Artigos do Portfólio

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
DE ANDRADE et al. (2021)	-	-	-	7	Não há estudos urbanos suficientes que se baseiem na realidade expressa pelas mídias digitais. Utilizando o Twitter, foi possível identificar atividades relacionadas com a chuva e as desigualdades intraurbanas e socioeconômicas.
ESFANDIARI et al. (2020)	7	6	9	7	Modelos de Aprendizagem de Máquinas tem o potencial de prever inundações e alertar os planejadores públicos, mas para isso é necessário testar e validar os dados. Neste estudo foram utilizados o método de Floresta Aleatória (FR) para definir o período de cheia.
JAHANGIR, REINEH e ABOLGHASEMI (2019)	36	43	36	7	Inundações é o desastre natural mais significativo pois causa mortalidade e perdas econômicas, através do método de Rede Neural Artificial (RNA), será possível compreender em qual escala acontecem esses desastres e instruir os planejadores urbanos.
NORALLAHI e SEYED KABOLI (2021)	5	4	6	7	O veloz crescimento urbano impermeabilizado, a má drenagem urbana e as mudanças climáticas ocasionam as inundações. Para compreender este fenômeno dramático é necessário produzir um mapeamento de suscetibilidade através de inúmeros dados. Três metodologias foram aplicadas a fim de comparar suas acurácias: Produção de Conjunto de Regras de Algoritmo Genérico, Entropia Máxima e Floresta Aleatória.
CHEN et al. (2020)	35	26	39	6,5	Estudo sobre a interação da população com uma área de delta, utilizando a estrutura conceitual DPSIR (força, pressão, estado, impacto e resposta). Métodos de Aprendizagem de Máquina foram usados a fim de evidenciar o risco de inundação da área.
HUSSAIN et al. (2021)	3	1	3	6,5	A abordagem escolhida para propor um mapa de vulnerabilidade foi o método de Processo de Hierarquia Analítica (AHP) com fatores de entrada criteriosos e padronizados,

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
					almejando compreender mais a área para manejar as inundações.
OKADA et al. (2021)	4	4	5	6,5	O uso de Aprendizagem de Máquina para prever inundações a avaliar avarias nas edificações. As imagens SAR foram usadas para dar suporte aos fatores de entrada dos métodos de Aprendizagem de Máquina.
RAHMATI et al. (2019)	22	22	26	6,5	Estudo com o objetivo de promover uma análise sobre a eficiência da Rede Neural Auto-organização para mapeamento de inundações urbanas e verificar sua acurácia comparado com a realidade.
RINCÓN, KHAN e ARMENAKIS (2018)	38	32	78	6,5	Devido ao aumento nas ocorrências de inundações, a pesquisa vis desenvolver mapas de suscetibilidade de inundações a partir do método de Processo de Hierarquia Analítica (AHP). Foram evidenciados 4 fatores de entrada para que o resultado seja simplificado e eficiente.
SAMANTA et al. (2018)	58	52	80	6,5	O mapa de suscetibilidade se faz imprescindível para gerir a questão da hidrologia urbana. Desta forma, o método de Razão de Frequência (FR) foi aplicado para produzir esse mapa tão importante.
TAMIRU e DINKA (2021)	-	-	-	6,5	A inovação deste estudo está na sua completa coleta de fatores de entrada com inúmeras informações e todos esses dados serão utilizados na produção de um mapa de risco de inundações através do método RNA.
TEHRANY et al. (2015)	348	312	434	6,5	O método de Máquina de vetor de suporte (SVM) tem sido amplamente usado para mapear áreas de risco de inundações, por isso o estudo almeja avaliar a acurácia e a capacidade de previsão do método.
WAQAS et al. (2021)	5	4	9	6,5	O estudo proposto tem por objetivo comparar quais dos métodos aplicados: AHP e FR são mais eficazes para mapear áreas com riscos de inundações possibilitando diretrizes mitigadoras e de adaptação.
ARMENAKIS et al. (2017)	34	30	50	6	Os dados de Observação da Terra (EO) são imprescindíveis para mapear riscos de inundações onde há população e

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
					desenvolvimento econômico. Os mapas especiais capazes de contribuir são feitos a partir da modelagem 3D das cidades e ilustram os resultados obtidos.
BRANDT et al. (2021)	1	1	2	6	As inundações em ambiente urbano ocorrem em todo o globo e com auxílio da modelagem e das ferramentas de planejamento será possível gerir o problema. O risco de inundação precisa ser considerado pelos planejadores urbanos utilizando sempre de novas ferramentas e inovações.
DARABI et al. (2019)	157	143	197	6	O mapa de risco foi produzido usando dados hidrológicos e hidráulicos, através da Aprendizagem de Máquina, a fim de evitar inundações. O método GARP e QUEST foram testados para obter sua acurácia, já que são testes recentes para evidenciar inundações e alagamentos em tecido urbano.
DOU et al. (2018)	22	19	32	6	Promover políticas urbanas para melhorar a cidade e sanar problemas com inundações. Para tal, é necessário mapear zonas problemáticas, utilizando, então o método multiparâmetro de risco de inundação.
MACALALAD et al. (2021)	-	-	-	6	As bacias hidrográficas precisam ser avaliadas sobre seus fluxos d'água, mas seu processo é difícil, sendo assim o modelo PBDH utiliza de sensoriamento remoto para colher os dados e enviá-los para análise.
MOHAMED e EL-RAEY (2020)	19	17	34	6	Há inúmeras questões que envolvem problemas hídricos, mas a preocupação mais urgente é quanto às inundações, dessa forma, é vital estudar quais são os locais com pontos vulneráveis.
SARKAR et al. (2021)	4	3	13	6	Os alagamentos frequentes durante as precipitações evidenciaram um estudo aprofundado sobre zonas susceptíveis a inundações e alagamentos, dessa forma, uma simulação Mike Urban foi aplicada para compreender como a cidade se comporta.

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
SKILODIMOU, BATHRELLOS e ALEXAKIS (2021)	8	8	9	6	Um estudo inovador quando questiona sobre riscos de inundação após queimadas em áreas urbanas. O método utilizado foi o Processo de Hierarquia Analítica (AHP) que tem por objetivo descobrir se há conexões entre inundações e queimadas.
TIEN BUI et al. (2019);	78	69	89	6	O objetivo do estudo é avaliar qual a performance de um modelo (Função de Crença Evidência) com o método Regressão Logística (LR) para a produção de mapas de suscetibilidade de inundações em áreas urbanas.
YEGANEH e SABRI (2014)	18	-	29	6	Este estudo visa compreender quais variáveis são mais significativas para desenvolver mapas de risco de inundações em áreas urbanizadas e possibilitando formas de propor a prevenção e a mitigação.
DI SALVO et al. (2018)	22	19	27	5,5	O estudo se propõe a, baseado em GIS, produzir o mapeamento de risco através da matriz de impacto potencial e somas de suscetibilidades.
DUY et al. (2017)	10	10	26	5,5	Mediante o intenso crescimento urbano e com as mudanças climáticas, faz-se necessário entender como as inundações e alagamentos ocorrem a fim de evita-las e mitiga-las utilizando teorias compactas.
GABRIELS, WILLEMS e VAN ORSHOVEN (2020)	-	-	-	5,5	A urbanização, com todos seus problemas de impermeabilização, evidencia um problema com as bacias hidrográficas e através dos métodos de Árvore de Regressão e Regressão Vetorial de Suporte, é possível compilar os dados e entender as vulnerabilidades da área em questão.
GIGOVÍĆ et al. (2017)	111	95	151	5,5	Para compreender de forma mais aprofundada as vulnerabilidades da área no âmbito das inundações e alagamentos, o método de Análise Multicritério baseado em GIS foi utilizado.
HOSSAIN e MENG (2020)	19	19	30	5,5	As inundações se tornaram um problema frequente e, por isso, devem ser mapeadas e avaliadas a fim de evitarem danos catastróficos. O mapeamento de

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
					risco utilizou dados GIS para analisar áreas conflituosas.
LEE et al. (2017)	126	117	164	5,5	O método Floresta Aleatória (RF) foi utilizado para produzir o mapa de suscetibilidade de inundações através de dados GIS, visando fornecer informações para planejadores urbanos.
MOHAMED (2021)	1	1	2	5,5	A intensidade e frequência da ocorrência de inundações se tornou insustentável e foi preciso desenvolver um mapa de vulnerabilidade, com o método Processo de Hierarquia Analítica (AHP), da área para compreender quais procedimentos promover no local.
TEHRANY et al. (2014)	128	117	174	5,5	O objetivo primário desta pesquisa é superar alguns problemas relacionados com a metodologia Regressão Logística para produzir o mapa de suscetibilidade de inundações para a área desejada.
YOUSSEF et al. (2016)	-	51	89	5,5	Depois de uma devastadora inundação, evidenciar as causas foi uma forte motivação para este estudo. Os fatores foram levantados e a fim de evitar novos problemas com inundações urbanas.
YOUSSEF, PRADHAN e SEFRY (2016)	-	91	132	5,5	Devido a problemas graves com inundações, o gerenciamento de áreas de risco torna-se fundamental. Dois métodos foram aplicados, Razão de Frequência e Regressão Logística, para que se pudesse comparar as acurácias.
CHOUBIN et al. (2019)	271	253	315	5	As cheias, quando efeito de desastre natural, causa muitos prejuízos para a natureza e para a vida humana. É imprescindível modelar esse fenômeno para evitar grandes perdas. Um mapa de suscetibilidade foi realizado juntamente com o mapa de inventário de inundações, a fim de compreender melhor a área.
DA SILVA et al. (2020)	10	6	17	5	As inundações são as mais catastróficas entre todos os desastres naturais, causando a perda de vidas, de infraestrutura

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
					e financeiras, por isso é fundamental formular políticas para gerir este grande problema urbano. Dessa forma, a modelagem baseada na Teoria da Utilidade Multiatributos foi aplicada, visando detectar inundações e garantir um melhor planejamento.
FAREGH e BENKHALED (2021)	-	-	1	5	Promover um mapa de risco que auxilie os planejadores urbanos a tomarem decisões é vital para evitar problemas socioeconômicos, por isso o método de análise multicritério foi implementado.
FRANCI et al. (2016)	27	20	41	5	Usando o sensoriamento remoto com uma análise multicritério é possível produzir o mapeamento de inundações, classificando a intensidade dessas inundações e proporcionando análises urbanas.
GARCIA, PICCOLO e BOHN (2021)	-	-	-	5	O objetivo central é estimar os riscos de inundações através de uma avaliação multicritério utilizando dados GIS para propor diretrizes e estratégias urbanísticas.
KARYMBALIS et al. (2021)	-	-	-	5	O mapeamento de risco de inundação é fundamental para tomadas de decisões de planejamento urbano. Dessa forma, o método AHP foi utilizada para prever como a área escolhida se comporta em relação à inundações.
LITTIDEJ e BUASRI (2019)	-	3	7	5	O uso do solo é muito importante para a bacia hidrográfica, por isso é importante propor modelos de Autônomos Celulares (CA) das áreas construídas para compreender a estabilidade dessas áreas em relação às inundações.
PANGALI SHARMA et al. (2021)	1	1	2	5	O estudo da geomorfologia está intrínseco com a qualidade do mapa de risco de inundações, principalmente a densidade da drenagem, relevo e a precipitação. Para isso foi proposto um estudo a fim de elaborar diretrizes.
SARHADI, SOLTANI e MODARRES (2012)	117	105	184	5	Utilizando o método HEC-RAS as várzeas não calibradas tornaram-se alvo de estudo. O mapa de vulnerabilidade de inundação da área foi comparado

Artigo	Citações (Scopus)	Citações (WoS)	Citações (Google)	Pontuação	Objetivo
					com o uso do solo para propor novas diretrizes.
TELLA e BALOGUN (2020)	8	7	13	5	Os mapas de suscetibilidade são dependentes dos dados inseridos, por esse fato vários métodos foram utilizados e os dados foram avaliados para que se pudesse comparar as acurácias de cada método (AHP e FAHP).

Fonte: Elaborado pela autora

Material Suplementar

<https://zenodo.org/records/10818968>

Fonte: Gallindo et al. (2024)

**APÊNDICE 2: Código para definir as coordenadas de cada ponto
mencionado nos relatos de imprensa**

```

1 # -*- coding: utf-8 -*-
2 """
3 Spyder Editor
4
5 This is a temporary script file.
6 """
7 #importar os pacotes necessarios
8 import pandas as pd
9 import numpy as np
10 from sklearn import preprocessing
11 from sklearn.model_selection import train_test_split
12 import sklearn.svm as svm
13 import sklearn.neural_network as neural_network
14 import sklearn.ensemble as ensemble
15 from sklearn import metrics
16 from sklearn import tree
17 import matplotlib.pyplot as plt
18 import seaborn as sns
19 import geopy as gp
20 import pycep_correios as pycep
21 from geopy.geocoders import Nominatim
22 from geopy.extra.rate_limiter import RateLimiter
23
24 #importar os dados
25 midias = pd.read_excel ('dadosTOTAIS.xlsx').dropna()
26 #tratamento preliminar dos dados
27
28
29 X = midias [['VEÍCULO', 'AUTOR', 'MANCHETE', 'SEÇÃO', 'TIPO',
30            'PONTOS', 'BAIRRO', 'CEP', 'RU', 'LINKS', 'ACESSO']]
31 X = X.reset_index()
32 #evidenciando os ceps
33 X['CEP'] = X['CEP'].astype(str)
34 address = []
35 for i in range(len(X['CEP'])):
36     address.append(pycep.get_address_from_cep(X['CEP'][i],
37                                             webservice=pycep.WebService.CORREIOS))
38
39 geolocator = Nominatim(user_agent="my_app")
40 #identificando o lat e long
41 location = []
42 for i in range(len(address)):
43     location.append(geolocator.geocode(address[i]['logradouro']+',
44                                         '+address[i]['cidade']+', '+address[i]['uf']+', Brazil'))
45
46 #extraíndo lat e long ignorando erros
47 latitude = []
48 longitude = []
49 for i in range(len(location)):
50     if location[i] is not None:
51         latitude.append(location[i].latitude)
52         longitude.append(location[i].longitude)
53     else:
54         latitude.append(np.nan)
55         longitude.append(np.nan)
56 #tabela midias3
57 X['latitude'] = latitude
58 X['longitude'] = longitude
59
60 X.to_csv('midias3.csv')

```

APÊNDICE 3: Código para aplicar os métodos de Machine Learning

```

1 # -*- coding: utf-8 -*-
2 """
3 Created on 28-03-2023

4 @author: IOG/JAAA
5 """

6 import pandas as pd
7 import numpy as np
8 from sklearn import preprocessing
9 from sklearn.feature_selection import RFE
10 from sklearn.model_selection import train_test_split
11 from sklearn import ensemble
12 from sklearn import metrics
13 from sklearn import svm
14 from sklearn import tree
15 from sklearn import neural_network
16 from sklearn import neighbors
17 from sklearn import linear_model
18 from sklearn.neighbors import KNeighborsClassifier
19 from sklearn.neighbors import NearestNeighbors
20 from sklearn.semi_supervised import SelfTrainingClassifier
21 import matplotlib.pyplot as plt
22 import seaborn as sns

23 #importar os dados
24 data=pd.read_csv('zonais 6.csv', sep=';').dropna()

25 ###
26 #numero aleatorio das simulacoes
27 rs = np.random.random_integers(0,1000)

28 # #tratamento preliminar dos dados
29 data.NUMPOINTS = np.where(data.NUMPOINTS >
    data.NUMPOINTS.quantile(1), data.NUMPOINTS.quantile(1).astype(int),
    data.NUMPOINTS)
30 data2 = data[(data.NUMPOINTS == 0) & (data['PRX
    CHmean']>1000)].sample(n=253, random_state=0)
31 data1 = data[data.NUMPOINTS != 0]
32 data1 = pd.concat([data2,data1])
33 data1 = data1.sample(frac=1).reset_index(drop=True)

34 y = data1.NUMPOINTS

35 X = data1[['SLOPEmean', 'NDVImean', 'PRX
    CHmean', 'LULCmajori', 'SPImean', 'TWImean', 'SOILSmajor', 'VEGmajorit',
    'MDEmean', 'FLOWmean', 'BHCGmean', 'DRAINAGEma', 'GEOTECmajo',
    'MDEVERT_me']] #14

36 #separacao dos datasets em sub-datasets de treinamento e teste
37 X_train, X_test, y_train, y_test = train_test_split(X, y,
    test_size=0.3, random_state=(0))

38 ###
39 #criar shape treino e validacao
40 Xmapa = data1[['X', 'Y', 'SLOPEmean', 'NDVImean', 'PRX
    CHmean', 'LULCmajori',
41 'SPImean', 'TWImean', 'SOILSmajor', 'VEGmajorit',
    'MDEmean', 'FLOWmean', 'BHCGmean',
42 'DRAINAGEma', 'GEOTECmajo', 'MDEVERT_me']] #14
43 X_trainmapa, X_testmapa, y_trainmapa, y_testmapa =

```

```

train_test_split(Xmapa, y, test_size=0.3, random_state=(0))

44 X_trainmapa.to_csv('train.csv')
45 X_testmapa.to_csv('test.csv')

46 #%%
47 #selecao e transformacao dos dados
48 min_max_scaler = preprocessing.MinMaxScaler().fit(X_train)
49 X = min_max_scaler.fit_transform(X)

50 X_train = min_max_scaler.fit_transform(X_train)
51 X_test = min_max_scaler.fit_transform(X_test)

52 #%%
53 #analise de importancia
54 n_features = 14 #escolher numero de features

55 estimator = svm.SVC(kernel= 'linear')
56 selector = RFE(estimator, n_features_to_select = n_features,
step=1)
57 selector = selector.fit(X, y)
58 print("classifier")
59 print(selector.support_)
60 print(selector.ranking_)
61 X = X[:,selector.support_]
62 X_train = X_train[:,selector.support_]
63 X_test = X_test[:,selector.support_]

64 #%%
65 #criacao e treinamento dos modelos

66 #SVM
67 mod1 = svm.SVC().fit(X_train, y_train)
68 #analise da acuracia
69 y_trained_mod1 = mod1.predict(X_train)
70 y_tested_mod1 = mod1.predict(X_test)
71 y_all_mod1 = mod1.predict(X)

72 print('MOD1')
73 accu_mod1_train = metrics.accuracy_score(y_train, y_trained_mod1)
74 print('Train')
75 print(accu_mod1_train)

76 accu_mod1_test = metrics.accuracy_score(y_test, y_tested_mod1)
77 print('Test')
78 print(accu_mod1_test)

79 accu_mod1_all = metrics.accuracy_score(y, y_all_mod1)
80 print('All')
81 print(accu_mod1_all)
82 print('')

83 #NEURAL
84 mod2 = neural_network.MLPClassifier(activation= 'relu',
random_state=(rs)).fit(X_train, y_train)
85 #analise da acuracia
86 y_trained_mod2 = mod2.predict(X_train)
87 y_tested_mod2 = mod2.predict(X_test)
88 y_all_mod2 = mod2.predict(X)

89 print('MOD2')

```

```

90 accu_mod2_train = metrics.accuracy_score(y_train, y_trained_mod2)
91 print('Train')
92 print(accu_mod2_train)

93 accu_mod2_test = metrics.accuracy_score(y_test, y_tested_mod2)
94 print('Test')
95 print(accu_mod2_test)

96 accu_mod2_all = metrics.accuracy_score(y, y_all_mod2)
97 print('All')
98 print(accu_mod2_all)
99 print('')

100 #RANDOM FOREST
101 mod3 =
    ensemble.RandomForestClassifier(random_state=(rs)).fit(X_train,
        y_train)
102 #analise da acuracia
103 y_trained_mod3 = mod3.predict(X_train)
104 y_tested_mod3 = mod3.predict(X_test)
105 y_all_mod3 = mod3.predict(X)

106 print('MOD3')
107 accu_mod3_train = metrics.accuracy_score(y_train, y_trained_mod3)
108 print('Train')
109 print(accu_mod3_train)

110 accu_mod3_test = metrics.accuracy_score(y_test, y_tested_mod3)
111 print('Test')
112 print(accu_mod3_test)

113 accu_mod3_all = metrics.accuracy_score(y, y_all_mod3)
114 print('All')
115 print(accu_mod3_all)
116 print('')

117 #K NEIGHBORS
118 mod4 = KNeighborsClassifier(n_neighbors=1).fit(X_train, y_train)
119 #analise da acuracia
120 y_trained_mod4 = mod4.predict(X_train)
121 y_tested_mod4 = mod4.predict(X_test)
122 y_all_mod4 = mod4.predict(X)

123 print('MOD4')
124 accu_mod4_train = metrics.accuracy_score(y_train, y_trained_mod4)
125 print('Train')
126 print(accu_mod4_train)

127 accu_mod4_test = metrics.accuracy_score(y_test, y_tested_mod4)
128 print('Test')
129 print(accu_mod4_test)

130 accu_mod4_all = metrics.accuracy_score(y, y_all_mod4)
131 print('All')
132 print(accu_mod4_all)
133 print('')

134 #TREES
135 mod5 =
    ensemble.ExtraTreesClassifier(random_state=(rs)).fit(X_train,
        y_train)

```

```

136 #analise da acuracia
137 y_trained_mod5 = mod5.predict(X_train)
138 y_tested_mod5 = mod5.predict(X_test)
139 y_all_mod5 = mod5.predict(X)

140 print('MOD5')
141 accu_mod5_train = metrics.accuracy_score(y_train, y_trained_mod5)
142 print('Train')
143 print(accu_mod5_train)

144 accu_mod5_test = metrics.accuracy_score(y_test, y_tested_mod5)
145 print('Test')
146 print(accu_mod5_test)

147 accu_mod5_all = metrics.accuracy_score(y, y_all_mod5)
148 print('All')
149 print(accu_mod5_all)
150 print('')
151 #analise da acuracia - desempenho

152 #REGRESSAO LINEAR
153 mod6 = linear_model.LogisticRegression().fit(X_train, y_train)
154 #analise da acuracia
155 y_trained_mod6 = mod6.predict(X_train)
156 y_tested_mod6 = mod6.predict(X_test)
157 y_all_mod6 = mod6.predict(X)

158 print('MOD6')
159 accu_mod6_train = metrics.accuracy_score(y_train, y_trained_mod6)
160 print('Train')
161 print(accu_mod6_train)

162 accu_mod6_test = metrics.accuracy_score(y_test, y_tested_mod6)
163 print('Test')
164 print(accu_mod6_test)

165 accu_mod6_all = metrics.accuracy_score(y, y_all_mod6)
166 print('All')
167 print(accu_mod6_all)
168 print('')

169 #CLASSIFICACAO LINEAR
170 mod7 = linear_model.RidgeClassifier().fit(X_train, y_train)
171 #analise da acuracia
172 y_trained_mod7 = mod7.predict(X_train)
173 y_tested_mod7 = mod7.predict(X_test)
174 y_all_mod7 = mod7.predict(X)

175 print('MOD7')
176 accu_mod7_train = metrics.accuracy_score(y_train, y_trained_mod7)
177 print('Train')
178 print(accu_mod7_train)

179 accu_mod7_test = metrics.accuracy_score(y_test, y_tested_mod7)
180 print('Test')
181 print(accu_mod7_test)

182 accu_mod7_all = metrics.accuracy_score(y, y_all_mod7)
183 print('All')
184 print(accu_mod7_all)
185 print('')

```

```

186 #CLASSIFICACAO SELF TRAINING
187 mod8 =
    SelfTrainingClassifier(base_estimator=estimator).fit(X_train,
y_train)
188 #analise da acuracia
189 y_trained_mod8 = mod8.predict(X_train)
190 y_tested_mod8 = mod8.predict(X_test)
191 y_all_mod8 = mod8.predict(X)

192 print('MOD8')
193 accu_mod8_train = metrics.accuracy_score(y_train, y_trained_mod8)
194 print('Train')
195 print(accu_mod8_train)

196 accu_mod8_test = metrics.accuracy_score(y_test, y_tested_mod8)
197 print('Test')
198 print(accu_mod8_test)

199 accu_mod8_all = metrics.accuracy_score(y, y_all_mod8)
200 print('All')
201 print(accu_mod8_all)
202 print('')

203 #GRADIANT BOOSTING
204 mod9 = ensemble.GradientBoostingClassifier(random_state=(rs),
n_estimators=14, learning_rate=0.1).fit(X_train, y_train)
205 #analise da acuracia
206 y_trained_mod9 = mod9.predict(X_train)
207 y_tested_mod9 = mod9.predict(X_test)
208 y_all_mod9 = mod9.predict(X)

209 print('MOD9')
210 accu_mod9_train = metrics.accuracy_score(y_train, y_trained_mod9)
211 print('Train')
212 print(accu_mod9_train)

213 accu_mod9_test = metrics.accuracy_score(y_test, y_tested_mod9)
214 print('Test')
215 print(accu_mod9_test)

216 accu_mod9_all = metrics.accuracy_score(y, y_all_mod9)
217 print('All')
218 print(accu_mod9_all)
219 print('')

220 #%%simulacao
221 X_sim = data[['SLOPEmean', 'NDVImean', 'PRX
    CHmean', 'LULCmajori', 'SPImean', 'TWImean',
222 'SOILSmajor', 'VEGmajorit', 'MDEmean', 'FLOWmean', 'BHCgmean',
223 'DRAINAGEma', 'GEOTECmajo', 'MDEVERT_me']]

224 X_sim = min_max_scaler.fit_transform(X_sim)[: ,selector.support_]

225 mod_chosen = mod9 #escolha o mod que teve o melhor desempenho

226 y_sim = mod_chosen.predict(X_sim)

227 data['inund1'] = y_sim

228 #plotar os resultados

```

```

229 fig1, axs = plt.subplots(1,2, figsize=(15,7),
    gridspec_kw={'width_ratios': [1.2, 1.5]})
230 (ax1, ax2) = axs
231 #observado
232 ax1.set_title('Observado')
233 sns.scatterplot(x=data1['X'], y=data1['Y'],
    hue=data1['NUMPOINTS'], ax=ax1, palette='coolwarm', alpha=0.8)
234 #ax1.invert_yaxis()

235 #simulado
236 ax2.set_title('Simulado')
237 sns.heatmap(data.pivot('Y','X','inund1'), ax=ax2, cmap='coolwarm',
    alpha=0.8)
238 ax2.invert_yaxis()

239 #fatores(preditores)
240 fig2, axs = plt.subplots(3,5, figsize=(35,21), sharex=True,
    sharey=True)
241 (ax1, ax2, ax3, ax4, ax5), (ax6, ax7, ax8, ax9, ax10), (ax11,
    ax12, ax13, ax14, ax15) = axs

242 sns.heatmap(data.pivot('Y','X','SLOPEmean'), ax=ax1,
    cmap='coolwarm', alpha=0.8)
243 ax1.invert_yaxis()

244 sns.heatmap(data.pivot('Y','X','NDVImean'), ax=ax2,
    cmap='coolwarm', alpha=0.8)
245 ax2.invert_yaxis()

246 sns.heatmap(data.pivot('Y','X','PRX CHmean'), ax=ax3,
    cmap='coolwarm', alpha=0.8)
247 ax3.invert_yaxis()

248 sns.heatmap(data.pivot('Y','X','LULCmajori'), ax=ax4,
    cmap='coolwarm', alpha=0.8)
249 ax4.invert_yaxis()

250 sns.heatmap(data.pivot('Y','X','SPImean'), ax=ax5,
    cmap='coolwarm', alpha=0.8)
251 ax5.invert_yaxis()

252 sns.heatmap(data.pivot('Y','X','TWImean'), ax=ax6,
    cmap='coolwarm', alpha=0.8)
253 ax6.invert_yaxis()

254 sns.heatmap(data.pivot('Y','X','SOILSmajor'), ax=ax7,
    cmap='coolwarm', alpha=0.8)
255 ax7.invert_yaxis()

256 sns.heatmap(data.pivot('Y','X','VEGmajorit'), ax=ax8,
    cmap='coolwarm', alpha=0.8)
257 ax8.invert_yaxis()

258 sns.heatmap(data.pivot('Y','X','MDEmean'), ax=ax9,
    cmap='coolwarm', alpha=0.8)
259 ax9.invert_yaxis()

260 sns.heatmap(data.pivot('Y','X','FLOWmean'), ax=ax10,
    cmap='coolwarm', alpha=0.8)
261 ax10.invert_yaxis()

```

```
262 sns.heatmap(data.pivot('Y','X','BHCgmean'), ax=ax11,  
    cmap='coolwarm', alpha=0.8)  
263 ax11.invert_yaxis()  
  
264 sns.heatmap(data.pivot('Y','X','DRAINAGEma'), ax=ax12,  
    cmap='coolwarm', alpha=0.8)  
265 ax12.invert_yaxis()  
  
266 sns.heatmap(data.pivot('Y','X','GEOTECmajo'), ax=ax13,  
    cmap='coolwarm', alpha=0.8)  
267 ax13.invert_yaxis()  
  
268 sns.heatmap(data.pivot('Y','X','MDEVERT_me'), ax=ax14,  
    cmap='coolwarm', alpha=0.8)  
269 ax14.invert_yaxis()  
  
270 sns.heatmap(data.pivot('Y','X','inund1'), ax=ax15,  
    cmap='coolwarm', alpha=0.8)  
271 ax15.invert_yaxis()  
272 data.to_csv('final5.csv')
```