



Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul
Metodologia Científica PPGTA



Felipe Gavioli Diniz

**FATORES NATURAIS E ANTRÓPICOS QUE
INFLUENCIAM NA PRESENÇA DE NITRATO NO
AQUÍFERO SERRA GERAL LIVRE NO PERÍMETRO
URBANO DE CAMPO GRANDE - MS**

**FATORES NATURAIS E ANTRÓPICOS QUE
INFLUENCIAM NA PRESENÇA DE NITRATO NO
AQUÍFERO SERRA GERAL LIVRE NO PERÍMETRO
URBANO DE CAMPO GRANDE - MS**

Discente: Felipe Gavioli Diniz

Orientadora: Sandra Garcia Gabas

Co-orientador: Jamil Alexandre Ayach Anache

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
PROGRAMA DE PÓS GRADUAÇÃO DE TECNOLOGIAS AMBIENTAIS

FOLHA DE APROVAÇÃO

FELIPE GAVIOLI DINIZ

**FATORES NATURAIS E ANTRÓPICOS QUE INFLUENCIAM NA
PRESENÇA DE NITRATO NO AQUÍFERO SERRA GERAL LIVRE
NO PERÍMETRO URBANO DE CAMPO GRANDE - MS**

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 Tecnologias Ambientais. Estudo realizado dentro da linha de pesquisa de poluição de águas subterrâneas.

Aprovado em: 04 de julho de 2024

Banca Examinadora

Sandra Garcia Gabas, Doutora em Engenharia Civil, Universidade Federal de Mato Grosso do Sul - UFMS

Claudia Luciana Varnier, Doutora em Hidrogeologia, Universidade de São Paulo - USP

Natalia de Souza Pelinson, Doutora em Engenharia Hidráulica e Saneamento, Universidade Federal de São Carlos, UFSCar

RESUMO

O objetivo deste trabalho foi analisar quais fatores antrópicos e naturais que influenciam na presença do nitrato nas águas subterrâneas em área urbana utilizando aprendizado de máquina. O local da pesquisa foi o perímetro urbano do município de Campo Grande, capital do estado do Mato Grosso do Sul, que está localizada na porção norte do município. A metodologia envolveu a utilização do *software* QGIS para a produção de mapas e a linguagem Python junto com as bibliotecas Scikit-learn, NumPy, SciPy e Matplotlib para análise dos dados e desenvolvimento dos modelos de aprendizado de máquina. Foram consideradas as seguintes variáveis espaciais como influenciadoras do nível de nitrato: distância até a rede de água e esgoto, zona urbana, densidade populacional por bairro, declividade, hipsometria, Índice de Vegetação por Diferença Normalizada (NDVI), geologia, profundidade do nível da água, características geotécnicas e proximidade de sistemas de drenagem. Nove modelos estatísticos do grupo de classificadores e sete do grupo de regressores foram adotados para analisar os dados coletados em 68 poços em 2018. Os resultados preditivos mostraram que a zona urbana e a declividade foram as variáveis com maior influência sobre o valor de nitrato em ambos os modelos, por outro lado, a geologia foi a variável com menor influência nos modelos classificadores e a distância da rede de água nos modelos regressores. Os resultados mostraram que os modelos classificadores tiveram um desempenho superior aos modelos regressores, com cinco modelos apresentando um resultado com o grupo treino variando entre 0.60 a 0.80 e o grupo teste entre 0.40 a 0.60, destacando o modelo MLP Classifier com a melhor performance. Os demais modelos classificadores apresentaram *overfitting*, um problema que também afetou o desempenho dos modelos regressores, com dois modelos apresentando o mesmo problema. Além disso, cinco modelos regressores contaram com um resultado negativo e apenas o modelo Logistic Regressor relevou um resultado aceitável com o grupo treino em 0.60 e com um valor aproximado a este no grupo teste. Por meio da estimativa da distribuição do nitrato, foi constatado que as maiores concentrações ocorrem em áreas com maiores declives e alta densidade populacional. Por outro lado, as menores concentrações são encontradas em regiões onde não há registro de redes de esgoto, com menor declive, baixa densidade populacional e onde o nível da água subterrânea é mais profundo.

Palavras chaves: Águas Subterrâneas; Aprendizado de máquina; Hidrogeologia.

ABSTRACT

The objective of this work was to analyze which anthropogenic and natural factors influence the presence of nitrate in groundwater in urban areas using machine learning. The research site was the urban perimeter of the municipality of Campo Grande, capital of the state of Mato Grosso do Sul, which is located in the northern portion of the municipality. The methodology involved the use of the QGIS software for the production of maps and the Python language together with the Scikit-learn, NumPy, SciPy and Matplotlib libraries for data analysis and development of machine learning models. The following spatial variables were considered as influencing the nitrate level: distance to the water and sewage network, urban area, population density per neighborhood, slope, hypsometry, Normalized Difference Vegetation Index (NDVI), geology, water level depth, geotechnical characteristics and proximity to drainage systems. Nine statistical models from the classifier group and seven from the regressor group were adopted to analyze the data collected from 68 wells in 2018. The predictive results showed that the urban area and slope were the variables with the greatest influence on the nitrate value in both models. On the other hand, geology was the variable with the least influence in the classifier models and the distance from the water network in the regressor models. The results showed that the classifier models outperformed the regressor models, with five models presenting a result with the training group ranging from 0.60 to 0.80 and the test group between 0.40 and 0.60, highlighting the MLP Classifier model with the best performance. The other classifier models presented overfitting, a problem that also affected the performance of the regressor models, with two models presenting the same problem. Furthermore, five regressor models had a negative result and only the Logistic Regressor model revealed an acceptable result with the training group at 0.60 and with a value close to this in the test group. By estimating the distribution of nitrate, it was found that the highest concentrations occur in areas with greater slopes and high population density. On the other hand, the lowest concentrations are found in regions where there is no record of sewage networks, with less slope, low population density and where the groundwater level is deeper.

Key-words: Groundwater; Machine learning; hydrogeology.

AGRADECIMENTOS

Primeiramente agradecer a minha família, especialmente minha mãe por ao longo da minha vida, ter me ensinado o princípio da educação e acompanhar todo este processo de mestrado de perto, além do meu irmão, com suas dicas e conselhos.

Agradecer a minha orientadora Sandra Garcia Gabas e meu co-orientador Jamil Alexandre Ayach Anache pela orientação, conhecimento compartilhado, dicas, sugestões sobre o futuro, e principalmente paciência durante a elaboração deste trabalho, um objetivo que eu levava comigo desde o início da minha graduação.

Além dos familiares e professores, não deixar de mencionar meus amigos, especialmente os amigos que ganhei durante a minha graduação, que me apoiaram deste o início desta caminhada, ouvindo meus desabafos e auxiliando sobre assuntos pertinentes.

O presente trabalho contou com ajuda da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, que durante 24 meses, disponibilizou uma bolsa para que eu pudesse ter dedicação exclusiva para a pesquisa.

E por último, agradecer o Programa de Pós-Graduação de Tecnologias Ambientais (PPGTA) e a estrutura da Universidade Federal de Mato Grosso do Sul (UFMS), que me disponibilizaram uma ótima estrutura e diversas experiências para o meu desenvolvimento e ganho de experiência ao longo desses dois anos.

LISTA DE FIGURAS

Figura 01 - Exemplo de aquíferos livres e confinados.....	15
Figura 02 - Classificação de aquíferos a partir da estrutura das rochas	17
Figura 03 - Mapa geológico e hidrogeológico do município de Campo Grande, MS.....	20
Figura 04 - Mapa de localização do perímetro urbano de Campo Grande, MS	27
Figura 05 - Etapas metodológicas do processo de aprendizado de máquina para predição da distribuição da concentração de nitrato em aquífero livre	30
Figura 06 - Distribuição espacial dos poços tubulares no perímetro urbano de Campo Grande, MS	36
Figura 07 - Mapa de localização dos poços tubulares usados para treinamento e teste em Campo Grande, MS.....	41
Figura 08 - Mapa de hipsometria de Campo Grande, MS	42
Figura 09 - Mapa de declividade de Campo Grande, MS	43
Figura 10 - Índice de Vegetação da Diferença Normalizada (NDVI) de Campo Grande, MS.....	44
Figura 11 - Geologia da área urbana de Campo Grande, MS	45
Figura 12 - Profundidade do nível de água da cidade de Campo Grande	46
Figura 13 - Carta da isoespessura do solo de Campo Grande, MS	47
Figura 14 - Distância da rede de drenagem/Hidrografia de Campo Grande, MS.....	48
Figura 15 - Distância da rede de água de Campo Grande, MS	49
Figura 16 - Distância da rede de esgoto de Campo Grande, MS	50
Figura 17 - Fossas dos lotes de Campo Grande, MS	51
Figura 18 - Zona urbana de Campo Grande, MS.....	52
Figura 19 - População por bairro da cidade de Campo Grande, MS	53
Figura 20 - Desempenho dos modelos classificadores	54
Figura 21 - Desempenho dos modelos regressores	54
Figura 22 - Estimativa da distribuição da concentração do nitrato nas águas subterrâneas de Campo Grande, MS.....	60
Figura 23 - Estimativa da área com a maior concentração do nitrato nas águas subterrâneas de Campo Grande, MS	62

LISTA DE TABELAS

Tabela 01 - Características dos fatores de entrada.....	28
Tabela 02 – Propriedades do Satélite Sentinel-2 (EngeSat)	32
Tabela 03 - Importância das variáveis espaciais sobre a presença de nitrato nas águas subterrâneas de Campo Grande, MS	56

SUMÁRIO

1. INTRODUÇÃO	11
2. REVISÃO BIBLIOGRÁFICA	14
2.1 Aquíferos	14
2.1.1 Tipos de aquíferos	14
2.2 Hidrogeologia de Campo Grande e Mato Grosso do Sul.....	17
2.3 Tipos de captações de águas subterrâneas.....	19
2.3.1 Poços	19
2.3.2 Poços tubulares	20
2.4 Fontes poluidoras das águas subterrâneas em áreas urbanas	22
2.5 O uso do aprendizado de máquina para estudos sobre águas subterrâneas	24
3 MATERIAL E MÉTODOS.....	26
3.1 Área de estudo	26
3.2 Procedimentos metodológicos.....	27
3.2.1 Coleta dos aspectos naturais	30
3.2.1.1 Declividade e hipsometria.....	30
3.2.1.2 NDVI.....	31
3.2.1.3 Geologia.....	33
3.2.1.4 Profundidade do nível d'água.....	33
3.2.1.5 Espessura dos solos.....	33
3.2.1.6 Distância para drenagem/hidrografia.....	33
3.2.2 Coleta dos aspectos antrópicos.....	34
3.2.2.1 Distância para rede de água e rede de esgoto.....	34
3.2.2.2 Distância para fossas	34
3.2.2.3 Zona urbana	34
3.2.2.4 População por bairro	35
3.2.3 Poços tubulares da área de estudo - Campo Grande, MS.....	35
3.2.4 Processamento e análise dos dados.....	36
3.2.4.1 Neural Network.MLP Classifier.....	38
3.2.4.2 Linear Model.SGD Classifier.....	39
3.2.4.3 Linear Model.Ridge Classifier	39
3.2.4.4 Self Training Classifier.....	39
3.2.4.5 K Neighbors Classifier	39
3.2.4.6 Random Forest Classifier	40
3.2.4.7 Extra Trees Classifier	40
3.2.4.8 Gradient Boosting Classifier	40
3.2.4.9 Support Vector Classification (SVC).....	40
3.2.5 Treinamento e validação dos modelos do aprendizado de máquina	40

4 RESULTADOS E DISCUSSÃO	42
4.1 Resultado das variáveis espaciais	42
4.1.1 Hipsometria	42
4.1.2 Declividade	42
4.1.3 NDVI	43
4.1.4 Geologia.....	43
4.1.5 Profundidade do nível de água.....	45
4.1.6 Carta da isoessura do solo.....	46
4.1.7 Distância para drenagem/hidrografia.....	47
4.1.8 Distância para rede de água e de esgoto.....	48
4.1.9 Distância entre as fossas.....	50
4.1.10 Zona urbana	51
4.1.11 População por bairro	52
4.2 Desempenho dos modelos.....	53
4.3 Importância dos fatores naturais e antrópicos na ocorrência de nitrato	56
4.4 Estimativa da distribuição da concentração de nitrato na área urbana de Campo Grande, MS	59
5 CONCLUSÃO	63
6 REFERÊNCIAS BIBLIOGRÁFICAS	65
7 APÊNDICES	78
Apêndice A - Informações sobre os processos deferidos de Outorga de Direito de Uso de Água Subterrânea dos poços tubulares da cidade de Campo Grande, MS	64
Apêndice B - Informações sobre as características dos poços tubulares e aquíferos da cidade de Campo Grande, MS	88
Apêndice C - Código para aplicar os métodos de Machine Learning	99

1. INTRODUÇÃO

Em diversos lugares no planeta, as águas subterrâneas surgem como um recurso de equilíbrio para ecossistemas aquáticos e terrestres, fundamental para o desenvolvimento econômico e social de países, graças à sua função de abastecer a população por diversos tipos de atividades com água de qualidade (Jamie, et al, 2022).

De acordo com Lall, Josset e Russo (2020), as águas subterrâneas refletem cerca de 40% da água usada para irrigação de culturas agrícolas e um 1/3 da água usada em indústrias, além de representar 99% de toda água líqüida doce no planeta. No Brasil, as águas subterrâneas exercem um papel importante no abastecimento de água potável dos centros urbanos, representando 43% do abastecimento total das cidades brasileiras (ANA, 2021; Cao, et al, 2022).

Para diagnosticar a qualidade de águas subterrâneas e certificar se a água está adequada para o consumo humano, alguns parâmetros são estabelecidos levando em consideração as características hidroquímicas e hidrogeológicas da região.

Segundo a portaria GM/MS nº 888 (2021), os parâmetros que determinam a qualidade para consumo humano são: turbidez, residual de desinfetante, cor, pH, fluoreto, gosto, odor, cianotoxinas, produtos para desinfecção, acrilamida, epícloridrina, cloreto de vinila, coliformes totais e *Escherichia coli*. A frequência para o monitoramento desses parâmetros varia a cada duas horas, semanalmente, mensalmente, trimestralmente e semestralmente, dependendo do parâmetro analisado (GM/MS 888, 2021).

O máximo permitido de cada contaminante para consumo humano, dessedentação animal, irrigação e recreação em águas subterrâneas é estabelecido pela Resolução N° 396/2008. A cada cinco anos, os órgãos públicos responsáveis pelo monitoramento da qualidade da água devem realizar uma caracterização mais completa, abrangendo mais parâmetros orgânicos, inorgânicos, agrotóxicos e microorganismos (CONSELHO NACIONAL DO MEIO AMBIENTE, 2009).

Até 2018, o município de Campo Grande contava com 146 poços tubulares profundos, usados para diversas finalidades. Esses poços captam a água de três sistemas aquíferos, sendo eles: Sistema Aquífero Bauru (SAB), Sistema Aquífero Guarani (SAG) e Sistema Aquífero Serra Geral (SASG). Dentre os parâmetros determinados pelo Instituto do Meio Ambiente do Mato Grosso do Sul - IMASUL, o que apresenta os maiores valores é o nitrato, demandando maior cuidado em relação aos outros parâmetros (MANZANO, 2018; CARDOSO, 2019).

O nitrato (NO_3) provém da combinação química entre nitrogênio e oxigênio, podendo estar presente no ar, no solo, nas águas superficiais e subterrâneas. Normalmente, é encontrado em ambientes rurais com uso intensivo de fertilizantes agrícolas e em cidades com um mal planejamento, estando relacionado a resíduos domésticos, dejetos humanos provenientes de esgoto e fossas sépticas, fezes de animais, oxidação de amônia e condições meteorológicas (BREUER, ET AL, 2008; BAIRD E CANN, 2011; HAMIDI, ET AL, 2023).

A alta concentração de nitrato nos recursos subterrâneos pode gerar problemas não apenas ao meio ambiente, mas também à saúde humana. Como o nitrato é um ânion poliatômico, sua presença no corpo humano pode acabar afetando a função da tireoide. Além disso, seus efeitos são relacionados à metemoglobinemia infantil e alguns tipos de câncer, mas não há estudos suficientes que comprovem essa relação (SHIUE, 2015; WARD, ET AL, 2018).

Graças ao avanço da tecnologia, estudar a distribuição não apenas do nitrato, mas de diversos poluentes nas águas subterrâneas, tornou-se mais prático e eficiente. A inteligência artificial (IA) e o seu subcampo, aprendizado de máquina, são regularmente usados para estudos sobre recursos hídricos; assim, abordagens convencionais como ferramenta de interpolação e modelos físicos de transporte de contaminantes baseados em equações são deixados de lado por não apresentarem resultados tão eficazes como os modelos sofisticados oriundos da inteligência artificial (Rodriguez-galiano, et al, 2014; Sajedi-hosseini, et al, 2018; Sarkar, et al, 2022).

Objetivo geral

O objetivo principal é compreender como diferentes fatores (humanos e naturais) podem afetar a presença de nitrato nas águas subterrâneas, com a aplicação de técnicas de aprendizado de máquina, para o município de Campo Grande, MS.

Objetivos específicos

Dentre os principais objetivos específicos, podem ser destacados:

- Analisar dos fatores antrópicos, que são relacionados às atividades humanas, como o uso do solo, a densidade populacional e os serviços de saneamento;
- Verificar os fatores naturais que influenciam a contaminação por nitrato, analisando aspectos físicos e ambientais da região;
- Identificar fatores naturais e antrópicos para entender sua relação com os níveis de nitrato;
- Mapear a distribuição das concentrações de nitrato, para identificar áreas críticas e padrões de contaminação

2. REVISÃO BIBLIOGRÁFICA

2.1 Aquíferos

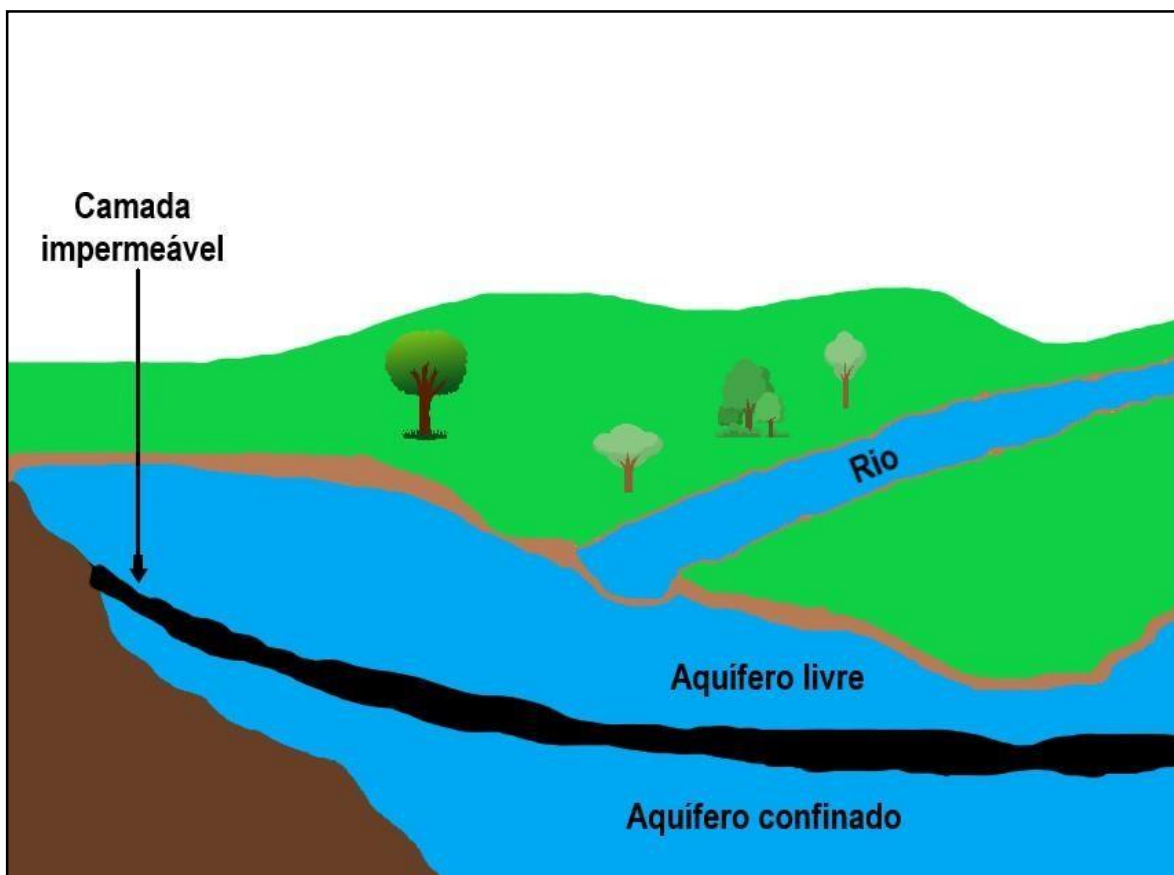
Na área da hidrogeologia, a palavra aquífero é um dos termos mais importantes, possuindo diversos significados dependendo do contexto de sua utilização. O aquífero é uma formação geológica permeável saturada, gerado entre os interstícios de rochas do tipo sedimentares, ígneas, metamórficas e em sedimentos inconsolidados, contendo uma grande quantidade de água que circulam em seu interior, apresentando grande potencial de transmissão de água em gradientes hidráulicos comuns (Freeze e Cherry, 1979; Feitosa e Manoel, 2008).

Essas são características que diferem os aquíferos dos aquícludes e aquíclados, os quais não apresentam um bom potencial para obtenção de água. Os aquícludes possuem uma estrutura saturada com uma quantidade de água considerável, porém é incapaz de transmitir essa água em condições normais. Os aquíclados são camadas menos permeáveis, desempenhando um papel de membrana que dificulta a produção de água por meio de poços (Freeze e Cherry, 1979).

2.1.2 Tipos de aquíferos

Segundo Feitosa e Manoel Filho (2008), a classificação dos aquíferos é feita a partir da pressão das águas em suas superfícies limítrofes e de sua capacidade em transmitir água por essas camadas limítrofes. Os aquíferos podem ser classificados em dois tipos, os livres e confinados (Figura 01).

Figura 01: Exemplo de aquíferos livres e confinados



Fonte: DINIZ (2023)

Os aquíferos livres, também chamados de freáticos, têm como limite superior a superfície de saturação, expostos à pressão atmosférica, possuindo uma estrutura inteiramente aflorante e uma recarga direta de suas reservas. Já os aquíferos confinados, também chamados de aquíferos sob pressão, são um tipo de aquífero em que a pressão exercida pela água em seu topo é superior à pressão atmosférica. (Feitosa e Manoel Filho, 2008; Manzione, 2015)

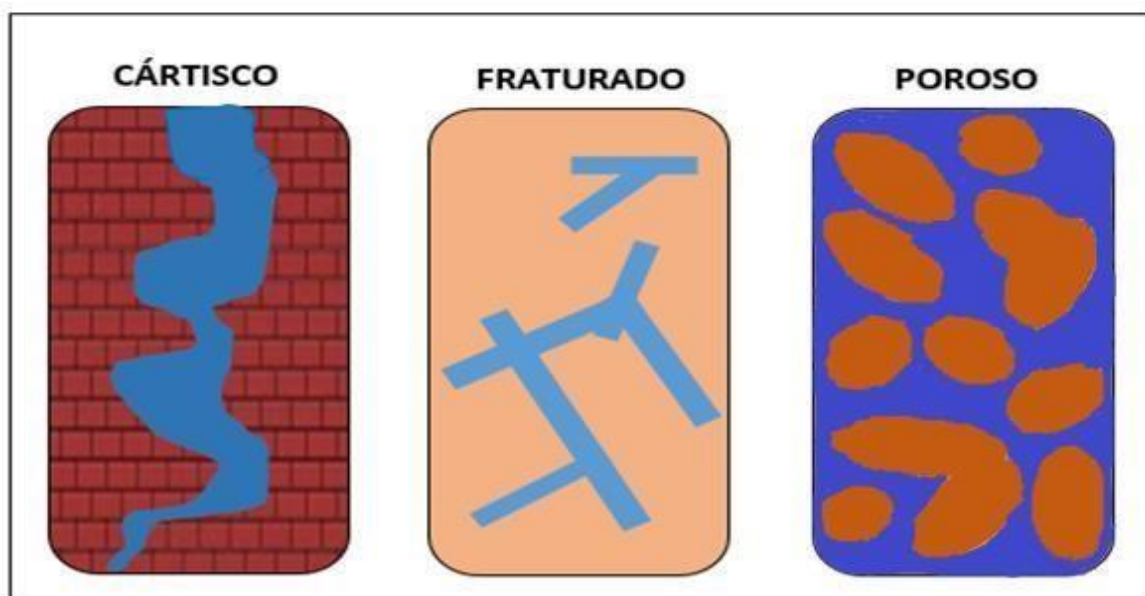
A recarga dos aquíferos confinados decorre de águas em excesso dos aquíferos livres ou de outros aquíferos confinados, que infiltram na camada devido à gravidade e em algumas áreas onde aflora à superfície, provendo da própria chuva. Os aquíferos livres e confinados podem ser divididos em drenantes e não drenantes, divididos de acordo com as características das suas camadas limítrofes (Feitosa e Manoel Filho, 2008; Manzione, 2015)

No aquífero livre, os drenantes são aqueles onde sua base é semipermeável e os não drenantes apresentam uma base totalmente impermeável. No aquífero confinado, os drenantes possuem uma estrutura semi-permeável, possibilitando o fluxo de águas no topo e em sua base, enquanto que nos aquíferos confinantes não drenantes sua estrutura detém uma estrutura totalmente impermeável (Feitosa e Manoel Filho, 2008).

Em alguns casos raros, há o surgimento dos aquíferos suspensos (Figura 02), um tipo de aquífero formado entre uma base inferior impermeável e uma base superior permeável, localizada entre o terreno e a superfície freática (Manziona, 2015).

Levando em consideração o tipo de porosidade das rochas, surge outra classificação para os aquíferos, divididos em cárstico, fraturado e poroso (Figura 02). O primeiro é constituído por rochas carbonáticas geradas a partir da precipitação química, no qual ocorre a dissolução de rochas, surgindo cavidades pela circulação da água; o fraturado é representado por rochas ígneas e metamórficas com fraturas e fendas, resultado de movimentos tectônicos. O transporte da água nesse tipo de aquífero é feito entre essas fraturas que vão diminuindo de acordo com a profundidade e; o poroso é formado por rochas sedimentares consolidadas e sedimentos inconsolidados e solos. A circulação da água acontece entre os poros e grãos existentes das rochas sedimentares (Rebouças 2013; Manziona, 2015).

Figura 02: Classificação de aquíferos a partir da estrutura das rochas



Fonte: DINIZ (2022)

2.2 Hidrogeologia de Campo Grande e Mato Grosso do Sul

A hidrogeologia do estado do Mato Grosso do Sul é dividida em dois tipos de rochas, as sedimentares que constituem os aquíferos porosos presentes na bacias sedimentares do Paraná e do Pantanal e as ígneas-metamórficas que constituem os aquíferos fraturados ou de fissuras, presentes em embasamentos cristalinos e em algumas regiões da bacia do Paraná (PERH-MS, 2010).

No território de Mato Grosso do Sul, há uma grande diversidade de aquíferos com um complexo sistema hidrogeológico, divididos entre as regiões hidrográficas do Paraná e do Paraguai. Esses são separados por oito unidades: Sistema Aquífero Cenozóico; Sistema Aquífero Bauru; Sistema Aquífero Serra Geral; Sistema Aquífero Guarani; Sistema Aquífero Aquidauana-Ponta Grossa; Sistema Aquífero Furnas; Sistema Aquífero Pré-cambriano Calcários e; Sistema Aquífero Pré-cambriano (PERH-MS, 2010).

Considerando o percentual por área de afloramento, os aquíferos com as maiores extensões são o Bauru (SAB) e o Cenozóico (SAC). O aquífero Bauru ocupa 37% do território de Mato Grosso do Sul, com boa parte da sua área localizada na Bacia Hidrográfica do Paraná, incluindo nesta área, uma pequena parcela do município de Campo Grande. Sua composição conta com rochas

sedimentares de dois grupos distintos, do Grupo Caiuá (Formação Santo Anastácio), da idade cretácia superior e do Grupo Bauru (Formações Vale do Rio do Peixe e Marília), uma formação mais jovem da idade neocretácia que está acima do Grupo Caiuá (Mato Grosso do Sul, 2008; PERH-MS, 2010; Uechi, Gabas e Lastoria, 2017).

O Aquífero Cenozóico, também chamado de Pantanal, está totalmente inserido na Bacia Hidrográfica do Paraguai, ocupando 27% da área do estado. Em sua composição estão presentes sedimentos arenosos finos de baixa compactação com uma matriz siltico-argilosa pertencentes a Bacia do Pantanal. Sua espessura média varia entre 20 a 30 metros, ultrapassando os 70 metros em algumas regiões (Mato Grosso do Sul, 2008; PERH-MS, 2010; Souza, 2013).

Outro aquífero que abrange um grande percentual de área é o Aquífero Guarani (SAG), considerado o maior reservatório de água subterrânea da América do Sul, abrangendo Argentina, Brasil, Paraguai e Uruguai. No Brasil, está situado em seis estados, sendo um deles Mato Grosso do Sul, ocupando uma área total de 213.200 km², representando 25% da área brasileira, dos quais, apenas cerca de 36.000 km² correspondem às áreas de afloramento (Gatsmans e Kiang, 2004).

No país, sua formação rochosa arenosa faz parte do Grupo Rosário do Sul e Pirambóia e da Formação botucatu, todas pertencentes a Bacia do Paraná. A espessura rochosa do Aquífero Guarani pode chegar aos 800 metros e no estado de Mato Grosso do Sul, essa espessura tem uma média de 600 metros, desempenhando um papel importante no abastecimento de cidades, incluindo Campo Grande (PERH-MS, 2010).

No estado Sul-Mato-Grossense o aquífero Serra Geral também exerce uma função importante no abastecimento, especialmente pela sua localidade, englobando em sua região de afloramento o território de Dourados e uma área significativa de Campo Grande. Além de ser uma fonte de abastecimento de água para sua população, o uso da sua água está atrelado a interesses econômicos, principalmente na agricultura (Lastoria, et al, 2006).

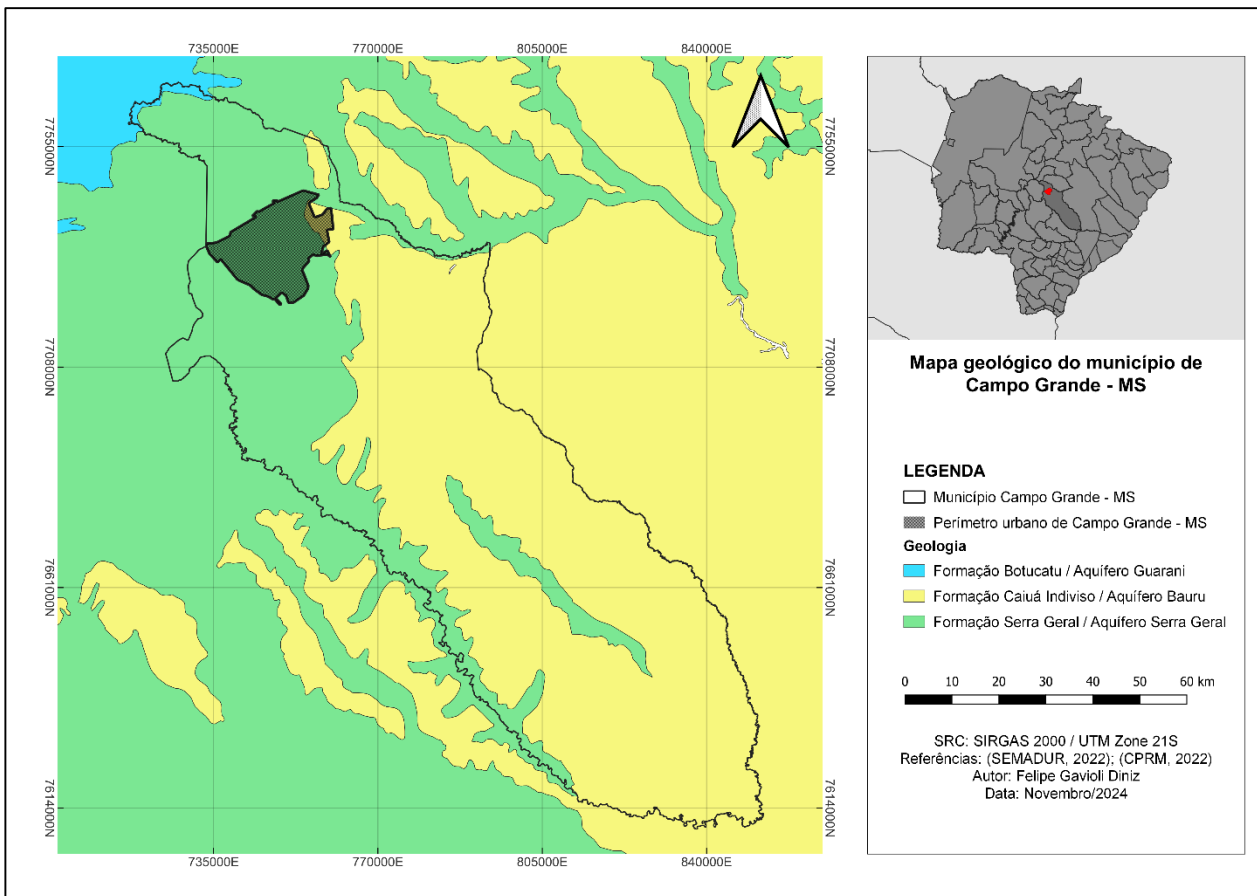
A estrutura do aquífero Serra Geral apresenta uma porosidade fissural e anisotrópica, onde o fluxo de água se concentra principalmente em superfícies de descontinuidade horizontal. Essas superfícies incluem as zonas de topo e base dos

derrames, bem como os contatos entre os derrames. Além disso, o fluxo também ocorre em superfícies verticais, nos planos de disjunção colunar que podem ser observados no interior dos derrames (Lastoria, et al, 2007).

A exploração da água no Serra Geral ocorre em áreas que a espessura dos sedimentos que pertencem ao Grupo Bauru é pequena, variando entre 150 a 200 metros, e em algumas áreas alcançando os 380 metros. A infiltração deste aquífero chega aos 8%, com 20% de sua reserva explotável considerada reserva renovavel (Mato Grosso do Sul, 2008; PERH-MS, 2010).

Ou seja, a hidrogeologia do municipio de Campo Grande é formada por três aquíferos: Aquífero Guarani, Aquífero Serra Geral e o Aquífero Bauru. O primeiro está presente em uma pequena parcela ao norte do municipio, o segundo está localizado mais ao oeste, englobando majoritariamente a área do perímetro urbano e por último o Aquífero Bauru, que diferente do Serra Geral, está localizado a leste e também no perímetro urbano de Campo Grande, mas em uma menor proporção, conforme Figura 03

Figura 03: Mapa geológico e hidrogeológico do município de Campo Grande, MS



Fonte: DINIZ (2022)

2.3 Tipos de captações de águas subterrâneas

2.3.1 Poços

A captação das águas subterrâneas contribui para regiões que sofrem com a falta de águas superficiais e para complementar o volume de água do sistema público de água em algumas cidades. Essa água pode ser usada em atividades industriais, agrícolas e diversos usos domésticos. A captação ocorre por meio do uso de poços, classificados em: poços escavados manualmente, poços com coletores com drenos horizontais simples e drenos radiais, galerias flutuantes e poços tubulares (Feitosa e Manoel Filho, 2008; São Paulo, 2008).

Profundidade dos aquíferos, equipamentos disponíveis para perfuração e condições de terrenos como o tipo de solo e a topografia do local são fatores que

influenciam na escolha de como a água subterrânea será captada, optando geralmente pelo uso dos poços tubulares (FEITOSA E MANOEL FILHO, 2008).

2.3.2 Poços tubulares

Os poços tubulares podem ser definidos como uma obra de engenharia que busca extrair a água subterrânea de aquíferos rasos e profundos através de uma perfuração vertical. Esse tipo de poço pode ser dividido em jorrante e não jorrante, determinado pelo nível potenciométrico em relação a superfície topográfica. A perfuração de um poço tubular deve seguir as regras da norma vigente da ABNT NBR de 2017 (Associação Brasileira de Normas Técnicas, 2017).

Os poços jorrantes estão presentes apenas em aquíferos confinados, visto que, nesse tipo de aquífero a água detém uma pressão maior que a da atmosfera, e quando da perfuração de um poço, a água ascende acima do nível da superfície terrestre, e estabilizando em uma posição, fenômeno conhecido como artesianismo. Já o poço tubular não jorrante ocorre quando a água não ultrapassa o nível da superfície terrestre, necessitando de uma bomba para bombear a água até a superfície, podendo ocorrer em aquíferos livres e confinados (Feitosa e Manoel Filho, 2008).

A perfuração de um poço tubular em Mato Grosso do Sul precisa seguir um procedimento estabelecido pelo IMASUL. Primeiramente, deve-se solicitar o processo de autorização de perfuração do poço juntamente ao IMASUL no sistema de Registros e Informações Estratégicas do Meio Ambiente (SIRIEMA), on-line. Em um segundo momento, é necessário seguir as exigências estabelecidas na Resolução SEMADE nº 21 de 2015, que exige o projeto do poço tubular já definido, seguindo as normas ABNT – NBR 12212 de 2017: Projeto de poço para captação de água subterrânea – Procedimento (Cardoso, 2019).

Após o envio dos documentos de disponibilidade hídrica e do empreendimento, eles são analisados, e em caso de deferimento, é expedido o documento que autoriza a perfuração do poço tubular, que desta vez, acompanha as regras estabelecidas na ABNT - NBR 12244 de 2006: Poço tubular – Construção de poço para captação de água subterrânea (Cardoso, 2019).

2.4 Fontes poluidoras das águas subterrâneas em áreas urbanas

Devido a sua situação de confinamento, as águas subterrâneas apresentam uma boa qualidade, representando uma das principais fontes de água potável no mundo, o que a torna atrativa. Entretanto, esses mananciais subterrâneos em áreas urbanas em países de todos os níveis de desenvolvimento são atingidos por poluentes gerados por atividades antrópicas, ocasionando diversos tipos de doenças em humanos (Schmoll, et al, 2006).

A contaminação das águas subterrâneas põe em risco não apenas a saúde da população local, como também dificulta o desenvolvimento socioeconômico com uma visão mais sustentável de uma cidade ou região. A maioria desses poluentes são gerados nas redes de esgoto no subsolo, principalmente aquelas com falta de manutenção, construção de aterros sanitários em terrenos inapropriados e do acúmulo de lixo domésticos e industriais sem uma coleta adequada, problemas comuns de cidades que apresentam uma rápida urbanização e sem um planejamento adequado (Wang, et al, 2023).

As redes de esgoto são planejadas para receber e transportar efluentes domésticos, hospitalares e industriais até as estações de tratamento de água. Porém, durante esse trajeto, os detritos e contaminantes carregados geram o processo de sedimentação, diminuindo a capacidade dos dutos. Esse efeito adverso pode provocar a decomposição anaeróbia dos materiais orgânicos, que influencia diretamente na corrosão da infraestrutura da rede coletora, provocando o aumento do odor (Huang, et al, 2021).

Outro fator de poluição ligado diretamente a contaminação dos aquíferos são os locais de disposição de resíduos sólidos urbanos (RSU) que, com o aumento gradativo das cidades e da população, tornaram-se comuns no espaço urbano. Porém, normalmente, esta disposição não é planejada adequadamente, não contando com um projeto de engenharia adequado e um controle ambiental, produzindo quantidades significativas de chorume e gases poluentes (Singh, et al, 2008; Morita, et al, 2021).

A intensidade da poluição das áreas de disposição de RSU decorre de alguns fatores como: quantidade de chorume produzido; pluviosidade da região, material

geológico do subsolo e gradiente hidráulico, ocorrendo principalmente em áreas fraturadas ou em regiões onde o lençol freático apresenta baixa profundidade (Singh, et al, 2008).

Ainda sobre os aterros sanitários, Casado et al (2015) e Al-Yaqout e Hamoda (2003) afirmam que resíduos domésticos, industriais e hospitalares são constantemente dispostos inadequadamente, liberando grandes quantidades de Magnésio (Mg), Ferro (Fe); Cálcio (Ca) e Zinco (Zn), além de altas taxas de pH, presente em chorumes de aterros antigos.

Como a rede de esgoto e o aterro sanitário, as fossas sépticas e negras são outras fontes de poluição em cidades, principalmente naquelas onde o processo de urbanização ocorre de forma tardia e lenta. São usadas para armazenar esgotos de residências sem acesso à rede coletora de esgoto, sendo considerada uma solução paliativa e extremamente prejudicial ao meio ambiente (Katz, Eberts e Kauffman, 2011; Gil Marques, et al, 2019).

Estudos comprovam que essas fossas atuam como uma das principais fontes de poluição de aquíferos rasos no espaço urbano, com a liberação de metais, nitrato e a redução do oxigênio dissolvido nas águas subterrâneas, infiltrando-se através do solo (Katz, Eberts e Kauffman, 2011; Gil Marques, et al, 2019).

Segundo Shivendra e Ramaraju (2015); Sowah, et al, (2017) e; Yang, et al (2017), a poluição por meio das fossas sépticas e negras ocorrem com mais intensidade em regiões suburbanas e com alta taxa de densidade demográfica, apresentando vazamentos por possuir uma estrutura inadequada; contribuindo, desta forma, para a dispersão dos contaminantes, configurando um desafio para os gestores da área do meio ambiente.

Em ambientes urbanos, outras fontes de poluição ganham destaque pela capacidade de contaminação dos recursos hídricos subterrâneos; são elas:

construção em não conformidade com normas técnicas e abandono de poços de captação de água; cemitérios e postos de gasolina em não conformidade com normas técnicas e a ausência de fiscalização; tanques enterrados contendo contaminantes, fezes de animais e produtos que apresentam riscos de corrosão (Santos e Câmara, 2002; Rao, Dinakar e Sun, 2022). A propagação dos poluentes nos corpos d'água subterrâneos é influenciada por diversos fatores naturais, como: profundidade do aquífero, composição da zona não saturada, declividade e (Leal, et al, 2012).

2.5 O uso do aprendizado de máquina para estudos sobre águas subterrâneas

Recentemente, graças ao avanço da tecnologia e da informática, o aprendizado de máquina passou por um período de intenso desenvolvimento, dispondo de novas técnicas e ferramentas. O avanço tecnológico contribuiu para que esse método de análise fosse adaptado e usado em diversos campos de pesquisa, ganhando destaque pela sua eficiência na previsão de dados e pela facilidade em ler dados complexos, atraindo o interesse de profissionais, estudantes e do público geral (Rodriguez-Galiano, et al, 2014; Knoll, Breuer e Bach, 2019; Fradkov, 2020)

A aplicação do aprendizado de máquina em estudos sobre recursos hídricos é algo frequente nos últimos anos, dada a complexidade em resolver problemas dessa área de estudo. Os modelos convencionais acabam sendo deixados de lado por serem caros, lentos e incompletos; em contrapartida, o aprendizado de máquina torna-se uma boa escolha buscando custo-eficiência (Huang, et al, 2021; Singha, et al, 2021)

Alguns modelos convencionais como os modelos conceituais e numéricos apresentam limitações técnicas, entre elas, a dificuldade de sustentar uma grande quantidade de parâmetros de entrada no processo de modelagem. Isso levou pesquisadores a investir na Inteligência Artificial (IA), uma ferramenta que apresenta enorme capacidade em lidar com grandes quantidades de dados (Agrawal, et al, 2021; Nordin, et al, 2021).

O aprendizado de máquina é uma ferramenta válida e eficiente para determinar a influência de fatores antrópicos e naturais sobre a poluição nas águas subterrâneas. Garante eficácia na hora de extrair dados preditivos e capacidade em prever a influência dos fatores sobre a qualidade da água em recursos naturais e conta com técnicas e recursos com alta dimensionalidade e relacionamentos não lineares. Tudo isso com a assistência de hardware e arquitetura computacional moderna (Agrawal, et al, 2021; Haggerty, et al, 2023).

Por ser de fácil interação e contar com um sistema de bibliotecas científicas já consolidada, faz com que sua escolha se torne mais atraente para o desenvolvimento de algoritmos e análise de dados. Neste cenário surge o Scikit-learn, uma biblioteca de aprendizado de máquina que fornece algoritmos eficientes com uma interface de fácil acesso e com uma programação imperativa (Pedregosa, et al, 2011).

O Scikit-learn oferece alguns recursos únicos e um conhecimento muito básico das aplicações de aprendizado de máquina (ajuste de modelo, previsão, validação cruzada, etc.). Além disso, essa biblioteca fornece suporte ao aprendizado supervisionado e não supervisionado e disponibiliza ferramentas para ajustar modelos, pré-processar dados, avaliar modelos e selecionar os modelos adequados (Scikit Learn, 2023).

Esses modelos podem ser classificados em classificadores e regressores. Os modelos classificadores e regressores apresentam diferenças na hora de prever uma resposta, embora, ambos utilizem dados de entrada já observados para o treinamento e para o teste. A classificação estima os novos dados (resultados) em forma de categoria, classe ou faixas, baseando-se nos dados de entrada. Já a regressão com o mesmo método, estima os novos dados com um valor numérico, sem considerar categorias. Portanto, variáveis categóricas terão um resultado melhor nos modelos classificadores e variáveis numéricas, um resultado superior nos modelos regressores (Lee e Shin, 2020).

Ainda sobre os modelos classificadores e regressores, Lantz, (2013) afirma que as técnicas de regressão são usadas para geração de rótulos de algumas observações com base em dados e informações já existentes. Sobre os modelos regressores, Soto (2013) afirma que a regressão é um método que busca identificar

relações entre variáveis independentes e dependentes por meio da estatística e a sua tarefa é buscar o ajuste de uma função para aproximar os dados de entrada com os dados de saída

Graças a suas técnicas avançadas, o aprendizado de máquina está cada vez mais aplicado na previsão de risco e perigos agravantes dentro dos estudos de ciências. Suas técnicas são aplicadas em campos da ciência agrícola, relacionados com o monitoramento dos solos e de culturas, e em estudos relacionados a hidrologia (Awais, et al, 2021).

Sobre a sua aplicação em estudos sobre a água pode vir a contribuir para solucionar problemas em: modelagem do nível freático, gestão de recursos hídricos, previsão de chuva, produção de sedimento, previsão de nível de água de lagos e reservatórios, previsão de enchentes e poluição de águas subterrâneas (Tahmasebi, et al, 2020; Ahmadi, et al, 2022).

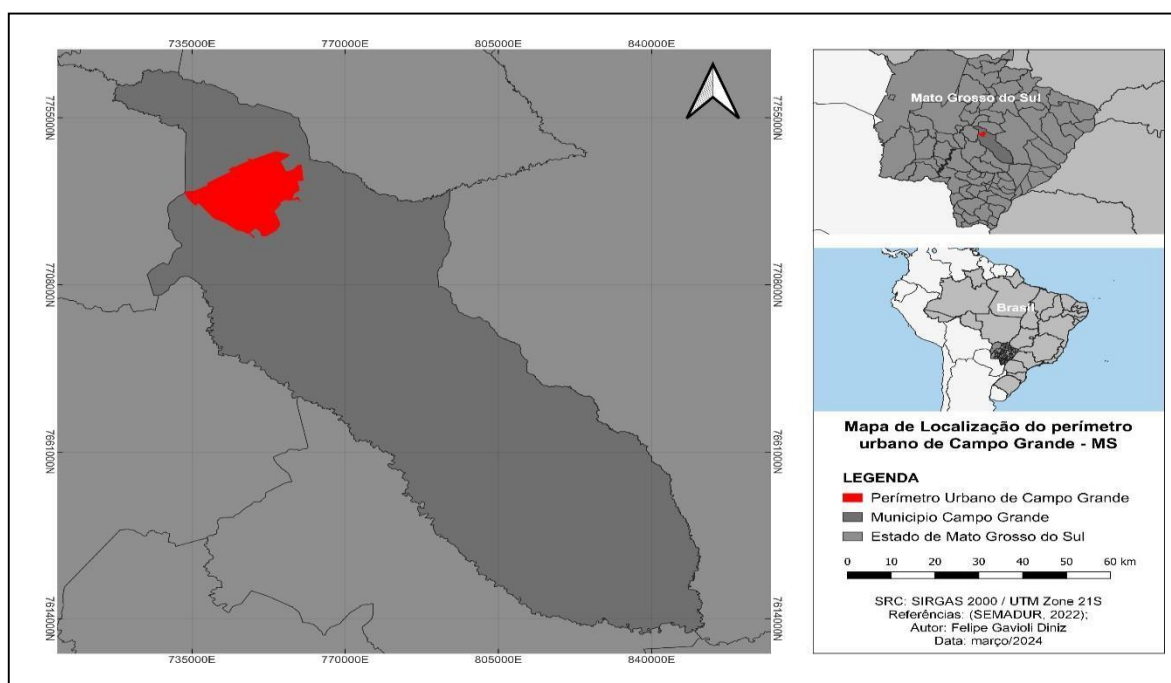
Sobre o aprendizado de máquina em estudos de poluição de águas subterrâneas, Raghavendra e Deka (2014) afirmam que a aplicação das técnicas do aprendizado de máquina pode servir para mapeamentos de contaminantes de águas subterrâneas, medição da toxicidade de poluentes em sistemas hídricos e na vulnerabilidade de aquíferos. Atualmente, dentro desta linha de estudos sobre poluição de águas subterrâneas, o aprendizado de máquina está comumente relacionado ao mapeamento de nitrato e a fonte de origem desse poluente (Ranson, et al, 2022; Awais, et al, 2021).

3. MATERIAL E MÉTODO

3.1 Área de estudo

O município de Campo Grande, capital do estado de Mato Grosso do Sul, local escolhido para a pesquisa, está situado na porção central do estado representando 2,26% da área total de Mato Grosso do Sul com 8.082,97km² (Figura 03). A sede do município de Campo Grande está localizada na sua porção norte, na divisa das águas das Bacias do Paraná e Paraguai, com uma extensão de 35.903,77ha, definida pelas coordenadas geográficas 20°28'13,4" latitude Sul e 54°37'25,9" longitude Oeste (Figura 04) (Campo Grande, 2022).

Figura 04: Mapa de localização do perímetro urbano de Campo Grande, MS



Fonte: DINIZ (2022)

O município de Campo Grande localiza-se sobre a bacia sedimentar do Paraná, constituída pelos litotipos do Grupo Caiuá Indiviso e pela Formação Botucatu e Formação Serra Geral pertencentes ao Grupo São Bento (Figura 04) (CPRM, 2006).

De acordo com o censo de 2022, a população de Campo Grande atingiu o número de 898.100 pessoas, não chegando a 916.001 pessoas, como estava estimado no censo de 2010. O crescimento absoluto foi de 111.303 pessoas entre 2010 a 2022 e a densidade demográfica para o mesmo ano foi de 111,11 hab/km². A taxa de urbanização ultrapassa os 98%, sendo o Aero Rancho o bairro mais populoso, localizado ao sul da cidade (Campo Grande, 2023; IBGE, 2024)

3.2 Procedimentos metodológicos

As variáveis espaciais definidas na pesquisa para avaliar a distribuição da poluição de nitrato nas águas subterrâneas foram: distância para rede de água e rede de esgoto, zona urbana, população por bairro, declividade, hipsometria, Índice de Vegetação de Diferença Normalizada - NDVI, geologia, profundidade do nível da água e distância para drenagem/hidrografia.

A escolha dessas variáveis teve como referência os seguintes autores: Tesoriero e Voss, (1997); Masetti, et al, (2008); Devic, Djordjevic e Sakan, (2014); Khatri e Tyagi, (2015); Krishna, et al, (2015) e; Knoll, Breuer e Bach, (2019). Todos os autores citados tratam essas variáveis como causadores ou possíveis causadores de poluição das águas subterrâneas, e para essa pesquisa, essas variáveis foram adaptadas, considerando as características da área de estudo e os equipamentos disponíveis.

A base bibliográfica deste trabalho foi obtida por meio de duas fontes: Portal de Periódicos da CAPES, com acesso gratuito pela Comunidade Acadêmica Federada - CAFE e Google Acadêmico. Para os 57 trabalhos internacionais usados no trabalho, foi usada a lista de base e coleções do Periódico da CAPES e para os trabalhos nacionais, a busca foi feita pelo Google Acadêmico, ferramenta virtual de pesquisa do Google para pesquisa da literatura acadêmica.

Alguns fatores foram adquiridos em órgãos públicos como na Agência Municipal de Meio Ambiente e Planejamento Urbano - PLANURB, Sistema Municipal de Indicadores de Campo Grande - SISGRAN, e em *sites* como United

States Geological Survey - USGS e Sentinel Hub EO Browser, além do processamento realizado no *software* QGIS versão 3.22.4 (Tabela 01).

Tabela 01: Características dos fatores de entrada

FATORES DE ENTRADA	FONTE	ESCALA OU RESOLUÇÃO ESPACIAL	
Declividade	USGS	Resolução espacial	30m
Espessura do solo	PLANURB	Escala	1:10.000
Fossas	Qgis	Escala	1:10.000
Geologia	PLANURB	Escala	1:10.000
Hidrografia	PLANURB	Escala	1:10.000
Hipsometria	USGS	Resolução espacial	30m
NVDI	Sentinel Hub EO Browser	Resolução espacial	10m
Rede de água	PLANURB	Escala	1:10.000
Rede de esgoto	PLANURB	Escala	1:10.000
População por bairro	SISGRAN	Escala	1:10.000
Profundidade/NA	PLANURB	Escala	1:10.000
Zona urbana	SISGRAN	Escala	1:10.000

Fonte: Elaborada pelo autor (2024)

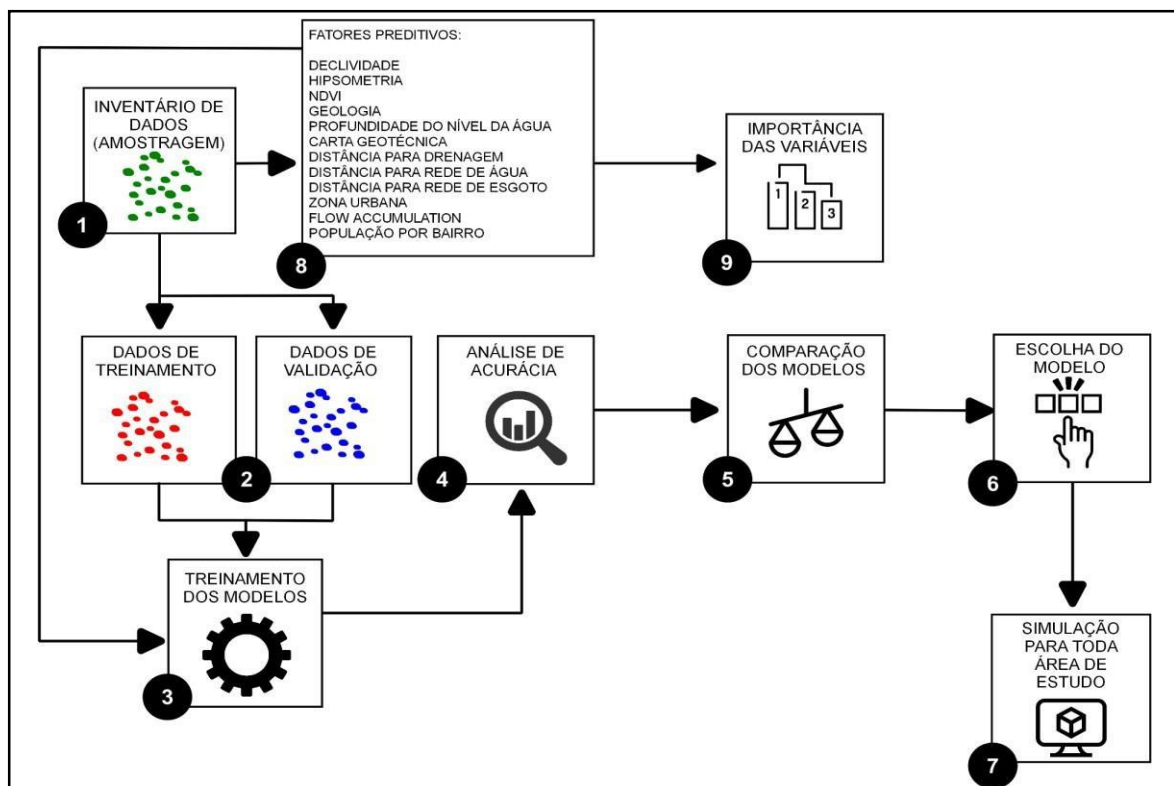
Os dados de nitrato foram obtidos por meio dos poços outorgados até 2018 da cidade de Campo Grande, MS, disponibilizados em Cardoso (2019). Dentro do perímetro urbano, a existência de 102 poços tubulares destinados a diversos usos, no qual, parte significativa apresentou mais de uma análise de nitrato. As informações sobre cada poço, incluindo coordenadas geográficas, altitude, profundidade, localização, sistema aquífero aflorante e captado, finalidade de uso e nitrato (mg/L) foram colocadas em uma tabela (Apêndice A e B) para melhor entendimento do assunto.

Para o tratamento e a análise dos dados, foi utilizada linguagem de programação Python, disponibilizada de forma gratuita. A Figura 05 apresenta as etapas desenvolvidas na linguagem de programação. O pacote científico usado nesta pesquisa foi o Scikit-learn (Pedregosa et al., 2011), um módulo Python de código, aberto, que engloba algoritmos clássicos de aprendizado de máquina.

Outros pacotes científicos em linguagem Python também foram utilizados, como, por exemplo, numpy (Harris, et al, 2020), scipy (Virtanen, et al, 2020) e matplotlib (Hunter, 2007). Essas bibliotecas fornecem soluções mais simples e eficientes para resolver problemas de aprendizagem de máquina, processamento e

visualização de dados (Pedregosa, 2011).

Figura 05: Etapas metodológicas do processo de aprendizado de máquina para predição da distribuição da concentração de nitrato em aquífero livre



Fonte: Elaborado pelo autor (2023)

3.2.1 Coleta dos aspectos naturais

Como fatores naturais que podem influenciar a concentração de nitrato nas águas subterrâneas em aquíferos livres, foram considerados: Declividade, Hipsometria, Índice de Vegetação da Diferença Normalizada (NDVI), Geologia, Profundidade do nível d'água, Espessura do solo e Distância para drenagem/hidrografia, descritos a seguir.

3.2.1.1 Declividade e hipsometria

O mapa de declividade e de hipsometria foram produzidos utilizando modelo digital de elevação ou de superfície (MDE ou MDS), elaborados a partir de dados SRTM. A missão Shuttle Radar Topography Mission (SRTM) foi realizada em fevereiro de 2000, usando o método interferometria SAR para coleta dos dados e

fornecendo dados da superfície terrestre entre os paralelos 60-N e 56-S (Berthier, et al, 2006).

A imagem SRTM foi obtida em 2022 pelo *site* United States Geological Survey (USGS), disponível em <https://earthexplorer.usgs.gov/>. Com a imagem SRTM baixada, usando o *software* Qgis versão 3.22.4, a primeira etapa foi desenvolver o mapa hipsométrico, uma etapa mais simples, pois o arquivo SRTM se encontra no formato Modelo Digital de Elevação (MDE). A única dificuldade foi ajustar a imagem considerando os limites da área urbana de Campo Grande e definir a sua simbologia, considerando o tipo de renderização, a interpolação e os gradientes de cores.

O mapa de declividade foi produzido usando o mesmo método empregado para o mapa hipsométrico, mas desta vez incluindo uma nova etapa. Essa etapa incluiu transformar o arquivo SRTM em declividade; para isso, foi acessada, na aba de ferramentas, a opção “declividade” e, em seguida, com uma nova página aberta, a opção “Declividade expressa em porcentagem (ao invés de graus)”. Essa opção faz com que a imagem de declividade produzida possa ter sua distribuição de classes baseada em porcentagem.

3.2.1.2 Índice de Vegetação da Diferença Normalizada (NDVI)

O Índice de Vegetação da Diferença Normalizada (NDVI) é um dos índices mais usados em estudos de sensoriamento remoto, principalmente os que tratam sobre monitoramento e análise de vegetação e culturas no campo. Neste trabalho, foi usado para ter uma percepção da distribuição da vegetação no perímetro urbano de Campo Grande.

O NDVI foi produzido no *software* Qgis 3.22.4, utilizando o satélite Sentinel 2, através da diferença e da soma das reflectâncias do infravermelho próximo e vermelho na calculadora raster, um complemento existente no Qgis (equação 01) (Cordeiro et al., 2017; Wojtowicz, et al, 2021).

$$\text{NDVI: NIR - RED / NIR + RED}$$

equação 01

NIR: Reflectância da banda do infravermelho próximo

RED: Reflectância da banda do vermelho

O Sentinel 2 é um satélite do Projeto Copernicus lançado no ano de 2015, possuindo uma resolução de alta qualidade e contendo um satélite multiespectral com resoluções espaciais que variam entre 10 m, 20 m e 60 m. No total, há treze bandas espectrais ricas em informações (Tabela 02), sendo ideal para estudos e monitoramento de culturas e uso do solo (Zhou, et al, 2022). O NDVI foi gerado a partir de uma imagem sentinel 2 da data 19/11/22, obtida gratuitamente no *site* Sentinel Hub EO Browser, disponível em <https://apps.sentinel-hub.com/eo-browser/>.

Tabela 02: Propriedades do Satélite Sentinel-2 (EngeSat)

Resolução	Banda	Nome da banda	Comprimento da onda central	Combinações de bandas
10 m	B02	Blue (azul)	490	Cor verdadeira RGB 04/03/02 Falsa cor 1 e 2 RGB 08/04/03 e 04/08/03
	B03	Green (verde)	560	
	B04	Red (vermelho)	665	
	B08	NIR (infravermelho próximo)	842	
20 m	B05	Red Edge 01	705	SWIR 1 RGB 12/11/8A
	B06	Red Edge 02	740	
	B07	Red Edge 03	783	
	B08A	Red Edge 04	865	

	B11	SWIR 1	1610	
	B12	SWIR 2	2190	
60 m	B01	Aerossol	443	-
	B09	Water Vapor	940	
	B10	Cirrus	1375	

Autor: EngSat (2022). Adaptado pelo autor

3.2.1.3 Geologia

Os dados geológicos da cidade de Campo Grande foram obtidos através da PLANURB via e-mail em janeiro de 2023. Esses dados estavam em formato shapefile com escala 1:10:000, pronto para o uso no *software* Qgis.

3.2.1.4 Profundidade do nível d'água

A profundidade do nível d'água foi definida usando de base as informações recolhidas na carta geotécnica, obtidas no *site* da prefeitura de Campo Grande, a partir da qual foi feito o *download* das informações em formato *shapefile* e o *download* do documento em formato pdf.

3.2.1.5 Espessura dos solos

A isoespessura do solo (Figura 14) foi obtida por meio da PLANURB, através de um e-mail em novembro de 2023 solicitando a carta de isoespessuras das Coberturas Sedimentares da Carta Geotécnica de Campo Grande em formato vetorial (Planurb, 2023).

3.2.1.6 Distância para drenagem/hidrografia

A distância da hidrografia foi medida no *Software* Qgis 3.22.4, usando o complemento "Proximidade (distância raster)". Anteriormente, os dados de

hidrografia que estavam em formato vetorial foram transformados em formato raster no complemento “Converter vetor para raster”, de modo que fossem usados para calcular a distância entre as drenagens.

3.2.2 Coleta dos aspectos antrópicos

Como fatores antrópicos, normalmente associados à contaminação de nitrato em águas subterrâneas, foram considerados: Distância para rede de água e rede de esgoto, Distância para fossas, Zona urbana e População por bairro, descritos a seguir.

3.2.2.1 Distância para rede de água e rede de esgoto

A distância para rede de água e esgoto foi obtida usando o mesmo método para distância de hidrografia do item 3.2.1.5. Nesta etapa, como se tratava da medição de dois dados distintos, primeiro foi produzida a distância para rede de água e posteriormente para a rede de esgoto.

3.2.2.2 Distância para fossas

A distância entre as fossas dos lotes de Campo Grande foi gerada usando o mesmo método de distância do item 3.2.1.5 e 3.2.2.1. Para o posicionamento hipotético das fossas, foram criados pontos representando as fossas existentes em Campo Grande, por meio da ferramenta “centroides” no *software* Qgis. Essa ferramenta cria uma nova camada vetorial de pontos, que corresponde ao ponto central dos polígonos de uma camada de entrada. Neste caso, a camada de entrada foi o shapefile dos lotes de Campo Grande e a camada de saída, ou seja, os centroides, as fossas de cada loteamento.

3.2.2.3 Zona urbana (Zoneamento)

Os dados da Zona Urbana foram obtidos através do banco de dados da prefeitura de Campo Grande (PMCG) pelo Sistema Municipal de Indicadores de Campo Grande (SISGRAN). Esses dados compõem a mancha urbana de Campo Grande dividindo-a em zonas definidas por critérios específicos, mostrados no item 4.1.9, podendo ser usados para se ter um entendimento do uso e ocupação do solo.

O download dos dados da Zona Urbana foi feito no formato shapefile juntamente com outros dados relacionados ao plano diretor da cidade.

3.2.2.4 População por bairro

Os dados da população por bairro foram usados para medir a densidade populacional de Campo Grande, sendo obtidos no mesmo endereço que os dados da Zona Urbana, o SISGRAN, que considera os números da população do censo de 2010. Para a produção do mapa, foi necessário criar uma coluna na tabela de atributos e gerar intervalos de classes de acordo com a quantidade de pessoas por bairro. Ao todo, foram criados quatro intervalos de classes, como pode ser visto na Figura 18.

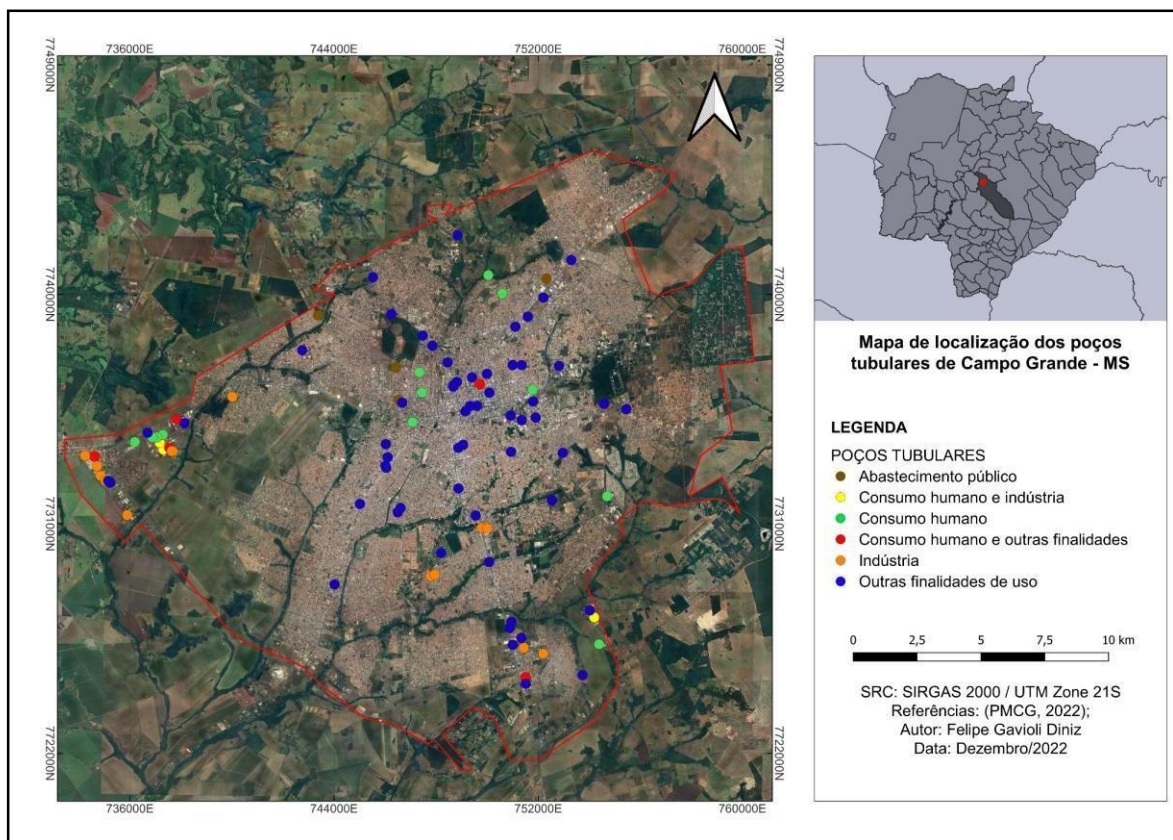
3.2.3 Poços tubulares da área de estudo - Campo Grande, MS

No município de Campo Grande, MS, foram outorgados 146 poços tubulares até outubro de 2018, espalhados por todo o município. Como o foco da pesquisa concentrou-se apenas no perímetro urbano, foram considerados para análise apenas os poços localizados na cidade de Campo Grande, descartando os poços da área rural.

Ao todo, dos 102 poços no perímetro urbano, foram considerados para a pesquisa 97 poços, descartando cinco poços, aqueles perfurados em dois tipos de aquíferos, em razão de que, esse tipo de poço pode alterar o resultado de nitrato, além do descarte dos poços sem dados de nitrato.

Dos 97 poços considerados para pesquisa, 95 localizam-se no Aquífero Serra Geral - SASG e 2 no Aquífero Bauru - SAB. Esses poços apresentaram diversas finalidades, alguns possuindo mais de um tipo de uso, como: indústria; abastecimento público; consumo humano; consumo humano e indústria; consumo humano e outras finalidades de uso (Figura 06). As informações de cada poço foram disponibilizadas nos apêndices A e B.

Figura 06: Distribuição espacial dos poços tubulares no perímetro urbano de Campo Grande, MS e tipo de uso.



Fonte: DINIZ (2022)

3.2.4 Processamento e análise dos dados

Com as variáveis espaciais estabelecidas, a primeira etapa do processamento dos dados foi utilizar o *software* QGIS versão 3.22.4 para rasterizar as variáveis (fatores preditivos) que estavam em formato vetorial. Isso ocorreu na ferramenta “Converter vetor para raster” e, em seguida, mesclar os dados das variáveis em formato raster com os dados dos 97 poços em formato vetorial. Para essa etapa, foi usada a ferramenta “Amostrar valores do raster” que desempenha a função de gerar uma nova camada vetorial contendo a camada de entrada do tipo vetorial (dados dos poços) e atributos obtidos das camadas com valores raster (dados das variáveis espaciais: fatores preditivos). Dessa forma, são importados os valores correspondentes para cada ponto gerando atributos desse vetor. Logo, para cada ponto (poço), o conjunto de dados contém um valor de concentração de nitrato e os valores amostrados das camadas de fatores preditivos.

O próximo passo foi salvar o arquivo no formato “Planilha de cálculo MS Office open XML [XLSX] e, em seguida, realizar a importação dos dados para o ambiente em linguagem de programação. Utilizando a linguagem Python, foi realizada a etapa “tratamento preliminar dos dados” que consistiu em excluir os poços em que os valores de nitrato superassem o valor do percentil 80%, em razão de serem valores muito distantes dos demais, prejudicando o desempenho da modelagem subsequente. Desta forma, foram excluídos 29 poços, resultando em 68 poços adequados para pesquisa e com os valores de concentração de nitrato não ultrapassando 5 mg/L.

Em seguida, foi realizada a etapa “análise de acurácia”, utilizada para determinar a qualidade de um parâmetro ou dado analisado. Neste trabalho, o desempenho dos modelos regressores e classificadores foram avaliados pelo valor do grupo treino, grupo teste e a totalidade. Ao todo foram usados 16 modelos estatísticos, sendo 9 classificadores e 7 regressores.

Os modelos testados no estudo preliminar sobre as variáveis espaciais que influenciam na presença de nitrato foram: Ensemble Random Forest Regressor; svm SVR; Tree Decision Tree Regression; Ensemble Gradient Boosting Regressor; svm SVC; Ensemble Random Forest Classifier; Ensemble Gradient Boosting Classifier; Ensemble Extra Trees Classifier; KNeighbors Classifier; Self Training Classifier; Linear model.Logistic Regression; Neural Network.MLP Regressor; Ensemble.Adaboost Regression; Neural Network MLP Classifier; Linear Model.SGD Classifier e; Linear Model.Ridge Classifier. No próximo item estão listados apenas os modelos classificadores que não apresentaram resultados negativos.

Ambos os modelos estão enquadrados no aprendizado supervisionado do machine learning, um tipo de aprendizado que possui um conjunto de entradas juntamente com o seu conjunto de saídas e busca estabelecer uma relação aproximada entre eles, criando uma hipótese (Russell e Norvig, 2002).

Os modelos classificadores e regressores apresentam diferenças na hora de prever uma resposta, embora, ambos utilizem dados de entrada já observados para o treinamento e para o teste. A classificação estima os novos dados (resultados) em forma de categoria, classe ou faixas, baseando-se nos dados de entrada, já a regressão com o mesmo método, estima os novos dados com um valor numérico,

sem considerar categorias. Portanto, variáveis categóricas terão um resultado melhor nos modelos classificadores e variáveis numéricas, um resultado superior nos modelos regressores (Lee e Shin, 2020).

Ainda sobre os modelos classificadores e regressores, Lantz, (2013) afirma que as técnicas de regressão são usadas para geração de rótulos de algumas observações com base em dados e informações já existentes. Sobre os modelos regressores, Soto (2013) afirma que a regressão é um método que busca identificar relações entre variáveis independentes e dependentes por meio da estatística e a sua tarefa é buscar o ajuste de uma função para aproximar os dados de entrada com os dados de saída. Outro procedimento realizado, foi a etapa de “análise de importância”, usada para determinar a influência de cada variável sobre o valor do nitrato dos poços tubulares. Foram criados dois rankings, um sobre a influência de cada variável usando os modelos regressores e outro usando os modelos classificadores

A última etapa realizada nesta pesquisa foi para produzir a estimativa da distribuição do nitrato nas águas subterrâneas de Campo Grande. O primeiro passo foi no ambiente Python, realizar o procedimento “simulação”, procedimento executado por 5 vezes, criando uma média da distribuição. Essa média foi salva em um documento formato “Planilha de cálculo MS Office open XML [XLSX]”.

Em seguida, esse documento foi aberto no *software* QGIS versão 3.22.4 e usado na ferramenta “criar camadas de pontos a partir de tabela”. Com os pontos vetorizados criados, contendo as informações da distribuição do nitrato, o último procedimento foi rasterizar estas informações na ferramenta “converter vetor para raster”, para então criar a distribuição do nitrato nas águas subterrâneas.

3.2.4.1 Neural Network.MLP Classifier

As redes neurais trabalham de forma interativa e, nos últimos anos, apresentou melhorias significativas no seu modelo graças a alguns avanços técnicos como: projetar estratégias e soluções eficazes para evitar sobreajuste. E, devido à sua capacidade atual de ajustar modelos de treinamento, sua aplicação em problemas que técnicas simples não conseguem mais realizar vem se tornando algo comum, como reconhecimento de fala e ajuste em problemas de classificação

(He, et al, 2015; Erdem e Bozkurt, 2021).

3.2.4.2 Linear Model.SGD Classifier

A Descida de Gradiente Estocástico (SGD) é um algoritmo considerado uma técnica de otimização, podendo ser usado como uma extensão dos classificadores lineares. Apresenta uma abordagem simples e didática, testado com eficiência em problemas de aprendizado de máquina esparsos e de larga escala, tipo de problema encontrado constantemente no processamento de linguagem e classificação de texto (Zhang, 2004 e Pedregosa, 2011).

3.2.4.3 Linear Model.Ridge Classifier

O Ridge Classifier é uma variante classificadora do Regressor Ridge, que primeiramente converte os alvos binários e, logo em seguida, trata o problema otimizando-o com uma tarefa de regressão. Ainda existem dúvidas sobre a técnica empregada no Ridge Classifier, mas o seu resultado garante uma precisão maior do que técnicas tradicionais disponíveis (Pedregosa, 2011).

3.2.4.4 Self Training Classifier

Esse modelo proporciona a chance de um classificador supervisionado trabalhar como classificador semi-supervisionado, usando dados não rotulados. Esse procedimento funciona pressupondo várias repetições pseudo-rótulos para dados que não possuem rótulos e, em seguida, os adiciona no conjunto de treinamento (Pedregosa, 2011).

3.2.4.5 K Neighbors Classifier

O algoritmo baseado em vizinhos funciona com a classificação sendo calculada com base na distância dos “vizinhos” mais próximos em relação ao ponto principal. Destacam-se dois tipos de técnicas baseadas em vizinhos, KNeighborsClassifier, normalmente a mais usada e a aplicada na pesquisa, além da RadiusNeighborsClassifier. Na KNeighborsClassifier, os valores são altamente dependentes dos dados e seus limites apresentam pouca distinção (Pedregosa,

2011).

3.2.4.6 : Random Forest Classifier

Floresta aleatória é um modelo no qual são construídas árvores através de amostras do método de substituição, também chamado de amostra Bootstrap. Nesse modelo, os valores considerados como “média” são usados para buscar a melhor precisão preditiva e controlar os excessos (Pedregosa, 2011).

3.2.4.7 : Extra Trees Classifier

Como o Rndom Florest, o algoritmo Extra-Trees baseia-se em árvores aleatórias, considerado um conjunto diversificado de classificadores e criado especialmente para gerar classificadores. Nesse algoritmo o resultado parcial é determinado pela previsão da média dos classificadores individuais (Pedregosa, 2011).

3.2.4.8 : Gradient Boosting Classifier

Esse algoritmo usado para classificação trabalha com a otimização de funções consideradas diferenciais e arbitrárias. Por causa da sua eficácia, pode ser aplicado em diversas áreas, como trabalhos envolvendo ecologia e pesquisas realizadas na internet (Pedregosa, 2011).

3.2.4.9 Support Vector Classification (SVC)

A interpretação do Vetor de Suporte de Classificação (SVC) é um algoritmo que busca o hiperplano sempre com a mesma distância entre duas classes. Lembrando que o SVC utiliza a função Kernel para criar superfícies não lineares (Bhavsar e Panchal, 2012; Pedregosa, 2011).

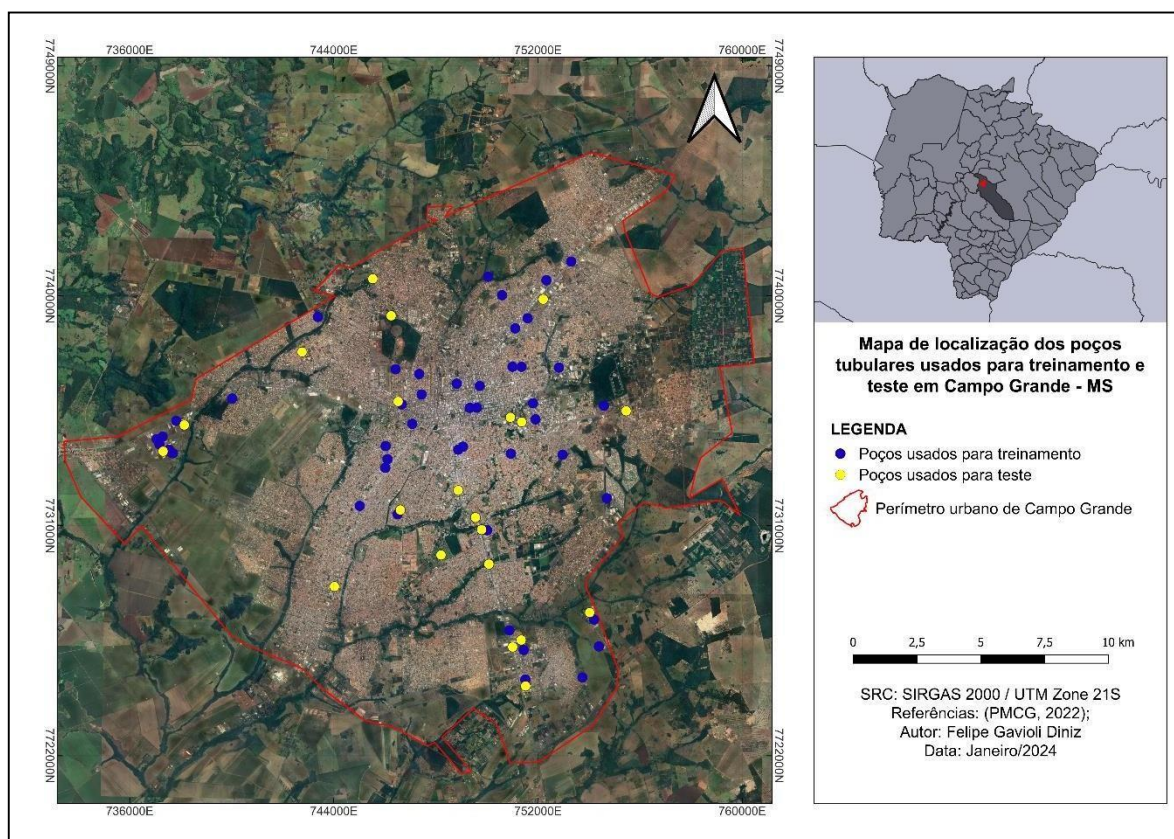
3.2.5 Treinamento e validação dos modelos do aprendizado de máquina

O conjunto de dados de concentração de nitrato nos poços de Campo Grande foram divididos em duas classes: a de treinamento (70% = 47) e a de teste (30% = 21) para avaliar o desempenho dos modelos estatísticos (Figura 07). Além de realizar a separação dos sub-datasets de treinamento e teste em 70/30, foram

executadas tentativas de 80/20. E na etapa de “tratamento preliminar dos dados”, houve a realização de testes com 95%, 90%, 85% e 80% percentis, no qual o último foi escolhido pelo desempenho mais satisfatório. Os modelos apresentaram pouca variação no resultado das classes treinamento e teste ao longo das cinco realizações.

Como já mostrado, esses modelos foram divididos em dois grupos, os classificadores e regressores, onde cada grupo apresentou um resultado da influência das variáveis espaciais sobre o valor de nitrato. O modelo com o melhor resultado foi utilizado para elaboração do mapa da estimativa da distribuição do nitrato nas águas subterrâneas de Campo Grande.

Figura 07: Mapa de localização dos poços tubulares usados para treinamento e teste em Campo Grande, MS



Fonte: Diniz (2024)

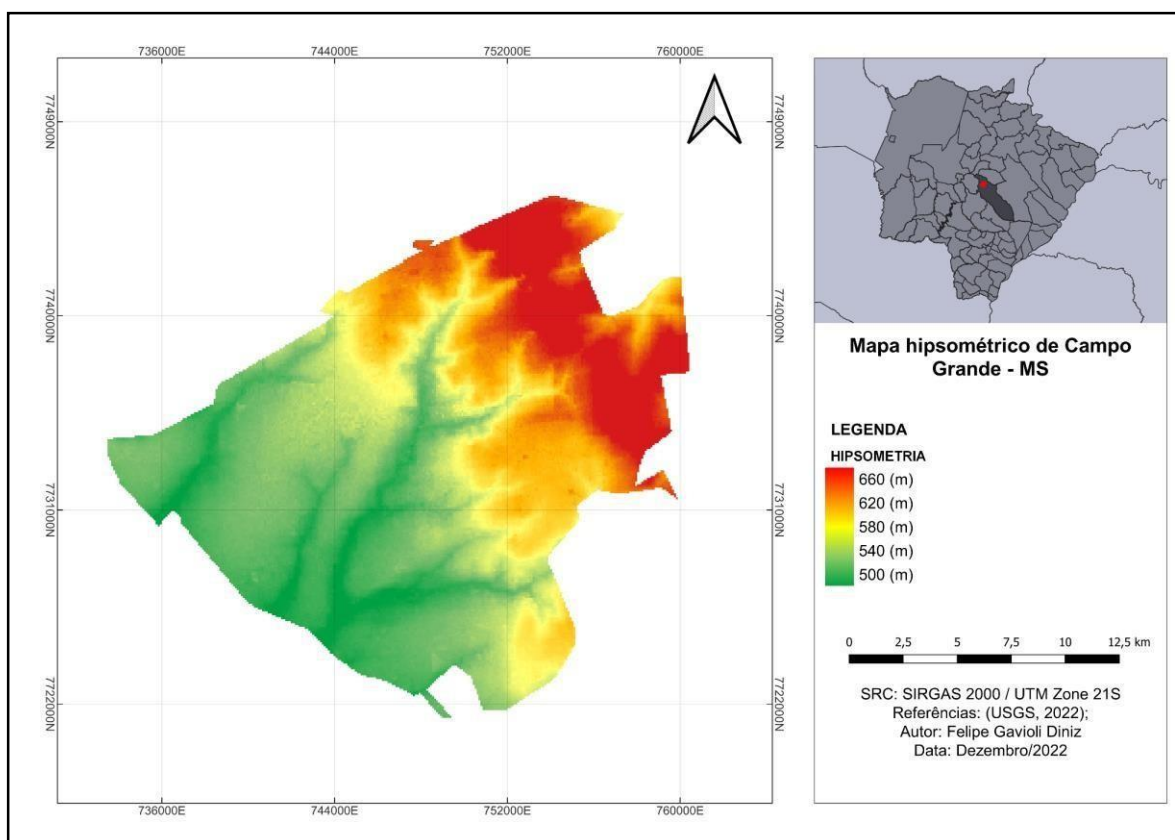
4. RESULTADOS E DISCUSSÃO

4.1 Variáveis espaciais

4.1.1 Hipsometria

A altitude de Campo Grande varia entre 500 a 680 metros (Figura 08), com uma média de 572 metros e as maiores altitudes concentrando-se na região norte do perímetro urbano. As menores altitudes ocorrem na região sul da cidade e em áreas de Depósitos Aluvionares nas margens de córregos.

Figura 08: Mapa de hipsometria de Campo Grande, MS



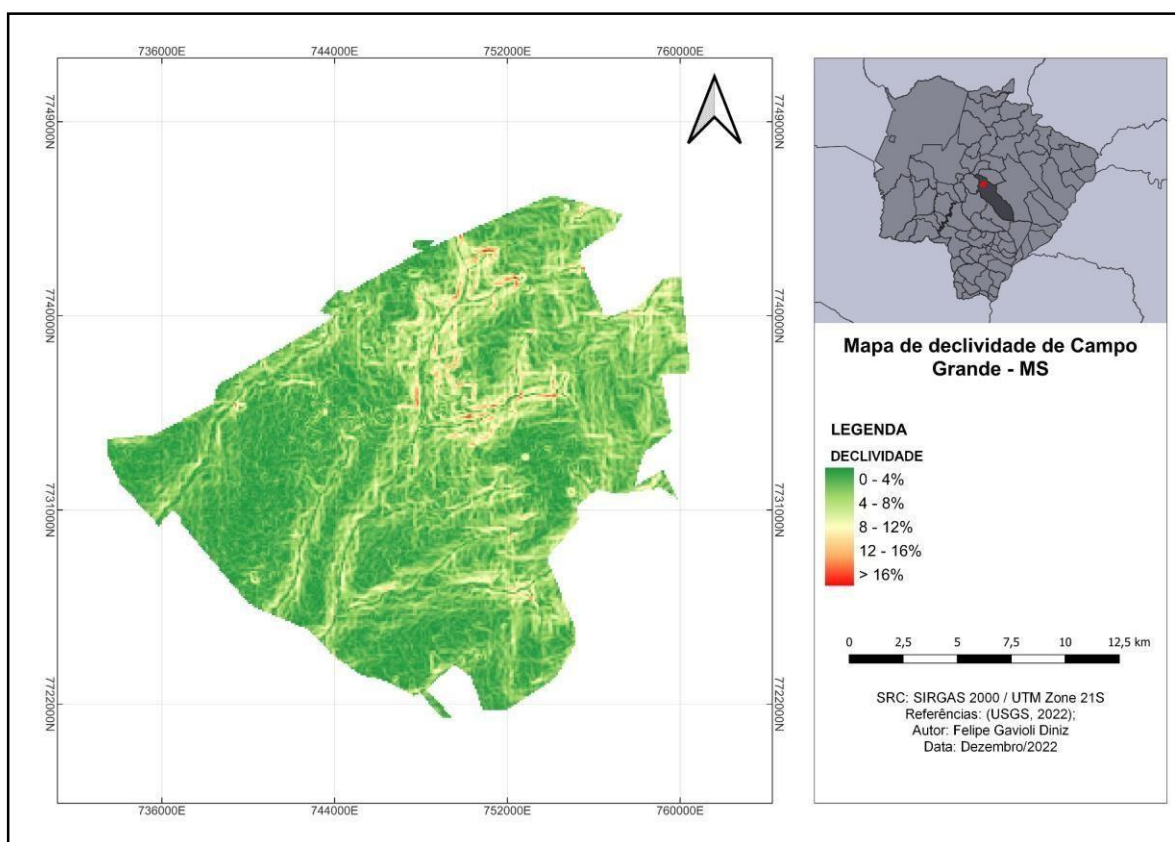
Fonte: DINIZ (2022)

4.1.2 Declividade

Nesta pesquisa, o mapa de declividade (Figura 09) teve como referência o sistema de classes de declividade da Embrapa (EMBRAPA, 1979). O mapa de

declividade de Campo Grande apresenta cinco classes: a classe com a maior área foi a classe 0 – 4%, que representa um relevo plano sem ondulações. As áreas com maiores declives, representada pela classe > 16%, encontram-se com menor ocorrência na cidade, concentrada na região norte e central, próxima aos corpos hídricos e onde há um aumento da altitude.

Figura 09: Mapa de declividade de Campo Grande, MS



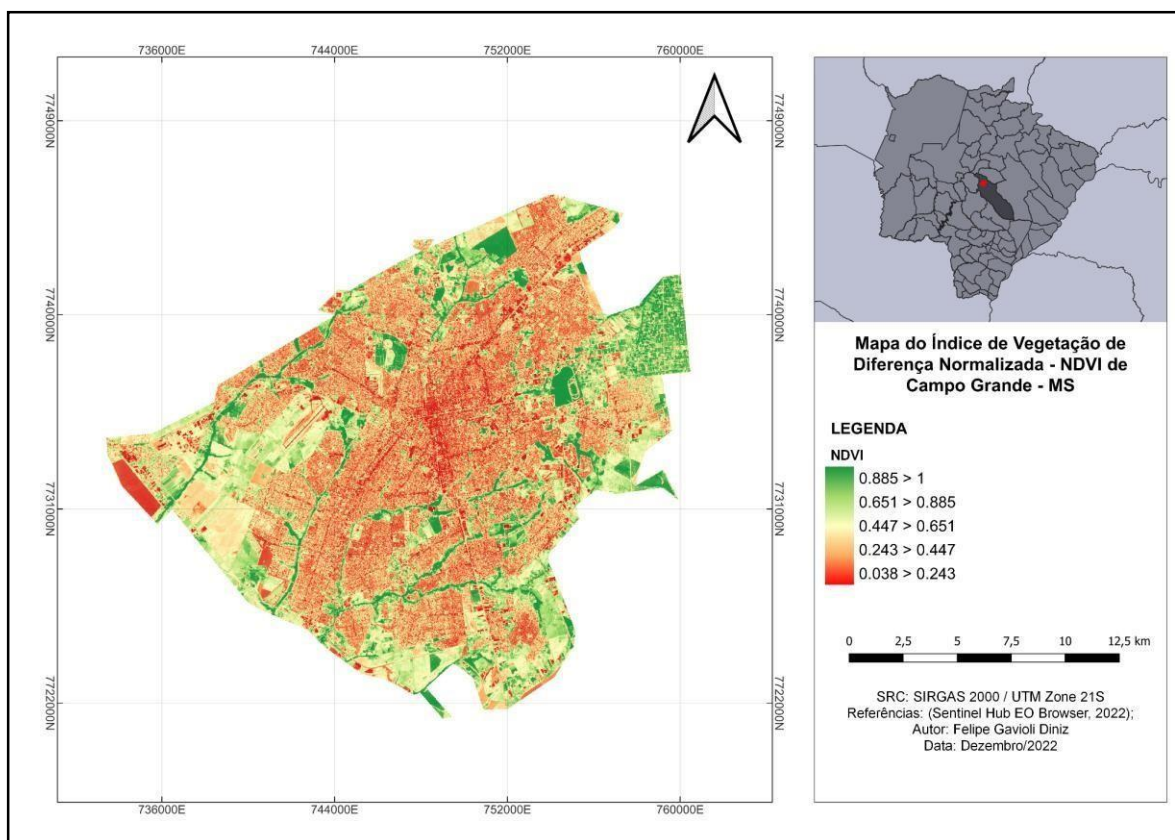
Fonte: DINIZ (2022)

4.1.3 NDVI

Os valores do NDVI (Figura 10) foram divididos automaticamente no Qgis em 5 intervalos de classes, com os valores variando entre +1 a 0. Os valores próximos a 1 representados pela cor verde indicam áreas com vegetações densas e saudáveis; em contrapartida, os valores próximos a 0 representados pela cor vermelha, indicam

áreas com vegetações prejudicadas, solo exposto e edificações, ou seja, áreas sem a produção de fitomassa e em alguns pontos com o solo impermeável. Nota-se que na região central de Campo Grande quase não há áreas com vegetação, isso pode ser justificado pela intensidade da urbanização na região e por concentrar um volume maior de edificações em comparação com outras regiões da cidade, visto que, as atividades do setor terciário estão condensadas na região central.

Figura 10: Índice de Vegetação da Diferença Normalizada (NDVI) de Campo Grande, MS



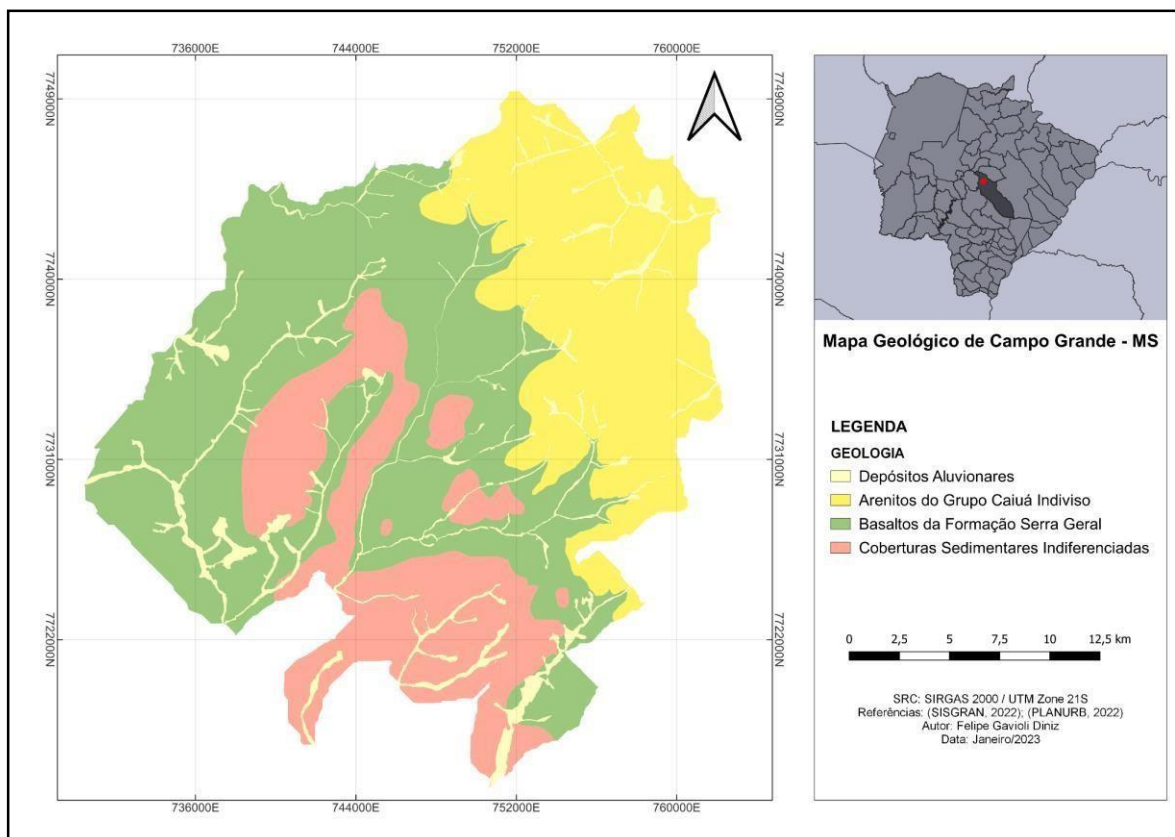
Fonte: DINIZ (2022)

4.1.4 Geologia

A geologia de Campo Grande (Figura 11) é composta por 4 classes: Arenitos do Grupo Caiuá Indiviso (27,6%); Basaltos da Formação Serra Geral (45,2%); Coberturas Sedimentares Indiferenciadas (21,4%) e Depósitos Aluvionares (5,8%).

Os dados geológicos de Campo Grande não abrangem apenas o perímetro urbano, como também a área de expansão urbana (Campo Grande, 2020).

Figura 11: Geologia da área urbana de Campo Grande, MS

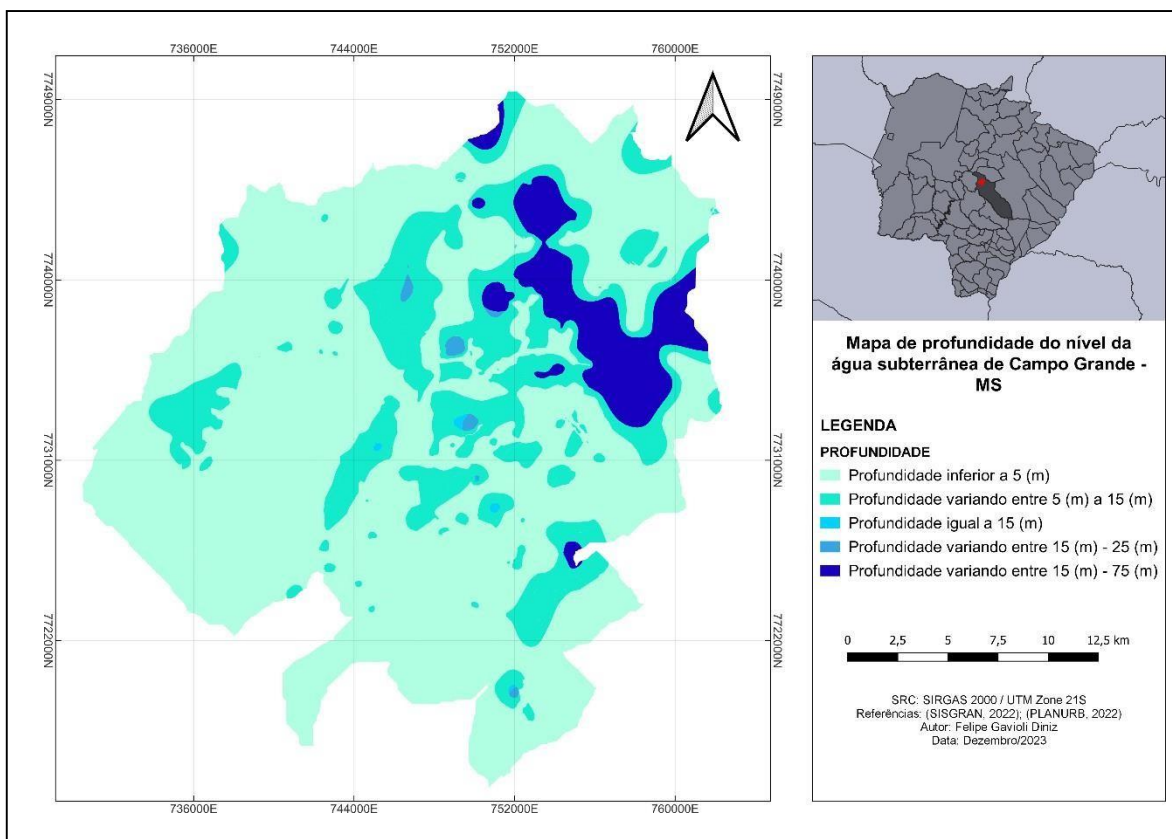


Fonte: DINIZ (2022)

4.1.5 Profundidade do nível de água

A profundidade do nível d'água (Figura 12) foi definida usando de base as informações recolhidas na carta geotécnica de Campo Grande elaborada em 2020 pela empresa Hidrosul Ambiental. Ao todo, o nível d'água foi dividida em 5 categorias: inferior a 5 metros (71,5%); variando entre 5 a 15 metros (21,6%); igual a 15 metros (0,1%); variando entre 15 a 25 metros (0,3%) e variando entre 15 a 75 metros (6,5%) (Campo Grande, 2020).

Figura 12: Profundidade do nível de água da cidade de Campo Grande

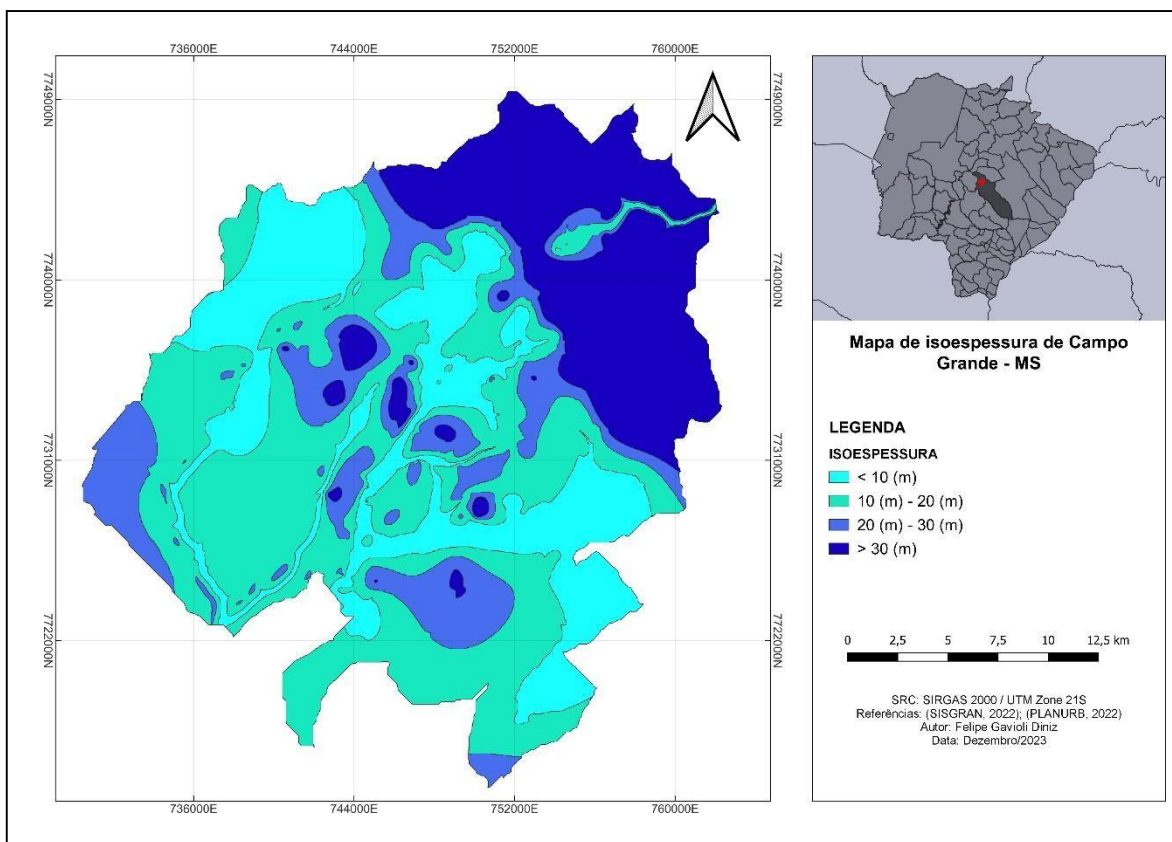


Fonte: DINIZ (2023)

4.1.6 Carta da isoesspesura do solo

O mapa de isoesspesuras (Figura 13) foi elaborado com base nos dados de profundidade das camadas geológicas de cada unidade litológica, disponibilizados pela PLANURB através de um email. A obtenção desses dados pela prefeitura foi feita por meio de ensaios de sondagem e de poços tubulares profundos em quatro intervalos: < 10 m (23,5%), de 10 a 20 m (39%), 20 a 30 m (15,2%) e > 30 m (22,3%).

Figura 13: Carta da isoespessura do solo de Campo Grande, MS

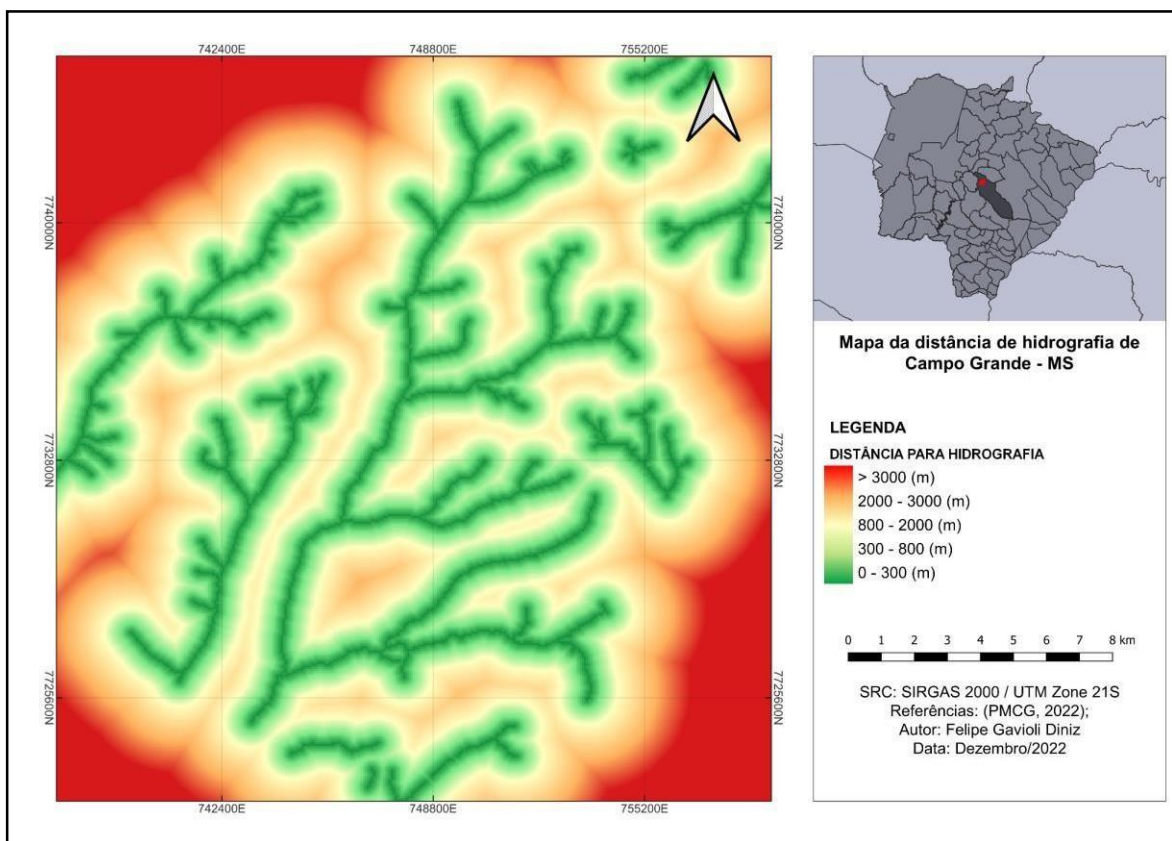


Fonte: DINIZ (2023)

4.1.7 Distância para drenagem/hidrografia

A distância da hidrografia de Campo Grande (Figura 14) varia com mais frequência entre 800 a 2000 metros, revelando uma drenagem espaçada e com bacias hidrográficas bem definidas ao longo da cidade.

Figura 14: Distância da rede de drenagem/Hidrografia de Campo Grande, MS

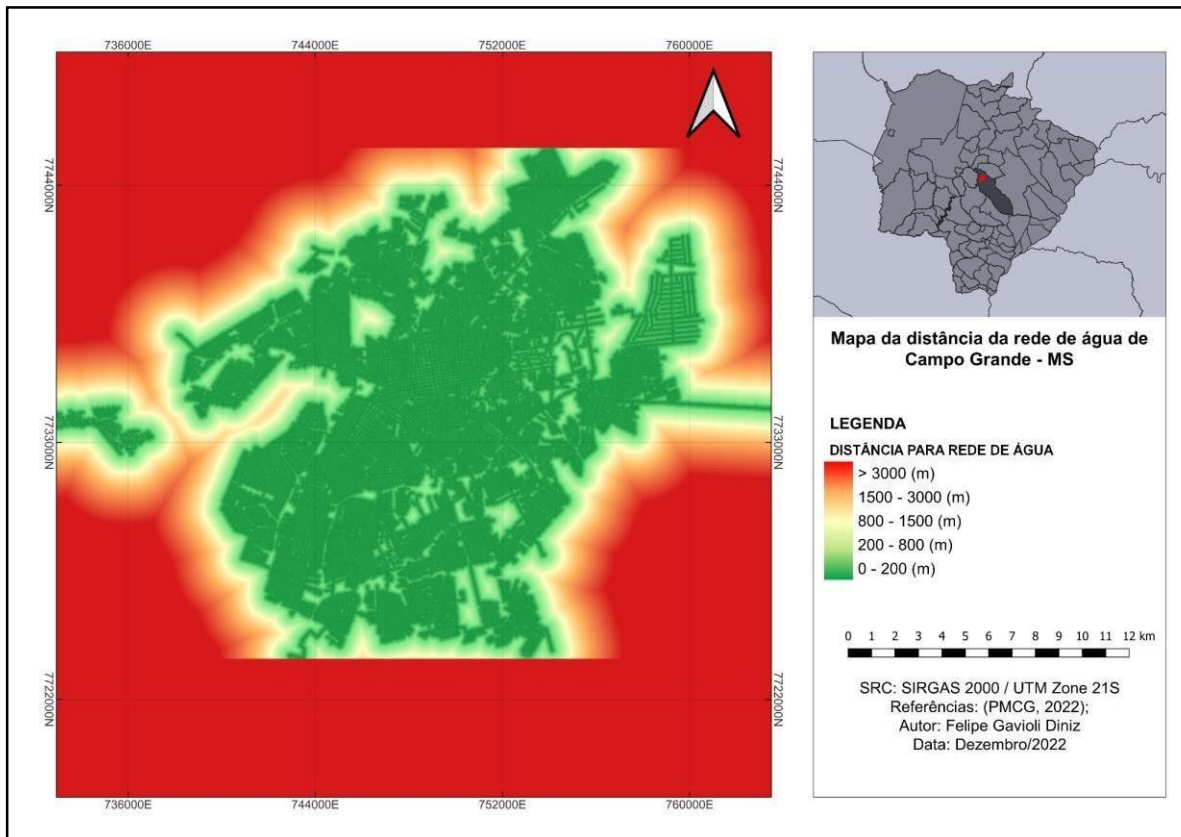


Fonte: DINIZ (2022)

4.1.8 Distância para rede de água e de esgoto

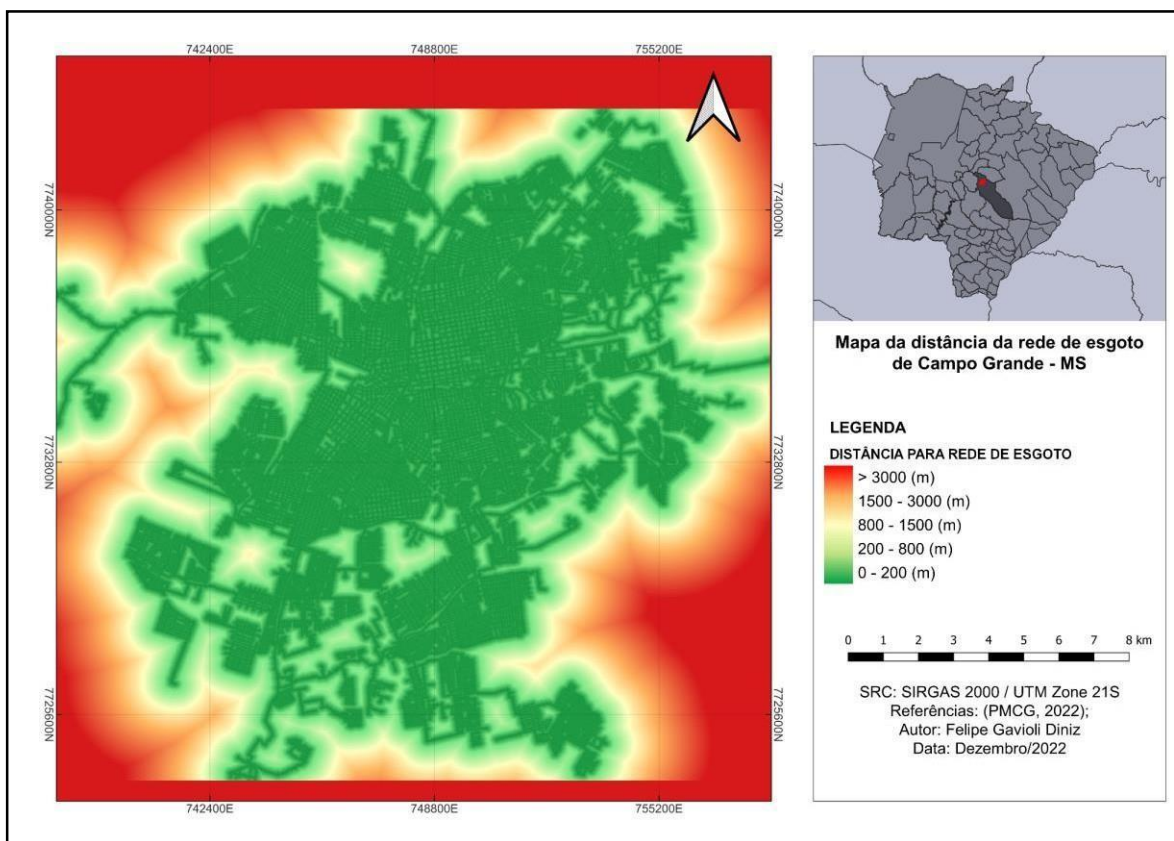
O resultado dessa análise revelou que a distância da drenagem de água (Figura 15) e esgoto (Figura 16) possuem semelhanças, com a distância variando de 0 a 200 metros em quase todo perímetro urbano, principalmente na região central da cidade. Isso pode estar relacionado com a expansão da cidade nas últimas décadas, que concentra uma população de classe alta e uma estrutura de drenagem já consolidada no centro da cidade, enquanto as áreas periféricas mais distantes do centro são carentes por uma infraestrutura básica como rede de drenagem de esgoto e água (Tucci, 2005).

Figura 15: Distância da rede de água de Campo Grande, MS



Fonte: DINIZ (2022)

Figura 16: Distância da rede de esgoto de Campo Grande, MS

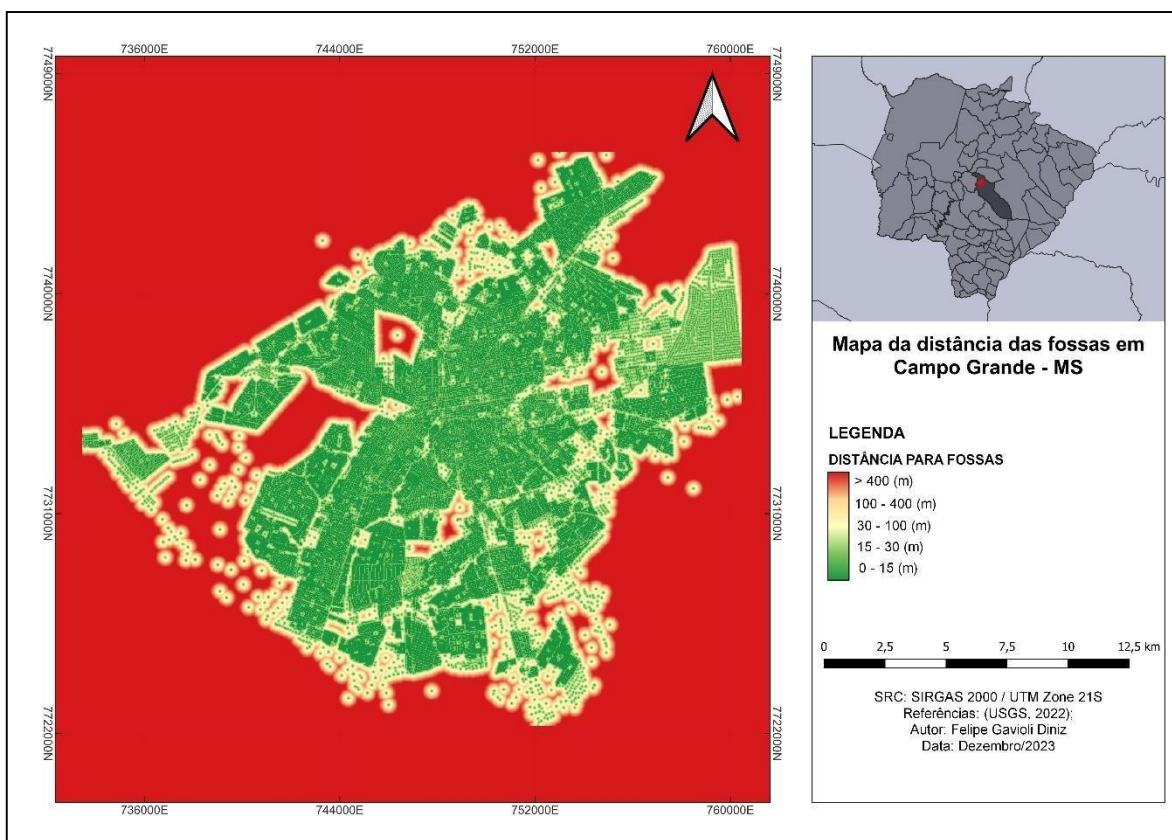


Fonte: DINIZ (2022)

4.1.9 Distância entre as fossas

Observando a Figura 17, nota-se que boa parte da distância das fossas entre os lotes estão concentradas em uma distância de até 30 metros, revelando a proximidade destas fossas. Além disso, toda região do perímetro urbano é composta por fossas, especialmente na região central da cidade, com as maiores distâncias presentes ao sul do perímetro urbano.

Figura 17: Distância entre as fossas dos lotes em Campo Grande, MS



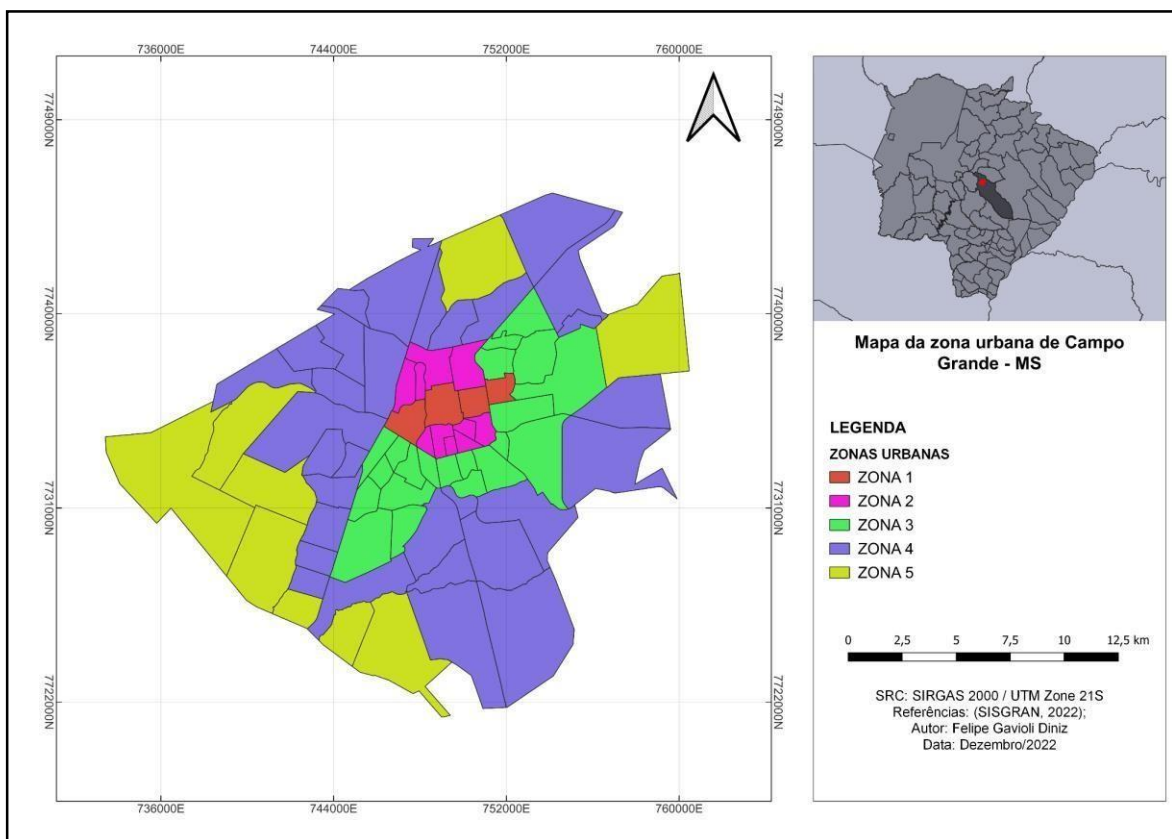
Fonte: DINIZ (2023)

4.1.10 Zona urbana

A zona urbana de Campo Grande é dividida em 5 zonas (Figura 18), divisão feita a partir de critérios como: categorias de uso permitidos (residencial, comercial, industrial, serviços, loteamento); índices e instrumentos urbanísticos aplicáveis à zona e eixo de adensamento (taxa de ocupação, índice de elevação e diversos coeficientes); lotes mínimos (m) (área, testada esquina e testada meio da quadra) e recuos mínimos (m) (frente, lateral e fundos).

Os critérios usados são determinados pela Lei Complementar n. 341/2018, estabelecendo para cada zona uma estrutura específica que influencia na organização espacial da cidade. Desse modo, cada zona possui suas próprias diretrizes que impactam no crescimento ou alterações no espaço urbano.

Figura 18: Zona urbana de Campo Grande, MS

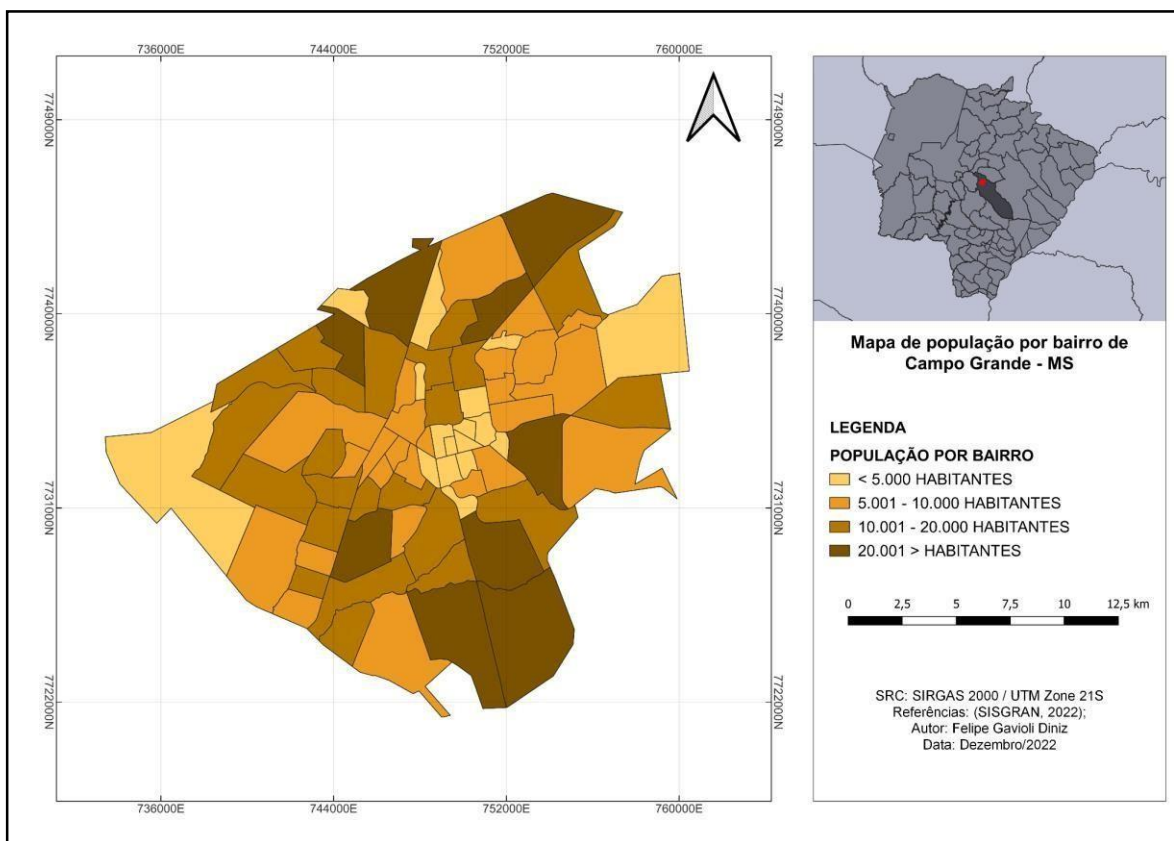


Fonte: DINIZ (2022)

4.1.11 População por bairro

A população por bairro de Campo Grande (Figura 19) foi dividida em quatro intervalos de classe, correspondendo: 5.000 habitantes ou menos (25%), entre 5.001 a 10.000 habitantes (32%), entre 10.001 a 20.000 habitantes (31%) e acima de 20.001 habitantes (12%). O bairro mais populoso de Campo Grande é o Aero Rancho com 36.057 pessoas, localizado na região Anhanduizinho, no sul do perímetro urbano. Por outro lado, o bairro com menor população é o TV Morena com 2.132 pessoas, localizado na região Bandeira, no centro da cidade (Campo Grande, 2022).

Figura 19: População por bairro da cidade de Campo Grande, MS

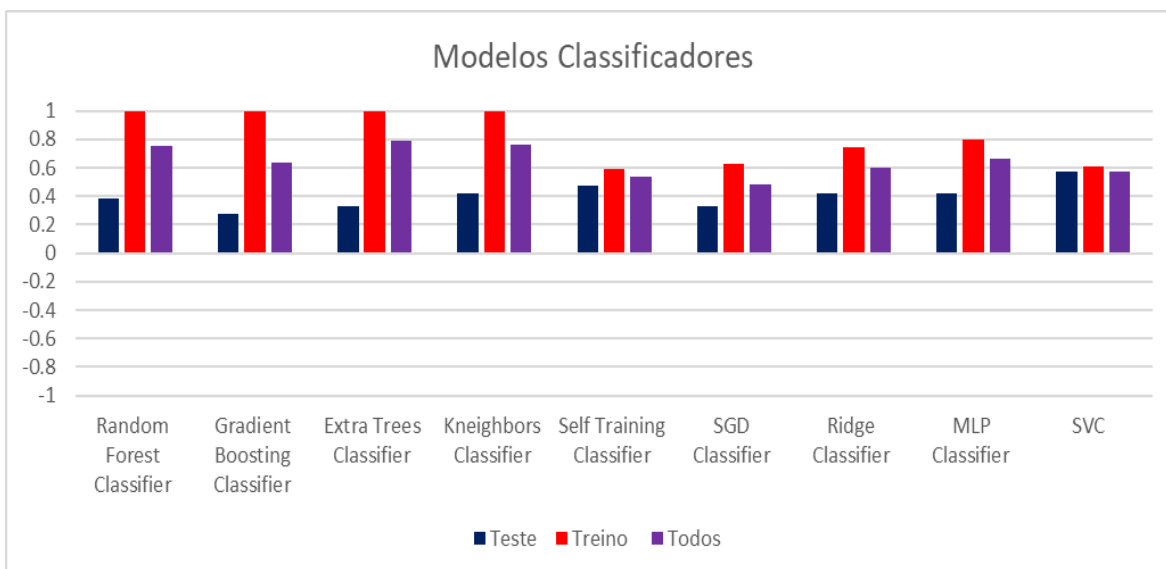


Fonte: DINIZ (2022)

4.2 Desempenho dos modelos

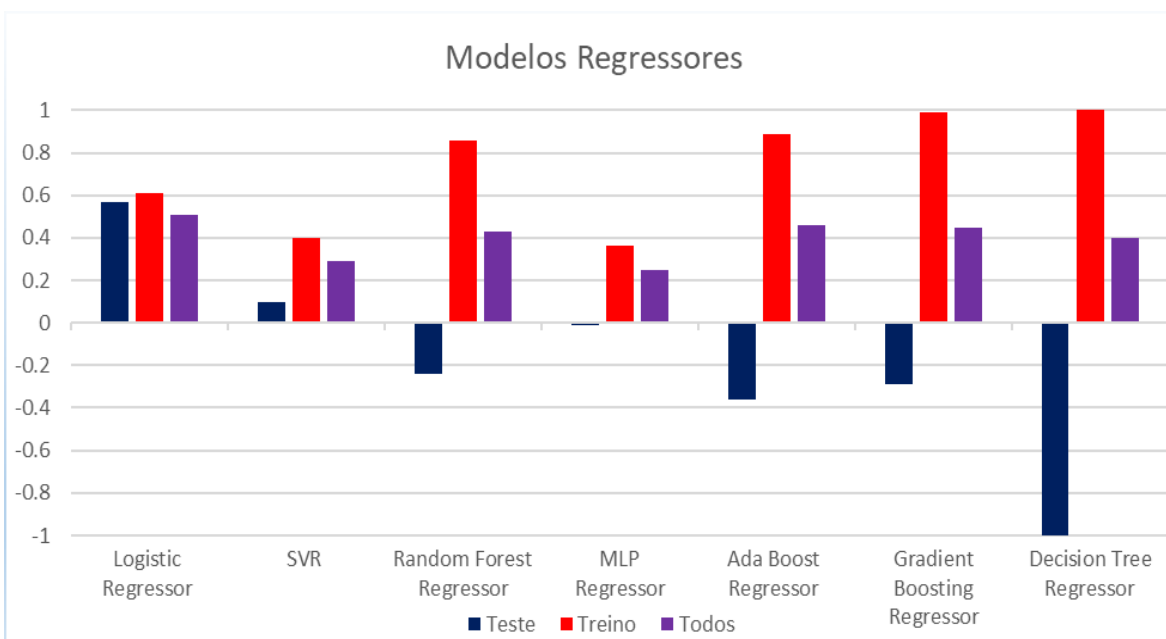
O resultado final da etapa “análise de acurácia” mostrou que os nove modelos classificadores apresentaram um desempenho superior aos sete modelos regressores, como pode ser observado nas Figuras 20 e 21.

Figura 20: Desempenho dos modelos classificadores



Fonte: DINIZ (2023)

Figura 21: Desempenho dos modelos regressores



Fonte: DINIZ (2023)

Os desempenhos dos modelos classificadores mostraram-se mais eficientes em relação aos modelos regressores, com cinco modelos apresentando uma eficiência razoável. Esses modelos foram: Self Training Classifier; SVC; SGD

Classifier; Ridge Classifier e; MLP Classifier, de modo que, o último destes modelos apresentou o melhor resultado. Essa diferença de eficiência entre os modelos impactou na importância das variáveis sobre o valor de nitrato nas águas subterrâneas, com algumas apresentando grandes discrepâncias no *ranqueamento* realizado utilizando os modelos classificadores e regressores.

Os demais modelos classificadores apresentaram *overfitting*, um problema no aprendizado de máquina onde o modelo ajusta-se perfeitamente no grupo treino, possuindo uma alta precisão, mas que, tem dificuldades de prever novas simulações por não funcionar corretamente no grupo teste; portanto, devem ser evitados (Ying, 2019).

Isso pode responder ao mau desempenho de alguns modelos regressores que apresentaram valores negativos no grupo de treino e *overfitting* ou resultados quase perfeitos no grupo de treinamento. A Figura 20 evidencia isso, mostrando que cinco modelos possuíram dados do grupo teste negativo, especialmente o Ada Boost Regressor, Gradient Boosting Regressor e Decision Tree Regressor, modelos próximos ou com *overfitting*. No grupo dos modelos regressores, apenas o Logistic Regressor apresentou um desempenho aceitável.

Esse resultado final ocorreu utilizando o percentual 80% (selecionando os poços com valores de nitrato abaixo desse valor), procedimento que cortou 20% dos poços existentes na etapa “tratamento preliminar” dos dados. Isso foi necessário, pois nas amostras existentes sobre o nitrato, presentes nos poços tubulares, havia a ocorrência de *outliers*.

Outliers consistem em dados discrepantes, ou seja, dados com valores distantes do conjunto de dados regulares. Sua presença no treinamento pode reduzir a acurácia do modelo, distorcendo os resultados. Desta forma, foram cortados os *outliers* para um melhor desempenho.

Mesmo com este corte, nenhum modelo, fora os que apresentaram *overfitting*, conseguiram passar de 0,70 no desempenho dos modelos, mostrando que o aprendizado de máquina não é uma boa ferramenta para prever contaminação (Paulheim e Meusel, 2015; Zhang e Luo, 2015), visto que os valores que puderam ser preditos pelo modelo estão abaixo de 10 mg/L, que é o limite

mínimo para classificar uma área como contaminada por nitrato e ser evitada para o consumo humano, de acordo com Resolução CONAMA nº. 396. (Brasil, 2008).

O desempenho dos modelos classificadores em relação aos modelos regressores nesta pesquisa pode estar relacionado com a sua capacidade em ordenar mediante a importância, influência ou relevância. Esse é o objetivo em um ranqueamento e não prever um valor contínuo, este sendo o propósito dos modelos regressores (Lima, 2022).

Esta parte evidencia que o aprendizado de máquina apresenta problemas para esse tipo de pesquisa, visto que, há grandes dificuldades em simular novos resultados e trabalhar com dados discrepantes. Por outro lado, atualmente suas vantagens estão relacionadas com a redução de custo e tempo, maior eficiência e aumento da produtividade, graças as inovações tecnológicas recentes que fizeram essa ferramenta evoluir.

4.3 Importância dos fatores naturais e antrópicos na ocorrência de nitrato

Com a etapa da “análise de importância” foi comparada a importância de cada variável espacial sobre o valor de nitrato nas águas subterrâneas de Campo Grande, MS, usando os modelos classificadores e regressores (Tabela 03).

Tabela 03: Importância das variáveis espaciais sobre a presença de nitrato nas águas subterrâneas de Campo Grande, MS

IMPORTÂNCIA	CLASSIFICADORES	REGRESSORES
1°	Declividade	Zona urbana
2°	Profundidade/NA	Geologia
3°	Rede de esgoto	Rede de esgoto
4°	População por bairro	Hidrografia
5°	Espessura do solo	Rede de água
6°	NDVI	Fossas
7°	Rede de água	Hipsometria

8°	Geologia	Declividade
9°	Zona urbana	Profundidade/NA
10°	Fossas	NDVI
11°	Hipsometria	População por bairro
12	Hidrografia	Espessura do solo

Fonte: DINIZ (2023)

A partir da Tabela 2, nota-se uma diferença considerável nos resultados da influência das variáveis sobre o valor de nitrato nas águas subterrâneas utilizando os modelos classificadores e regressores. Como o modelo classificador apresentou uma eficiência superior ao modelo regressor, o ranqueamento do modelo classificador passou maior confiabilidade nos resultados.

Nos modelos classificadores a variável com maior influência foi a declividade, enquanto, nos modelos regressores essa mesma variável ficou em oitavo lugar. A declividade está relacionada com a poluição das águas subterrâneas por meio da inclinação do terreno. Quanto maior a inclinação, maior será a concentração do volume de água em um determinado local e conseqüentemente, a infiltração de água com solutos para o solo.

Neste processo há o transporte de poluentes originários de várias atividades antrópicas, além de sedimentos resultantes do processo de erosão. Os poluentes e os sedimentos são depositados nas áreas do entorno, terrenos mais baixos com um declive moderado ou totalmente plano. Juntamente com a profundidade do nível da água, geologia vulnerável e espessura do solo, facilitam a infiltração desses poluentes para as águas subterrâneas Martinez, (2010) e Yadav, et al, (2023).

A segunda variável com maior influência no ranqueamento dos classificadores foi a profundidade do nível da água. Segundo Lake, et al. (2003) e Bridha e Elango (2015), áreas com o nível freático próximo da superfície são consideradas mais vulneráveis para poluição dos recursos hídricos subterrâneos, sendo atribuídos valores maiores em uma classificação de importância como nesta pesquisa.

Observando a Figura 13 sobre a profundidade do nível da água em Campo

Grande, nota-se que quase 72% das águas subterrâneas estão localizadas a menos de 5 metros da superfície, tornando-se áreas mais suscetíveis a poluição. No entanto, apenas 6,5% das águas subterrâneas estão entre 15 e 75 metros de profundidade, indicando uma área menos suscetível a poluição em relação a profundidade, sobretudo na região nordeste da cidade.

A única variável com a mesma importância em ambos os modelos foi a rede de esgoto, ficando em terceiro no ranqueamento. Como falado anteriormente, a rede de esgoto é uma das principais fontes poluidoras das águas subterrâneas, devido a sua estrutura que pode apresentar falhas e vazamentos, representando uma ameaça real aos aquíferos rasos não confinados (Sowah, et al, 2017).

Essa rede de esgoto, composta em grande maioria por água e matéria orgânica em decomposição, é fonte de metais pesados, como o mercúrio e o chumbo, e de nitrato. Esses metais pesados contaminam a água, tornando-se prejudiciais não apenas a saúde humana, mas também ao desenvolvimento socioeconômico e a qualidade ambiental (McArthur, et al, 2012; Li, et al, 2021).

Ao consumir essa água contaminada pelo nitrato e outros tipos de poluentes, a pessoa pode ser infectada e contrair doenças, sendo as mais comuns, a leptospirose, cólera e a neoplasia gástrica, além de infecções bacterianas. Grupos mais vulneráveis como bebês e mulheres grávidas estão mais suscetíveis a estas doenças (McArthur, et al, 2012; Khatri e Tyagi, 2015; Li, et al, 2021).

A variável com maior impacto nos modelos regressores foi a zona urbana, e que nos modelos classificadores ficou em nono, apresentando uma grande discrepância, comprovando a diferença da qualidade dos resultados dos modelos classificadores e regressores. Isso pode ser explicado pelo fato de praticamente todos os poços estar em área urbana.

Outras variáveis que apresentaram divergências nos resultados foram: a espessura do solo, quinto lugar nos modelos classificadores e em último nos regressores; hidrologia, segundo lugar nos regressores e em último nos

classificadores; geologia, segundo lugar na análise usando os modelos regressores e apenas em oitavo usando os modelos classificadores.

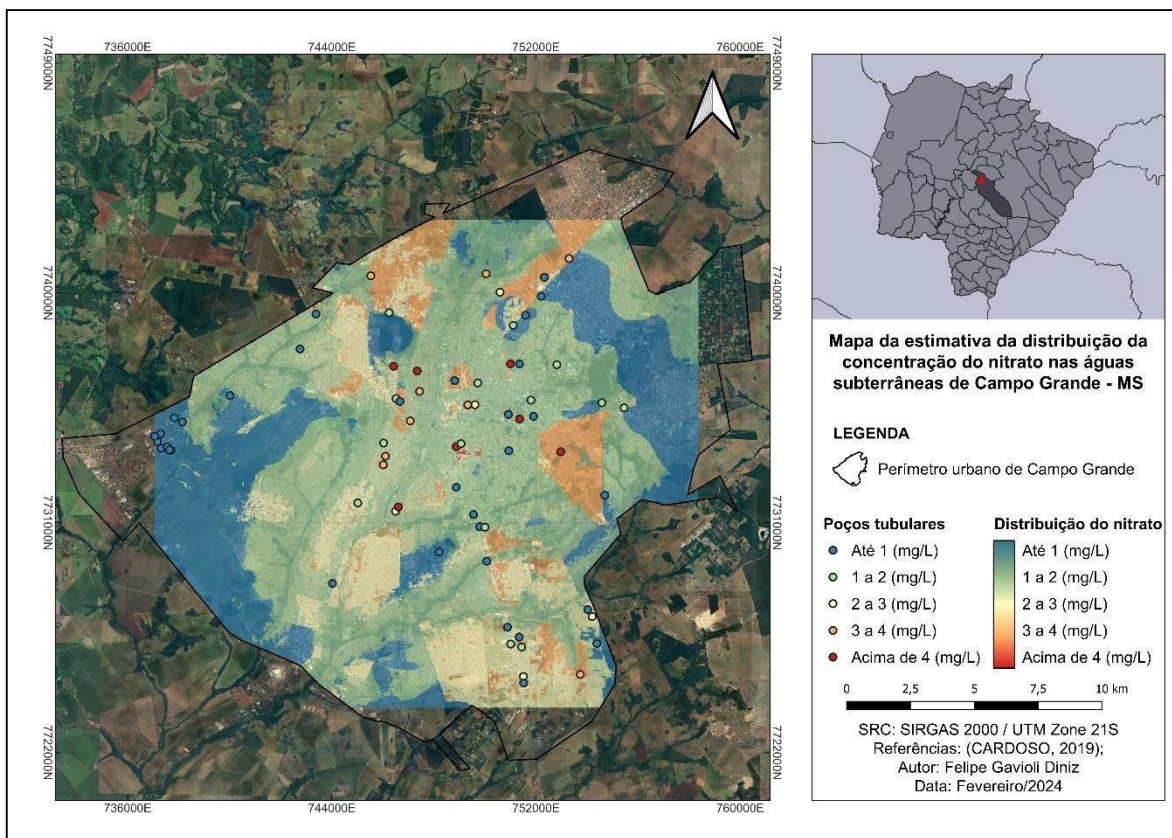
No geral, os fatores naturais tiveram maior importância nos modelos classificadores, exceto a distância para hidrografia, geologia e hipsometria que tiveram maior relevância nos modelos regressores. Enquanto isso, todos os fatores antrópicos, salvo a rede de esgoto que se mostrou com um resultado igual nos dois modelos, e a população por bairro que se destacou nos classificadores, apresentaram relevância superior nos modelos regressores.

Futuramente, visando melhorar e complementar pesquisas desta mesma linha de estudo, outros parâmetros, que não foram abordados nesta pesquisa, podem ser investigados. O escoamento superficial e os lineamentos estruturais surgem como os principais parâmetros a serem utilizados, pois estão relacionados diretamente com a declividade, e desta forma, podem contribuir para compreender porque esse parâmetro teve grande influência sobre o valor de nitrato usando os modelos classificadores.

4.4. Estimativa da distribuição da concentração de nitrato na área urbana de Campo Grande, MS

Com a etapa “simulação” foi realizada a estimativa da distribuição do nitrato nas águas subterrâneas de Campo Grande, MS (Figura 22).

Figura 22: Estimativa da distribuição da concentração do nitrato nas águas subterrâneas de Campo Grande, MS



Fonte: DINIZ (2024)

Analisando a Figura 10 sobre a declividade de Campo Grande e a Figura 23 sobre a estimativa da distribuição do nitrato, nota-se que uma parcela das áreas com as maiores concentrações de nitrato, com um valor entre 3 e 4 mg/L está localizada na região norte, em torno das áreas com os maiores declives da cidade. Isso corrobora que a declividade foi uma das variáveis que mais influenciou sobre a presença do nitrato nas águas subterrâneas graças ao processo de escoamento superficial, deposição e infiltração desses poluentes até a chegada nas águas subterrâneas.

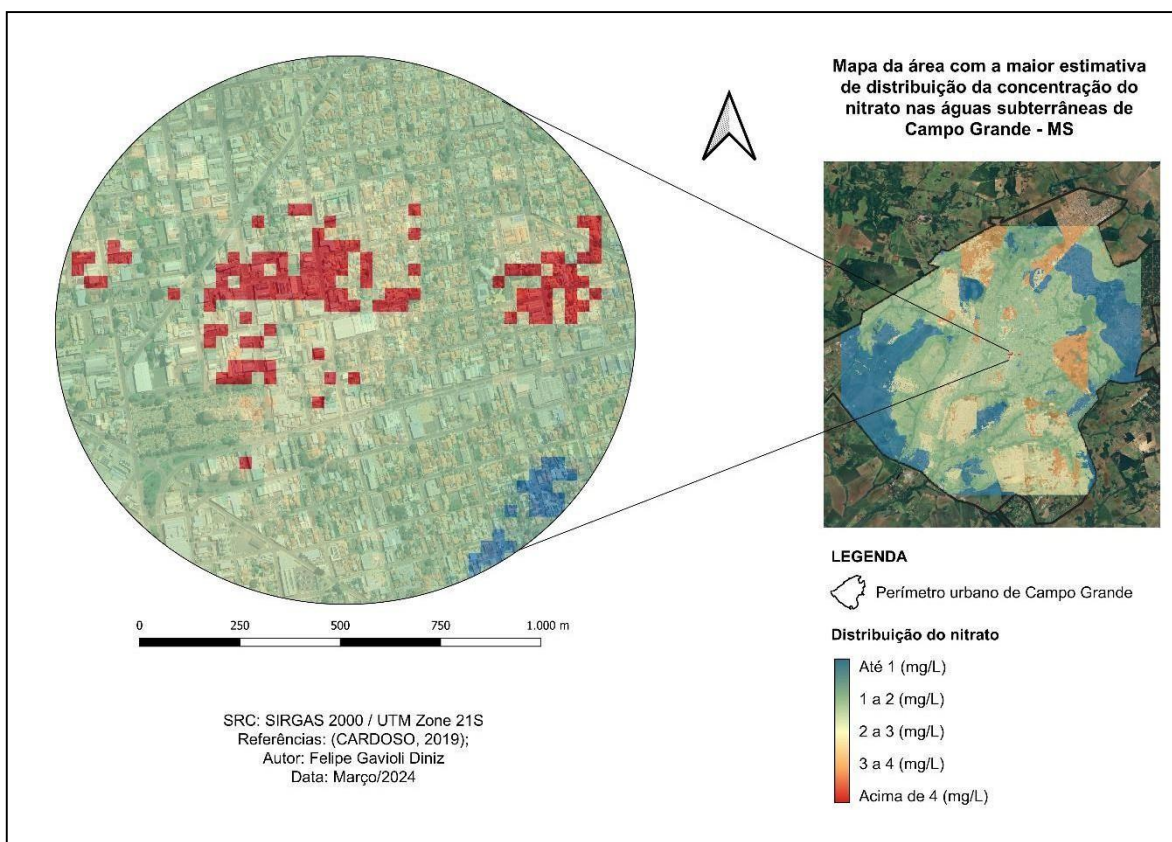
Outras regiões que apresentaram uma concentração de nitrato entre 3 e 4 mg/L foram ao leste, no bairro do Tiradentes e ao sul da cidade, nos bairros do Centro Oeste, Moreninhas e Universitário. Esses bairros estão entre os mais populosos da cidade, com uma população acima dos 20.000 habitantes, ratificando

a importância da variável população por bairro que ficou em quarto no ranqueamento sobre a importância das variáveis sobre o valor de nitrato nas águas subterrâneas. Contudo, em outros bairros com a população superior a 20.000 habitantes, como o Aero Rancho e Santo Amaro, a quantidade de nitrato nas águas subterrâneas foi de 2 a 3 mg/L.

Sobre isto, Marques Arsênio et al (2018) afirmam que bairros com uma alta densidade de pessoas pode agravar o risco de contaminação das águas subterrâneas de aquíferos não confinados e semiconfinados, mediante a exploração, por meio de poços tubulares com diversas finalidades de uso, especialmente para o abastecimento. Nestas regiões de alta densidade, a urbanização é rápida e apresenta uma série de riscos e desafios, entre eles a construção de uma infraestrutura básica inadequada e uma coleta de resíduos ineficiente. (Muhammad e Zhonghua, 2014; La Vigna, 2022).

Paralelamente, a única região que apresentou uma faixa de nitrato acima de 4 mg/L, por volta de 5mg/L na cidade de Campo Grande foi na região central da cidade, onde localiza-se a Avenida Rui Barbosa (Figura 23). Essa região apresenta um grande fluxo populacional, atividades e comércios e uma rede de esgoto densa, motivos que podem justificar esta taxa elevada de nitrato em comparação a outras regiões.

Figura 23: Estimativa da área com a maior concentração do nitrato nas águas subterrâneas de Campo Grande, MS



Fonte: DINIZ (2024)

Por outro lado, as menores concentrações de nitrato foram encontradas na região leste e oeste da cidade de Campo Grande. Na região oeste, há áreas de poucos declives, poços tubulares com poucas concentrações de nitrato, além de ser uma região quase sem nenhuma urbanização. Nela está localizada o aeroporto internacional de Campo Grande e o bairro Núcleo Industrial, um bairro com pouca população e moradia fixa, não tendo rede de esgoto ampla e apresentando vários vazios urbanos.

Já na região leste de Campo Grande, a baixa concentração de nitrato está associada com o nível baixo da profundidade do nível da água. Analisando a Figura 22, nota-se uma grande mancha na cor azul, indicando uma área com pouca concentração de nitrato, de até 1 mg/L. Na Figura 13, percebe-se que está mesma

mancha está presente no mapa sobre a profundidade do nível d'água, uma região onde a água se encontra entre 15 e 75 metros abaixo da superfície. Isso reforça o que já existe na literatura, onde regiões com maiores profundidades, há menor ocorrência de nitrato e de outros poluentes por estarem mais protegidas de possíveis poluidores.

5. CONCLUSÃO

Esta pesquisa buscou identificar quais fatores antrópicos e naturais que influenciam na presença do nitrato nas águas subterrâneas da cidade de Campo Grande, MS utilizando o aprendizado de máquina. Para ajudar a responder esses objetivos, foram estabelecidos objetivos específicos visando facilitar o entendimento de pesquisas desta natureza, sobretudo com o uso do aprendizado de máquina.

O processo de identificar quais fatores influenciam a presença de um poluente nas águas subterrâneas pode ser dividido em várias etapas. Nesta pesquisa, a primeira etapa foi a caracterização espacial de cada variável, respondendo o que foi determinado no primeiro objetivo específico. Os dados obtidos desta caracterização, apresentados por mapas, contribuiu não apenas a entender a estrutura da cidade, mas também ajudar a responder outras perguntas e objetivos deste trabalho.

Para responder o segundo objetivo específico, foi feita a compilação das informações de cada poço utilizado nesta pesquisa por meio de tabelas, contendo coordenadas geográficas, finalidade de uso, sistema aquífero aflorando e captado de cada poço e os valores de nitrato. Isso permitiu descobrir a quantidade de nitrato em cada poço utilizado na pesquisa, e usar estas informações de nitrato para contribuir em outros objetivos. Anterior a isso, buscando melhorar a qualidade dos dados de nitrato, foi realizado no Python a exclusão dos poços que não continham dados de nitrato ou que foram perfurados em mais de um aquífero.

O ranqueamento realizado sobre a importância de cada fator utilizado, após os resultados mostrarem que os modelos classificadores apresentam um desempenho superior aos regressores, serviu para responder o terceiro objetivo específico. Foi constatando que a declividade e a profundidade do nível da água são atributos espaciais que juntamente com a rede de esgoto, influem sobre o valor de nitrato nas águas subterrâneas.

Esses dados foram assegurados pela estimativa da distribuição do nitrato, mostrando que as áreas com as maiores concentrações de nitrato nas águas subterrâneas de Campo Grande são aquelas em torno das áreas com maiores declives.

Conclui-se que estas áreas com as maiores concentrações de nitrato não são recomendáveis para construção de poços, apesar de ainda estarem dentro do limite estabelecido Resolução CONAMA nº. 396. Entretanto, são áreas que podem apresentar agravamento nos níveis de nitrato, por conta das características naturais e antrópicas e caso não houver um estudo de redução desse poluente.

Normalmente, técnicas para redução de poluentes em águas subterrâneas são dispendiosas, evitando ser aplicadas em grandes escalas. Técnicas como a desnitrificação química e biológica vem sendo bastante aplicada para conter o crescimento do nitrato em águas subterrenêas, porém a opção pela técnica mais adequada depende das características do aquífero, do custo da operação e da qualidade da água a ser tratada.

Além disso, o aprendizado de máquina não apresentou um desempenho satisfatório para esta pesquisa. O seu mal desempenho pode estar atrelado ao método aplicada na pesquisa, carecendo de utilização de modelos mais robustos ou de qualidade dos dados fornecidos que não os mais adequados para análises hidrogeológicas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, Purushottam et al. Exploring artificial intelligence techniques for groundwater quality assessment. **Water**, v. 13, n. 9, p. 1172, 2021.

AHMADI, Arman et al. Groundwater level modeling with machine learning: a systematic review and meta-analysis. **Water**, v. 14, n. 6, p. 949, 2022.

ANA. **ATLAS Brasil**. Vol 1. Brasília, 2010. Disponível em: <https://www.aris.sc.gov.br/uploads/revista/2735/YXwFOHqcDQSDQEA2bJfi2KfS2Pt-Binw.pdf>. Acesso em: 10 mai. 2022.

ALMEIDA SILVA, Caroline Aparecida; DE OLIVEIRA, Jefferson Nascimento. O USO DO SERRA GERAL NO ESTADO DO MATO GROSSO DO SUL. **Águas Subterrâneas**, 2011.

AL-YAQOUT, A. F.; HAMODA, M. F. Evaluation of landfill leachate in arid climate—a case study. **Environment international**, v. 29, n. 5, p. 593-600, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 12212: **Projeto de poço tubular para captação de água subterrânea - Procedimento**: Referências. Rio de Janeiro, p. 17. 2017.

AWAIS, Muhammad et al. Assessing nitrate contamination risks in groundwater: a machine learning approach. **Applied Sciences**, v. 11, n. 21, p. 10034, 2021.

BAIRD, C.; CANN, M. **Química ambiental**. 4. ed. Porto Alegre: Bookman, 2011.

BERTHIER, Etienne et al. Biases of SRTM in high-mountain areas: Implications for the monitoring of glacier volume changes. **Geophysical Research Letters**, v. 33, n. 8, 2006.

BHAVSAR, Himani; PANCHAL, Mahesh H. A review on support vector machine for data classification. **International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)**, v. 1, n. 10, p. 185-189, 2012.

BRASIL. **Portaria GM/MS Nº 888, de 4 de maio de 2021**. Altera o Anexo XX da Portaria de Consolidação GM/MS nº 5, de 28 de setembro de 2017, para dispor sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade. Diário Oficial da União. Brasília, DF, Edição 85, p.127. 07 mai. 2021. Seção 1.

BRASIL. Ministério do Meio Ambiente. Conselho Nacional do Meio Ambiente – CONAMA. **Resolução nº 396, de 03 de abril de 2008**: Dispõe sobre a classificação e diretrizes ambientais para o enquadramento das águas subterrâneas. Brasília, 2008.

BREUER, Lutz et al. Current concepts in nitrogen dynamics for mesoscale catchments. **Hydrological Sciences Journal**, v. 53, n. 5, p. 1059-1074, 2008.

BRINDHA, Karthikeyan; ELANGO, L. Cross comparison of five popular groundwater pollution vulnerability index approaches. **Journal of Hydrology**, v. 524, p. 597-613, 2015.

CAO, Xu et al. Impacts of anthropogenic groundwater recharge (AGR) on nitrate dynamics in a phreatic aquifer revealed by hydrochemical and isotopic technologies. **Science of The Total Environment**, p. 156187, 2022.

CAMPO GRANDE. Agência Municipal do Meio Ambiente e Planejamento Urbano (Org.) **Diagnóstico Ambiental da Bacia do Córrego Bandeira**. Campo Grande, 2017. Disponível em: <http://www.campogrande.ms.gov.br/planurb/wpcontent/uploads/sites/18/2020/08/DIAG-BANDEIRA-FINAL.pdf>. Acesso em: 04 out. 2022

CAMPO GRANDE. Instituto Municipal de Planejamento Urbano (Org.). **Bacia Hidrográfica Córrego Bandeira: é bom saber**. Campo Grande, 2018. Disponível em: <http://www.campogrande.ms.gov.br/planurb/wp-content/uploads/sites/18/2018/11/Folheto-Bandeira.pdf>. Acesso em: 03 out. 2022.

CAMPO GRANDE. Instituto Municipal de Planejamento Urbano (Org.) **Perfil Socioeconômico de Campo Grande**. ed 30. Campo Grande, 2023. Disponível em: <https://prefcg-repositorio.campogrande.ms.gov.br/wp-content/uploads/sites/18/2023/08/PERFIL-2023-PDF-SITE-2.pdf>. Acesso em: 03 mar. 2024

CAMPO GRANDE. Prefeitura de Campo Grande e Hidrosul Ambiental Serviços Geológicos Ltda. (Org.) **Carta Geotécnica de Campo Grande**: Resumo executivo, 2020. Disponível: <https://www.campogrande.ms.gov.br/>. Acesso em: 19 dez. 2022.

CARDOSO, Juarez Marques. **AVALIAÇÃO DA QUALIDADE DA AGUA SUBTERRANEA DE POÇOS TUBULARES OUTORGADOS NO MUNICIPIO DE CAMPO GRANDE – MS**. 2019. 229 f. Dissertação (mestrado) – Curso de Gestão e Regulação de Recursos Hídricos, Unesp, Ilha Solteira, 2019.

CASADO, Ismael et al. Use of electrical tomography methods to determinate the extension and main migration routes of uncontrolled landfill leachates in fractured areas. **Science of the Total Environment**, v. 506, p. 546-553, 2015.

CONSELHO NACIONAL DO MEIO AMBIENTE. **Resolução Nº 420**, de 28 de dezembro de 2009. Dispõe sobre critérios e valores orientadores de qualidade do solo quanto à presença de substâncias químicas e estabelece diretrizes para o gerenciamento ambiental de áreas contaminadas por essas substâncias em decorrência de atividades antrópicas. Brasil, 2009.

CORDEIRO, Ana Paula Assumpção et al. REGIÕES HOMOGÊNEAS DE VEGETAÇÃO UTILIZANDO A VARIABILIDADE DO NDVI. **Ciência Florestal**, Santa Maria, v. 27, n. 3, p. 883-896, set. 2017.

CPRM - SERVIÇO GEOLÓGICO DO BRASIL. **Mapa geológico do estado do Mato Grosso do Sul**. 2006. Mapa, color. Escala 1: 1.000.000. Programa Geologia do Brasil.

DEVIC, Gordana; DJORDJEVIC, Dragana; SAKAN, Sanja. Natural and anthropogenic factors affecting the groundwater quality in Serbia. **Science of the Total Environment**, v. 468, p. 933-942, 2014.

EngeSat. **Sentinel-2**. 2022. Disponível em: <http://www.engesat.com.br/sentinel-2/>. Acesso em: 02 de dez. 2022.

ERDEM, Ebru; BOZKURT, Ferhat. A comparison of various supervised machine learning techniques for prostate cancer prediction. **Avrupa Bilim ve Teknoloji Dergisi**, n. 21, p. 610-620, 2021.

FEITOSA, F. A. C.; MANOEL FILHO, J. **Hidrogeologia: conceitos e aplicações**. 3. ed. Rio de Janeiro: CPRM, LABHID-UFPE, 2008. 812 p.

FREEZE, R. Allan; CHERRY, John A. Groundwater Prentice-Hall Inc. **Eaglewood Cliffs, NJ**, 1979. 604 p.

GASTMANS, Didier; KIANG, Chang Hung. Avaliação da hidrogeologia e hidroquímica do Sistema Aquífero Guarani (SAG) no estado de Mato Grosso do Sul. **Águas Subterrâneas**, v. 19, n. 1, 2005.

GILMARQUES, Carlos Henrique et al. Evolução espacial e temporal da contaminação por nitrato no aquífero urbano de Urânia (SP). **Águas subterrâneas**, 2019.

HAGGERTY, Ryan et al. Application of machine learning in groundwater quality modeling-A comprehensive review. **Water Research**, p. 119745, 2023.

HAMIDI, Mohammad Daud et al. Spatial estimation of groundwater quality, hydrogeochemical investigation, and health impacts of shallow groundwater in Kabul city, Afghanistan. **Sustainable Water Resources Management**, v. 9, n. 1, p. 20, 2023.

HARRIS, Charles R. et al. Array programming with NumPy. **Nature**, v. 585, n. 7825, p. 357-362, 2020.

HE, Kaiming et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **Proceedings of the IEEE international conference on computer vision**. 2015. p. 1026-1034.

HUANG, Ruixing et al. Machine learning in natural and engineered water systems. **Water Research**, v. 205, p. 117666, 2021.

HUNTER, John D. Matplotlib: A 2D graphics environment. **Computing in science & engineering**, v. 9, n. 03, p. 90-95, 2007.

IBGE. **Cidades e estados**, 2024. Disponível em: <https://www.ibge.gov.br/cidades-e-estados/ms/campo-grande.html> . Acesso em: 03 mar. 2024

JAMEI, Mehdi et al. Computational assessment of groundwater salinity distribution within coastal multi-aquifers of Bangladesh. **Scientific Reports**, v. 12, n. 1, p. 1-28, 2022.

KHATRI, Nitasha; TYAGI, Sanjiv. Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas. **Frontiers in life science**, v. 8, n. 1, p. 23-39, 2015.

MANZANO, Liliane Maia Tcacenco. **MONITORAMENTO DOS SISTEMAS AQUÍFEROS UTILIZADOS NO ABASTECIMENTO PÚBLICO DO MUNICÍPIO DE CAMPO GRANDE, MS**. 113 f. Dissertação (Mestrado) - Curso de Geologia Ambiental, Universidade Federal do Paraná, Curitiba, 2018.

MANZIONE, Rodrigo Lilla. **Águas Subterrâneas**. Paco Editorial, 2015. 389 p.

MARQUES ARSÊNIO, André et al. Mitigation potential of sanitation infrastructure on groundwater contamination by nitrate in Maputo. *Sustainability*, v. 10, n. 3, p. 858, 2018.

MARTÍNEZ, Leidy Luz García. **Distribuição de poluição difusa por sedimentos urbanos em áreas impermeáveis em Porto Alegre. 2010. 90 f.** Tese de Doutorado. Dissertação (Mestrado). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.

MASETTI, Marco et al. Spatial and statistical assessment of factors influencing nitrate contamination in groundwater. **Journal of environmental management**, v. 86, n. 1, p. 272-281, 2008.

MATO GROSSO DO SUL. **Referências Ambientais e Sócio-Econômicas para Gestão do Território do Estado do Mato Grosso do Sul**: uma contribuição ao zoneamento ecológico-econômico do Mato Grosso do Sul. Vol. 1. 2008. 206 p.

MCARTHUR, J. M. et al. Waste-water impacts on groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal Basin and Red River Basin, Vietnam. **Science of the Total Environment**, v. 437, p. 390-402, 2012.

MORITA, Alice KM et al. Pollution threat to water and soil quality by dumpsites and non-sanitary landfills in Brazil: A review. **Waste Management**, v. 131, p. 163-176, 2021.

MUHAMMAD, Akhtar Malik; ZHONGHUA, Tang. Municipal solid waste and its relation with groundwater contamination in Lahore, Pakistan. **Research Journal of Applied Sciences, Engineering and Technology**, v. 7, n. 8, p. 1551-1560, 2014.

NORDIN, Nur Farahin Che et al. Groundwater quality forecasting modelling using artificial intelligence: A review. **Groundwater for Sustainable Development**, v. 14, p. 100643, 2021.

KATZ, Brian G.; EBERTS, Sandra M.; KAUFFMAN, Leon J. Using Cl/Br ratios and other indicators to assess potential impacts on groundwater quality from septic systems: a review and examples from principal aquifers in the United States. **Journal of Hydrology**, v. 397, n. 3-4, p. 151-166, 2011.

KRISHNA, R. et al. Groundwater vulnerability to pollution mapping of Ranchi district using GIS. **Applied water science**, v. 5, p. 345-358, 2015.

LAKE, Iain R. et al. Evaluating factors influencing groundwater vulnerability to nitrate pollution: developing the potential of GIS. **Journal of environmental management**, v. 68, n. 3, p. 315-328, 2003.

LASTORIA, Giancarlo et al. Hidrogeologia da formação serra geral no Estado de Mato Grosso do Sul. **Águas Subterrâneas**, v. 20, n. 1, 2006.

LASTORIA, Giancarlo et al. EVIDÊNCIAS DA CONECTIVIDADE HIDRÁULICA ENTRE OS SISTEMAS AQUÍFEROS SERRA GERAL E GUARANI NO ESTADO DE MATO GROSSO DO SUL E ASPECTOS AMBIENTAIS CORRELACIONADOS. **Águas Subterrâneas**, 2007.

LALL, Upmanu; JOSSET, Laureline; RUSSO, Tess. A snapshot of the world's groundwater challenges. **Annual review of environment and resources**, v. 45, p. 171-194, 2020.

LA VIGNA, Francesco. Urban groundwater issues and resource management, and their roles in the resilience of cities. **Hydrogeology Journal**, v. 30, n. 6, p. 1657-1683, 2022.

LEAL, Ramos et al. Assessing the inconsistency between groundwater vulnerability and groundwater quality: the case of Chapala Marsh, Mexico. **Hydrogeology journal**, v. 20, n. 3, p. 591, 2012.

LEE, In; SHIN, Yong Jae. Machine learning for enterprises: Applications, algorithm selection, and challenges. **Business Horizons**, v. 63, n. 2, p. 157-170, 2020.

LI, Peiyue et al. Sources and consequences of groundwater contamination. **Archives of environmental contamination and toxicology**, v. 80, p. 1-10, 2021.

LIMA, Deborah Lopes Correia. **MAPEAMENTO DE ÁREAS ALAGÁVEIS NA BACIA AMAZÔNICA UTILIZANDO O CLASSIFICADOR RANDOM FOREST A PARTIR DE DADOS EXTRAÍDOS DO MDE-SRTM**. 2022. 107 f. Dissertação (Mestrado), INPE, São José dos Campos, 2022.

PAULHEIM, Heiko; MEUSEL, Robert. A decomposition of the outlier detection problem into a set of supervised learning problems. **Machine Learning**, v. 100, p. 509-531, 2015.

PEDREGOSA, Fabian et al. Scikit-learn: Machine learning in Python. **the Journal of machine Learning research**, v. 12, p. 2825-2830, 2011.

PERH-MS. Plano Nacional de Recursos Hídricos do Estado do Mato Grosso do Sul. Secretaria do Estado do Meio Ambiente, do Planejamento, da Ciência e da

Tecnologia. **Instituto do Meio Ambiente do Mato Grosso do Sul**. Campo Grande – MS. Editora: UEMS, 2010. 194 p.

PLANURB. DADOS DE ESPESSURA DO SOLO. Mensagem recebida por: <https://mail.google.com/mail/u/0/#search/planurb/FFNDWLHpmgLPFFzjHqrkvMlftPgzNBjs> . Recebido em: 22 de nov. 2023.

RAGHAVENDRA, Sujay. N; DEKA, Paresh Chandra. Support vector machine applications in the field of hydrology: a review. **Applied soft computing**, v. 19, p. 372-386, 2014.

RANSOM, Katherine Marie et al. Machine learning predictions of nitrate in groundwater used for drinking supply in the conterminous United States. **Science of The Total Environment**, v. 807, p. 151065, 2022.

RAO, N. Subba; DINAKAR, A.; SUN, L. Estimation of groundwater pollution levels and specific ionic sources in the groundwater, using a comprehensive approach of geochemical ratios, pollution index of groundwater, unmix model and land use/land cover—A case study. **Journal of Contaminant Hydrology**. v. 248, p. 103990, 2022.

REBOUÇAS, A. C. Águas subterrâneas. In: GIAMPÁ, C.E.Q; GONÇALES, V.G. **Águas Subterrâneas e poços tubulares profundos**. 2 ed. rev. e atual. São Paulo: Oficina de textos, 2013. 496 p.

RODRIGUEZ-GALIANO, Victor et al. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). **Science of the Total Environment**, v. 476, p. 189-206, 2014.

Russell, Stuart, and Peter Norvig. 2002. *Artificial Intelligence: A Modern Approach*. Second. Prentice Hall.

SAJEDI-HOSSEINI, Farzaneh et al. A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination. **Science of the total environment**, v. 644, p. 954-962, 2018.

SANTOS, Thereza Christina Carvalho; CÂMARA, João Batista Drummond. **Geo Brasil 2002: perspectivas do meio ambiente no Brasil**. Brasília, DF: IBAMA, 2002., 2002.

SÃO PAULO. **As águas subterrâneas do Estado de São Paulo**. São Paulo, 2008. p. 104-104. Disponível em: <http://www.leb.esalq.usp.br/leb/disciplinas/Fernando/leb1440/Aula%208/As%20aguas%20subterraneas%20do%20Estado%20de%20Sao%20Paulo.pdf>. Acesso em: 13 ago. 2022.

SARKAR, Soumyajit et al. Predicting regional-scale elevated groundwater nitrate contamination risk using machine learning on natural and human-induced factors. **ACS ES&T Engineering**, v. 2, n. 4, p. 689-702, 2022.

SCHMOLL, Oliver et al. (Ed.). Protecting groundwater for health: managing the quality of drinking-water sources. **World Health Organization**, 2006. 689 p.

SHIUE, Ivy. Urinary thiocyanate concentrations are associated with adult cancer and lung problems: US NHANES, 2009–2012. **Environmental Science and Pollution Research**, v. 22, n. 8, p. 5952-5960, 2015.

SINGH, Umesh Kumar et al. Assessment of the impact of landfill on groundwater quality: a case study of the Pirana *site* in western India. **Environmental monitoring and assessment**, v. 141, n. 1-3, p. 309-321, 2008.

SINGHA, Sudhakar et al. Prediction of groundwater quality using efficient machine learning technique. **Chemosphere**, v. 276, p. 130265, 2021.

SOUZA, Aline Assunção. **O Aquífero Cenozóico em São Gabriel do Oeste-MS**. 2013.

SOWAH, Robert A. et al. Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States. **Water research**, v. 108, p. 330-338, 2017.

SPYDER, Home - Spyder IDE. Disponível: <<https://www.spyder-ide.org/>>. Acesso em: 24 de jan. de 2023.

TAHMASEBI, Pejman et al. Machine learning in geo-and environmental sciences: From small to large scale. **Advances in Water Resources**, v. 142, p. 103619, 2020.

TANG, Jiayi et al. Multiple pollutants in groundwater near an abandoned Chinese fluorine chemical park: concentrations, correlations and health risk assessments. **Scientific reports**, v. 12, n. 1, p. 1-9, 2022.

TENG, YanGuo et al. Risk assessment framework for nitrate contamination in groundwater for regional management. **Science of the Total Environment**, v. 697, p. 134102, 2019.

TESORIERO, Anthony J.; VOSS, Frank D. Predicting the probability of elevated nitrate concentrations in the Puget Sound Basin: implications for aquifer susceptibility and vulnerability. **Groundwater**, v. 35, n. 6, p. 1029-1039, 1997

TUCCI, Carlos EM. **Gestão de águas pluviais urbanas**. Programa de Modernização do Setor Saneamento, Secretaria Nacional de Saneamento Ambiental, Ministério das Cidades, 2005.

UECHI, Denise Aguenta; GABAS, Sandra Garcia; LASTORIA, Giancarlo. ANÁLISE COMPARATIVA DA CONCENTRAÇÃO DE METAIS PESADOS NO AQUÍFERO BAURU EM MATO GROSSO DO SUL E EM SÃO PAULO. In: **RECURSOS HÍDRICOS: GESTÃO, PLANEJAMENTO E TÉCNICAS EM PESQUISA**. Editora Científica Digital, 2021. p. 48-57.

UR REHMAN, Khalil et al. Ecological risk assessment of heavy metals in vegetables irrigated with groundwater and wastewater: the particular case of Sahiwal district in Pakistan. **Agricultural Water Management**, v. 226, p. 105816, 2019.

VIRTANEN, Pauli et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. **Nature methods**, v. 17, n. 3, p. 261-272, 2020.

YADAV, Brijesh et al. Application of analytical hierarchical process, multi-influencing factor, and geospatial techniques for groundwater potential zonation in a semi-arid region of western India. **Journal of Contaminant Hydrology**, v. 253, p. 104122, 2023.

YANG, Yun-Ya et al. Micropollutants in groundwater from septic systems: Transformations, transport mechanisms, and human health risk assessment. **Water research**, v. 123, p. 258-267, 2017.

YING, Xue. An overview of overfitting and its solutions. In: **Journal of physics: Conference series**. IOP Publishing, 2019. p. 022022.

WANG, Shou et al. Hydrochemical evolution characteristics, controlling factors, and high nitrate hazards of shallow groundwater in a typical agricultural area of Nansi Lake Basin, North China. **Environmental Research**, p. 115430, 2023.

WARD, Mary H. et al. Drinking water nitrate and human health: an updated review. **International journal of environmental research and public health**, v. 15, n. 7, p. 1557, 2018.

WÓJTOWICZ, Andrzej et al. A random forest model for the classification of wheat and rye leaf rust symptoms based on pure spectra at leaf scale. **Journal of Photochemistry and Photobiology B: Biology**, v. 223, p. 112278, 2021.

ZHANG, Tong. Solving large scale linear prediction problems using stochastic gradient descent algorithms. In: **Proceedings of the twenty-first international conference on Machine learning**. 2004. p. 116.

ZHANG, Kai; LUO, Minxia. Outlier-robust extreme learning machine for regression problems. **Neurocomputing**, v. 151, p. 1519-1527, 2015.

ZHOU, Xin-Xing et al. Research on remote sensing classification of fruit trees based on Sentinel-2 multi-temporal imageries. **Scientific Reports**, v. 12, n. 1, p. 1-14, 2022.

APÊNDICES

Apêndice A: Informações sobre os processos deferidos de Outorga de Direito de Uso de Água Subterrânea dos poços tubulares da cidade de Campo Grande, MS

Poço	Coordenadas geográficas Latitude (SIRGAS 2000)	Coordenadas geográficas Longitude (SIRGAS 2000)	Declaração de Uso de Recursos Hídricos (DURH)	Processo	Localização: Área rural e Área urbana	Método da análise de água	Finalidade de Uso	Nitrato (mg/L)
1	-20° 28' 52.05"	-54° 45' 15.27"	DURH000348	0000044/2016	Área urbana	SMEWW. 22, 2012	Indústria	0,65
2	-20° 26' 59.59"	-54° 37' 45"	DURH001528	0000087/2016	Área urbana	SMEWW. 21, 2005 e 22, 2012	Consumo Humano	4,9
3	-20° 28' 8.80"	-54° 43' 1.45"	DURH000739	0000106/2016	Área urbana	SMEWW. 22, 2012 e ISSO, 9308	Outras Finalidades de Uso	0,18
4	-20° 28' 52.76"	-54° 45' 2.28"	DURH006233	0000160/2016	Área urbana	SMEWW. 21, 2005 e 22, 2012	Consumo Humano e Outras Finalidades de Uso	5,84
5	-20° 29' 27"	-54° 36' 50"	DURH000996	0000174/2016	Área urbana	SMEWW. 22, 2012. USEPA. CETESB	Outras Finalidades de Uso	<0,10
6	-20° 25' 46.81"	-54° 35' 19.64"	DURH006562	0000195/2016	Área urbana	SMEWW. 22, 2012 e ISSO 9308	Outras Finalidades de Uso	0,07

7	-20° 32' 24"	-54° 35' 38"	DURH006348	0000208/2016	Área urbana	SMEWW. 22, 2012 e ISSO 9308	Outras Finalidades de Uso	0,11 E NI
8	-20° 28' 3.27"	-54° 37' 53.44"	DURH006334	0000212/2016	Área urbana	SMEWW. 22, 2012	Consumo Humano	3,2 E NI
9	-20° 26' 25.63"	-54° 37' 27.81"	DURH004893	0000234/2016	Área urbana	SMEWW. 22, 2012 e ISSO 9308	Outras Finalidades de Uso	6
10	-20° 25' 17.81"	-54° 35' 54.61"	DURH006480	0000241/2016	Área urbana	SMEWW. 22, 2012 e ISSO 9308	Consumo Humano	3
11	-20° 27' 13.66"	-54° 36' 23.11"	DURH001462	0000242/2016	Área urbana	SMEWW. 21, 2005 e 22, 2012	Consumo Humano e Outras Finalidades de Uso	NI E 1,2
12	-20° 31' 17.08"	-54° 37' 20.26"	DURH002573	0000250/2016	Área urbana	SMEWW. 20, 1998; 21, 2005; 22, 2012 e ABNT	Indústria	8
13	-20° 31' 19.21"	-54° 37' 25.80"	DURH002685	0000251/2016	Área urbana	SMEWW. 20, 1998; 21, 2005; 22, 2012 e ABNT	Indústria	8
14	-20° 24' 54.63"	-54° 36' 13.81"	DURH002398	0000329/2016	Área urbana	SMEWW. 22, 2012	Consumo Humano	3,8 E 2,1
15	-20° 29' 23.63"	-54° 44' 47.12"	DURH001608	0000330/2016	Área urbana	SMEWW. 22, 2012 e ISO	Indústria	0,02
16	-20° 29' 21.32	-54° 44' 45.14"	DURH001607	0000331/2016	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano	0,03

17	-20° 28' 59.60"	-54° 38' 29.01"	DURH003499	0000332/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 3,9
18	-20° 29' 1.96"	-54° 38' 28.05"	DURH000715	0000333/2016	Área urbana	SMEWW. 22, 2012 e EPA	Outras Finalidades de Uso	NI E 5,1
19	-20° 27' 42.72"	-54° 33' 4.68"	DURH006801	0000334/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 1,9
20	-20° 29' 5.00"	-54° 44' 58.90"	DURH006349	0000359/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Indústria	0,61
21	-20° 27' 48.39"	-54° 36' 42.15"	DURH006795	0000363/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 6,6
22	20° 33' 26.44"	-54° 35' 15.46"	DURH000224	0000383/2016	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano e	2,5
23	-20° 26' 46.57"	-54° 37' 7.09"	DURH004103	0000387/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	13,67
24	-20° 29' 40.26"	-54° 34' 43.19"	DURH003429	0000390/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	230,8
25	-20° 28' 27.43"	-54° 43' 44.60"	DURH001582	0000395/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Consumo Humano	<0,10
26	-20° 27' 34.63"	-54° 35' 10.81"	DURH002175	0000403/2016	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	1,6 E NI

27	-20° 26' 33.77"	-54° 40' 23.72"	DURH000037	0000416/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,6
28	-20° 26' 12.80"	-54° 37' 41.41"	DURH000246	0000418/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 5,5
29	-20° 27' 36.63"	-54° 33' 34.81"	DURH002783	0000423/2016	Área urbana	SMEWW. 22, 2012 e EPA	Outras Finalidades de Uso	1,05
30	-20° 27' 0.43"	-54° 36' 13.66"	DURH003285	0000447/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 8,4
31	-20° 24' 59.33"	-54° 38' 49.58"	DURH001638	0000463/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 3,7
32	-20° 26' 48.6"	-54° 35' 26.9"	DURH002680	0000472/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,32
33	-20° 27' 38.63"	-54° 38' 7.81"	DURH004202	0000582/2016	Área urbana	SMEWW. 22, 2012 e EPA	Outras Finalidades de Uso	0,4
34	-20° 31' 0.29"	-54° 36' 7.07"	DURH004589	0000587/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	1
35	-20° 27' 25.57"	-54° 37' 41.37"	DURH005565	0000602/2016	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano	3,3
36	-20° 28' 44.53"	-54° 43' 18.22"	DURH002278	0000617/2016	Área urbana	SMEWW. 22, 2012; USEPA	Indústria	0,5

37	-20° 27' 5.22"	-54° 36' 33.91"	DURH007089	0000628/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 9,2
38	-20° 32' 56.02"	-54° 34' 52.13"	DURH004681	0000639/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Indústria	8,5
39	-20° 28' 44.90"	-54° 43' 16.40"	DURH000380	0000671/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Indústria	0,2 E 0,9
40	-20° 25' 59.73"	-54° 35' 36.21"	DURH001448	0000673/2016	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	2
41	-20° 29' 34.19"	-54° 33' 28.98"	DURH001346	0000687/2016	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano	0,03
42	20° 29' 17.34"	-54° 44' 54.00"	DURH005016	0000691/2016	Área urbana	SMEWW. 22, 2012	Indústria	<0,1; 0,2 E 0,5
43	-20° 25' 48.88"	-54° 40' 3"	DURH004259	0000695/2016	Área urbana	SMEWW. 22, 2012	Abastecimento Público	<1
45	-20° 27' 11.23"	-54° 36' 54.16"	DURH000216	0000696/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 0,3
45	-20° 30' 1.22"	-54° 36' 26.12"	DURH007602	0000720/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,16
46	-20° 32' 42.93"	-54° 33' 36.70"	DURH006910	0000726/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Consumo Humano	<0,10 E NI

47	-20° 32' 45.34"	-54° 35' 33.31"	DURH007681	0000743/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	1,22
48	-20° 28' 34.03"	-54° 43' 34.68"	DURH007808	0000781/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Consumo Humano e	0,152
49	20° 26' 48.67"	-54° 35' 39.14"	DURH003534	0000802/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	4,81
50	-20° 32' 8.81"	-54° 33' 43.75"	DURH007865	0000806/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Consumo Humano e	NI E 2,9
51	20° 28' 31.94"	-54° 38' 29.02"	DURH003533	0000808/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 2
52	-20° 29' 52.99"	-54° 38' 7.84"	DURH003537	0000810/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 4,1
53	-20° 29' 58.41"	-54° 38' 11.55"	DURH003752	0000811/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 2,8
54	-20° 28' 48.61"	-54° 38' 26.18"	DURH003536	0000818/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 3,3
55	-20° 27' 58.84"	-54° 35' 25.73"	DURH007892	0000827/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	4,53
56	-20° 28' 39.33"	-54° 35' 39.71"	DURH003885	0000839/2016	Área urbana	SMEWW. 22, 2012. EPA	Outras Finalidades de Uso	0,7

57	-20° 26' 48.91"	-54° 34' 36.47"	DURH007914	0000840/2016	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	1,4
58	-20° 27' 24.44"	-54° 36' 9.94"	DURH007939	0000857/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 5,5
59	-20° 30' 17.16"	-54° 36' 10.22"	DURH002405	0000864/2016	Área urbana	SMEWW. 22, 2012	Indústria	1,2
60	-20° 30' 16.59"	-54° 36' 17.58"	DURH002408	0000865/2016	Área urbana	SMEWW. 22, 2012	Indústria	0,9
61	20° 28' 3.70"	-54° 43' 12.46"	DURH007952	0000871/2016	Área urbana	SMEWW. 22, 2012	Consumo Humano e	0,2
62	-20° 28' 40.37"	-54° 43' 21.69"	DURH006103	0000874/2016	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano e	0,05
63	-20° 28' 39.64"	-54° 34' 29.63"	DURH007967	0000892/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	5
64	-20° 29' 23.69"	-54° 44' 43.51"	DURH000602	0000893/2016	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	NI E 0,07
65	-20° 32' 48.87"	-54° 35' 18.40"	DURH007988	0000900/2016	Área urbana	SMEWW. 22, 2012	Indústria	1,4
66	-20° 30' 49.51"	-54° 37' 12.00"	DURH001367	0000922/2016	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	0,9

67	-20° 27' 53.14	-54° 35' 40.93"	DURH001850	0000943/2016	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	ND E 0,001
68	-20° 28' 31.33"	-54° 36' 44.40"	DURH003535	0000956/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	1,87
69	-20° 24' 4.29"	-54° 36' 56.18"	DURH005375	0000967/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 10
70	-20° 33' 22.69"	-54° 33' 58.41"	DURH005403	0000968/2016	Área urbana	SMEWW. 21, 2005. 22, 2012	Outras Finalidades de Uso	NI E 3,8
71	-20° 28' 35.15"	-54° 36' 50.82"	DURH003883	0000972/2016	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	4,6 E NI
72	-20° 28' 42.65"	-54° 43' 29.69"	DURH007880	0000996/2017	Área urbana	SMEWW. 22, 2012 e ISO	Consumo Humano e	0,04 E NI
73	-20° 28' 27.39"	-54° 43' 38.47"	DURH005666	00001012/2017	Área urbana	SMEWW. 22, 2012	Consumo Humano	0,3
74	-20° 28' 23.94"	-54° 43' 30.64"	DURH000492	0001013/2017	Área urbana	SMEWW. 22, 2012	Consumo Humano	0,3
75	-20° 27' 41.78"	-54° 36' 36.03"	DURH000491	00001056/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	3,6
76	20° 27' 19.80"	-54° 35' 11.64"	DURH007427	00001057/2017	Área urbana	SMEWW. 22, 2012. USEPA.	Consumo Humano	39,94 E NI

77	-20° 24' 33.73"	-54° 34' 22.00"	DURH008467	00001108/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	3,4
78	-20° 27' 55.19"	-54° 35' 6.93"	DURH008611	00001161/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,2 E NI
79	-20° 29' 48.54"	-54° 39' 2.80"	DURH008740	0001184/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	2
80	-20° 29' 26.07"	-54° 44' 40.74"	DURH008856	0001214/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,2
81	-20° 33' 34.68"	-54° 35' 15.02"	DURH008862	0001238/2017	Área urbana	SMEWW. 22, 2012 e ISO	Outras Finalidades de Uso	0,06
82	-20° 27' 17.30"	-54° 36' 59.10"	DURH008913	0001240/2017	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	7,5
83	-20° 27' 15.70"	-54° 36' 58.40"	DURH008914	0001240/2017	Área urbana	SMEWW. 22, 2012	Outras Finalidades de Uso	6
84	-20° 27' 34.25"	-54° 41' 57.41"	DURH001325	0001341/2017	Área urbana	SMEWW. 22, 2012 e ISO	Indústria	0,05
85	-20° 27' 35"	-54° 38' 13"	DURH001026	0001367/2017	Área urbana	SMEWW. 22, 2012, 23, 2017	Abastecimento Público	1,916 ; 3 E 2
86	-20° 32' 16.02"	-54° 35' 35.62"	DURH009183	0001371/2017	Área urbana	SMEWW. 22, 2012. USEPA. CETESB	Outras Finalidades de Uso	140,72 E NI

87	-20° 32' 36.27"	-54° 35' 21.95"	DURH004787	0001395/2017	Área urbana	SMEWW. 23, 2017. ISO	Outras Finalidades de Uso	0,4 E 0,16
88	-20° 31' 31.55"	-54° 39' 35.46"	DURH009285	0001399/2017	Área urbana	SMEWW. 22, 2012. ISO	Outras Finalidades de Uso	0,2
89	-20° 27' 41.43"	-54° 36' 26.39"	DURH009303	0001432/2017	Área urbana	SMEWW. 22, 2012. ISO	Outras Finalidades de Uso	2,2
90	-20° 26' 54"	-54° 38' 17"	DURH000521	0001461/2017	Área urbana	SMEWW. 23, 2017. CETESB	Abastecimento Público	4,77 E 7
91	-20° 32' 0.01"	-54° 33' 49.92"	DURH009592	0001510/2017	Área urbana	SMEWW. 22, 2012. ISO	Outras Finalidades de Uso	0,06
92	-20° 24' 58"	-54° 34' 55"	DURH001523	0001611/2017	Área urbana	SMEWW. 23, 2017. CETESB	Abastecimento Público	< 0,003 E NI
93	-20° 30' 6.82"	-54° 44' 17.06"	DURH009970	0001663/2017	Área urbana	SMEWW. 22, 2012. ISO	Indústria	0,13
94	-20° 28' 21.57"	-54° 43' 51.21"	DURH10111	0001705/2017	Área urbana	SMEWW. 22, 2012. ISO	Outras Finalidades de Uso	0,1
95	-20° 28' 33.60"	-54° 44' 8.27"	DURH005376	0001891/2017	Área urbana	SMEWW. 22, 2012. EPA.	Consumo Humano	0,2 E NI
96	-20° 25' 46.02"	-54° 38' 24.08"	DURH010860	0002050/2017	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	1,2

97	-20° 25' 22.23"	-54° 34' 58.94"	DURH004268	0002523/2018	Área urbana	SMEWW. 22, 2012. USEPA.	Outras Finalidades de Uso	< 0,1
----	-----------------	-----------------	------------	--------------	-------------	----------------------------	------------------------------	-------

Apêndice B: informações sobre as características dos poços tubulares e aquíferos da cidade de Campo Grande, MS

Poço	Profundidade	Diâmetro (Pol.)	Altitude	Litologia - topo para base (m)	Aquíferos conforme perfil do poço (topo a base em metros)	Sistema aquífero aflorante	Sistema aquífero captado	Aquífero: PP e PT	PI, PR E NI	Nível estático (m)
1	140	8	541	0 a 30 SL; 30 a 140 BS	30 a 140 - SASG	SASG	SASG	PP	PR	37,28
2	48	4	569	NI	SASG	SASG	SASG	PP	NI	7,34
3	48	4	516	NI	SASG	SASG	SASG	PP	NI	5
4	43,2	6	541	0 a 35 SL; 35 a 43,20 BS	35 a 43,20 - SASG	SASG	SASG	PP	PI	1,19
5	104	6	561	0 a 22 SL; 22 a 103 BS	22 a 103 - SASG	SASG	SASG	PP	PR	12,35

6	150	6	645	0 a 14 SL; 14 a 150 BS	14 a 150 SASG	SASG	SASG	PP	PR	8
7	40	4	540	0 a 10 SL; 10 a 40 BS	10 a 40 - SASG	SASG	SASG	PP	PR	2,35
8	80	6	554	0 a 18 SL; 18 a 80 BS	18 a 80 - SASG	SASG	SASG	PP	PR	8,6
9	103	4	550	0 a 10 SL; 10 a 103 BS	10 a 103 - SASG	SASG	SASG	PP	PI	3
10	89	6	...	0 a 7 SL; 7 a 89 BS	7 a 89 - SASG	SASG	SASG	PP	PR	5
11	65	6	579	0 a 6 SL; 6 a 65 BS	6 a 65 - SASG	SASG	SASG	PP	PI	5,28
12	59,5	6	545	0 a 24,5 SL; 24,5 a 59,5 BS	24,5 a 59,5 - SASG	SASG	SASG	PP	PR	15
13	66	6	543	0 a 32 SL; 32 a 66 BS	32 a 66 - SASG	SASG	SASG	PP	PR	12
14	80	6	583	0 a 6 SL; 6 a 80 BS	6 a 80 - SASG	SASG	SASG	PP	PI	21
15	17	4	538	0 a 12 SL; 12 a 17 BS	12 a 17 - SASG	SASG	SASG	PP	PI	1,91
16	14	4	539	0 a 12 SL; 12 a 14 BS	12 a 14 - SASG	SASG	SASG	PP	PI	2,12

17	25	4	545	0 a 22 SL; 22 a 25 BS	22 a 25 - SASG	SASG	SASG	PP	PI	3,56
18	26	4	544	0 a 22 SL; 22 a 26 BS	22 a 26 - SASG	SASG	SASG	PP	PI	5,65
19	56	6	646	0 a 56 AC	0 a 56 - SAB	SAB	SAB	PP	PI	6,09
20	100	6	540	0 a 18 SL; 18 a 100 BS	18 a 100 - SASG	SASG	SASG	PP	PR	15,23
21	145	6	552	0 a 7 SL; 7 a 145 BS	7 a 145 - SASG	SASG	SASG	PP	PI	2,65
22	150	6	559	0 a 42 SL; 42 a 150 BS	42 a 150 - SASG	SASG	SASG	PP	PI	26,4
23	102	6	558	0 a 14 SL; 14 a 102 BS	14 a 102 - SASG	SASG	SASG	PP	PR	8
24	100	6	558	0 a 20 SL; 20 a 100 BS	20 a 100 - SASG	SASG	SASG	PP	PI	8
25	126	6	532	0 a 40 SL; 40 a 126 BS	40 a 126 - SASG	SASG	SASG	PP	PR	6
26	40	6	514	0 a 12 SL; 12 a 40 BS	12 a 40 - SASG	SASG	SASG	PP	PI	3
27	90	6	554	0 a 29 SL; 29 a 90 BS	29 a 90 - SASG	SASG	SASG	PP	PI	5,91

28	100	6	579	0 a 4 SL; 4 a 100 BS	4 a 100 - SASG	SASG	SASG	PP	PI	8,8
29	33	4	636	0 a 12 SL; 12 a 33 BS	12 a 33 - SASG	SASG	SASG	PP	PI	3
30	40	4	593	0 a 4 SL; 4 a 40 BS	4 a 40 - SASG	SASG	SASG	PP	PI	4,4
31	117	6	608	0 a 4 SL; 4 a 117 BS	4 a 117 - SASG	SASG	SASG	PP	PI	44
32	115	6	625	NI	SASG	SASG	SASG	SI	NI	2,45
33	60	4	545	0 a 12 SL; 12 a 60 SG	12 a 60 - SASG	SASG	SASG	PP	PI	3
34	50	4	547	0 a 12 SL; 12 a 50 SG	12 a 50 - SASG	SASG	SASG	PP	PI	1,97
35	299	6	570	0 a 4 SL; 4 a 299 BS	4 a 299 - SASG	SASG	SASG	PP	PI	18
36	84	6	522	0 a 84 BS	0 a 84 - SASG	SASG	SASG	PP	PI	3,84
37	110	6	600	0 a 7 SL; 7 a 110 BS	7 a 110 - SASG	SASG	SASG	PP	PI	3,43
38	180	6	571	0 a 48 SL; 48 a 180 BS	48 a 180 SASG	SASG	SASG	PP	PR	5

39	30	6	630	0 a 10 SL; 10 a 30 BS	10 a 30 SASG	SASG	SASG	PP	PI	3
40	90	6	436	0 a 6 SL; 6 a 90 BS	6 a 90 SASG	SASG	SASG	PP	PI	4
41	120	6	620	0 a 18 SL; 18 a 120 BS	18 a 120 SASG	SASG	SASG	PP	PI	5
42	80	6	541	0 a 24 SL; 24 a 80 BS	24 a 50 - SASG	SASG	SASG	PP	PR	24,06
43	217	10	564	0 a 24 SL; 24 a 134 BS; 134 a 217 ABT	134 a 217 - SAG	SASG	SASG	PT; PP	PR	69
44	42,05	4	566	0 a 12 SL; 12 a 42,05 BS	12 a 42,05 - SASG	SASG	SASG	PP	PI	4,55
45	20	4	558	0 a 10 SL; 10 a 20 BS	10 a 20 - SASG	SASG	SASG	PP	PI	1,94
46	96	6	592	0 a 20 SL; 20 a 96 BS	20 a 96 - SASG	SASG	SASG	PP	PR	6,02
47	130	6	552	0 a 30 SL; 30 a 130 BS	30 A 130 - SASG	SASG	SASG	PP	PI	5,2
48	38	4	530	0 a 4 SL; 4 a 36 A argiloso; 36 a 38 BS;	4 a 38 - SASG	SASG	SASG	PP	PR	4,23
49	140	6	625	0 a 12 SL; 12 a 140 BS	12 a 140 - SASG	SASG	SASG	PP	PI	2,55

50	40	4	582	0 a 6 SL; 6 a 40 BS	6 a 40 - SASG	SASG	SASG	PP	PI	2,76
51	45	6	554	0 a 20 SL; 20 a 36 Argila; 36 a 45 BS	36 a 45 - SASG	SASG	SASG	PP	PI	5,53
52	40	6	523	0 a 36 SL; 36 a 40 BS	36 a 45 - SASG	SASG	SASG	PP	PI	7,1
53	120	6	522	0 a 38 SL; 38 a 120 BS	38 a 120 - SASG	SASG	SASG	PP	PI	6,15
54	37,2	4	548	0 a 33 SL; 33 a 37,2 BS	33 a 37,2 - SASG	SASG	SASG	PP	PI	1,8
55	40	4	577	0 a 12 SL; 12 a 40 BS	12 a 40 - SASG	SASG	SASG	PP	PI	5,33
56	55	6	528	0 a 15 SL; 15 a 80 BS	15 a 80 - SASG	SASG	SASG	PP	PI	2,84
57	70	6	608	0 a 15 SL; 15 a 70 BS	15 a 70 - SASG	SASG	SASG	PP	PI	2,82
58	206	6	588	0 a 18 SL; 18 a 200 BS	18 a 200 - SASG	SASG	SASG	PP	PR	8,8
59	85	6	NI	0 a 14 SL; 14 a 85 BS	14 a 85 - SASG	SASG	SASG	PP	PI	40

60	92	6	555	0 a 12 SL; 12 a 92 BS	12 a 92 - SASG	SASG	SASG	PP	PI	75
61	130	6	520	0 a 20 SL; 20 a 130 BS	20 a 130 - SASG	SASG	SASG	PP	PR	8,2
62	60	6	525	0 a 22 SL; 22 a 60 BS	22 a 60 - SASG	SASG	SASG	PP	PI	6,4
63	120	8	624	0 a 30 SL; 30 a 120 BS	30 a 120 - SASG	SASG	SASG	PP	PI	5,2
64	130	6	539	0 a 30 SL; 30 a 130 BS	30 A 130 - SASG	SASG	SASG	PP	PR	10
65	69	6	559	0 a 12 SL; 12 a 69 BS	12 a 69 - SASG	SASG	SASG	PP	PI	6,09
66	100	6	553	0 a 8,15 SL; 8,15 a 100 BS	8,15 a 100 - SASG	SASG	SASG	PP	PI	7,61
67	102	6	549	0 a 10 SL; 10 a 102 BS	10 a 102 - SASG	SASG	SASG	PP	PI	28,4
68	100	6	560	0 a 15 SL; 15 a 100 BS	15 a 100 - SASG	SASG	SASG	PP	PI	10,15
69	120	6	643	0 a 7 SL; 7 a 120 BS	15 a 120 - SASG	SASG	SASG	PP	PI	9

70	110	6	596	0 a 12 SL; 12 a 110 BS	12 a 110 - SASG	SASG	SASG	PP	PI	15,23
71	35	6	555	0 a 12 SL; 12 a 35 BS	12 a 35 - SASG	SASG	SASG	PP	PI	2,73
72	50	6	527	0 a 24 SL; 24 a 50 BS	24 a 50 - SASG	SASG	SASG	PP	PI	10,11
73	130	6	530	0 a 40 SL; 40 a 130 BS	40 a 130 - SASG	SASG	SASG	PT; PP	PR	8,52
74	102	6	527	0 a 36 SL; 36 a 102 BS	36 a 102 - SASG	SASG	SASG	PT; PP	PR	20,2
75	120	6	564	0 a 7 SL; 7 a 120 BS	7 a 120 SASG	SASG	SASG	PP	PI	12
76	240	6	587	0 a 6 AC; 6 a 240 BS	6 a 240 - SASG	SASG	SASG	PP	PR	14,59
77	36	4	667	0 a 34 AC; 34 a 36 BS	0 a 36 - SAB	SAB	SAB	PP	PR	3
78	80	6	600	0 a 12 AC; 12 a 80 BS	12 a 80 - SASG	SASG	SASG	PP	PR	5,2
79	100	8	540	0 a 10 SL; 10 a 100 BS	10 a 100 - SASG	SASG	SASG	PP	PI	5,13
80	22,12	4	539	0 a 22 SL	0 a 22 - SASG	SASG	SASG	PP	PI	2,64

81	100	6	560	0 a 40 SL; 40 a 100 BS	40 a 100 - SASG	SASG	SASG	PT; PP	PI	9,47
82	120	6	554	0 a 24 SL; 24 a 120 BS	24 a 120 - SASG	SASG	SASG	PP	PR	50,1
83	130	6	555	0 a 26 SL; 26 a 130 BS	26 a 130 SASG	SASG	SASG	PP	PR	55
84	200	6	535	0 a 20 SL; 20 a 200 BS	0 a 20 SL; 20 a 200 BS	SASG	SASG	PP	PI	40
85	100	6	570	0 A 9 SL; 9 a 100 BS	9 a 100 SASG	SASG	SASG	PP	PR	40
86	102	6	534	0 a 24 SL; 24 a 102 BS	24 a 102 - SASG	SASG	SASG	PP	PI	5,2
87	96,5	6	552	0 a 36 SL; 36 a 96,5 BS	36 a 96,5 SASG	SASG	SASG	PP	PR	12
88	10	4	511	0 a 10 SL	0 a 10 SASG	SASG	SASG	PP	PI	0,56
89	102	6	574	0 a 2 SL; 2 a 102 SG	2 a 102 - SASG	SASG	SASG	PP	PI	23,3
90	99	6	582	0 a 8 SL; 8 a 99 BS	8 a 99 SASG	SASG	SASG	PP	PR	25

91	110	6	573	0 a 8 SL; 8 a 110 BS	8 a 110 SASG	SASG	SASG	PP	PI	5,35
92	67	8	657	0 a 15 SL; 15 a 67 SG	15 a 67 - SASG	SASG	SASG	PP	PR	14
93	123	6	515	0 a 18 SL; 18 a 123 SG	18 a 123 SASG	SASG	SASG	PP	PI	11,8
94	120	6	533	0 a 38 SL; 38 a 120 BS	38 a 120 - SASG	SASG	SASG	PP	PI	7,5
95	122	6	537	0 a 18 SL; 18 a 122 BS	18 a 122 SASG	SASG	SASG	PP	PR	30,6
96	80	6	621	0 a 80 BS	0 a 80 - SASG	SASG	SASG	PP	PI	2,75
97	76	6	NI	0 a 20 SL; 20 a 76 BS	20 a 76 SASG	SASG	SASG	PP	PI	4,6

APÊNDICE C: Código para aplicar os métodos de Machine Learning

```

1  # -*- coding: utf-8                                     -*-
2  """
3  Created on Fri Feb  3 15:49:56 2023
4
5  @author: Not
6  """
7
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.ensemble as ensemble
13 from sklearn import metrics
14 from sklearn import svm
15 from sklearn import tree
16 from sklearn import neural_network
17 from sklearn.feature_selection import RFE
18 from sklearn import svm
19 from sklearn import neighbors
20 from sklearn import linear_model
21 from sklearn.neighbors import KNeighborsClassifier
22 from sklearn.neighbors import NearestNeighbors
23 from sklearn.semi_supervised import SelfTrainingClassifier
24 import matplotlib.pyplot as plt
25 import seaborn as sns
26
27 #importar os dados
28 amostra = pd.read_excel('informacoes5.xlsx').dropna()
29
30 #numero aleatorio das simulacoes
31 rs = 1#np.random.random_integers(0,1000)
32
33 #tratamento preliminar dos dados
34 # amostra.NITRATO = np.where(amostra.NITRATO > amostra.NITRATO.quantile(0.75), 4,
35 amostra.NITRATO)
36 # amostra.NITRATO = np.where((amostra.NITRATO <= amostra.NITRATO.quantile(0.75)) &
37 (amostra.NITRATO > amostra.NITRATO.quantile(0.5)), 3, amostra.NITRATO)
38 # amostra.NITRATO = np.where((amostra.NITRATO <= amostra.NITRATO.quantile(0.5)) &
39 (amostra.NITRATO > amostra.NITRATO.quantile(0.25)), 2, amostra.NITRATO)
40 # amostra.NITRATO = np.where(amostra.NITRATO < amostra.NITRATO.quantile(0.25), 1,
41 amostra.NITRATO)
42 amostra.NITRATO = np.where(amostra.NITRATO > amostra.NITRATO.quantile(0.80), np.nan,
43 amostra.NITRATO)
44 # amostra.DECLI = np.where(amostra.DECLI > amostra.DECLI.quantile(0.95), np.nan,
45 amostra.DECLI)
46 # amostra.HIPSO = np.where(amostra.HIPSO > amostra.HIPSO.quantile(0.95), np.nan,
47 amostra.HIPSO)
48 # amostra.NDVI = np.where(amostra.NDVI > amostra.NDVI.quantile(0.95), np.nan,
49 amostra.NDVI)
50 # amostra.AGUA = np.where(amostra.AGUA > amostra.AGUA.quantile(0.95), np.nan,
51 amostra.AGUA)
52 # amostra.HIDRO = np.where(amostra.HIDRO > amostra.HIDRO.quantile(0.95), np.nan,
53 amostra.HIDRO)
54 # amostra.ESGOTO = np.where(amostra.ESGOTO > amostra.ESGOTO.quantile(0.95), np.nan,
55 amostra.ESGOTO)
56 amostra = amostra.dropna()
57
58 y = amostra.NITRATO
59
60 X = amostra[['GEOLOGIA', 'POP_BAIRRO', 'ZONA', 'DECLI', 'HIPSO', 'NDVI', 'AGUA',
61 'HIDRO', 'ESGOTO', 'profundidade_NA', 'fossas', 'ISO']]
62 #X = amostra[['DECLI', 'HIPSO', 'NDVI', 'AGUA', 'HIDRO', 'ESGOTO', 'profundidade_NA',

```

```

'fossas']]
51
52 #separacao dos datasets em sub-datasets de treinamento e teste
53 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state
=(0))

54
55 #selecao e transformacao dos dados
56 min_max_scaler = preprocessing.MinMaxScaler().fit(X_train)
57 X = min_max_scaler.fit_transform(X)
58

59 X_train = min_max_scaler.fit_transform(X_train)
60 X_test = min_max_scaler.fit_transform(X_test)
61 X2 = X
62 X2_train = X_train
63 X2_test = X_test
64 y2_train = y_train.astype(int)
65 y2_test = y_test.astype(int)
66 y2 = y.astype(int)
67
68 #mapa treinamento e teste
69
70 Xmapa = amostra[['x','y','GEOLOGIA', 'POP_BAIRRO', 'ZONA', 'DECLI', 'HIPSO', 'NDVI',
'AGUA', 'HIDRO', 'ESGOTO', 'profundidade_NA', 'fossas','ISO']]
71 #X = amostra[['DECLI', 'HIPSO', 'NDVI', 'AGUA', 'HIDRO', 'ESGOTO', 'profundidade_NA',
'fossas']]
72
73 #separacao dos datasets em sub-datasets de treinamento e teste
74 X_train_mapa, X_test_mapa, y_train_mapa, y_test_mapa = train_test_split(Xmapa, y,
test_size=0.3, random_state=(0))
75
76 X_train_mapa.to_excel('treinamento.xlsx')
77 X_test_mapa.to_excel('teste.xlsx')
78
79
80 #analise de importancia
81
82 n_features =12
83
84 estimator = svm.SVR(kernel= 'linear')
85 selector = RFE(estimator, n_features_to_select = n_features, step=1)
86 selector = selector.fit(X, y)
87 print("regressor")
88 print(selector.support_)
89 print(selector.ranking_)
90 X = X[:,selector.support_]
91 X_train = X_train[:,selector.support_]
92 X_test = X_test[:,selector.support_]
93
94
95 estimator = svm.SVC(kernel= 'linear')
96 selector = RFE(estimator, n_features_to_select = n_features, step=1)
97 selector = selector.fit(X2, y2)
98 print("classifier")
99 print(selector.support_)
100 print(selector.ranking_)
101 X2 = X2[:,selector.support_]
102 X2_train = X2_train[:,selector.support_]
103 X2_test = X2_test[:,selector.support_]
104
105
106
107 #criacao e treinamento dos modelos
108 modl = ensemble.RandomForestRegressor(random_state=(rs)).fit(X_train, y_train)
109
110 #analise da acuracia - desempenho
111

```

```

112 y_trained_mod1 = mod1.predict(X_train)
113 y_tested_mod1 = mod1.predict(X_test)
114 y_all_mod1 = mod1.predict(X)
115
116 print('MOD1')
117 accu_mod1_train = metrics.r2_score(y_train, y_trained_mod1)
118 print('Train')
119 print(accu_mod1_train)
120
121 accu_mod1_test = metrics.r2_score(y_test, y_tested_mod1)
122 print('Test')
123 print(accu_mod1_test)
124
125 accu_mod1_all = metrics.r2_score(y, y_all_mod1)
126 print('All')
127 print(accu_mod1_all)

128 print('')
129
130 #####
131 mod2 = svm.SVR().fit(X_train, y_train)
132
133 #analise da acuracia
134 y_trained_mod2 = mod2.predict(X_train)
135 y_tested_mod2 = mod2.predict(X_test)
136 y_all_mod2 = mod2.predict(X)
137
138 print('MOD2')
139 accu_mod2_train = metrics.r2_score(y_train, y_trained_mod2)
140 print('Train')
141 print(accu_mod2_train)
142
143 accu_mod2_test = metrics.r2_score(y_test, y_tested_mod2)
144 print('Test')
145 print(accu_mod2_test)
146
147 accu_mod2_all = metrics.r2_score(y, y_all_mod2)
148 print('All')
149 print(accu_mod2_all)
150 print('')
151
152 #####
153 mod3 = tree.DecisionTreeRegressor(random_state=(rs)).fit(X_train, y_train)
154
155 #analise da acuracia
156 y_trained_mod3 = mod3.predict(X_train)
157 y_tested_mod3 = mod3.predict(X_test)
158 y_all_mod3 = mod3.predict(X)
159
160 print('MOD3')
161 accu_mod3_train = metrics.r2_score(y_train, y_trained_mod3)
162 print('Train')
163 print(accu_mod3_train)
164
165 accu_mod3_test = metrics.r2_score(y_test, y_tested_mod3)
166 print('Test')
167 print(accu_mod3_test)
168
169 accu_mod3_all = metrics.r2_score(y, y_all_mod3)
170 print('All')
171 print(accu_mod3_all)
172 print('')
173
174 #####
175 mod4 = neural_network.MLPRegressor(random_state=(rs)).fit(X_train, y_train)
176
177 #analise da acuracia
178 y_trained_mod4 = mod4.predict(X_train)

```

```

179 y_tested_mod4 = mod4.predict(X_test)
180 y_all_mod4 = mod4.predict(X)
181
182 print('MOD4')
183 accu_mod4_train = metrics.r2_score(y_train, y_trained_mod4)
184 print('Train')
185 print(accu_mod4_train)
186
187 accu_mod4_test = metrics.r2_score(y_test, y_tested_mod4)
188 print('Test')
189 print(accu_mod4_test)
190
191 accu_mod4_all = metrics.r2_score(y, y_all_mod4)
192 print('All')
193 print(accu_mod4_all)
194 print('')
195
196 #####

197 mod5 = ensemble.AdaBoostRegressor(random_state=(rs)).fit(X_train, y_train)
198
199 #analise da acuracia

200 y_trained_mod5 = mod5.predict(X_train)
201 y_tested_mod5 = mod5.predict(X_test)
202 y_all_mod5 = mod5.predict(X)
203
204 print('MOD5')
205 accu_mod5_train = metrics.r2_score(y_train, y_trained_mod5)
206 print('Train')
207 print(accu_mod5_train)
208
209 accu_mod5_test = metrics.r2_score(y_test, y_tested_mod5)
210 print('Test')
211 print(accu_mod5_test)
212
213 accu_mod5_all = metrics.r2_score(y, y_all_mod5)
214 print('All')
215 print(accu_mod5_all)
216 print('')
217
218 #####

219 mod6 = ensemble.GradientBoostingRegressor(random_state=(rs)).fit(X_train, y_train)
220
221 #analise da acuracia
222 y_trained_mod6 = mod6.predict(X_train)
223 y_tested_mod6 = mod6.predict(X_test)
224 y_all_mod6 = mod6.predict(X)
225
226 print('MOD6')
227 accu_mod6_train = metrics.r2_score(y_train, y_trained_mod6)
228 print('Train')
229 print(accu_mod6_train)
230
231 accu_mod6_test = metrics.r2_score(y_test, y_tested_mod6)
232 print('Test')
233 print(accu_mod6_test)
234
235 accu_mod6_all = metrics.r2_score(y, y_all_mod6)
236 print('All')
237 print(accu_mod6_all)
238 print('')
239
240 #####

241 mod7 = svm.SVC().fit(X2_train, y2_train)
242
243 #analise da acuracia

```

```

244 y2_trained_mod7 = mod7.predict(X2_train)
245 y2_tested_mod7 = mod7.predict(X2_test)
246 y2_all_mod7 = mod7.predict(X2)
247
248 print('MOD7')
249 accu_mod7_train = metrics.accuracy_score(y2_train, y2_trained_mod7)
250 print('Train')
251 print(accu_mod7_train)
252
253 accu_mod7_test = metrics.accuracy_score(y2_test, y2_tested_mod7)
254 print('Test')
255 print(accu_mod7_test)
256
257 accu_mod7_all = metrics.accuracy_score(y2, y2_all_mod7)
258 print('All')
259 print(accu_mod7_all)
260 print('')
261
262 #####
263 mod8 = ensemble.RandomForestClassifier(random_state=(rs)).fit(X2_train, y2_train)
264
265 #analise da acuracia
266 y2_trained_mod8 = mod8.predict(X2_train)
267 y2_tested_mod8 = mod8.predict(X2_test)
268 y2_all_mod8 = mod8.predict(X2)
269
270 print('MOD8')
271 accu_mod8_train = metrics.accuracy_score(y2_train, y2_trained_mod8)
272
273 print('Train')
274 print(accu_mod8_train)
275
276 accu_mod8_test = metrics.accuracy_score(y2_test, y2_tested_mod8)
277 print('Test')
278 print(accu_mod8_test)
279
280 accu_mod8_all = metrics.accuracy_score(y2, y2_all_mod8)
281 print('All')
282 print(accu_mod8_all)
283 print('')
284
285 #####
286 mod9 = neural_network.MLPClassifier(activation='relu', random_state=(rs)).fit(
X2_train, y2_train)
287
288 #analise da acuracia
289 y2_trained_mod9 = mod9.predict(X2_train)
290 y2_tested_mod9 = mod9.predict(X2_test)
291 y2_all_mod9 = mod9.predict(X2)
292
293 print('MOD9')
294 accu_mod9_train = metrics.accuracy_score(y2_train, y2_trained_mod9)
295 print('Train')
296 print(accu_mod9_train)
297
298 accu_mod9_test = metrics.accuracy_score(y2_test, y2_tested_mod9)
299 print('Test')
300 print(accu_mod9_test)
301 accu_mod9_all = metrics.accuracy_score(y2, y2_all_mod9)
302 print('All')
303 print(accu_mod9_all)
304 print('')
305
306 #####
307 mod10 = ensemble.GradientBoostingClassifier(random_state=(rs)).fit(X2_train, y2_train)
308
309 #analise da acuracia
310 y2_trained_mod10 = mod10.predict(X2_train)

```

```

310 y2_tested_mod10 = mod10.predict(X2_test)
311 y2_all_mod10 = mod10.predict(X2)
312
313 print('MOD10')
314 accu_mod10_train = metrics.accuracy_score(y2_train, y2_trained_mod10)
315 print('Train')
316 print(accu_mod10_train)
317
318 accu_mod10_test = metrics.accuracy_score(y2_test, y2_tested_mod10)
319 print('Test')
320 print(accu_mod10_test)
321 accu_mod10_all = metrics.accuracy_score(y2, y2_all_mod10)
322 print('All')
323 print(accu_mod10_all)
324 print('')
325
326 #####
327 mod11 = ensemble.ExtraTreesClassifier(random_state=(rs)).fit(X2_train, y2_train)
328
329 #analise da acuracia
330 y2_trained_mod11 = mod11.predict(X2_train)
331 y2_tested_mod11 = mod11.predict(X2_test)
332 y2_all_mod11 = mod11.predict(X2)
333
334 print('MOD11')
335 accu_mod11_train = metrics.accuracy_score(y2_train, y2_trained_mod11)
336 print('Train')
337 print(accu_mod11_train)
338
339 accu_mod11_test = metrics.accuracy_score(y2_test, y2_tested_mod11)
340 print('Test')
341 print(accu_mod11_test)
342 accu_mod11_all = metrics.accuracy_score(y2, y2_all_mod11)
343 print('All')
344 print(accu_mod11_all)
345 print('')
346
347 #####
348 mod12 = KNeighborsClassifier(n_neighbors=1).fit(X2_train, y2_train)
349
350 #analise da acuracia
351 y2_trained_mod12 = mod12.predict(X2_train)
352 y2_tested_mod12 = mod12.predict(X2_test)
353 y2_all_mod12 = mod12.predict(X2)
354
355 print('MOD12')
356 accu_mod12_train = metrics.accuracy_score(y2_train, y2_trained_mod12)
357 print('Train')
358 print(accu_mod12_train)
359
360 accu_mod12_test = metrics.accuracy_score(y2_test, y2_tested_mod12)
361 print('Test')
362 print(accu_mod12_test)
363 accu_mod12_all = metrics.accuracy_score(y2, y2_all_mod12)
364 print('All')
365 print(accu_mod12_all)
366 print('')
367
368 #####
369 mod14 = SelfTrainingClassifier(base_estimator=estimator).fit(X2_train, y2_train)
370
371 #analise da acuracia
372 y2_trained_mod14 = mod14.predict(X2_train)
373 y2_tested_mod14 = mod14.predict(X2_test)
374 y2_all_mod14 = mod14.predict(X2)
375
376 print('MOD14')
377 accu_mod14_train = metrics.accuracy_score(y2_train, y2_trained_mod14)

```

```

378 print('Train')
379 print(accur_mod14_train)
380
381 accur_mod14_test = metrics.accuracy_score(y2_test, y2_tested_mod14)
382 print('Test')
383 print(accur_mod14_test)
384 accur_mod14_all = metrics.accuracy_score(y2, y2_all_mod14)
385 print('All')
386 print(accur_mod14_all)
387 print('')
388
389
390 #####
391 mod15 = linear_model.SGDClassifier().fit(X2_train, y2_train)
392
393 #analise da acuracia
394 y2_trained_mod15 = mod15.predict(X2_train)
395 y2_tested_mod15 = mod15.predict(X2_test)
396 y2_all_mod15 = mod15.predict(X2)
397
398 print('MOD15')
399 accur_mod15_train = metrics.accuracy_score(y2_train, y2_trained_mod15)
400 print('Train')
401 print(accur_mod15_train)
402
403 accur_mod15_test = metrics.accuracy_score(y2_test, y2_tested_mod15)
404 print('Test')
405 print(accur_mod15_test)
406 accur_mod15_all = metrics.accuracy_score(y2, y2_all_mod15)
407 print('All')
408 print(accur_mod15_all)
409 print('')
410
411
412 #####
413 mod16 = linear_model.LogisticRegression().fit(X2_train, y2_train) 414
415 #analise da acuracia
416 y2_trained_mod16 = mod16.predict(X2_train)
417 y2_tested_mod16 = mod16.predict(X2_test)
418 y2_all_mod16 = mod16.predict(X2)
419
420 print('mod16')
421 accur_mod16_train = metrics.accuracy_score(y2_train, y2_trained_mod16)
422 print('Train')
423 print(accur_mod16_train)
424
425 accur_mod16_test = metrics.accuracy_score(y2_test, y2_tested_mod16)
426 print('Test')
427 print(accur_mod16_test)
428 accur_mod16_all = metrics.accuracy_score(y2, y2_all_mod16)
429 print('All')
430 print(accur_mod16_all)
431 print('')
432
433 #####
434 mod17 = linear_model.RidgeClassifier().fit(X2_train, y2_train)
435
436 #analise da acuracia
437 y2_trained_mod17 = mod17.predict(X2_train)
438 y2_tested_mod17 = mod17.predict(X2_test)
439 y2_all_mod17 = mod17.predict(X2)
440
441 print('mod17')
442 accur_mod17_train = metrics.accuracy_score(y2_train, y2_trained_mod17)
443 print('Train')
444 print(accur_mod17_train)
445

```



```

446 accu_mod17_test = metrics.accuracy_score(y2_test, y2_tested_mod17)
447 print('Test')
448 print(accu_mod17_test)
449 accu_mod17_all = metrics.accuracy_score(y2, y2_all_mod17)
450 print('All')
451 print(accu_mod17_all)
452 print('')
453
454
455 %%
456 #plot dos fatores
457 simulacao = pd.read_csv('simulacao5.csv', sep=';').dropna()
458 cgr = pd.read_excel('cgr.xlsx')
459
460 # fig1, axs = plt.subplots(4,3, figsize=(18,15), sharex=True, sharey=True)
461 # (ax1, ax2, ax3), (ax4, ax5, ax6), (ax7, ax8, ax9), (ax10, ax11, ax12) = axs
462
463 # sns.heatmap(simulacao.pivot('Y','X','NIVEL11'), ax=ax1, cmap='Reds')
464 # ax1.invert_yaxis()
465 # ax1.set_title('Water Table')
466 # sns.heatmap(simulacao.pivot('Y','X','GEOLOGIA11'), ax=ax2, cmap='Reds')
467 # ax2.invert_yaxis()
468 # ax2.set_title('Geology')
469 # sns.heatmap(simulacao.pivot('Y','X','POP_BAIRRO'), ax=ax3, cmap='Reds')
470 # ax3.invert_yaxis()
471 # ax3.set_title('Population')
472 # sns.heatmap(simulacao.pivot('Y','X','ZONA1'), ax=ax4, cmap='Reds')
473 # ax4.invert_yaxis()
474 # ax4.set_title('Geotechnical class')
475 # sns.heatmap(simulacao.pivot('Y','X','DECLI1'), ax=ax5, cmap='Reds')
476 # ax5.invert_yaxis()
477 # ax5.set_title('Slope gradient')
478 # sns.heatmap(simulacao.pivot('Y','X','HIPS01'), ax=ax6, cmap='Reds')
479 # ax6.invert_yaxis()
480 # ax6.set_title('Elevation')
481 # sns.heatmap(simulacao.pivot('Y','X','NDVI1'), ax=ax7, cmap='Reds')
482 # ax7.invert_yaxis()
483 # ax7.set_title('NDVI')
484 # sns.heatmap(simulacao.pivot('Y','X','AGUA1'), ax=ax8, cmap='Reds')
485 # ax8.invert_yaxis()
486 # ax8.set_title('Distance to water pipe')
530

```