



Mercator - Revista de Geografia da UFC  
ISSN:  
ISSN: 1984-2201  
edantas@ufc.br  
Universidade Federal do Ceará  
Brasil

# FRAGILIDADE AMBIENTAL POR ALGORITMOS DE MACHINE LEARNING

**Marcelo Pereira de Souza, Cristão**

**Augusto Pereira Silva, Lucas**

**Vieira Veloso, Gustavo**

**Esdras Leite, Estruturas**

**Inácio Fernandes Filho, Elpídio**

FRAGILIDADE AMBIENTAL POR ALGORITMOS DE MACHINE LEARNING

Mercator - Revista de Geografia da UFC, vol. 21, núm. 1, pp. 1-17, 2022

Universidade Federal do Ceará

Disponível em: <https://www.redalyc.org/articulo.oa?id=273674020019>

# FRAGILIDADE AMBIENTAL POR ALGORITMOS DE MACHINE LEARNING

*Cristão Marcelo Pereira de Souza*

*cmppsgeografia@gmail.com*

*Lucas Augusto Pereira Silva*

*mppsgeografia@gmail.com*

*Gustavo Vieira Veloso*

*mppsgeografia@gmail.com*

*Estruturas Esdras Leite*

*cmppsgeografia@gmail.com*

*Elpídio Inácio Fernandes Filho*

*mppsgeografia@gmail.com*



Acceso abierto diamante

## Abstract

O avanço dos modelos preditivos por Algoritmos de Machine Learning (ML) associados aos dados ambientais possibilita o aprimoramento dos modelos de fragilidade ambiental, que são ferramentas essenciais para a tomada de decisão. Este estudo teve como objetivo derivar uma predição de fragilidade ambiental por meio do teste de ML associado a covariáveis ambientais no estado de Minas Gerais. Utilizamos variáveis físico-ambientais (solo, geologia, clima, relevo) com peso de fragilidade para os atributos e cálculo da média para obtenção de um modelo de Fragilidade Ambiental Potencial (PFE). Posteriormente, extraímos os valores do PFE para uma grade de 4.800 pontos, que foi utilizada para gerar uma nova predição por ML denominada PEFML. Essa previsão foi baseada no teste de cinco algoritmos e um conjunto de 105 covariáveis ambientais. Os resultados indicaram que a predição do PEFML com melhor desempenho foi o modelo Random Forest ( $R^2$  0,59 e RMSE 0,47), indicando um predomínio do baixo nível de fragilidade ambiental. Os modelos PFE e PEFML apresentam fortes correlações (0,7 Pearson); no entanto, o PEFML tem correlações mais fortes com outros dados ambientais. Portanto, a predição do PEFML é um modelo robusto que captura informações de covariáveis e apresenta padrões espaciais coerentes.

**Keywords:** Modelo de Fragilidade Ambiental, Previsão Espacial, Floresta aleatória, Planejamento Ambiental.

## Resumo

O avanço de modelos preditivos por Algoritmos de Aprendizado de Máquina (ML) associados à dados ambientais possibilita o aprimoramento de modelos de fragilidade ambiental, os quais são importantes ferramentas para tomada de decisão. O objetivo do estudo foi derivar uma predição de fragilidade ambiental, testando ML associado a covariáveis ambientais no estado de Minas Gerais. Utilizamos diversas variáveis físico-ambientais (solo, geologia, clima, relé) com peso de fragilidade para os atributos e cálculo da média obtida do modelo de Fragilidade Ambiental Potencial (PFE). Posteriormente, extraímos os valores de PFE para uma nota de 4.800 pontos e utilizamos para gerar uma nova predição por ML, denominada PEFML. A predição foi com teste de cinco algoritmos e conjunto de 105 covariáveis ambientais. Comparamos os dois modelos de fragilidade ambiental (PEF e PEFML), inclusive com outros dados de riscos/vulnerabilidade/fragilidade. Os resultados indicam que a predição do PEFML de melhor desempenho foi o modelo Random Forest ( $R^2$  0,59 e RMSE 0,47), indicando predomínio do baixo nível de fragilidade. Os modelos de fragilidade PEF e PEFML têm forte correlação (0,7 Pearson), porém, PEFML possui correlações mais fortes com outros dados ambientais. Portanto, a predição de PEFML é um modelo robusto que capta informações de covariáveis e possui padrões espaciais coerentes

**Palavras-chave:** Modelos de fragilidade ambiental, Predição espacial, Floresta aleatória, Planejamento Ambiental.

## Resumen

O avanço dos modelos preditivos por meio de Algoritmos de Machine Learning (ML) associados a dados ambientais possibilita aprimorar os modelos de fragilidade ambiental, ferramentas fundamentais para a tomada de decisão. O objetivo deste estudo foi derivar uma predição de fragilidade ambiental por meio do teste ML associado a covariáveis ambientais no estado de Minas Gerais. Variáveis físico-ambientais (solo, geologia, clima, relevo) com peso de fragilidade para os atributos e cálculo da média foram utilizadas para obtenção de um modelo de Fragilidade Ambiental Potencial (PFE). Posteriormente, extraímos os valores do PFE para uma grade de 4800 pontos, que foi utilizada para gerar uma nova predição de ML, denominada PEFML. Essa predição foi baseada no teste de cinco algoritmos e um conjunto de 105 covariáveis ambientais. Os resultados indicaram que a predição do PEFML com melhor desempenho foi o modelo Random Forest ( $R^2$  0,59 e RMSE 0,47), indicando predomínio do baixo nível de fragilidade ambiental. Os modelos de PFE e PFML apresentam fortes correlações (0,7 Pearson); no entanto, o PEFML tem correlações mais fortes com outros dados ambientais. Portanto, a predição do PEFML é um modelo robusto que captura informações de covariáveis e apresenta padrões espaciais coerentes.

**Palabras clave:** Modelos de fragilidade ambiental, Previsão Espacial, Floresta aleatória, Planejamento Ambiental.

## INTRODUÇÃO

A modelagem espacial da fragilidade ambiental é uma ferramenta crucial para a gestão territorial e conservação dos recursos naturais. No Brasil, um modelo de fragilidade ambiental amplamente utilizado por Ross (1994) baseia-se na teoria da ecodinâmica da paisagem (Tricart, 1977). Os estudos aplicam esse modelo em duas categorias: (i) fragilidade ambiental potencial (PFE), baseada em dados de declividade, climáticos e individuais; (ii) Fragilidade ambiental emergente, que agrega dados do uso da terra e demonstra áreas das Comunidades Associadas com Atividades Antrópicas. Estudos têm constantemente refinado os modelos de fragilidade ambiental e adaptações metodológicas das propostas propostas; um exemplo é a incorporação de um componente geológico (Crepani et al., 2001; Spröl e Ross, 2006; Franco et al., 2011; Cruz et al., 2017; Campos et al., 2019; Costa et al., 2020). Além disso, as novas metodológicas atendem à tendência futura de usar cada vez mais métodos quantitativos e análises robustas em estudos de geociências (Murray et al., 2009; Padarian et al., 2020). Nesse cenário, estudos têm aplicado novas modificações na modelagem de fragilidade ambiental, por exemplo, lógica fuzzy, redes neurais, análise multicritério e redes bayesianas (Spörl et al., 2011; Campos et al., 2019; Costa et al., 2020; Amorim et al., 2021).

No âmbito de métodos robustos de análise, tem havido um rápido crescimento no uso de inteligência artificial, especialmente com algoritmos de Machine Learning (ML) em estudos de modelagem em vários campos das geociências (Bergen et al., 2019; Gomes et al., 2019; Souza et al., 2020; Silva et al., 2023), mais ainda pouco aplicado a estudos de fragilidade ambiental. A vantagem do ML é aprender padrões complexos para prever dados espaço-temporais usando várias fontes de dados, por exemplo, covariáveis ambientais (Kuhn e Johnson, 2013; Bergen et al., 2019; Padarian et al., 2020). Especificamente neste aspecto, alguns algoritmos de ML podem ter vantagem sobre outros modelos preditivos, como a interpolação de krigagem, que possui geoestatística em sua formulação e exige que os dados inseridos tenham dependência espacial, e às vezes esse padrão de distribuição espacial e inexistente em dados ambientais (Wang et al., 2020; Souza et al., 2022). Além disso, em relação aos covariáveis, estudos já sugerem que a inserção de novos dados (covariáveis) para auxiliar a modelagem da fragilidade ambiental é potencialmente promissora, enquanto a vulnerabilidade ambiental pode estar relacionada a diversos fatores físico-ambientais (Cruz et al., 2017; Amorim et al., 2021); no entanto, a abordagem com ML ainda não foi adequadamente testada para prever modelos de fragilidade ambiental. a vulnerabilidade ambiental pode estar relacionada a diversos fatores físico-ambientais (Cruz et al., 2017; Amorim et al., 2021); no entanto, a abordagem com ML ainda não foi adequadamente testada para prever modelos de fragilidade ambiental. a vulnerabilidade ambiental pode estar relacionada a diversos fatores físico-ambientais (Cruz et al., 2017; Amorim et al., 2021); no entanto, a abordagem com ML ainda não foi adequadamente testada para prever modelos de fragilidade ambiental.

Atualmente, com os avanços da geoinformação, existe uma vasta oferta de informações que funcionam como covariáveis na modelagem usando ML, que podem servir para melhorar os modelos de fragilidade ambiental. Por exemplo, o modelo digital de elevação (MDE) gera diversas covariáveis vinculadas à geomorfologia (Sena et al., 2020); a manipulação de bandas espectrais de imagens de satélite fornece índices de vegetação (Dias et al., 2021); modelos climáticos desenhados para o globo são frequentemente atualizados (Hijmans et al., 2005); e dados categoricos sobre informações ambientais são frequentemente disponibilizados na forma de um banco de dados geográfico (Heineck et al., 2003; UFV et al., 2010). Estudos em diversas áreas mostraram claros benefícios dos incrementos Hijmans et al., 2005; Gomes et al., 2019; Souza et al., 2022). A predição espacial por ML é especialmente necessária quando a determinada área de estudo apresenta uma grande quantidade de dados espaciais com diversos aspectos físicos ambientais (geodiversidade), onde estudos basados em análises qualitativas não são suficientes ou demorados, ou métodos de predição más simples não dão conta do dados bem (Bergen et al., 2019; Padarian et al., 2020; Souza et al., 2022). No Brasil, uma área de geodiversidade no estado de Minas Gerais, com contextos geotécnicos variados (Machado e Silva, 2010; Costa, 2021), condicionando diversos aspectos geomorfológicos, como superfícies achatadas por dissecção, cadeias montanhosas pertencentes a

contextos orogênicos e a pedodiversidade como a única mais intemperizada do mundo (Ker, 1997; Silva et al., 2018). Além disso, possui um arcabouço geocológico complexo (ou seja, presença de Cerrado, Mata Atlântica e Caatinga) decorrentes paleoclimáticos (Ab'sáber, 1970). Portanto, o avanço dos estudos de fragilidade ambiental é conectar vários aspectos ambientais, considerando a heterogeneidade ambiental, e o ML tem um papel fundamental nesse esforço. Este estudo tem como objetivo testar algoritmos de aprendizado de máquina para prever um novo modelo de Fragilidade Ambiental Potencial (PFE), demonstrando que é covariáveis são potencialmente explicativas para os níveis de fragilidade no estado de Minas Gerais.

## MATERIAIS E MÉTODOS

### ÁREA DE ESTUDO

O estado de Minas Gerais está localizado no sudeste do Brasil entre  $-23^{\circ}0'$  a  $-14^{\circ}0'$  S e  $-51^{\circ}0'$  a  $-40^{\circ}0'$  W (Figura 1). Com base no tipo de clima de Köppen, existem tipos de clima na região: Cwb (clima temperado com inverno seco e verão moderadamente frio), Cwa (clima temperado com inverno e verão secos), Aw (clima tropical de savana com estação seca de inverno) e As (clima úmido semitropical). O arco geológico é marcado por quatro grandes províncias: (i) a Província de São Francisco, com rochas cristalinas, frequentemente cobertas por sedimentos metassedimentares (Neoproterozóico); (ii) a Província da Mantiqueira possui maciços e morros desenvolvidos em rochas graníticas/granitóidas e metamórficas (Proterozóicas); (iii) a Província do Tocantins compra as faixas granítica/granitoide e xistosas (Proterozóicas) no tabuleiro do Cráton do São Francisco; e (iv) a Bacia Sedimentar do Paraná, com rochas máficas cobertas por arenitos do Cretáceo, formando extensas planícies (Ab'sáber, 1970; Heineck et al., 2003). O estado é marcado pela presença de três importantes biomas – Mata Atlântica, Cerrado e Caatinga, além de zonas de transição denominadas ecótonos (Ab'sáber, 1970).

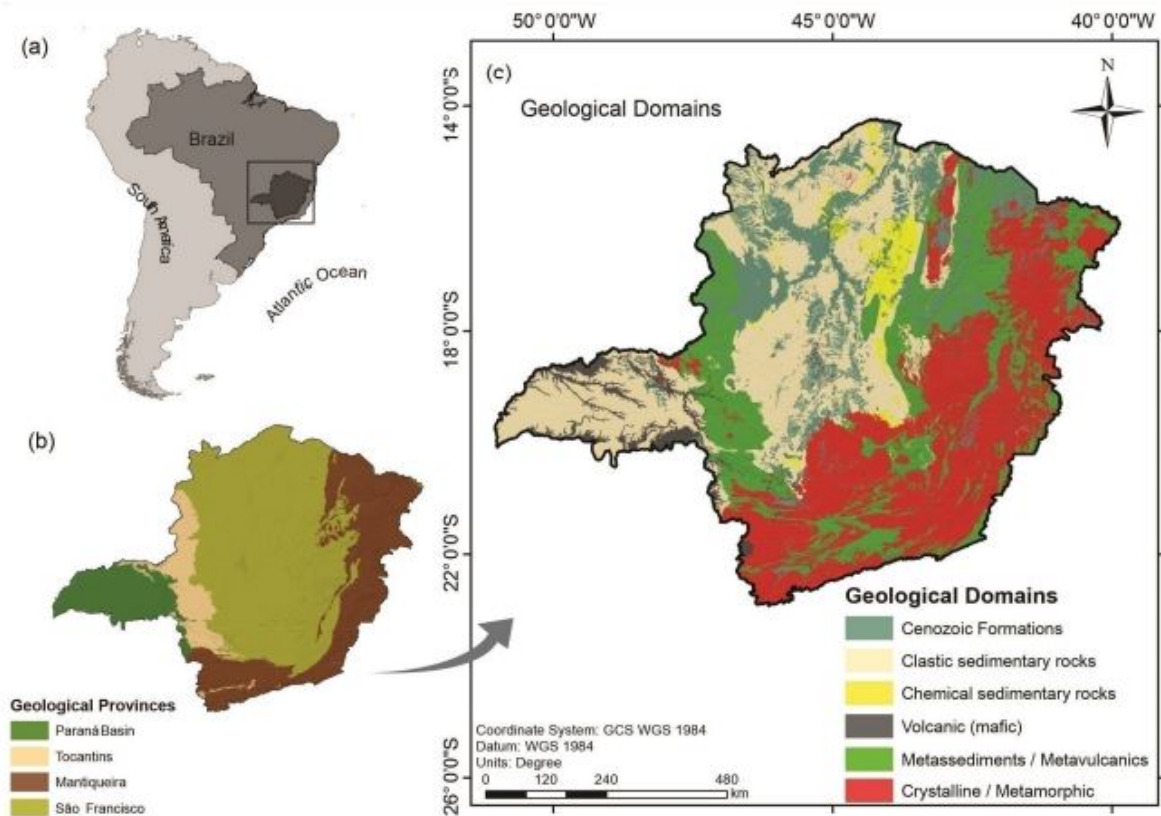


Figure 1 - (a) Location of Minas Gerais state, (b) geological provinces of Minas Gerais, (c) Geological domains of the state of Minas Gerais

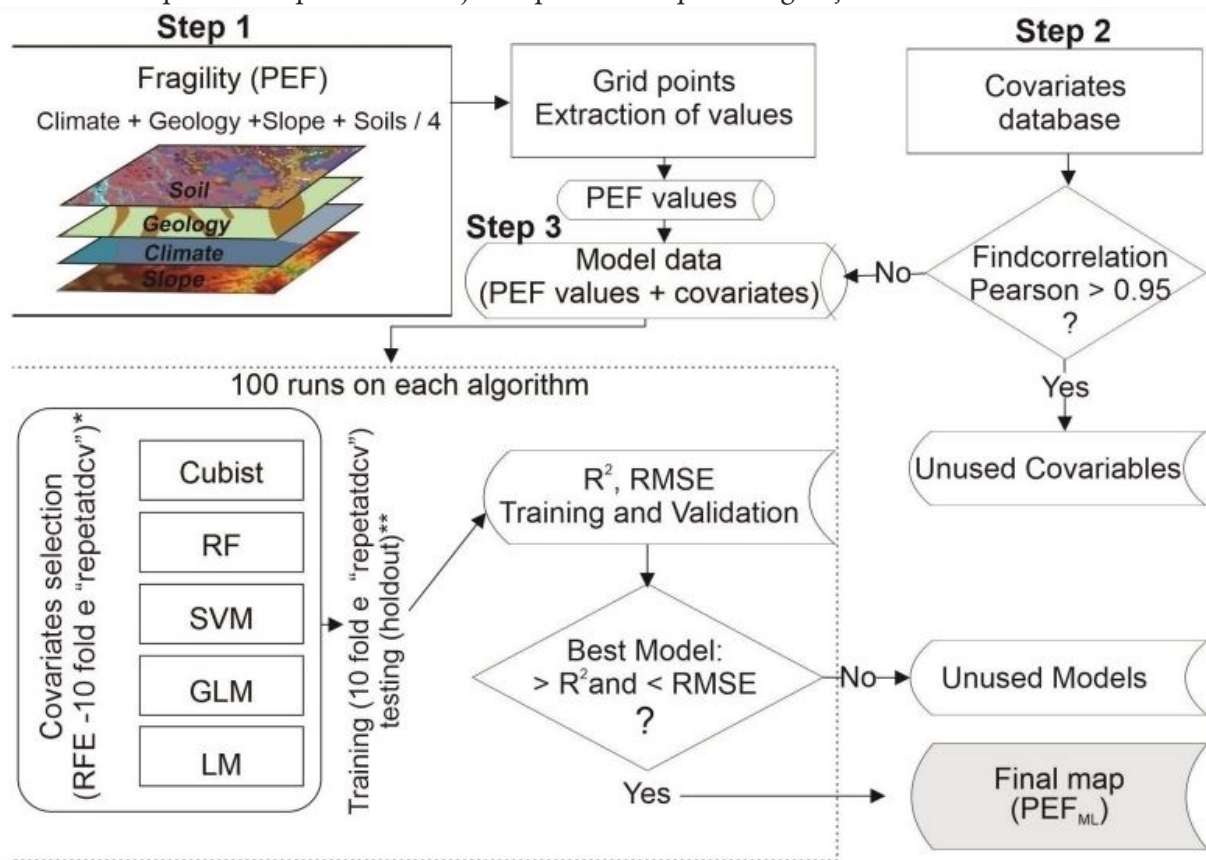
#fig1en.jpg

Figura 1 - (a) Localização do estado de Minas Gerais, (b) províncias geológicas de Minas Gerais, (c) domínios geológicos do estado de Minas Gerais

## PROCEDIMENTOS METODOLÓGICOS

Para a análise da fragilidade ambiental, montou-se um quadro metodológico sintetizado na Figura 2, com procedimentos executados no software R (Rcore, 2023). Anteriormente, aplicou-se um modelo de Fragilidade Ambiental Potencial (PFE), que consiste na inserção de variáveis físico-ambientais (Ross, 1994; Crepani et al., 2001; Spörl et al., 2011). As variáveis inseridas foram clima, geologia, relevo e solo, com pesos atribuídos à classe de cada variável. Nessa etapa de ponderação, selecionamos valores (pesos) de fragilidade ambiental disponíveis em estudos anteriores e os atribuímos como classes (Ross, 1994; Crepani et al., 2001; Franco et al., 2011; Spörl et al., 2011; Cruz et al., 2017; Campos et al., 2019; Amorim et al., 2021). Aplicamos uma análise multicritério (Analytical Hierarchy Process - AHP) para definir a importância relativa de cada variável em relação à fragilidade ambiental, conforme proposto por Amorim et al (2021). Finalmente, aplicamos uma sobreposição das variáveis usando a equação  $PEF = C + G + R + S / 4$ ; onde PEF: Fragilidade Ambiental Potencial, C: Clima, G: Geologia, R: Revezamento (encosta), S: Solo. Ao final, obtém-se um resultado resultante da média aritmética dos valores de fragilidade registrados nas aulas. As ordens de fragilidade ambiental obtidas são classificadas em 1: muito baixa, 2: baixa, 3: média, 4: forte, 5: muito forte. Revezamento, S: Solo. Ao final, obtém-se um resultado resultante da média aritmética dos valores de fragilidade registrados nas aulas. As ordens de fragilidade ambiental obtidas são classificadas em 1: muito baixa, 2: baixa, 3: média, 4: forte, 5: muito forte. Revezamento, S: Solo. Ao final, obtém-se um resultado resultante da média aritmética dos valores de fragilidade registrados nas aulas. As ordens de fragilidade ambiental obtidas são classificadas em 1: muito baixa, 2: baixa, 3: média, 4: forte, 5: muito forte.

Do mapa do PFE (etapa anterior), extraímos os valores de fragilidade ambiental do mapa para uma nota de 4.800 pontos distribuídos aleatoriamente com distância mínima de 3 km. Esses pontos constituem a principal variável para determinar os níveis de fragilidade ambiental usando a técnica do algoritmo Machine Learning (ML); e o resultado desse procedimento é a geração de um novo mapa, aqui denominado Potential Environmental Fragility by Machine Learning (PEFML). No entanto, para auxiliar a previsão por ML, montamos um banco de dados de covariáveis em formato raster, estruturado em resolução de 1x1 km: cinquenta e cinco dados espaciais do WorldClim (Hijmans et al., 2005); 42 covariáveis geomorfológicas da Shuttle Radar Topographic Mission (SRTM) (USGS, 2023), extraídas utilizando o software SAGA (Olaya e Conrad, 2009; Sena et al., 2020); Um dado de espectrometria gama e um de gravimetria (Heineck et al., 2003), um NDVI calculado a partir de imagens de satélite do sensor MODIS (Moderate Resolution Imaging Spectroradiometer) de agosto de 2019 (USGS, 2023), que é o mais seco e mostra a diferença mais significativa entre fitofisionomias. As variáveis solo, geologia, clima, não foram incluídas para evitar previsões invejadas, pois faziam parte da geração PEF.



#fig2en.jpg

Figura 2 - Fluxograma com a sequência: determinação de entrada (PFE); criação de banco de dados de covariáveis, seleção de covariáveis importantes com função RFE; teste de algoritmos para previsão de PEF com 100 execuções. Ao final, foi selecionado o algoritmo mais precisa para modelar e mapear o PEFML (Potential Environmental Fragility by Machine Learning PEFML). RF: Floresta Aleatória; SVM: Máquina de Suporte Vetorial; GLM: Modelos lineares generalizados; LM: Modelo linear. R2: R-quadrado; RMSE: Raiz do Erro Quadrático Médio

A etapa anterior gera uma estrutura de dados com valores de variáveis (PEF) e valores de covariáveis, mas o número excessivo de covariáveis também favorece a criação de modelos superestimados (Gomes et al., 2019; Padarian et al., 2020). Portanto, aplicamos um corte de correlação para eliminar covariáveis altamente correlacionadas, pois elas têm uma contribuição semelhante para explicar a distribuição de uma variável analisada. A função usada foi encontrar correlação, parametrizada pelo critério de corte de correlação >0,95 Pearson (Kuhn e Johnson, 2013). Além disso, aplicamos uma importante ferramenta de

seleção de covariáveis usando uma ferramenta amplamente utilizada em modelos de previsão, chamada Recursive Feature Elimination (RFE), que também evita previsões superestimadas. O RFE é um método de seleção de características inversas no conceito central baseado na eliminação de covariáveis sem importância (Kuhn e Johnson, 2013). Após remover a covariável menos importante, o RFE reajuste o modelo com um conjunto menor de covariáveis, reiniciando o processo para eliminar a covariável menos importante; e este processo é repetido várias vezes até a estabilização em que o nível de precisão determinado por R-quadrado ( $R^2$ ) não diminui. Este processo é realizado com base apenas no dataset de treinamento que foi 75% das amostras, enquanto os outros 25% são utilizados no processo de teste (holdout-test). No aprendizado de máquina, a fase de teste é um processo para avaliar o desempenho de um modelo treinado com um conjunto de teste (ou seja, 25% das amostras), que não foi visto pelo modelo na fase de treinamento.

A partir dos dados ajustados às covariáveis selecionadas pelo RFE, estas foram usadas para prever a fragilidade ambiental da área de estudo, testando diferentes algoritmos. Além disso, para cada algoritmo, o processo foi repetido em 100 execuções usando o subconjunto apropriado de covariáveis indicado pelo RFE. Na etapa de previsão, foram utilizados cinco algoritmos de aprendizado de máquina: Cubist (Quinlan, 1992), Generalized linear models-GLM (Hastie e Tibshirani, 1987), Linear Model Regression – LM (Faraway, 2016), Random Forest – RF (Breiman, 2001) e Support Vector Machine – SVM (Cortes e Vapnik, 1995). Esses modelos foram testados em diversos estudos (Brungard et al., 2015; Longe, 2016; Morellos et al., 2016; Gomes et al., 2019; Souza et al., 2022; Silva et al., 2023). Além disso, o uso de diferentes algoritmos é essencial para avaliar as limitações relacionadas à predição da variável-alvo por algoritmos que possuem diferentes rotinas estatísticas (Kuhn e Johnson, 2013; Padarian et al., 2020). Consideramos, na avaliação do algoritmo mais precisa, as métricas da fase de teste (25% das amostras) e avaliamos o efeito de overfitting comparando com os dados de treinamento (75% das amostras), usando como métricas: R-quadrado -  $R^2$  e raiz do erro quadrático médio – RMSE.

Para observar a variação entre o PEFML e o método original (PEF), uma nova grade de pontos foi criada com 4,800 pontos aleatórios, com posterior extração dos valores PEFML e PEF. Além disso, essa análise estatística foi disponibilizada para Minas Gerais (Mapas: Grau de conservação, Risco de erosão, Vulnerabilidade natural, Prioridade de conservação, Vulnerabilidade do solo, Erodibilidade). Esses mapas fazem parte do trabalho compilado do zoneamento agroecológico de Minas Gerais disponibilizados em dados matriciais (raster), com intervalos de 1 a 5, faixa compatível com o PEFML

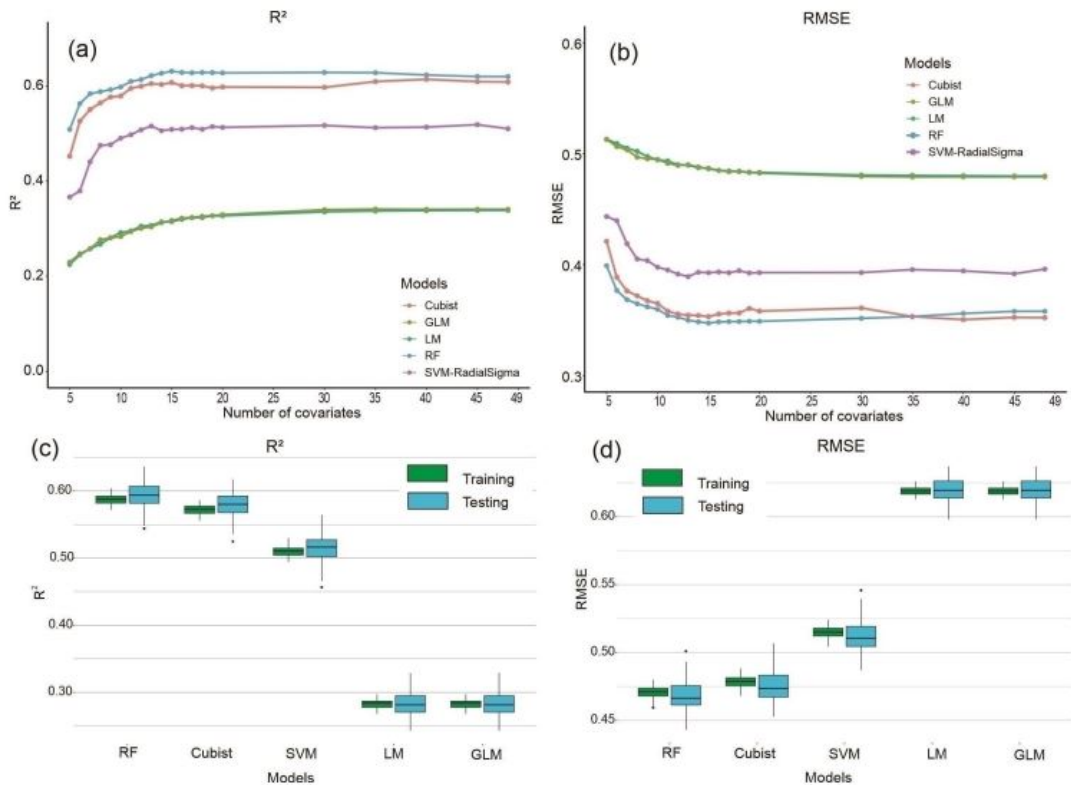
mapa (Scolforo et al., 2008). Sobre esses dados, aplicamos a análise de correlação de Pearson e a análise de cluster K-means para observar as relações entre esses mapas e os clusters de similaridade (Sena et al., 2020). Para fornecer uma visão espacial das correlações, selecionamos os mapas de falésias e vulnerabilidade ambiental com base em Scolforo et al (2008) e aplicamos um cálculo de subtração usando PEFML como referência, de acordo com a equação: Mapas de relação = PEFML - EZMi. Onde PEFML : Mapa de Fragilidade Ambiental Potencial por Machine Learning, EZM: Mapa da zona ambiental, i: mapa específico do EZM.

## RESULTADOS

### COVARIADAS AMBIENTAIS E SELEÇÃO DE MODELOS

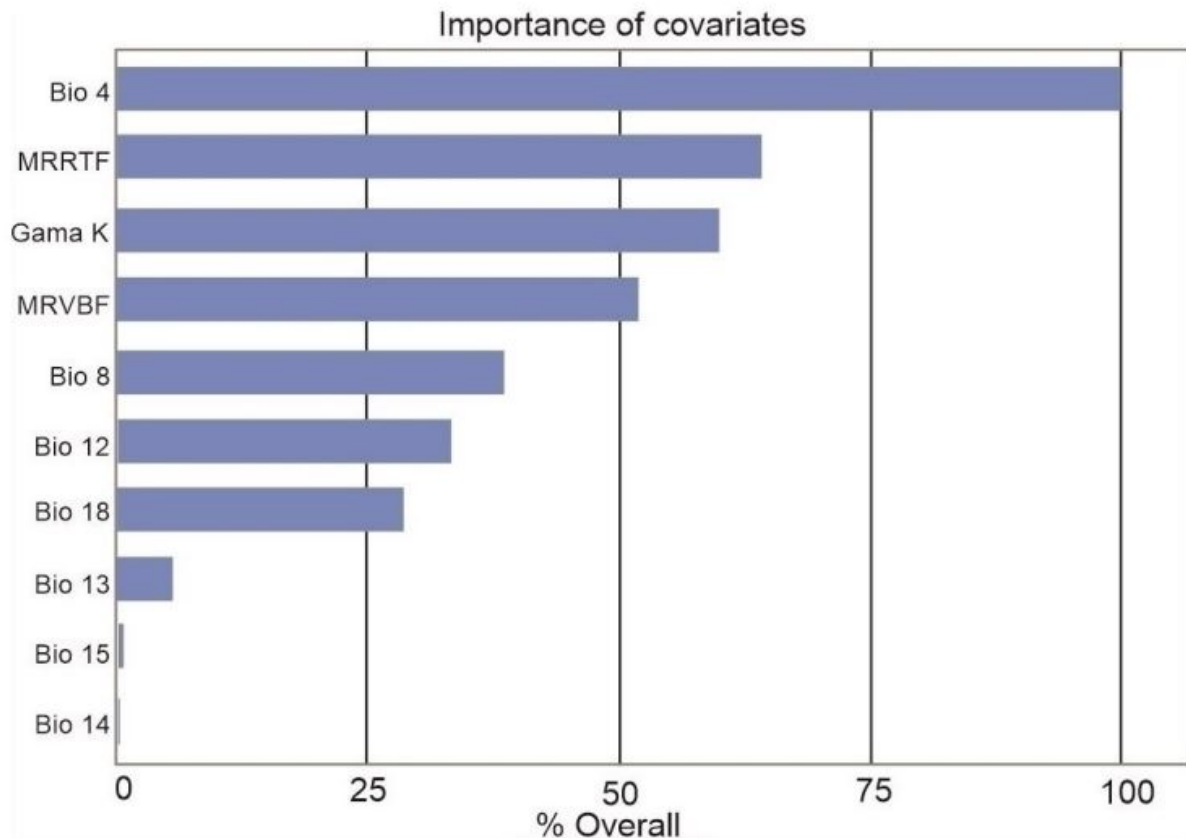
O processo preditivo do PEFML envolve cinco algoritmos e banco de dados de covariáveis com dados climáticos, topográficos, geoquímicos e de vegetação, com níveis de correlação de 0,95 nessas covariáveis. Ao aplicar a função RFE para selecionar covariáveis importantes, observamos que aproximadamente doze covariáveis é o número máximo para gerar  $R^2$  mais alto e menor RMSE nos algoritmos utilizados (Figura 3 a,b). O algoritmo Random Forests apresentou o melhor desempenho na seleção do RFE, selecionando apenas 10 covariáveis classificadas por nível de importância (Figura 4). De acordo com o ranking (% global), os dados bioclimáticos do WorldClim foram predominantes e mais significativos. Dentre as variáveis topográficas, apenas MRRTF (MRTF: multiresolution ridge-top planeness index), que indica

posições planas em áreas de alta altitude, e MRVBF (MRVBF: multiresolution valley bottom planeness index), que mostra superfícies planas no fundo do vale, foram incluídas na modelagem. Além dessas, uma covariável de espectrometria gama do elemento K ocupou o terceiro nível de importância. Esta covariável, além de indicar a espectrometria de raios gama para o potássio (40 K), também pode ser correlacionada com um certo volume do elemento na superfície. Portanto, Random Forest foi selecionado por precaução, pois de 100 execuções gerou  $R^2 > 0.61$  (treinamento 0.61, teste 0.59)



#fig3en.jpg

Figura 3 - Desempenho dos modelos: (a, b) evolução das métricas R-quadrado (R<sup>2</sup>), raiz do erro quadrático médio (RMSE) na seleção da covariável usando Recursive Feature Elimination. c, d) Métricas de precisão R<sup>2</sup> e erro RMSE, respectivamente, em 100 execuções usando covariáveis selecionadas por RFE. RF: Floresta Aleatória; SVM: Máquina de Suporte Vetorial; GLM: Modelos lineares generalizados; LM: Modelo linear. R<sup>2</sup>: R-quadrado; RMSE: Raiz do Erro Quadrático Médio



#fig4en.jpg

Figura 4 - A importância relativa das covariáveis, dada pelo percentual geral. Bio 4: tempero de temperatura (desvio do padrão \* 100), MRRTF: índice planicida de múltiplas resoluções, faixa K: elemento K de espectrometria gama, MRVBF: índice de planicídio multiresolução do fundo do vale, Bio 4: tempero de temperatura, Bio 8: temperatura média do trimestre mais quente, Bio 12: precipitação anual, BIO18: precipitação do trimestre mais longo, Bio 13: precipitação do mais seco, BIO15: precipitação do mais seco (coeficiente de variação), BIO14: precipitação do mais seco

## PREVISÃO DE FRAGILIDADE AMBIENTAL

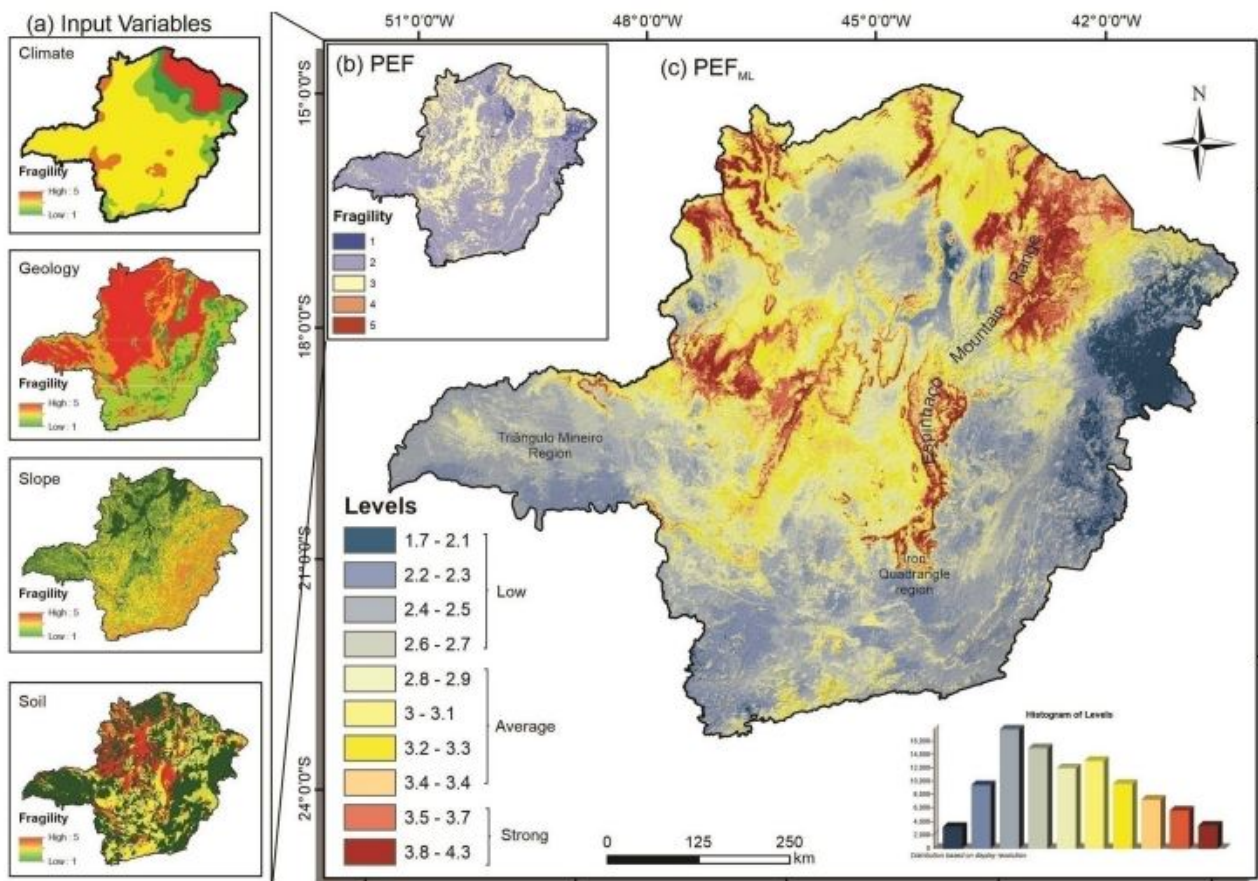
O mapa de Fragilidade Ambiental Potencial utilizando ML (PEFML) mostra padrões espaciais semelhantes com o mapa de fragilidade ambiental derivado de vários postes superiores (PEF) - (Figura 5). Essa similaridade atesta a robustez das previsões do ML em aprender como ocorre a distribuição de uma variável, e esse aprendizado foi feito com novas covariáveis (Figura 4), portanto, sem adicionar as variáveis que deram origem ao mapa de fragilidade ambiental pelo método de Ross (ou seja, variáveis mostradas na Figura 5a).

No entanto, a previsão usando o algoritmo de RF mostrou a desvantagem por normalizar valores que possuem baixa expressividade territorial, assim classes de fragilitye abaixo de 1.7 e acima de 4.3 foram eliminadas, incluindo assim a classe 1 (muito baixa) e 5 (muito forte). Portanto, a alternativa foi aplicar uma reclassificação dos níveis, de acordo com (1.7 a 2.7: baixo, 2.8 a 3.4: médio, 3.5 a 4: forte). Esse critério de separação prioriza classe de fragilitye média, que é dominante no mapa do PEF (Figura 5 b).

Apesar de normalizar os valores, a predição do ML gerou um mapa com feições mais detalhadas, pois a predição captura as feições das covariáveis inseridas e no criou grandes zonas homogêneas, indicando de forma coerente áreas de fragilidade ambiental (Figura 5 b,c). Portanto, a distribuição espacial do PEFML apresenta padrões espaciais que denotam a influência de várias variáveis de entrada recomendadas em estudos de fragilidade ambiental (Ross, 1994; Spröl e Ross, 2006; Cruz et al., 2017).

Por exemplo, a previsão capta a influência da geomorfologia no aumento do nível de fragilidade ambiental em áreas de maior declividade no contexto da região central, onde há relâmpagos de forte declínio (por exemplo, Serra do Espinhaço), e em áreas íngremes, como as planícies da região noroeste. Fatores geológicos e edáficos também contribuem para áreas de média e alta fragilidade, principalmente em áreas com maior variação pedagógica e geológica, como as regiões central e norte do estado (Iglesias e Uhlein, 2009; Silva et al., 2018). Além disso, a região norte tem adicionado fator de condições climáticas de baixa precipitação para contribuir para maiores níveis de fragilidade; Esta parte do estado inclui o limite do semiárido brasileiro.

A baixa fragilidade ambiental predomina na maior parte do estado, principalmente em áreas localizadas a leste. Essa configuração também atesta que o PEFMLmap é espacialmente consistente, uma vez que, em áreas localizadas a leste, a maioria dos vários pontos de entrada no modelo PEF também apresenta baixos níveis de fragilidade ambiental, com exceção do relé cabelo, pois há um relé do domínio Mares de Morros, prevalecendo maiores declividades (Figura 5a). Como áreas de baixa fragilidade ambiental também possuem extensão na região norte, onde apenas uma geologia variável apresenta maiores níveis de fragilidade, portanto, outras variáveis são responsáveis por essa atenuação. Esse efeito de atenuação se repete na região oeste (região do Triângulo Mineiro), devido à geologia expressar alta fragilidade ambiental; Enquanto isso, o revezamento do declínio e a predominância de solos profundos influenciam na redução de valores.



#fig5en.jpg

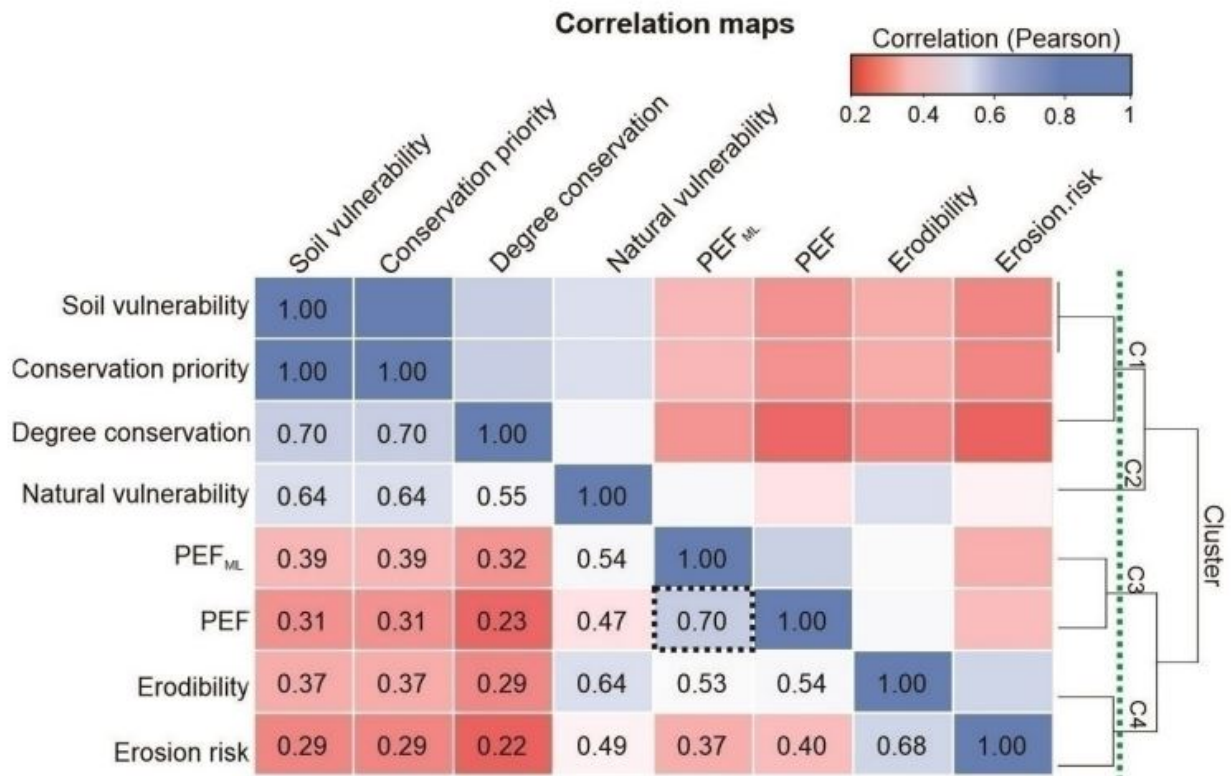
Figura 5 - (a) Entrada de variáveis com valores de fragilidade para gerar um mapa de Fragilidade Ambiental Potencial, (b) Mapa de Fragilidade Ambiental Potencial por superposição de variáveis ambientais, (c) Mapa de Fragilidade Ambiental Potencial gerado por Machine Learning utilizando o modelo de Floresta Aleatória (100 corridas) (PEFML), e histograma com distribuição de classes

## COMPARAÇÃO ESTATÍSTICA DE MAPAS

Além da semelhança do padrão espacial entre o mapa PEF pelo método convencional e o mapa PEFML derivado do aprendizado de máquina, também houve uma forte correlação estatística entre esses mapas (Pearson 0,70); e esses dois mapas estão no mesmo cluster (C3) (Figura 6), confirmando o nível de similaridade. Dessa forma, comparamos os mapas de fragilidade ambiental das metodologias (PFE e PEFML) com outras informações ambientais sobre risco e vulnerabilidade ambientais disponíveis para o estado de Minas Gerais (Scolforo et al., 2008). O PEFML mapa sempre mostrou uma correlação superior com mapas de risco e vulnerabilidade, indicando que é potencial mais explicativo de outros fatores ambientais: vulnerabilidade do solo, prioridade de conservação, grau de conservação, vulnerabilidade natural, erodibilidade e risco de erosão. Por outro lado, o mapa PEF apresentou valores fracos em todas as correlações com esses mapas (Figura 6).

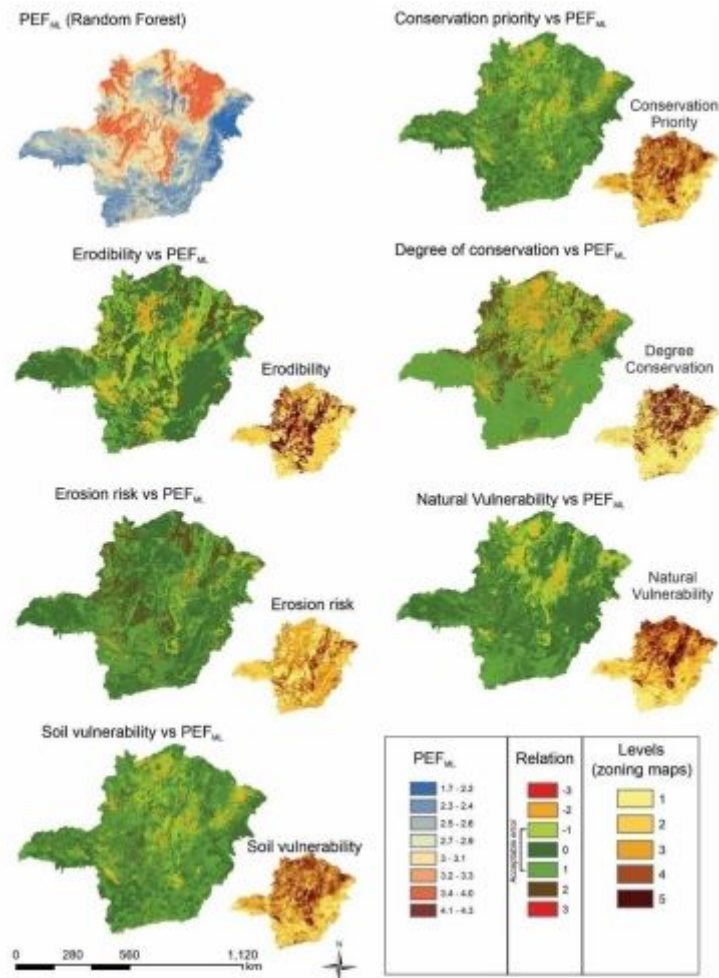
A comparação espacial entre o PEFML o mapa com os mapas de risco e vulnerabilidade apresentou correspondência entre os mapas ao indicar áreas com grau de fragilidade semelhante; mesmo considerando que o mapa PEFML possui um intervalo menor de classes (1,7 a 4,3). Portanto, na maior área do estado, os resultados mostram que a variação dos valores de fragilidade ambiental não existe.

PEFML com outros mapas está em uma faixa baixa (1 e -1 para cada classe de fragilitye). Essa baixa variação atesta que o mapa PEFML pode captar informações sobre risco/vulnerabilidade ambiental de fatores isolados, sugerindo ser um modelo mais completo.



#fig6pt.jpg

Figura 6 - Correlação entre os valores de fragilitye usando o método de álgebra de mapas (PEF) e o algoritmo de aprendizado de máquina (PEFML). Além disso, correlação com mapas de risco e vulnerabilidade usando dados da base de zoneamento digital de Minas Gerais (Scolforo et al., 2008). (Todos os valores com nível de significância alfa = 0,05; linha verde para separação dos agrupamentos (C1 a C4); correlação dos quadrados de preto entre PFE e PEFML



#fig7pt.jpg

Figura 7 - Mapas da relação entre o mapa de fragilidade ambiental (PEFML) (canto superior esquerdo), com cada mapa de vulnerabilidade e risco ambiental produzido a partir da zona econômico-ecológica do estado de Minas Gerais. Pequeno mapa em cada canto inferior é o mapa de risco e vulnerabilidade.

## DISCUSSÃO

### DESEMPENHO DO MODELO

A predição espacial de fragilitye usando vários ML associados a conjuntos de dados de covariáveis é um procedimento de análise espacial robusto e mostra diferentes desempenhos estatísticos de acordo com o algoritmo selecionado. O algoritmo mais eficiente para escolher covariáveis para predição PEFML foi o RF, com a ferramenta RFE ajustada a este algoritmo, e seu algoritmo lida bem com dados de alta dimensão e permite relações não lineares entre preditores (Breiman, 2001; Gomes et al., 2019). Ressaltamos a vantagem de usar a função RFE na fase de treinamento, eliminando covariáveis que não melhoraram o desempenho da predição, e essa remoção cria uma predição mais más simples e não superestimada, obedecendo ao princípio da parcimônia na modelagem (Brungard et al., 2015 ). o R2e quanto as métricas RMSE foram satisfatórias (Figura 3c,d), principalmente quando considerados os valores das fases de treinamento e teste, que foram semelhantes, indicando que na fase de execução do algoritmo, o modelo treinou e testou satisfatoriamente (ou seja, sem overfitting efeito) e o RF tem uma vantagem em mitigar esse efeito adverso (Breiman, 2001; Were et al., 2015). Dados de previsão ambiental, especialmente para grandes áreas, raramente geram R 2 acima de 0,70. A limitação de métricas altas decorre de fatores de baixa resolução das covariáveis, falta de covariáveis estritamente ligadas à variável analisada e dados que não

possuem relações lineares simples (Malone et al., 2009; Gomes et al., 2019; Padarian et al., 2019). al., 2020; Souza et al., 2022).

Considerando que o RF realizou o treinamento usando as covariáveis indicadas pelo RFE, algumas dessas covariáveis são potencialmente explicativas dos níveis de fragilidade ambiental em Minas Gerais (Figura 4). Em geral, as covariáveis selecionadas têm alguma relação com as variáveis de entrada para determinar o PEF (ou seja, clima, geologia, relevo, solo). Portanto, os dados do WorldClim foram predominantes e de maior importância, especialmente os bioclimáticos, que são mais significativos na explicação das tendências climáticas (Hijmans et al., 2005; Gomes et al., 2019). As informações topográficas do SRTM também compõem as covariáveis selecionadas para predição; Fazem parte desta seleção MRRTF e MRVBF que enfatizam áreas planas em altitude ou vales. Essas covariáveis topográficas estão correlacionadas com processos hidrológicos de erosão e deposição (Gallant e Dowling, 2003), associados a áreas de fragilidade ambiental. A terceira covariável mais significativa foram os dados gama-spectrometria do elemento potássio (K gama-spectrometria). O potássio é um componente dos minerais (feldspato, biotita e muscovita), presente em níveis proeminentes em algumas rochas e isolado com baixo desenvolvimento pedogenético e/ou apenas eutrófico (Guevara et al., 2018). Portanto, a seleção da gama-spectrometria de K foi eficiente porque combina duas informações (solo e geologia), que são variáveis para a modelagem do PEF. presente em níveis proeminentes em algumas rochas e solos com baixo desenvolvimento pedogenético e/ou apenas eutróficos (Guevara et al., 2018). Portanto, a seleção da gama-spectrometria de K foi eficiente porque combina duas informações (solo e geologia), que são variáveis para a modelagem do PEF. presente em níveis proeminentes em algumas rochas e solos com baixo desenvolvimento pedogenético e/ou apenas eutróficos (Guevara et al., 2018). Portanto, a seleção da gama-spectrometria de K foi eficiente porque combina duas informações (solo e geologia), que são variáveis para a modelagem do PEF.

## FRAGILIDADE AMBIENTAL DO ESTADO DE MINAS GERAIS

Em relação à distribuição espacial do PEFMLvalores, há uma conjunção de fatores que atuam para atenuar ou aumentar os níveis de fragilidade ambiental. Diversas regiões com maior fragilidade ambiental estão associadas a áreas de maior declividade, onde predomina o processo de morfologia, esse aspecto é marcado no contexto da Serra do Espinhaço ou escarpas de planície, feição geomorfológica típica de Minas Gerais (Callisto et al., 2016; Costa, 2021). Por exemplo, na Serra do Espinhaço, há ambientes sustentados pela interdependência de processos dinâmicos entre tipos de vegetação, clima, rugosidade, declividade e vazão hidrológico. Um exemplo são os campos rochosos ferruginosos "ecossistemas de canga"

Fatores geológicos também contribuem para o aumento da fragilidade ambiental, principalmente em áreas com grande variação litológica, onde as paisagens são esculpidas por processos diferenciais de nudez. Por exemplo, na área do Quadrilátero Ferrífero predominam rochas resistentes ao processo de rebaixamento vertical (downwearing). Além disso, ainda existem áreas que têm sido afetadas pelos processos de retração lateral das escarpas (backwearing), devido à erosão de litótipos mais frágeis, que formam a base das escarpas (Salgado et al., 2006). Os processos evolutivos são semelhantes aos associados à erosão diferencial envolvendo outros tipos litológicos, que têm sido registrados em outras áreas do estado, criando relés inclinados (Simões et al., 2020; Souza et al., 2020; Costa, 2021). Outro contexto geológico que frequentemente apresenta maiores níveis de fragilidade em áreas de rochas carbonáticas, particularmente em uma zona de geomorfologia cárstica, que tende a apresentar diversos problemas de vulnerabilidade devido à presença de fraturas e cavidades produzidas por processos cársticos (Pessoa et al., 2020). A atenuação da fragilidade ambiental em áreas de rochas carbonáticas deve-se à presença de solos profundos (Latosolos) e/ou solos relativamente eutróficos (Latosolos - Kandic) – (Ker, 1997), criando mais ambientes com predominância de processos morfogenéticos e pedogênicos. Além disso, as condições geológicas e topográficas que indicam um alto grau de fragilidade ambiental podem ser uma contribuição adicional para o fator climático, que parece ser o caso da região norte do estado, principalmente por estar em áreas sujeitas à desertificação (Barros et al., 2018). Portanto

Os menores níveis de fragilidade ambiental devem-se à presença de rochas graníticas e metamórficas (gnaiesses) na província geológica da Mantiqueira (Figura 1), sistemas rochosos com baixa vulnerabilidade (Cruz et al., 2017). Apenas as profundas, localizadas em áreas de revezamento inclinado, também contribuem para atenuar a fragilidade ambiental, pois estão apenas mais na paisagem e o processo de pedogênese é preponderante (Ker, 1997; Nunes et al., 2001). No entanto, os níveis de fragilidade ambiental podem aumentar a condição de alta declividade, associada à presença apenas daqueles que são baixos, como os Cambissolos (Nunes et al., 2001). Os valores de baixa e média fragilidade também estão sendo estagnados no oeste do estado, envolvendo grande parte da Baía do Paraná. Nessa região predomina com morfologia plana, profunda e de baixa fertilidade (Ker, 1997). Apenas os distróficos são apresentados, os atributos físicos são apresentados, e os processos de pedogênese suplantam os morfogênese (Motta et al., 2002). Portanto, os valores de fragilidade obedecem à configuração do terreno; apenas em áreas de maior declividade é que os valores de fragilidade aumentam (Martins e Rodrigues, 2012). No entanto, os níveis de fragilidade ambiental podem aumentar a condição de alta declividade, associada à presença apenas daqueles que são baixos, como os Cambissolos (Nunes et al., 2001). Os valores de baixa e média fragilidade também estão sendo estagnados no oeste do estado, envolvendo grande parte da Baía do Paraná. Nessa região predomina com morfologia plana, profunda e de baixa fertilidade (Ker, 1997). Apenas os distróficos são apresentados, os atributos físicos são apresentados, e os processos de pedogênese suplantam os morfogênese (Motta et al., 2002). Portanto, os valores de fragilidade obedecem à configuração do terreno; apenas em áreas de maior declividade é que os valores de fragilidade aumentam (Martins e Rodrigues, 2012).

## MAPAS DA CORRELAÇÃO

O modelo de fragilidade por superposição de mapas (variáveis) (Ross, 1994; Spröl e Ross, 2006), amplamente aplicado em estudos ambientais, continua sendo um método eficiente para identificar áreas de fragilidade ambiental (Spröl e Ross, 2006; Franco et al., 2011; Spörl et al., 2011; Martins e Rodrigues, 2012; Campos et al., 2019; Anjinho et al., 2021). No entanto, o modelo de fragilidade ambiental por aprendizado de máquina (PEFML), aplicado neste estudo, também se mostrou um método confiável, incluindo resultados altamente correlacionados com o método original (Figura 6). Além disso, o PEFML mapa foi mais correlacionado com outros dados sobre fragilidade, vulnerabilidade e risco ambiental para o estado de Minas Gerais (Scolforo et al., 2008). Essa correlação mais alta sugere que a previsão ML é um modelo mais explicativo de outros fatores ambientais não presente na modelagem PEF, que usa apenas variáveis sobrepostas com pesos atribuídos. Presumivelmente, essa maior correlação é uma contribuição do número de variáveis selecionadas na etapa da função RFE no teste do modelo, devido aos métodos preditivos de ML, resultado do mapa e influenciado pelas informações de várias variáveis que auxiliam nas previsões (Brungard et al., 2015; Gomes et al., 2019; Souza et al., 2022).

Os resultados apresentados indicam um ganho metodológico na modelagem de potenciais fragilidades ambientais utilizando ML. Basicamente, o ML é um método de análise robusto, e modelos complexos tendem a produzir previsões mais precisas do que modelos mais simples, pois além de testar diferentes algoritmos bem desenvolvidos no campo da ciência estatística, a previsão ML tem em sua estrutura a incorporação de covariáveis que auxilia na previsão (Kuhn e Johnson, 2013; Brungard et al., 2015; Morellos et al., 2016; Souza et al., 2022). Portanto, a inserção de covariáveis novas que possam explicar a distribuição espacial de áreas potencialmente frágeis também é um fator essencial, especialmente quando se considera que a vulnerabilidade ambiental está associada a uma multiplicidade de variáveis físicas e antrópicas (Cruz et al., 2017),

## CONCLUSÕES

O Modelo de Fragilidade Ambiental Potencial obtido pelo algoritmo de aprendizado de máquina (PEFML) é estatisticamente semelhante ao Modelo de Fragilidade Ambiental Potencial (PFE) adquirido pela superposição de diversos ambientes físicos.

O modelo Random Forest foi o mais eficiente na predição do PEFML, utilizando um conjunto de covariáveis significativas, com níveis de desempenho satisfatório na fase de validação (teste de  $R^2$  0,59 e RMSE 0,47). Nível de correlação com outros fatores de risco e vulnerabilidade ambiental, sendo um mapa mais explicativo de outros fatores que influenciam a fragilidade ambiental.

O PEFML indica níveis de fragilidade derivados de efeitos topográficos, geológicos, climáticos e pedológicos. Áreas de alta fragilidade estão associadas à geomorfologia montanhosa e à presença de escarpas nas planícies.

Destaca-se que a área de estudo é possível em escala regional, com uma variedade de projetos em pequena escala cartográfica. Portanto, replicar o modelo (PEFML) ao nível de pequenas bacias hidrográficas com validação em campo é uma forma de ratificar a eficiência da modelagem.

## REFERÊNCIAS

- AB'SÁBER, uma província geológica e domínios morfoclimáticos no Brasil. São Paulo: Universidade de São Paulo, Instituto de Geografia, 1970. 26 págs. Anos 1970
- AMORIM, AT, LOPES, ERN, SOUSA, JAP, SILVA, RCFD, SOUZA, JC, LOURENÇO, RW Fragilidade ambiental geomorfométrica de uma bacia hidrográfica: uma abordagem espacial multicrítica. *Monitoramento e Avaliação Ambiental*, v.193, p.850. 2021.
- ANJINHO, PDS, BARBOSA, MAGA, COSTA, CW, MAUAD, FF Análise da Fragilidade Ambiental no Planejamento do Uso do Solo em Bacias de Drenagem de Reservatórios: Um Estudo de Caso de Bacia Brasileira. *Política de Uso da Terra*, v.100, p.104946. 2021.
- BARROS, K.O., RIBEIRO, C.A.A.S., MARCATTI, G.E., LORENZON, A.S., CASTRO, N.L.M., DOMINGUES, G.F., CARVALHO, J.R., SANTOS, A.R. Cadeias de Markov e autômatos celulares para previsão de ambientes sujeitos à desertificação. *Revista de Gestão Ambiental*, v.225, p.160-167. 2018.
- BERGEN, K.J., JOHNSON, P.A., DE HOOP, M.V., BEROZA, G.C. Aprendizado de máquina para descoberta orientada a dados em geociência da Terra sólida. *Ciência*, v.363, p.eaau0323. 2019.
- BREIMAN, L. Florestas Aleatórias. *Aprendizado de Máquina*, v.45, p.5-32. 2001.
- BRUNGARD, C.W., BOETTINGER, J.L., DUNIWAY, M.C., WILLS, S.A., EDWARDS, T.C. Aprendizado de máquina para prever classes de solos em três paisagens semiáridas. *Geoderma*, v.239, p.68-83. 2015.
- CALLISTO, M., GONÇALVES, J.F., LIGEIRO, R. Recursos Hídricos nos Campos Rupestres da Serra do Espinhaço. In: G.W. Fernandes (Orgs.) *Ecologia e Conservação de Campos de Montanha no Brasil*. Cham: Editora Internacional Springer, 2016. p. 87-102.
- CAMPOS, J.A., AIRES, U.R.V., SILVA, D.D.D., CALIJURI, M.L. Fragilidade ambiental e dinâmica da cobertura vegetal no Parque Estadual da Lapa Grande, Mg, Brasil *Anais da Academia Brasileira de Ciências*, v.91, p.e20170940. 2019.
- CORTES, C., VAPNIK, V. Redes de Vetores de Suporte. *Aprendizado de Máquina*, v.20, p.273-297. 1995.
- COSTA, E.M., DOS ANJOS, L.H.C., PINHEIRO, H.S.K., GELSLEICHTER, Y.A., MARCONDES, R.A.T. Redes de crenças espaciais bayesianas: uma abordagem participativa para o mapeamento da vulnerabilidade ambiental no Parque Nacional do Itatiaia, Brasil. *Ciências da Terra Ambiental*, v.79, p.359. 2020.
- COSTA, L.R.F. Considerações Sobre as Macrounidades Geomorfológicas do Estado de Minas Gerais–Brasil. *William Morris Davis – Jornal de Geomorfologia*, v.2, p.11-18. 2021.
- CREPANI, E., MEDEIROS, J.D., HERNANDEZ FILHO, P., FLORENZANO, T.G., DUARTE, V., BARBOSA, C.C.F. Sensoriamento Remoto e Geoprocessamento Aplicados à Zona Ecológico-Econômica e à Ordem Territorial. São José dos Campos: INPE 2001. 124. 2001.
- CRUZ, B.B., MANFRÉ, L.A., RICCI, D.S., BRUNORO, D., APPOLINARIO JR, L., QUINTANILHA, J.A. Marco de fragilidade ambiental para sistemas de abastecimento de água: um estudo de caso na área da macrometrópole paulista (Se Brasil). *Ciências da Terra Ambiental*, v.76, p.441-453. 2017.
- DIAS, S.H.B., FILGUEIRAS, R., FERNANDES FILHO, E.I., ARCANJO, G.S., SILVA, G.H.D., MANTOVANI, E.C., CUNHA, F.F.D. Evapotranspiração de Referência do Brasil Modelada com Técnicas de Aprendizado de Máquina e Sensoriamento Remoto. *PLOS ONE*, v.16, p.e0245834. 2021.
- FARAWAY, J. *Modelos lineares com R*. Chapman Hall. Nova York: Chapman e Hall, 2016. 270. 2016.

- FRANCO, G.B., MARQUES, E.A.G., GOMES, R.L., CHAGAS, C.S., SOUZA, C.M.P., BETIM, L.S. Fragilidade Ambiental da Bacia Hidrográfica do Rio Almada-Bahia. *Revista de Geografia*, v.28, p.187 - 205. 2011.
- GALANTE, J.C., DOWLING, T.I. Um índice multirresolução da planicidade do fundo do vale para mapear áreas de deposição. *Pesquisa de Recursos Hídricos*, v.39, 2003.
- GOMES, L.C., FARIA, R.M., SOUZA, E., VELOSO, G.V., SCHAEFER, C.E.G.R., FERNANDES FILHO, E.I. Modelagem e Mapeamento de Estoques de Carbono Orgânico do Solo no Brasil. *Geoderma*, v.340, p.337-350. 2019.
- GUEVARA, Y.Z.C., SOUZA, J.J.L.L.D., VELOSO, G.V., VELOSO, R.W., ROCHA, P.A., ABRAHÃO, W.A.P., FERNANDES FILHO, E.I. Valores de Referência de Qualidade do Solo para a Bacia do Rio Doce. *Revista Brasileira de Ciência do Solo*, v.42, p.1-16. 2018.
- HASTIE, T., TIBSHIRANI, R. Modelos aditivos generalizados: algumas aplicações. *Revista da Associação Americana de Estatística*, v.82, p.371-386. 1987.
- HEINECK, C.A., LEITE, C.A.S., SILVA, M., VIEIRA, V.S. Mapa Geológico do Estado de Minas Gerais, Escala 1: 1.000.000. Belo Horizonte: Convênio COMIG/CPRM, v.1, 2003.
- HIJMANS, R.J., CAMERON, S.E., PARRA, J.L., JONES, P.G., JARVIS, A. Superfícies climáticas interpoladas de altíssima resolução para áreas terrestres globais. *Revista Internacional de Climatologia*, v.25, p.1965-1978. 2005.
- IGLESIAS, M., UHLEIN, A. Estratigrafia do Grupo Bambuí e coberturas fanerozoóicas no Vale do Rio São Francisco, norte de Minas Gerais. *Revista Brasileira de Geociências*, v.39, p.256-266. 2009.
- KER, J.C. Latossolos do Brasil: Uma Revisão. *Revista Geonomos*, v.5, p.17 - 40. 1997.
- KUHN, M., JOHNSON, K. Modelagem Preditiva Aplicada. Nova Iorque: Springer, 2013. 600. 2013.
- MALONE, B.P., MCBRATNEY, A.B., MINASNY, B., LASLETT, G.M. Mapeamento de Funções de Profundidade Contínua de Armazenamento de Carbono no Solo e Capacidade de Água Disponível. *Geoderma*, v.154, p.138-152. 2009.
- MARTINS, T.I.S., RODRIGUES, S.C. Análise e Mapeamento dos Graus de Fragilidade Ambiental da Bacia do Médio-Baixo Curso do Rio Araguari, Minas Gerais. *Caderno de Geografia (Geografia Cadern)*, v.22, 2012.
- MORELLOS, A., PANTAZI, X.-E., MOSHOU, D., ALEXANDRIDIS, T., WHETTON, R., TZIOTZIOS, G., WIEBENSOHN, J., BILL, R., MOUAZEN, A.M. Previsão baseada em aprendizado de máquina do nitrogênio total, carbono orgânico e teor de umidade do solo usando espectroscopia Vis-NIR. *Engenharia de Biosistemas*, v.152, p.104-116. 2016.
- MOTTA, P.E., CURTI, N., FRANZMEIER, D.P. Relação de solos e superfícies geomórficas no cerrado brasileiro. In: P. Oliveira, R.J. Marquês (Orgs.) *Os Cerrados do Brasil: Ecologia, História Natural de uma Savana Neotropical*. Nova Iorque: Columbia University Press, 2002. p. 13-32.
- MURRAY, A.B., LAZARUS, E., ASHTON, A., BAAS, A., COCO, G., COULTHARD, T., FONSTAD, M., HAFF, P., MCNAMARA, D., PAOLA, C., PELLETIER, J., REINHARDT, L. Geomorfologia, complexidade e a ciência emergente da superfície da Terra. *Geomorfologia*, v.103, p.496-505. 2009.
- NUNES, W.A.G.A., KER, J.C., SCHAEFER, C.E.G.R., FERNANDES FILHO, E.I., GOMES, F.H. Relação Solo-Paisagem-Material de Origem e Gênese de alguns Solos no Domínio do "Mar De Morros", Minas Gerais *Revista Brasileira de Ciência do Solo*, v.25, p.341-354. 2001.
- OLAYA, V., CONRAD, O. Capítulo 12 Geomorfometria em Saga. In: T. Hengl, H.I. Reuter (Eds.) *Desenvolvimentos em Ciência do Solo*. Elsevier, 2009. p. 293-308.
- PADARIAN, J., MINASNY, B., MCBRATNEY, A.B. Machine Learning and Soil Sciences: A Review Aided by Machine Learning Tools. *SOLO*, v.6, p.35-52. 2020.

- PESSOA, P., ATMAN, D., KIMURA, G. Problemas ambientais no Karst da Lagoa Santa. In: A. S. Auler, P. Pessoa (Orgs.) Lagoa Santa Karst: a icônica região cárstica do Brasil. Cham: Editora Springer International, 2020. p. 283-303.
- QUINLAN, J.R. Aprendendo com Aulas Contínuas. (Eds.), Conferência conjunta australiana sobre inteligência artificial, pp. 343-348. 1992.
- RCORE, T. R: Uma Linguagem e Ambiente para Computação Estatística. Viena, Áustria (2016). <http://www.R-project.org/> Data de acesso: 15 Jan 2023. 2023.
- ROSS, J.L.S. Análise Empírica da Fragilidade de Ambientes Naturais Antropizados. Revista do Departamento de Geografia, v.8, p.63-74. 1994.
- SALGADO, A.A.R., BRAUCHER, R., COLIN, F., NALINI, H.A., VARAJÃO, A.F.D.C., VARAJÃO, C.A.C. Taxas de Denudação do Quadrilátero Ferrífero (Minas Gerais, Brasil): Resultados Preliminares de Medidas de Fluxos de Solutos em Rios e Cosmogênicos Produzidos in Situ 10be. Revista de Exploração Geoquímica, v.88, p.313-317. 2006.
- SCOLFORO, J.R., OLIVEIRA, A., CARVALHO, L.D., MARQUES, J.J.G., LOUZADA, J.N., MELLO, C., PEREIRA, J.R., REZENDE, J.B., VALE, L.C.C. Zona Ecológico-Econômica de Minas Gerais. Lavras: UFLA, 2008. 2008.
- SENA, N.C., VELOSO, G.V., FILHO, E.I.F., FRANCELINO, M.R., SCHAEFER, C.E.G.R. Análise de Atributos do Terreno em Diferentes Resoluções Espaciais para Aplicação de Mapeamento Digital de Solos no Sudeste do Brasil. Geoderma Regional, p.e00268. 2020.
- SILVA, L.A.P., SOUZA, C.M.P., SILVA, C.R., BOLFE, E.L., ROCHA, A.M. Projeção dos Impactos das Mudanças Climáticas na Produtividade Primária Líquida da Amazônia Legal – Brasil Caderno de Geografia, v.33, p.1-22. 2023.
- SILVA, L.C.L., OLIVEIRA, F.S., RAMOS, V.D.V., SCHAEFER, C.E.G.R. Pedodiversidade No Estado De Minas Gerais-Brasil. Caderno de Geografia (Geografia), v.28, p.18-38. 2018.
- SIMÕES, P.L., VALADÃO, R.C., OLIVEIRA, C.V., OLIVEIRA, F.S. Uso de Atributos Pedológicos na Integração da Geração de Superfícies Geomorfológicas Escalonadas da Borda Oeste do Planalto do Espinhaço Meridional / Minas Gerais – Brasil. Revista Brasileira de Geomorfologia, v.21, 2020.
- SOUZA, C.M.P., FIGUEIREDO, N.A., COSTA, L.M., VELOSO, G.V., ALMEIDA, M.I.S., FERREIRA, E.J. Algoritmo de Aprendizado de Máquina na Predição de Índices Geomórficos para Avaliação da Influência da Estrutura da Paisagem em Sistemas Fluviais, Sudeste - Brasil. Revista Brasileira de Geomorfologia, v.21, p.363-378. 2020.
- SOUZA, C.M.P., VELOSO, G.V., MELLO, C.R., RIBEIRO, R.P., SILVA, L.A.P., LEITE, M.E., FERNANDES FILHO, E.I. Previsão Espaço-Temporal da Erosividade das Chuvas por Machine Learning, Minas Gerais, Brasil. Geocarto Internacional, v.37, p.1-19. 2022.
- SPÖRL, C., CASTRO, E., LUCHIARI, A. Aplicação de Redes Neurais Artificiais na Construção de Modelos de Fragility Ambiental. Revista do Departamento de Geografia, v.21, p.113-135. 2011.
- SPÖRL, C., ROSS, J.L.S. Análise Comparativa da Fragilidade Ambiental com Aplicação de Três Modelos. GEOUSP: Espaço e Tempo (Online), p.39-49. 2006.
- TRICART, J. Ecodinâmica. Rio de Janeiro: IBGE, 1977. 91. 1977.
- UFV, CETEC, UFLA, FEAM. Mapa de Solos do Estado de Minas Gerais - Escala 1: 500.000. pág. 49. 2010.
- URRIBAGO-OSPINA, L.M., JARDIM, C.M., RIVERA-FERNÁNDEZ, G., KOZOVITS, A.R., LEITE, M.G.P., MESSIAS, M.C.T.B. Conhecimento ecológico tradicional na gestão de ecossistemas ferruginosos: lições para diversificar o uso da terra. Meio Ambiente, Desenvolvimento e Sustentabilidade, v.23, p.2092-2121. 2021.
- USGS. Earthexplorer Disponível em: Data de acesso: 27 de janeiro de 2023. 2023.

- WANG, Z., SHI, W., ZHOU, W., LI, X., YUE, T. Comparação de transformações log-ratio aditivas e isométricas combinadas com modelos de aprendizado de máquina e krigagem de regressão para mapeamento de frações granulométricas do solo. *Geoderma*, v.365, p.114214. 2020.
- FORAM, K., BUI, D.T., DICK, Ø.B., SINGH, B.R. Uma Avaliação Comparativa de Regressão Vetorial de Suporte, Redes Neurais Artificiais e Florestas Aleatórias para Prever e Mapear Estoques de Carbono Orgânico do Solo em uma Paisagem Afromontana. *Indicadores Ecológicos*, v.52, p.394-403. 2015.