

XXVI SIMPÓSIO BRASILEIRO DE RECURSOS HIDRÍCOS

MODELAGEM PREDITIVA DA SUSCETIBILIDADE À FORMAÇÃO DE RAVINAS NA UGRHI 05 USANDO *RANDOM FOREST*

Luca C. D. Marote¹; Mateus Domingos²; Igor P. S. Santana³; Aleksey Y. Sheshukov⁴ & Murilo Cesar Lucas⁵

Abstract: This study aims to develop a predictive model of gully formation susceptibility in the Water Resources Management Unit 05 (UGRHI 05), located between the States of São Paulo and Minas Gerais, encompassing major river basins and areas with intense agricultural and urban activity. The methodology involved the identification and inventory of erosive features through visual interpretation of satellite imagery, generation of balanced samples (gullies and non-gullies), selection of conditioning geo-environmental factors, and application of the Random Forest (RF) algorithm. The factors used included topographic, hydrological, geological, geomorphological, and land use/land cover attributes, all adjusted to a 30-meter spatial resolution. The RF model achieved high accuracy (above 95%) and AUC = 99.5%, with land use/land cover standing out as the most influential predictive factor (55.2%), followed by elevation, geology, and SPI. The resulting susceptibility map enabled the identification of areas at high risk of gully formation, which were later validated using high-resolution imagery, confirming the model's ability to recognize both active gullies and less advanced stages of land degradation. The results demonstrate the effectiveness of this approach in supporting land-use planning and soil conservation strategies, highlighting its potential as a preventive environmental management tool in erosion-prone regions.

Resumo: Este estudo tem como objetivo elaborar um modelo preditivo da suscetibilidade à formação de ravinas na Unidade de Gerenciamento de Recursos Hídricos 05 (UGRHI 05), localizada entre os Estados de São Paulo e Minas Gerais, abrangendo importantes bacias hidrográficas e áreas com intensa atividade agrícola e urbana. A metodologia envolveu a identificação e inventário de feições erosivas por interpretação visual de imagens de satélite, geração de amostras balanceadas (ravinas e não-ravinas), seleção de atributos geoambientais condicionantes e aplicação do algoritmo *Random Forest* (RF). Os atributos utilizados foram topográficos, hidrológicos, geológicos, geomorfológicos e de uso e cobertura do solo. Todos eles foram ajustados para resolução espacial de 30 metros. O modelo RF obteve alta acurácia (> 95%) e AUC=99,5%, com destaque para o uso e cobertura do solo como principal fator preditivo (55,2%), seguido pela elevação, geologia e SPI. O mapa de suscetibilidade gerado permitiu identificar áreas com elevada suscetibilidade de formação de ravinas. Os resultados foram validados com imagens de alta resolução, evidenciando a capacidade do modelo em reconhecer ravinas ativas e em formação. As descobertas mostram a eficácia do uso de aprendizado de máquina para dar suporte ao planejamento territorial e ações de conservação do solo, sendo uma ferramenta promissora para a gestão ambiental.

1) Afiliação: Faculdade de Tecnologia, Unicamp, R. Paschoal Marmo, 1888, Jardim Nova Itália, 13.484-332, Limeira – SP, i184698@dac.unicamp.br

2) Afiliação: Faculdade de Tecnologia, Unicamp, R. Paschoal Marmo, 1888, Jardim Nova Itália, 13.484-332, Limeira – SP, m291018@dac.unicamp.br

3) Afiliação: Faculdade de Tecnologia, Unicamp, R. Paschoal Marmo, 1888, Jardim Nova Itália, 13.484-332, Limeira – SP, i198643@dac.unicamp.br

4) Afiliação: Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, KS, United States of America, ashesh@ksu.edu

5) Afiliação: Faculdade de Tecnologia, Unicamp, R. Paschoal Marmo, 1888, Jardim Nova Itália, 13.484-332, Limeira – SP, mlucas@unicamp.br

Palavras-Chave – Aprendizado de máquina; Conservação do solo; Segurança hídrica

INTRODUÇÃO

Dentre as diversas formas de degradação ambiental, a erosão do solo destaca impactos profundos na biodiversidade e fertilidade do solo, que comprometem os serviços ecológicos fundamentais para a produção de alimentos e propiciam o carreamento de sedimentos e o assoreamento de corpos hídricos e reservatórios. Além disso, a perda de carbono orgânico do solo e a redução na capacidade de sequestro de carbono atmosférico direcionam ainda mais à ocorrência de mudanças climáticas e alterações hidrológicas globais devido à erosão do solo (Panagos et al., 2022; Schwambach et al., 2024).

Os principais atributos que aumentam a susceptibilidade de erosão do solo estão associados à capacidade erosiva da precipitação, ao uso e ocupação da terra, às características topográficas e o tipo de solo do local. Conforme algumas combinações desses atributos, os padrões erosivos do escoamento superficial se intensificam gradativamente até formar sulcos e pequenos canais, cuja evolução atinge estágios de erosão permanentes com profundidades de 0,5 até 30 metros, classificando-os como ravinas e voçorocas. Esse tipo de padrão erosivo intensifica ainda mais o fluxo de sedimentos e a perda de solo, de modo a acentuar os impactos nos recursos hídricos e na agricultura (Douglas-Mankin et al., 2020).

Nesse contexto, diversos estudos têm sido direcionados a realizar previsões e estimativas da erosão do solo e do fluxo de sedimentos em ravinas e voçorocas, incluindo a aplicação de complexos modelos hidrológicos de base física, tais como WEPP (*Water Erosion Prediction Project*), REGEM (*Revised Ephemeral Gully Erosion Model*), entre outros (Douglas-Mankin et al., 2020). No entanto, tais modelos requerem uma grande quantidade e detalhamento de dados medidos em campo e em laboratório, que são escassos e difíceis de serem obtidos, em sua maior parte, o que limita o uso desses modelos em muitas áreas de estudo (Mohebzadeh et al., 2022).

Por outro lado, os dados de sensoriamento remoto são eficientes na obtenção de muitos dos atributos que levam à erosão do solo, incluindo variáveis topográficas, classificações de ocupação da terra, propriedades de solos e formações geológicas, além de índices de erosividade da chuva. Tais atributos podem ser obtidos a partir de processamento de modelos digitais de elevação e imagens de satélite, dados de estações pluviométricas, entre outros arquivos georreferenciados (De Faria Godoi et al., 2021; Mohebzadeh et al., 2022; Oliveira-Roza et al., 2024). Neste sentido, a aplicação de algoritmos de aprendizado de máquina em variáveis geoambientais relacionadas à erosão do solo tem ganhado destaque na literatura para identificar a suscetibilidade de ocorrência de ravinas e voçorocas em paisagens ao redor do mundo.

Por exemplo, Pal et al. (2022) aplicaram os modelos *Random Forest* (RF), *Gradient Boost*, *Extreme Gradient Boost* e *Support Vector Machine* em uma Bacia Hidrográfica ao leste da Índia, identificando 18 atributos relacionados à erosão do solo para investigar sua predição da suscetibilidade de ocorrer ravinas e voçorocas. Na China, Liu et al. (2025) combinaram três diferentes modelos em um *ensemble* para avaliar 25 variáveis na probabilidade de ocorrência das ravinas. Han et al. (2022) realizaram uma predição semelhante em áreas agrícolas de Illinois (EUA), identificando 15 atributos mais representativos dos 28 atributos totais, por meio do modelo de Entropia Máxima (MaxEnt). No Brasil, Lana et al. (2022) compararam os modelos RF, Naïve Bayes, Logistic Regression e Artificial Neural Network em duas bacias hidrográficas de Minas Gerais, utilizando como atributos 15 variáveis geoambientais condicionantes. Os resultados evidenciaram que o melhor desempenho em identificar as feições erosivas e o melhor mapa de suscetibilidade foi atribuído ao modelo RF. Filho et al. (2024) avaliaram a probabilidade de ocorrência de ravinas na Bacia

Hidrográfica do Rio Pirai (RJ), utilizando os modelos *Classification and Regression* (CART), *XGBoost*, RF, e SVM, por meio de 14 atributos ambientais, majoritariamente topográficos.

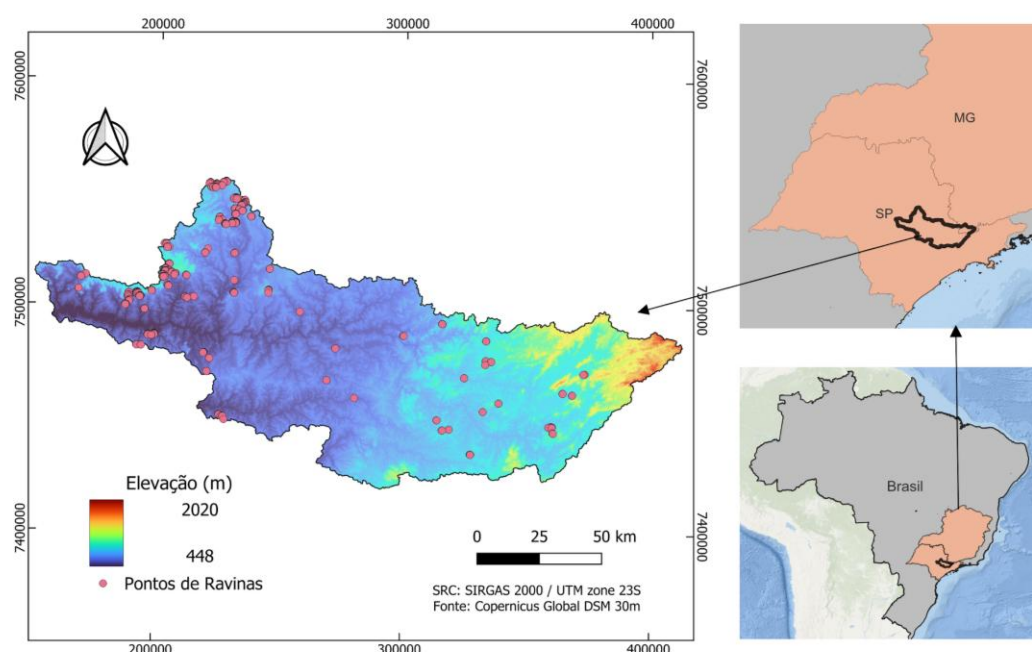
Dessa forma, o objetivo deste estudo é utilizar o modelo de aprendizado de máquina RF para identificar a suscetibilidade à formação de ravinas na Unidade de Gerenciamento de Recursos Hídricos 05.

ÁREA DE ESTUDO

A Unidade de Gerenciamento de Recursos Hídricos 05 (UGRHI 05) engloba três bacias hidrográficas principais, cuja rede de drenagem corresponde aos rios Piracicaba, Capivari, Jundiaí e seus afluentes. O território das bacias da UGRHI 05 compreende 71 municípios do estado de São Paulo e outros cinco municípios de Minas Gerais, totalizando uma população de aproximadamente 5,85 milhões de habitantes e uma área de 15.377 km². As principais cidades localizadas na UGRHI 05 são Campinas, Piracicaba, Jundiaí, Limeira, Americana, Sumaré, Santa Bárbara d'Oeste e Rio Claro (Basso et al., 2017).

A UGRHI 05 possui uma intensa atividade econômica voltada para a agricultura e transporte de cargas, além da crescente e expressiva taxa de urbanização (Comitês PCJ, 2020). A ocorrência de feições erosivas como ravinas e voçorocas é notável, de forma que sua identificação e avaliação de susceptibilidade são pertinentes para o planejamento territorial e práticas de conservação do solo. O mapa a seguir (Figura 1) ilustra a área da UGRHI 05, incluindo os pontos nos locais identificados com ravinas.

Figura 1 – Mapa de localização da área de estudo, destacando os pontos em que há presença de ravinas.



METODOLOGIA

O desenvolvimento deste trabalho seguiu o seguinte fluxo metodológico: i) identificação das feições erosivas na área de estudo e produção de um inventário de amostras; ii) levantamento de

atributos geoambientais condicionantes à erosão; iii) produção da base de dados para treino e teste; iv) avaliação da importância das variáveis para a classificação; v) avaliação da performance do RF na classificação das amostras; vii) aplicação do modelo treinado na área de estudo e avaliação da sua predição.

Inventário de amostras de ravinas

Inicialmente, as feições erosivas de ravinas e voçorocas foram identificadas visualmente pelas imagens de satélite de alta resolução disponibilizadas pelos serviços do Google Satellite no *software* QGIS 3.36.3, de modo a registrar amostras espacialmente diversificadas ao longo da área de estudo e capturar as diversidades geográficas e ambientais presentes (Figura 1). Tendo em vista o propósito de gerar um modelo capaz de identificar áreas suscetíveis a processos erosivos avançados, optou-se por registrar somente os “canais” erosivos notavelmente profundos e largos (acima de 3 metros de largura), embora a resolução do modelo digital de elevação utilizado não tenha sido suficiente para avaliar as profundidades reais destas feições.

Para treinar o algoritmo de aprendizado de máquina adequadamente, realizou-se um levantamento proporcional de amostras em áreas onde não ocorrem ravinas e feições erosivas, considerando a variabilidade topográfica e de cobertura do solo representativas da área de estudo. Além disso, as amostras de áreas sem erosão foram tomadas a uma distância mínima de 2 km das amostras positivas, com a pretensão de evitar o treinamento equivocado do modelo na distinção de áreas suscetíveis, tal como feito por Lana et al. (2022).

Para tanto, foram construídos polígonos georreferenciados sobre os locais amostrados, transformando-os, posteriormente, em *pixels* de 30 metros de resolução com presença e ausência de ravinas. Assim, utilizando as bibliotecas “Rasterio” e “Pandas” da linguagem Python, foram sorteadas 672 amostras de cada classe para compor uma base balanceada (Mohebzadeh et al., 2022).

Atributos geoambientais

A definição de atributos geoambientais representativos dos processos erosivos é fundamental para elaborar modelos e mapas de suscetibilidade assertivos, embora não haja um protocolo padronizado na literatura para definição destes atributos. De modo geral, os estudos feitos em diferentes áreas do mundo consideram atributos topográficos (primários e secundários), hidrológicos, antrópicos, pedológicos, geológicos, geomorfológicos e de cobertura do solo (Lana et al., 2022; Mohebzadeh et al., 2022).

Dessa forma, as variáveis geoambientais consideradas neste estudo foram selecionadas com base na sua recorrência em outros estudos reunidos por Mohebzadeh et al. (2022). Devido à influência topográfica na dinâmica do escoamento superficial e dos processos erosivos, foram considerados diferentes atributos derivados do modelo digital de elevação (Copernicus Global DSM – 30 m), sendo estes primários, como elevação, declividade, curvatura plana, curvatura de perfil e direção de fluxo, e outros secundários, como o Índice de Potência de Fluxo (SPI, do inglês, *Stream Power Index*) e o Índice de Umidade Topográfica (*Topographic Wetness Index*, TWI) (Filho et al., 2024).

Como fator hidrológico, considerou-se a erosividade da precipitação pelo Índice de Fournier Modificado (MFI, do inglês, *Modified Fournier Index*), uma vez que este índice é amplamente explorado nos estudos sobre erosão do solo no Brasil para calibrar o fator de erosividade (R) da Equação Universal de Perda de Solo (USLE), a partir de dados de precipitação diários (e.g., Almagro

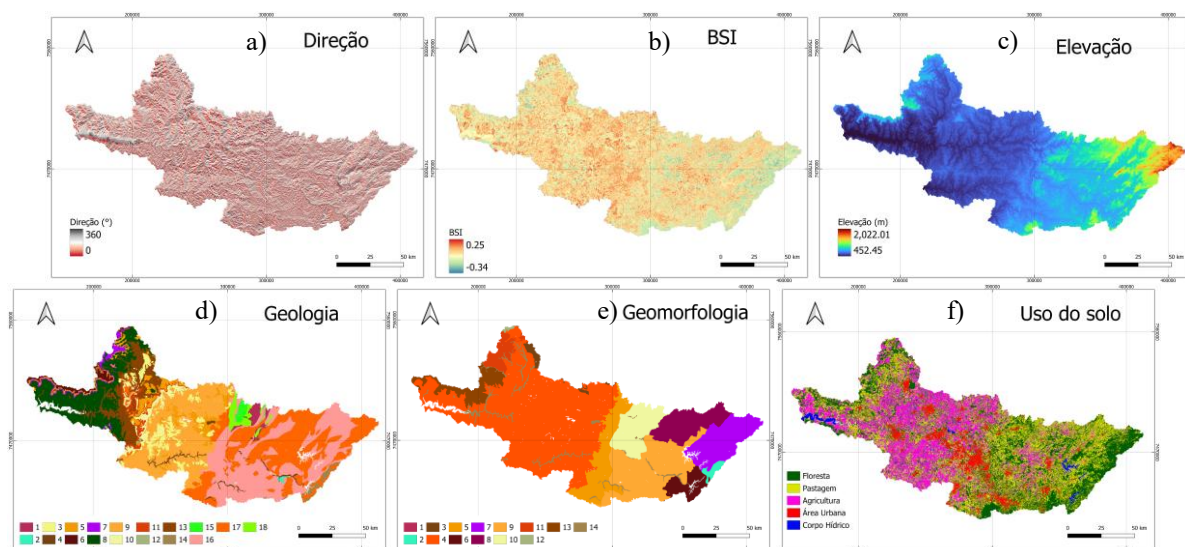
et al., 2017; Schwambach et al., 2024). Neste estudo, o valor de MFI foi calculado pelos dados de precipitação em grade (10 km^2) gerados por Xavier *et al.* (2016), utilizando a média de 21 anos de precipitação (2003 a 2023), ressaltando que este conjunto de dados foi atualizado em março de 2025.

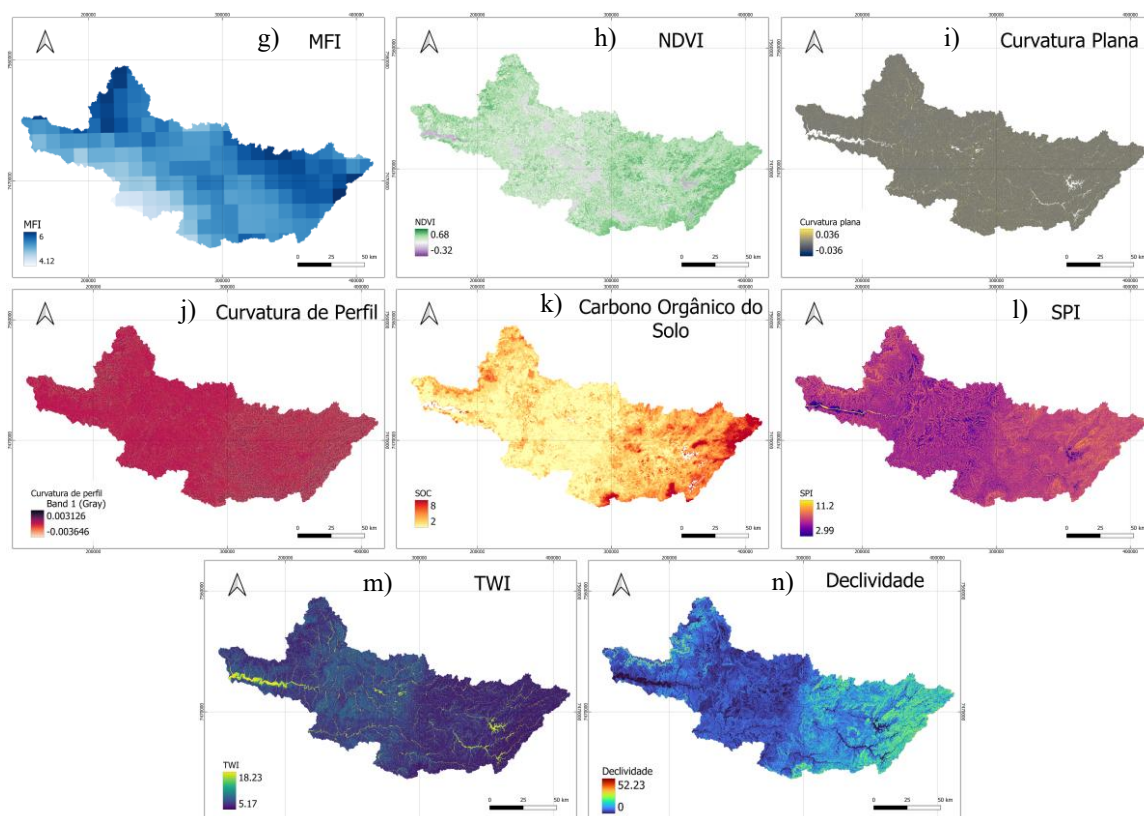
Os atributos geológicos e geomorfológicos foram considerados neste estudo para alimentar o modelo de classificação, tendo em vista sua influência nas propriedades físico-químicas dos solos, que podem estar intimamente ligadas à erosão e formação de ravinas. Tais atributos categóricos foram adquiridos do Banco de Dados e Informações Ambientais (BDiA) do Instituto Brasileiro de Geografia e Estatística (IBGE). As divisões geológicas e geomorfológicas presentes na UGRHI 05 totalizam 18 e 14 diferentes categorias, respectivamente.

As características dos solos consideradas como atributos condicionantes à erosão incluíram: a concentração de Carbono Orgânico do Solo (*Soil Organic Carbon*, SOC), em 5g/kg , que está disponível nos catálogos da plataforma Google Earth Engine (GEE) pelo OpenLandMap na resolução de 250 m , além do Índice de Solo Exposto (*Bare Soil Index*, BSI), calculado pelas bandas das imagens do satélite Sentinel-2 (resolução de 10 m), também obtidas pelo GEE. Além disso, o uso do solo foi obtido pela classificação ESA World Cover (2021), que utiliza imagens do Sentinel-1 e Sentinel-2, e o Índice de Vegetação por Diferença Normalizada (*Normalized Difference Vegetation Index*, NDVI) foi calculado no GEE pelas bandas do Sentinel-2 (2024), ambos nas resoluções de 10 m .

A partir da coleção de todos os atributos geoambientais supracitados (Figura 2), foi necessário realizar o ajuste e alinhamento de todas as resoluções para coincidir com a unidade de georreferenciamento de 30 m dos arquivos derivados do MDE e dos pixels do inventário de amostras. Desse modo, foi possível capturar os valores exatos de todos os atributos nos pixels amostrados, utilizando as bibliotecas “Rasterio” e “Pandas” da linguagem de programação Python, de forma completar a base de dados para o treinamento e validação do modelo.

Figura 2 – Atributos geoambientais considerados para a área de estudo: a) direção do escoamento superficial; b) BSI; c) elevação do terreno; d) geologia; e) geomorfologia; f) uso e ocupação do solo; g) MFI; h) NDVI; i) curvatura plana; j) curvatura de perfil; k) carbono orgânico do solo; l) SPI; m) TWI; n) declividade.





É importante ressaltar que foi feita a avaliação da multicolinearidade entre as variáveis independentes, calculando-se o Fator de Inflação da Variância (*Variance Inflation Factor*, VIF), uma vez que a correlação destas variáveis está relacionada ao ganho ou prejuízo de desempenho dos modelos. Considerando o limite de $VIF \leq 10$ para descartar a multicolinearidade entre as variáveis, conforme descrito por (Liu et al., 2025), foram calculados os valores de VIF para cada variável. Embora os resultados tenham sido omitidos deste trabalho, nenhuma das variáveis descritas neste estudo apresentou VIF acima do limite aceitável.

Treinamento e validação do modelo RF

Após o preparo da base de dados com as variáveis independentes (atributos geoambientais) e a variável-alvo dependente (classificação binária de ravinas e não-ravinas), utilizou-se a biblioteca “*Scikit-learn*” da linguagem Python para separar a base de dados em conjuntos de treino (70%) e teste (30%), conforme usualmente adotado por outros autores na literatura (e.g., Mohebzadeh et al., 2022). Em seguida, identificou-se a importância das variáveis pela ferramenta *feature importance* do RF, sendo este um modelo amplamente utilizado para treinar de dados geoambientais relacionados à erosão, devido a sua robustez e desempenho, além de não necessitar de processamentos adicionais das variáveis contínuas e categóricas (Lana et al., 2022).

A otimização dos hiperparâmetros componentes do RF foi feita pela ferramenta *GridSearchCV*, variando o número de árvores (10, 20, 50, 100, 200 e 500), a profundidade máxima das árvores (3, 5, 10, 20 e 50), o mínimo de amostras para a divisão dos nós (2, 5 e 10) e o mínimo de amostras na folha (1, 2 e 4). Assim, para evitar o viés e a variância do método na otimização do modelo, adotou-se a validação cruzada no conjunto de treino, isto é, dividindo-o em k partes (neste

caso, adotou-se $k = 10$), de modo que $k - 1$ divisões são utilizadas para o treinamento e a divisão restante é utilizada para validação, totalizando k diferentes interações do modelo com os dados (Lana et al., 2022). Por fim, a validação final do modelo foi feita no conjunto de teste separado no início.

As métricas consideradas para avaliar a classificação do RF foram adotadas com base no número de acertos e erros das classes positivas e negativas (*True Positive*, TP; *True Negative*, TN; *False Positive*, FP; e *False Negative*, FN). Assim, foram obtidas as métricas de acurácia, precisão, recall e f1-score, cujas equações estão descritas a seguir (Filho et al., 2024).

Elaboração do mapa de suscetibilidade

Após o treinamento, otimização, validação do RF na base de dados, a elaboração do mapa de suscetibilidade foi feita aplicando-se o modelo para classificar e obter a probabilidade de cada pixel ser considerado positivo (ravina) na área de estudo. Para tanto, utilizou-se as bibliotecas “Rasterio”, “Os” e “JenksNaturalBreaks” da linguagem Python, dividindo os valores de probabilidade em 5 categorias de suscetibilidade, de acordo com as quebras naturais de agrupamento dos dados, uma vez que determina arranjo com a menor variância possível dos valores em diferentes classes (Lana et al., 2022).

RESULTADOS E DISCUSSÃO

O modelo de classificação das amostras teve seus parâmetros otimizados em 50 árvores com profundidade máxima igual a 20, o mínimo de duas amostras para a divisão dos nós e uma amostra na folha. Assim, o desempenho desta configuração de RF mostrou resultados excelentes na predição das amostras do conjunto de treino, cujas métricas consideradas evidenciaram valores acima de 97% e baixo desvio padrão (<5%). Um desempenho apreciável da classificação também foi obtido para o conjunto de teste (>95%) (Quadro 1).

Quadro 1 – Métricas de classificação do modelo nos conjuntos de treino e teste.

Métrica	Treinamento	Teste
Acurácia	0,977 ± 0,011	0,960
F1-score	0,977 ± 0,011	0,960
AUC	0,997 ± 0,002	0,995

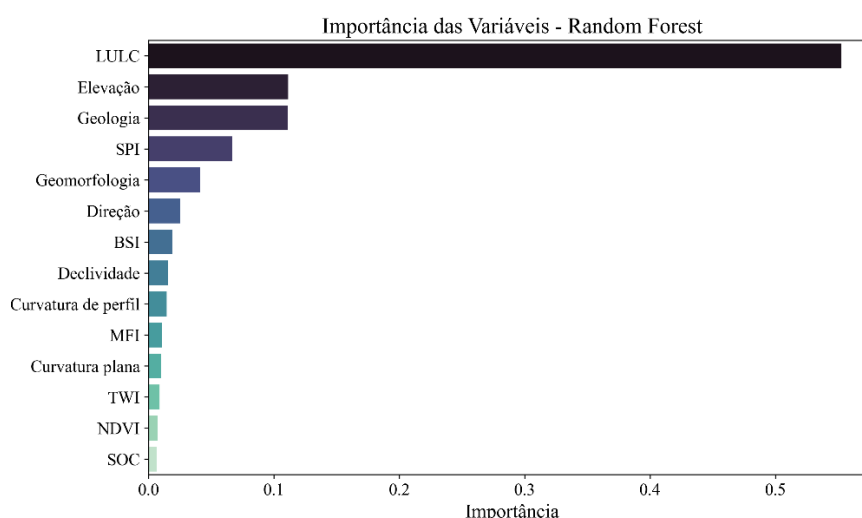
A partir da confiabilidade atrelada ao modelo treinado e otimizado, os resultados obtidos na avaliação da importância das variáveis demonstraram que a maior influência (55,2%) na classificação das amostras é atribuída ao tipo de uso do solo (LULC), cujas categorias (floresta, agricultura, pastagem/vegetação rasteira, área urbana e corpo hídrico) demonstram que as ravinas dependem primordialmente de um fator antrópico, possivelmente devido à vulnerabilidade das áreas com menor proteção e cobertura do solo, como pastagens e vegetações rasteiras, além das áreas agrícolas.

Na sequência, a elevação e a geologia do local apresentaram importâncias moderadas (aproximadamente 11%) na ocorrência das ravinas, o que pode estar atribuído ao maior número de amostras positivas estarem na porção Oeste da área de estudo, onde a elevação é menor e mais próxima do rio principal da Bacia Hidrográfica, e a geologia apresenta Arenito Piramboia e Arenitos de coloração avermelhada. O SPI demonstrou ter 6,6% de importância, uma vez que está associado à maior probabilidade de desenvolvimento de canais pelo poder de fluxo do escoamento superficial.

As categorias geomorfológicas apresentaram baixa importância (4,2%), seguidas da direção do fluxo (2,5%), índice de solo exposto (1,9%), declividade (1,57%) e curvatura de perfil (1,45%).

As demais variáveis apresentaram relevância muito baixa na ocorrência das feições erosivas, sendo 1,06% correspondente à erosividade da precipitação e as demais (curvatura plana, TWI, NDVI e a concentração de carbono orgânico do solo) com importância inferior a 1%. A imagem a seguir ilustra a comparação entre as importâncias de cada variável (Figura 3).

Figura 3 – Importância dos atributos geoambientais para a ocorrência de ravinas na área de estudo.

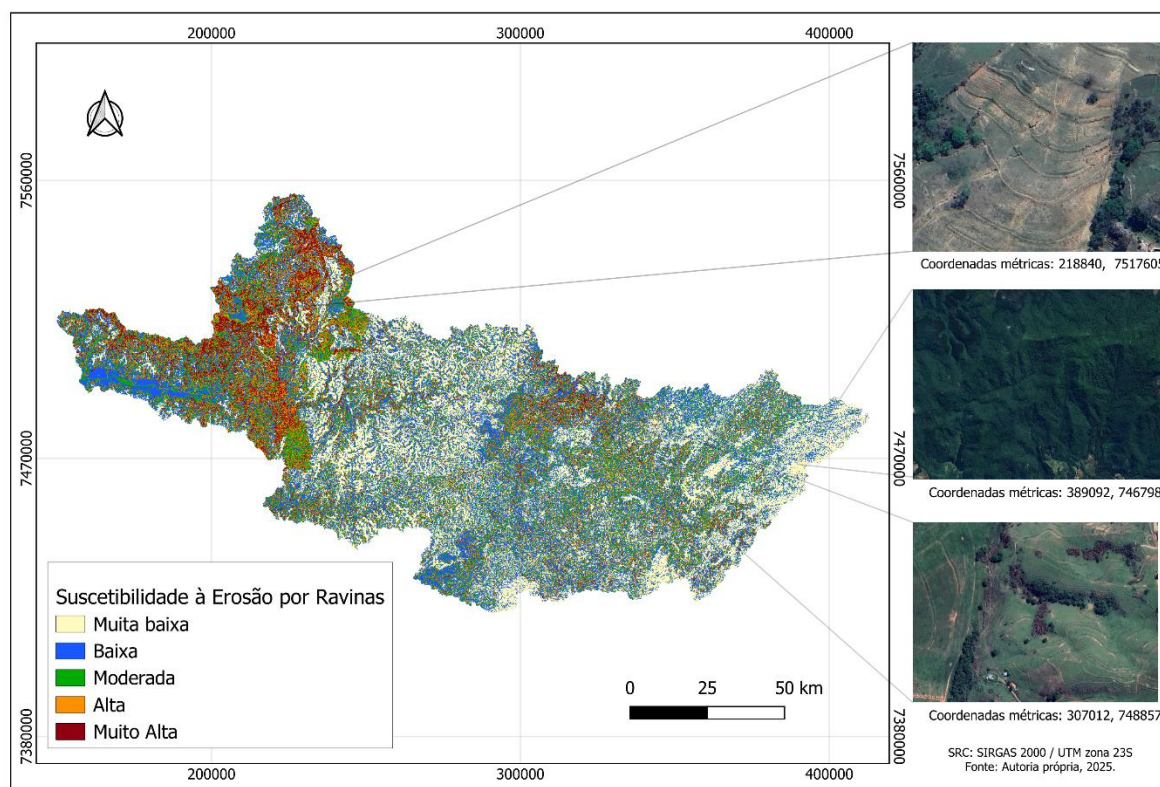


Ao ser confrontado com os padrões geoespaciais da área de estudo, o modelo RF identificou a probabilidade de cada pixel ser suscetível à formação de ravinas, classificando-os em 5 classes de acordo com as quebras naturais dos valores. Dessa forma, obteve-se um mapa de suscetibilidade com resolução de 30 m, destacando algumas áreas com elevado risco de haver erosão, estas que necessitam de maior atenção a respeito de práticas de proteção e conservação dos solos contra os efeitos erosivos do escoamento superficial.

A fim de realizar uma validação visual das áreas indicadas com suscetibilidade muito alta à formação de ravinas, as imagens de Google Satellite do software QGIS 3.36.3 foram aproximadas e investigadas para verificar a real situação da erosão nestes locais. Dessa forma, foi possível afirmar que o modelo capturou eficientemente os locais onde há expressiva concentração de ravinas, tanto mais avançadas, quando em estágios anteriores, como algumas porções de solo degradado que ainda não demonstram as “rachaduras” características das amostras positivas. Esta ampla capacidade do modelo em avaliar, não somente os canais erosivos mais avançados, mas também a suscetibilidade à erosão em estágios antecessores, é muito útil na ótica do planejamento territorial e das práticas protecionistas, tendo em vista que estes locais merecem atenção particular para evitar a evolução do transporte de sedimentos e o estabelecimento de ravinas e voçorocas permanentes.

Analogamente, as áreas identificadas com suscetibilidade muito baixa também puderam ser validadas nas imagens de satélite, de modo que correspondem, em sua maior parte, às áreas urbanizadas (com impermeabilização do solo) e de florestas densas, ambas com cobertura que dificultam a erosão pluvial e o escoamento superficial. O mapa de suscetibilidade está disposto na Figura 4, incluindo duas imagens de satélite capturadas em locais com suscetibilidade muito alta (ravinas expressivas) e outra imagem de um local classificado com suscetibilidade muito baixa (floresta densa).

Figura 4 – Mapa de suscetibilidade à erosão por ravinas da área de estudo. Três imagens foram adicionadas a partir das imagens de satélite do Google Earth.



CONCLUSÃO

O presente estudo constatou uma aplicação satisfatória do modelo de aprendizado de máquina Random Forest na identificação da suscetibilidade à ocorrência de erosão por ravinas e voçorocas para a Unidade de Gerenciamento de Recursos Hídricos 05, correspondente às Bacias Hidrográficas dos rios Piracicaba, Capivari e Jundiá. A classificação das amostras pelo modelo demonstrou métricas excelentes, atingindo acurácia e f1-score de 96,0% e AUC = 99,5%, destacando seu ótimo desempenho em distinguir as amostras positivas e negativas utilizadas para treino e validação, bem como a capacidade de identificar o fator antrópico de uso do solo como o mais importante para a distinção de áreas mais ou menos suscetíveis.

Portanto, tendo em vista os impactos diretos nos recursos hídricos atrelado a esse tipo de erosão, como o carregamento acentuado de sedimentos e assoreamento dos corpos hídricos, o mapa de suscetibilidade produzido demonstra alta utilidade em termos de planejamento territorial em políticas públicas, pois pode ser consultado na tomada de decisões para estabelecer práticas contra a degradação do solo e dos recursos hídricos.

AGRADECIMENTOS

Este trabalho contou com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, do Conselho Nacional de

Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento via Projeto CNPq Universal (Processo nº 403292/2023-9) e do Fundo de Apoio ao Ensino, à Pesquisa e à Extensão (FAEPEX) da UNICAMP (Processo nº 3111/23). O primeiro autor agradece à CAPES pela concessão de bolsa de mestrado e aos demais envolvidos na elaboração desta pesquisa: Bruna Carolina Rodrigues, Guilherme Henrique Rodrigues e Michel Santos Queiroz (Programa de Pós-Graduação em Tecnologia da Unicamp).

REFERÊNCIAS

- ALMAGRO, A., OLIVEIRA, P. T. S., NEARING, M. A., & HAGEMANN, S. (2017). "Projected climate change impacts in rainfall erosivity over Brazil". *Scientific Reports*, 7(1).
- DE FARIA GODOI, R.; RODRIGUES, D. B. B.; BORRELLI, P.; TARSO, P.; OLIVEIRA, S. (2021). "High-resolution soil erodibility map of Brazil". *Science of the Total Environment*, v. 781, 146673.
- DOUGLAS-MANKIN, K. R.; ROY, S. K.; SHESHUKOV, A. Y.; BISWAS, A.; GHARABAGHI, B.; BINNS, A.; RUDRA, R.; SHRESTHA, N. K.; DAGGUPATI, P. (2020). "A comprehensive review of ephemeral gully erosion models". *Catena*, v. 195, 104901.
- FILHO, J. da P. M.; GUERRA, A. J. T.; CRUZ, C. B. M.; JORGE, M. do C. O.; BOOTH, C. A. (2024). "Machine Learning Models for the Spatial Prediction of Gully Erosion Susceptibility in the Pirai Drainage Basin, Paraíba Do Sul Middle Valley, Southeast Brazil". *Land*, v. 13, n. 10, 1665.
- HAN, J.; GUZMAN, J. A.; CHU, M. L. (2022). "Gully erosion susceptibility considering spatiotemporal environmental variables: Midwest U.S. region". *European Journal of Remote Sensing*, v. 43, 101196.
- LANA, J. C.; CASTRO, P. de T. A.; LANA, C. E. (2022). "Assessing gully erosion susceptibility and its conditioning factors in southeastern Brazil using machine learning algorithms and bivariate statistical methods: A regional approach". *Geomorphology*, v. 402, 108159.
- LIU, C.; FAN, H.; WANG, Y. (2025). "Predictive performance and uncertainty analysis of ensemble models in gully erosion". *International Soil and Water Conservation Research*, v. 13, n. 2, p. 319.
- MOHEBZADEH, H.; BISWAS, A.; RUDRA, R.; DAGGUPATI, P. (2022). "Machine Learning Techniques for Gully Erosion Susceptibility Mapping: A Review". *Geosciences*, v. 12, n. 12, 0429.
- OLIVEIRA-ROZA, M. P.; CECÍLIO, R. A.; TEIXEIRA, D. B. S.; MOREIRA, M. C.; ALMEIDA, A. Q.; XAVIER, A. C.; ZANETTI, S. S. (2024). "Rainfall Erosivity over Brazil: A Large National Database". *Data*, v. 9, n. 10, 0120.
- PAL, S.; PAUL, S.; DEBANSHI, S. (2022). "Identifying sensitivity of factor cluster based gully erosion susceptibility models". *Environmental Science and Pollution Research*, v. 29, pp. 90964–90983, 22063-3.
- PANAGOS, P.; BORRELLI, P.; MATTHEWS, F.; LIAKOS, L.; BEZAK, N.; DIODATO, N.; BALLABIO, C. (2022). "Global rainfall erosivity projections for 2050 and 2070". *Journal of Hydrology*, v. 610, 127865.
- SCHWAMBACK, D.; AMORIM BRANDÃO, A. R.; BERTOTTO, L. E.; BERNDTSSON, R.; ZHANG, L.; WENDLAND, E.; PERSSON, M. (2024). "Quantifying Soil Loss in the Brazilian Savanna Ecosystem: Current Rates and Anticipated Impact of Climate Changes". *Land Degradation & Development*, v. 35, pp 5786-5803, 5331.
- XAVIER, A. C., KING, C. W. & SCANLON, B. R. "Daily gridded meteorological variables in Brazil (1980-2013)", *International Journal of Climatology*. v.36(6), pp. 2644-2659.