



MAPEAMENTO DE SUSCEPTIBILIDADE ÀS EROSÕES LINEARES EM ÁREAS RURAIS NO MUNICÍPIO DE TEODORO SAMPAIO (SP) E MUNICÍPIOS ADJACENTES ATRAVÉS DO USO DE MACHINE LEARNING

João Osvaldo Rodrigues Nunes ¹
Glauber Verner Firmino ²

RESUMO

As erosões lineares são processos intensos de degradação dos solos. Pesquisas de mapeamento de suscetibilidade às erosões lineares (GESM) baseados em aprendizado de máquina permitem estimar o potencial e a probabilidade de uma dada paisagem ser acometida por esses processos, bem como compreender seus padrões espaciais. Apesar de seu potencial, ainda se encontram pouco difundidas no Brasil. A presente pesquisa realizou o mapeamento de suscetibilidade às erosões lineares em áreas rurais no extremo Oeste Paulista, que abrangem áreas no município de Teodoro Sampaio (SP) e partes menores de municípios adjacentes (Mirante do Paranapanema, Euclides da Cunha Paulista, Marabá Paulista e Presidente Epitácio), as quais apresentam grande quantidade de voçorocas e ravinas. A metodologia se baseou na 1) elaboração de um mapa inventário de erosões lineares e no mapeamento temático dos fatores ambientais; 2) seleção de atributos, separação do conjunto de dados em treino e teste na proporção 70-30%; 3) modelagem da suscetibilidade a partir dos algoritmos Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM) e Redes Neurais Artificiais Multilayer Perceptron (MLP); e 4) avaliação e seleção de modelo de maior desempenho. Além disso, foram realizados trabalhos de campo sistemáticos para a análise da paisagem e para a elaboração e validação dos mapas temáticos produzidos. Os valores de desempenho indicaram desempenho excelente ($AUC = 0,9$ a $1,0$) para os modelos MLP ($0,93$) e RF ($0,92$), e muito bom ($0,8$ a $0,9$) para os modelos LR ($0,89$) e SVM ($0,89$). Constataram-se menores valores de susceptibilidade em áreas de formação florestal, padrões de relevo de menores declividades e densidade de drenagem e maior convexidade. Maiores valores ocorreram em áreas de pastagem e proximidade de áreas urbanas, relevos côncavos, planícies e alvéolos, vertentes e rupturas de declive. Os setores norte e nordeste, caracterizados por padrão de relevo associado ao grupo litológico Bauru, com maior declividade e densidade de drenagem, concentraram maiores níveis de susceptibilidade e de registro de erosões. Através dos resultados, foi possível caracterizar e mapear características ambientais da paisagem e elaborar um mapa de suscetibilidade às erosões lineares, permitindo a compreensão da morfodinâmica através dos diferentes arranjos da paisagem.

INTRODUÇÃO

As erosões lineares (sulcos, ravinas, voçorocas) são processos destrutivos do solo, gerando problemas socioambientais como destruição de infraestruturas, assoreamento

¹ Professor Titular do Departamento de Geografia da Universidade Estadual Paulista - SP, joao.o.nunes@unesp.br;

² Pesquisador do IBGE. Doutor em Geografia pela Universidade Estadual Paulista - SP, glauber.firmino@unesp.br.



fluvial e perda de solo fértil, resultando em crescentes prejuízos (Guerra; Soares da Silva; Botelho, 2015; Kuhn et al., 2023; Scremin; Kemerich, 2016; Valentin; Poesen; Li, 2005; Weill; Pires Neto, 2007).

Recentemente, mapeamentos de suscetibilidade a erosões lineares (GESM) baseados em aprendizado de máquina (ML) têm se destacado. Esses métodos, que empregam algoritmos computacionais orientados a dados, oferecem quantificação, comparação de desempenho e análise de relações complexas entre variáveis qualitativas e quantitativas (Aboutaib et al., 2023; Arabameri et al., 2020b; Han; Guzman; Chu, 2022; Mohebzadeh et al., 2022; Pereira; Rodrigues, 2022).

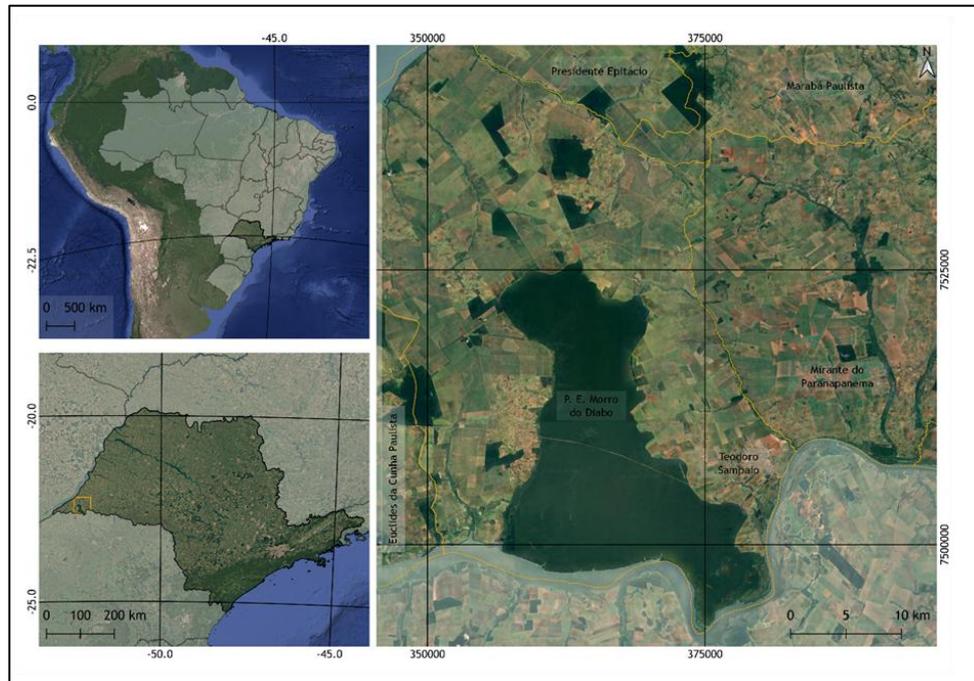
Apesar do potencial preditivo e da replicabilidade para o planejamento ambiental, essas metodologias são pouco difundidas no Brasil, com alguns exemplos pontuais nas regiões Sudeste e Centro-Oeste (Bouramtane et al., 2022; Lana; Castro; Lana, 2022; Marques Filho et al., 2024).

Nesse contexto, esta pesquisa objetiva avaliar e mapear a suscetibilidade a erosões lineares no Extremo Oeste de São Paulo, abrangendo Teodoro Sampaio-SP e municípios adjacentes. Para isso, são aplicados algoritmos de ML como Random Forest, Logistic Regression, Support Vector Machine e Artificial Neural Networks Multilayer Perceptron, visando contribuir para o planejamento e gestão ambiental local, dada a recorrente formação de ravinas e voçorocas na região.

METODOLOGIA (OU MATERIAIS E MÉTODOS)

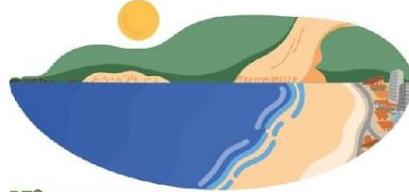
A área de pesquisa possui 2.262 km² de área e se localiza no Extremo Oeste Paulista (Figura 1), abrangendo a maior parte do município de Teodoro Sampaio-SP, e partes menores de municípios adjacentes (Mirante do Paranapanema, Euclides da Cunha Paulista, Marabá Paulista e Presidente Epitácio).

Figura 1. Localização da área de estudo no Extremo Oeste Paulista, Estado de São Paulo, Brasil.



O presente estudo foi elaborado a partir de três etapas:

1. Preparação de dados geospaciais: elaboração de mapas de fatores geoambientais selecionados com base na literatura e sua correlação com a ocorrência de erosões lineares. Incluiu a criação de um mapa inventário de erosões a partir de revisão bibliográfica, uso de imagens de satélite e dados secundários, processamento de dados topográficos e morfométricos (CopDEM30), interpretação aerofotogramétrica, e trabalhos de campo para levantamento e validação. Dados do IPT (2012) sobre erosões lineares foram atualizados através do mapeamento geomorfológico.
2. Processamento dos dados e modelagem
 - a. Seleção de atributos: Matriz de Correlação e Análise de Multicolinearidade.
 - b. Separação aleatória de dataset em treino e teste, com proporção de 70-30%.



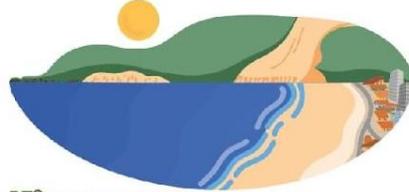
15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

- c. Modelagem por emprego de Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM) e Redes Neurais Artificiais Multilayer Perceptron (ANN-MLP), com otimização de hiperparâmetros.
 - d. Avaliação de precisão por indicadores de Accuracy, Precision, Recall, F-1 Score, Kappa Cohen e Área sob a Curva (AUC-ROC), com representação gráfica (AUC-ROC) e Matriz de Confusão.
3. Elaboração dos mapas temáticos: Criação dos mapas de suscetibilidade no QGIS, com adequação de simbologia, layout de impressão e inclusão de elementos cartográficos (legenda, seta norte, escala, projeção etc.).

A seleção das variáveis no presente projeto, descritas no Quadro 1, teve como critérios a disponibilidade dos dados de sensoriamento remoto em escala/resolução adequada, a adequação à área estudada e aos objetivos propostos na pesquisa e sua importância para a elaboração do mapa de suscetibilidade (Figura 2).

Quadro 1. de seleção de variáveis utilizada na elaboração do mapa de susceptibilidade a erosão.

Tipos de variáveis	Nome das variáveis	Resolução/Escala	Fonte de Dados
Atributos topográficos primários	Elevação	30m	CopernicusDEM
	Orientação das vertentes		
	Declividade		
	Rugosidade (<i>Roughness</i>)		
	Curvatura horizontal		
	Curvatura vertical		
	Índice de Convergência (Convergence Index)		
Atributos topográficos secundários	Índice de Poder de Escoamento (<i>Stream Power Index - SPI</i>)	30m	CopernicusDEM
	Índice de Heterogeneidade do Terreno (<i>Terrain Rudgness Index TRI</i>)		
	Índice de Umidade Topográfica (<i>Topographic Wetness Index - TWI</i>)		

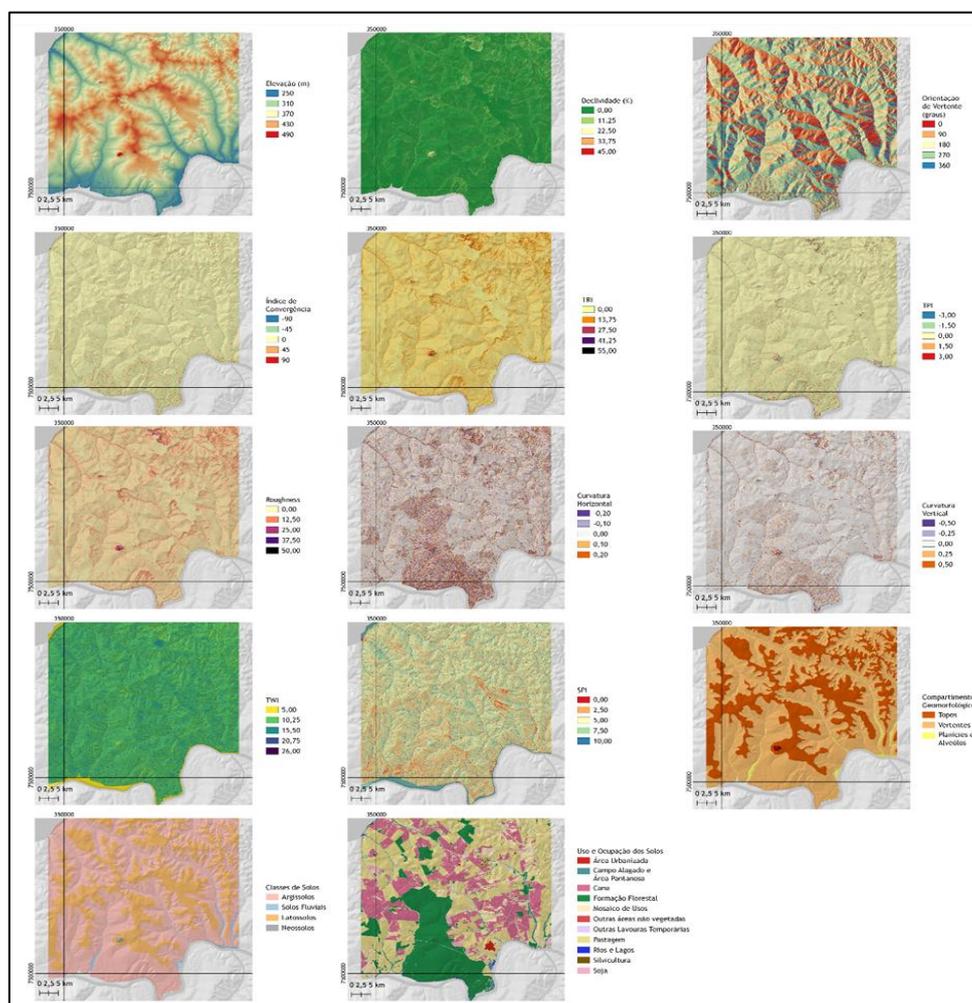


15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

	Índice de Posição Topográfica (<i>Topographic Position Index - TPI</i>)		
	Compartimentos geomorfológicos	1:25.000	Interpretação estereoscópica de análgfo georreferenciado (GE Pro + StereoPhoto Maker + QGIS)
Fator antropogênico	Uso do solo e cobertura vegetal	30m	MapBiomias
Outros fatores	Classes de solos	1:25.000	Elaboração própria a partir de interpretação estereoscópica associada à declividade, dados de campo e levantamentos em escalas menos detalhadas

Elaborado pelos autores (2025).

Figura 2. Mapa das variáveis utilizada na elaboração do mapa de susceptibilidade a erosão.



Elaborado pelos autores (2025).



REFERENCIAL TEÓRICO

Recentemente, o aprendizado de máquina (ML) tem se destacado internacionalmente para o mapeamento de suscetibilidade a erosões lineares (Han; Guzman; Chu, 2022; Mohebzadeh et al., 2022; Pereira; Rodrigues, 2022). O ML, que emprega modelos estatísticos para realizar tarefas sem programação explícita (Mahesh, 2020), oferece vantagens como quantificação de desempenho, análise de dados complexos e não-lineares, integração de dados diversos e dispensa de hipóteses prévias sobre as relações entre variáveis (Aboutaib et al., 2023; Arabameri et al., 2020a; Mohebzadeh et al., 2022). Apesar disso, estas técnicas são pouco difundidas no Brasil, com estudos pontuais em Minas Gerais (Lana; Castro; Lana, 2022), Mato Grosso do Sul (Bouramtane et al., 2022) e Rio de Janeiro (Marques Filho et al., 2024), contrastando com a concentração de publicações em países como Irã, China e Índia.

A suscetibilidade a erosões é controlada por fatores geoambientais, incluindo atributos topográficos, pedológicos, uso do solo e precipitações (Bashir et al., 2017; Conoscenti et al., 2014; Mohebzadeh et al., 2022; Poesen et al., 2003; Rădoane; Rădoane, 2017; Valentin; Poesen; Li, 2005; Weill; Pires Neto, 2007). Embora Mohebzadeh et al. (2022) destaquem topografia, hidrologia e fatores antropogênicos como mais influentes e variáveis como elevação, declividade e uso do solo sejam amplamente empregadas, a ordem de importância das variáveis não é consensual entre algoritmos e contextos geográficos (Bouramtane et al., 2022; Marques Filho et al., 2024; Mohebzadeh et al., 2022). Essa variação é observada tanto em estudos nacionais (Lana et al., 2022) quanto internacionais em diversas paisagens (Azareh et al., 2019; Hitouri et al., 2022; Pourghasemi et al., 2017; Rahmati et al., 2016, 2022; Liu et al., 2023; Wang et al., 2022; Conoscenti et al., 2014; Chuma et al., 2023).

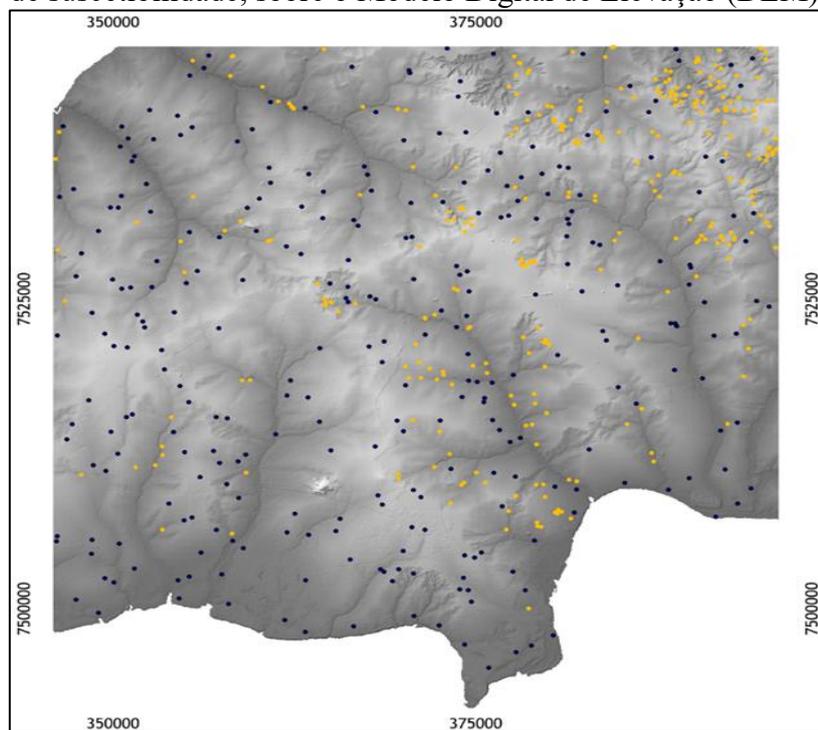
O Random Forest (RF) é amplamente utilizado e performático em estudos de suscetibilidade (Amiri et al., 2019; Bouramtane et al., 2022; Garosi et al., 2019; Gayen et al., 2019; Lana; Castro; Lana, 2022; Marques Filho et al., 2024; Rahmati, 2017; Were et al., 2023), devido sua capacidade de lidar com grandes dados e relações não-lineares (Lana; Castro; Lana, 2022). No entanto, Redes Neurais Artificiais Multilayer Perceptron (ANN), Support Vector Machine (SVM), Frequency Ratio (FR), Logistic Regression (LR) e Weight of Evidence (WofE) também demonstraram bom desempenho (Conoscenti

et al., 2014; Gayen et al., 2019; Pourghasemi et al., 2017; Rahmati, 2017; Zabihi et al., 2018).

RESULTADOS E DISCUSSÃO

Os dados obtidos no mapeamento geomorfológico tiveram seus centroides computados e foram adicionados ao arquivo *shapefile* de erosões. Os pontos indicando ausência de erosões foram gerados aleatoriamente, em um tamanho amostral equivalente ao número de pontos de localização das erosões lineares encontram-se representados na Figura 3.

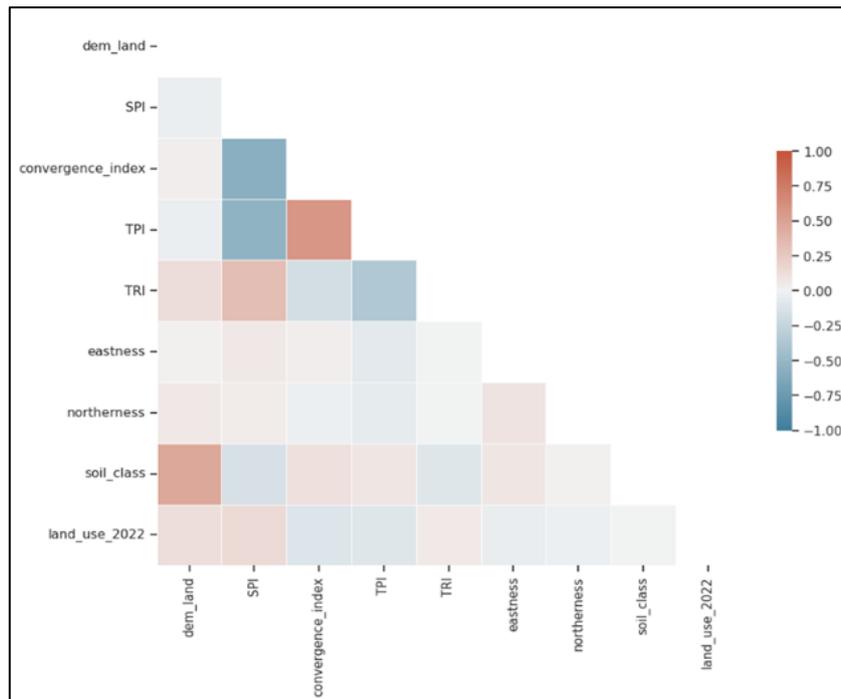
Figura 3. Pontos de erosões lineares (amarelo) e amostras geradas aleatoriamente indicando pontos de ausência de erosões lineares (azul), para treino e teste do modelo de suscetibilidade, sobre o Modelo Digital de Elevação (DEM).



Elaborado pelos autores (2024).

A Matriz de Correlação de Pearson foi utilizada para remover variáveis preditivas com correlação acima de 0.75 ou abaixo de -0.75. Foram eliminadas: TWI (correlacionado ao SPI); Curvaturas Vertical e Horizontal (correlacionadas ao TPI); e Declividade e Rugosidade (correlacionadas ao TRI). As variáveis restantes estão representadas na Figura 4.

Figura 4. Matriz de Correlação de Pearson.



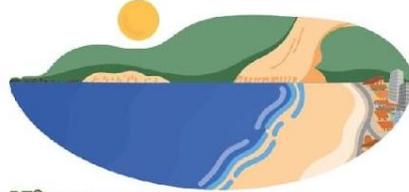
Elaborado pelos autores (2025).

Em relação à importância de variáveis, nos quatro modelos, os dados de TRI, Índice de Convergência e de Uso da Terra compuseram os três atributos de maior importância.

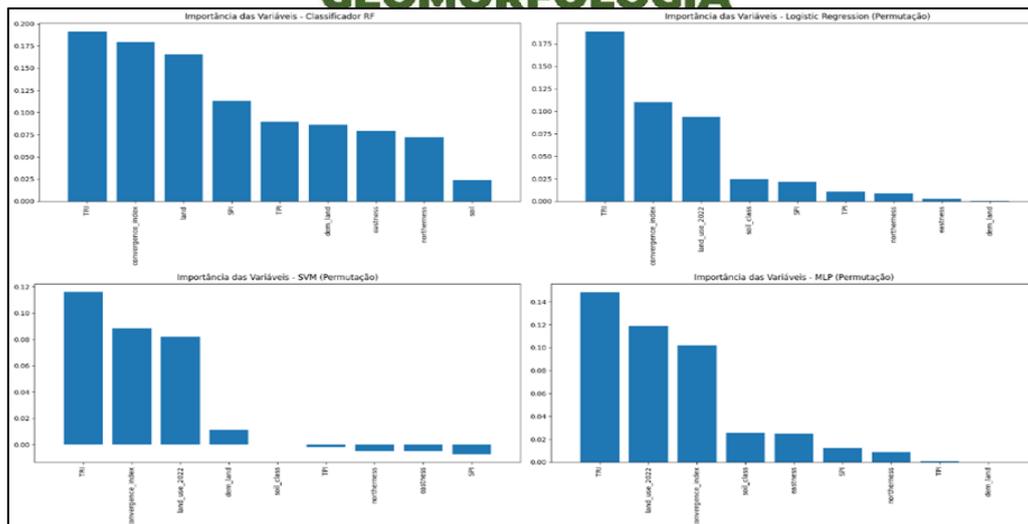
Os valores de AUC-ROC (Figura 5 e Figura 6) para avaliação do desempenho dos modelos demonstraram maior qualidade no algoritmo de Multilayer Perceptron, obtendo $AUC = 0.93$. Os demais algoritmos resultaram nos seguintes valores: Random Forest ($AUC = 0.92$), Logistic Regression ($AUC = 0.89$), Support Vector Machine ($AUC = 0.89$).

Através da classificação citada por Mohebzadeh et al. (2022), para avaliação do desempenho de modelos a partir dos valores AUC-ROC, considera-se o desempenho do modelo LR e SVM como muito bom (0.8 a 0.9) e dos algoritmos MLP e RF como excelente (0.9 a 1.0).

Figura 5. Valores de AUC-ROC para avaliação do desempenho dos modelos.

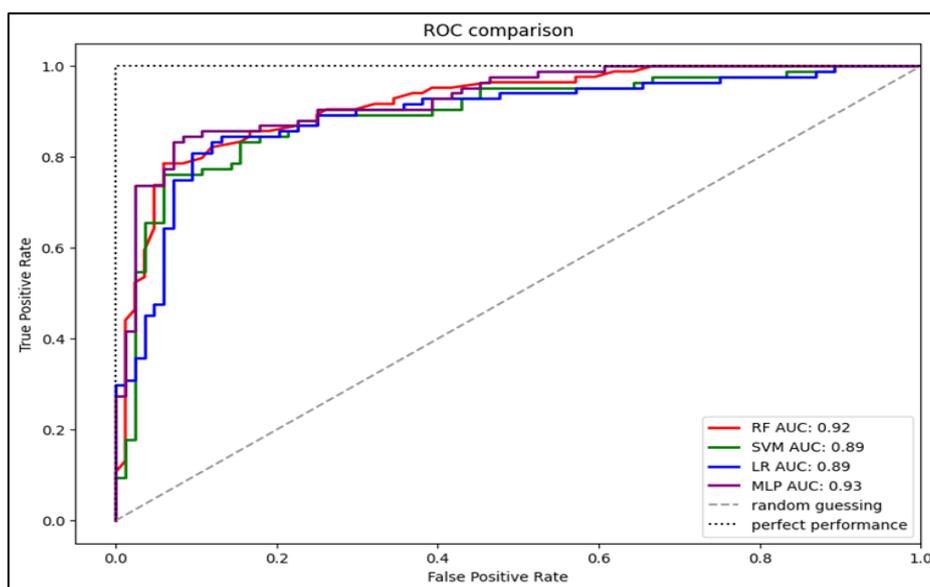


15º SIMPÓSIO NACIONAL DE GEOMORFOLOGIA



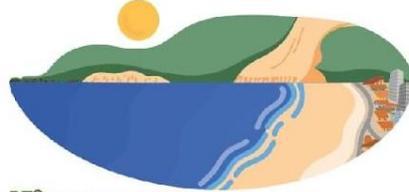
Elaborado pelos autores (2025).

Figura 6. Valores de AUC-ROC, com destaque para os algoritmos MLP e RF.



Elaborado pelos autores (2025).

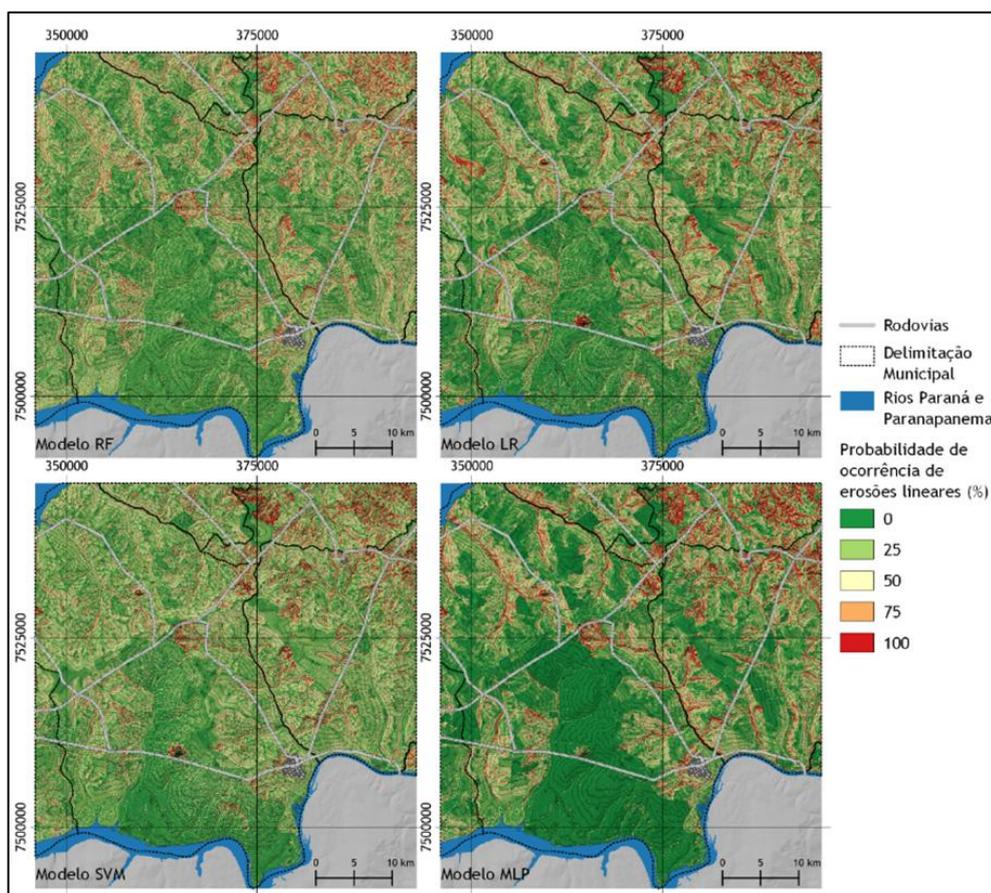
Os scores de Accuracy foram: Multilayer Perceptron (MLP) = 0.89, Random Forest (RF) = 0.85, Logistic Regression (LR) = 0.84 e Support Vector Machine (SVM) = 0.82. Quanto ao coeficiente Kappa Cohen, obtiveram-se: MLP = 0.77, RF = 0.70, LR = 0.68 e SVM = 0.63, todos classificados como substanciais (0.6 a 0.8) (MOHEBZADEH et al., 2022). Ambos os indicadores demonstram desempenho superior do modelo baseado no algoritmo MLP.



15º SIMPÓSIO NACIONAL DE GEOMORFOLOGIA

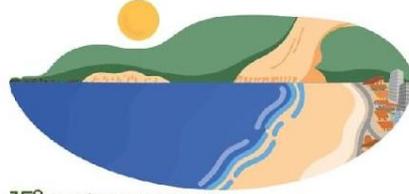
Os valores de suscetibilidade representam a probabilidade de ocorrência de erosões lineares por pixel, calculados pela função `predict_proba` do Stack Classifier (Scikit-Learn, Python), que gera a probabilidade de predição binária (0 e 1). Após exportados para GeoTIFF, esses valores permitiram a elaboração do Mapa de Suscetibilidade de cada modelo (Figura 7).

Figura 7. Mapa de Suscetibilidade de cada modelo.



Elaborado pelos autores (2025).

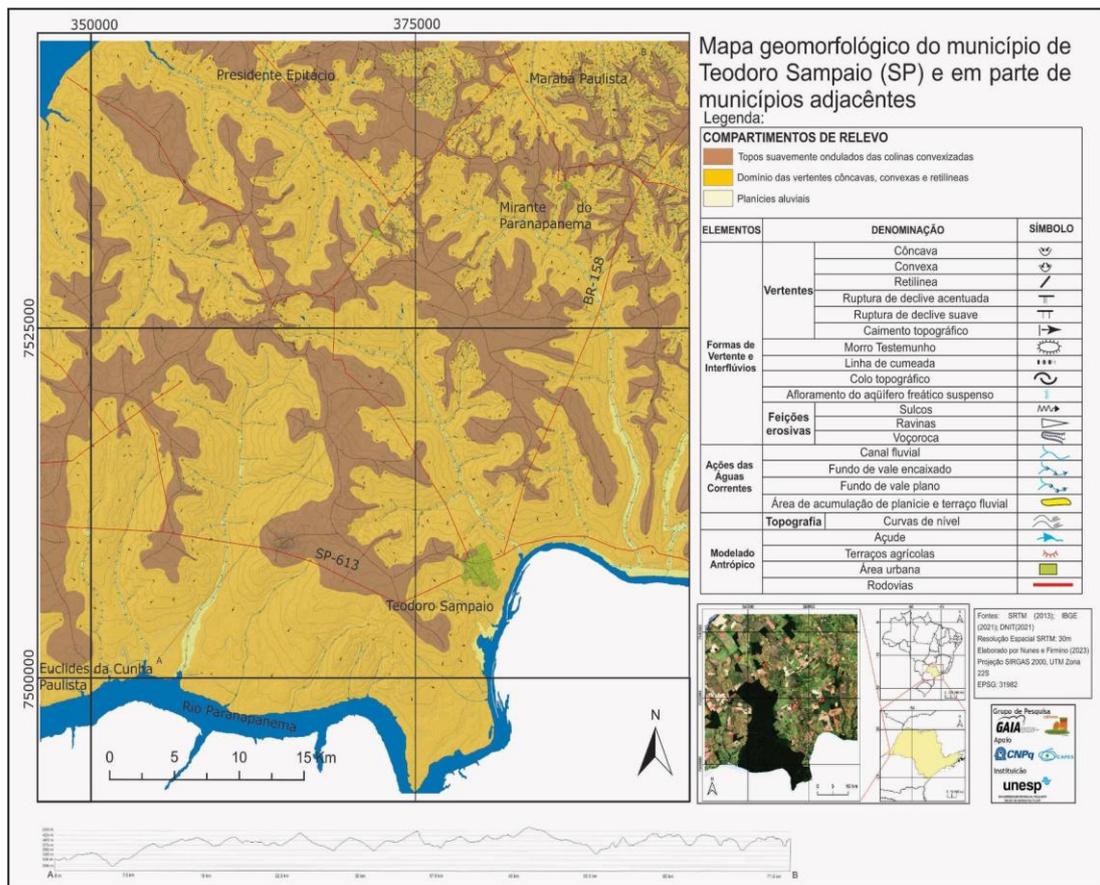
A análise do mapa do modelo MLP, juntamente com mapas temáticos, com destaque para o de geomorfologia (Figura 8) e o gráfico de importância de atributos, revelou padrões espaciais entre características físico-ambientais e o nível de suscetibilidade. Áreas de baixa a muito baixa suscetibilidade (0-25%) associam-se a formação florestal, algumas áreas de agricultura (cana e soja), relevo de baixa declividade e rugosidade, índice de convergência positivo (relevos convexos) e topos com Latossolos. A média suscetibilidade (25-50%) relaciona-se majoritariamente à agricultura (cana e



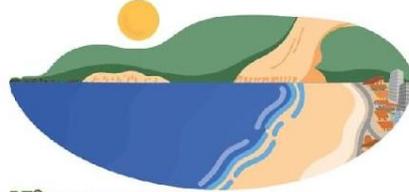
15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

soja), índices de convergência próximos a zero e rugosidade média, e domínio de vertentes com Argissolos e Neossolos Litólicos. Já a alta a muito alta suscetibilidade (50-100%) está ligada a pastagens e áreas urbanas, índices de convergência negativos (relevos côncavos, que aumentam o escoamento superficial), planícies e alvéolos, vertentes com solos hidromórficos e Argissolos, além de rupturas de declive e cabeceiras em anfiteatro (Figura 9).

Figura 8. Mapa geomorfológico da área de estudo.

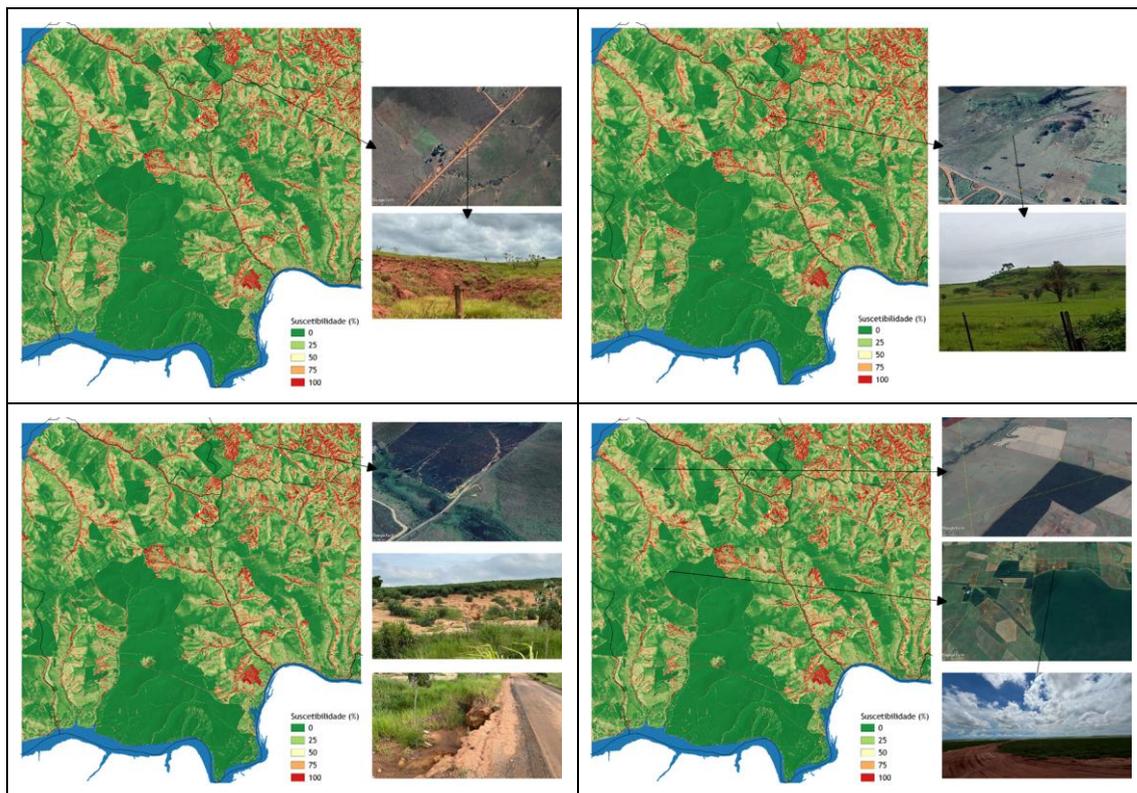


Elaborado pelos autores (2024).



15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

Figura 9. Mapa do modelo MLP (redes neurais) de susceptibilidade a erosão linear.



Elaborado pelos autores (2025).

A aplicação de RF, LR, SVM e ANN validou as vantagens do ML para mapeamento de suscetibilidade a erosões (Aboutaib et al., 2023; Arabameri et al., 2020a; Mohebzadeh et al., 2022), destacando sua capacidade com dados complexos e não-lineares, além de quantificar o desempenho dos modelos. Apesar do RF ser comumente o mais citado (Amiri et al., 2019; Bouramtane et al., 2022; Garosi et al., 2019; Gayen et al., 2019; Lana; Castro; Lana, 2022; Marques Filho et al., 2024; Rahmati, 2017; Were et al., 2023), as Redes Neurais Artificiais Multilayer Perceptron apresentaram desempenho ligeiramente superior neste estudo.

A alta importância do índice de convergência ressalta o papel das concavidades na concentração do fluxo superficial, corroborando Botelho; Silva (2007) e Zanatta; Lupinacci; Boin (2019), o que torna canais e cabeceiras de drenagem críticas. O uso e cobertura do solo foram atributos cruciais, evidenciando a influência antrópica nos níveis de erosão; pastagens mostraram maior suscetibilidade, enquanto áreas florestais (como



PE Morro do Diabo e EE Mico Leão Preto) reforçam a importância de práticas vegetativas para prevenção de desastres (Zonta et al., 2012).

CONSIDERAÇÕES FINAIS

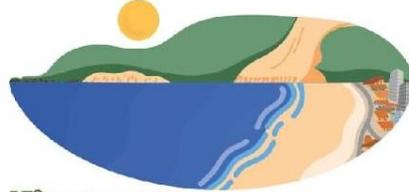
A metodologia empregada permitiu avaliar e mapear a suscetibilidade a erosões lineares no extremo Oeste Paulista, confirmando que os níveis de suscetibilidade dependem de uma combinação de fatores geoambientais topográficos, geomorfológicos, pedológicos e antropogênicos. O uso dos algoritmos RF, LR, SVM e MLP possibilitou o emprego de variáveis qualitativas e quantitativas, incluindo índices geomorfométricos precisos do CopDEM30, que se mostraram úteis para estudos geomorfológicos. A calibração e avaliação do desempenho dos modelos demonstraram boa acurácia para todas as ferramentas, com destaque para RF e, principalmente, MLP.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

REFERÊNCIAS

- BASHIR, S. et al. Soil and water conservation. Faisalabad: University of Agriculture, 2017.
- BISONG, E. Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners. 1. ed. Berkeley, CA: Apress, 2019.
- BOTELHO, R. G. M.; SILVA, A. S. Bacia Hidrográfica e Qualidade Ambiental. *In*: Reflexões sobre a Geografia Física no Brasil. 2. ed. [S.l.]: Bertrand Brasil, 2007. p. 153–192.
- CHAN, K. Y. et al. Deep neural networks in the cloud: Review, applications, challenges and research directions. *Neurocomputing*, v. 545, p. 126–327, 2023.
- PERROTTA, et al., M. M. Geologia e recursos minerais do estado de São Paulo. [S.l.]: CPRM, 2006.
- CREPANI, Edison et al. Sensoriamento remoto e geoprocessamento aplicados ao zoneamento ecológico-econômico e ao ordenamento territorial. [S.l.]: Inpe São José dos Campos, 2001.
- KERTZMAN, Fernando Facciolla et al. Mapa de erosão do Estado de São Paulo. *Revista do Instituto Geológico*, v. Volume Especial, 1995.
- KNISEL, W. G. CREAMS: A Field Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. [S.l.]: U.S. Department of Agriculture, Science and Education Administration, 1980.
- MOHEBZADEH, H. et al. Machine Learning Techniques for Gully Erosion Susceptibility Mapping: A Review. *Geosciences*, v. 12, n. 12, p. 429, 2022.



15º SIMPÓSIO NACIONAL DE
GEOMORFOLOGIA

- NEARING, MA; NICKS, AD. Evaluation of the Water Erosion Prediction Project (WEPP) model for hillslopes. In: *Modelling soil erosion by water*. [S.l.]: Springer, 1998. p. 43–53.
- OLIVEIRA, M. A. T. Processos Erosivos e Preservação de Áreas de Risco de Erosão por Voçorocas. In: GUERRA, A. J. T.; SOARES DA SILVA, A.; BOTELHO, R. G. M. (Orgs.). *Erosão e conservação dos solos: conceitos, técnicas e aplicações*. Rio de Janeiro (RJ): Bertrand Brasil, 2015. p. 340.
- PELOGGIA, A. U. G. Conceitos fundamentais da análise de terrenos antropogênicos: o estudo da agência geológico-geomorfológica humana e de seus registros. *Revista do Instituto Geológico*, v. 40, n. 1, p. 1–17, 2019.
- PEREIRA, J. S.; RODRIGUES, S. C. Erosão por voçorocas: estado da arte. In: CARVALHO JUNIOR, O. A. et al. (Org.). *Revisões de Literatura da Geomorfologia Brasileira*. Brasília: UnB, 2022.
- RĂDOANE, M.; RĂDOANE, N. Gully erosion. In: *Landform Dynamics and Evolution in Romania*. [S.l.: S.n.]. p. 371–396.
- ROSE, C. W. Research Progress on Soil Erosion Processes and a Basis for Soil Conservation Practices. In: *Soil Erosion Research Methods*. 2nd. ed. [S.l.]: Routledge, 1994. p. 22.
- ROSSI, M.; NALON, M. A.; KANASHIRO, M. M.. Atlas de Suscetibilidades dos Solos do Estado de São Paulo. São Paulo: SIMA/IPA, 2022.
- TRICART, Jean. *Principes et méthodes de la Géomorphologie*. 1. ed. Paris: Masson, 1965.
- WEILL, M. A. M.; PIRES NETO, A. G. P. Erosão e assoreamento. In: SANTOS, R. F. dos (Org.). *Vulnerabilidade Ambiental: desastres naturais ou fenômenos induzidos?* Brasília, DF: MMA, 2007. p. 192.
- WISCHMEIER, Walter H.; SMITH, Dwight David. *Predicting rainfall erosion losses: a guide to conservation planning*. [S.l.]: Department of Agriculture, Science and Education Administration, 1978.
- WOODWARD, D. E. Method to predict cropland ephemeral gully erosion. *CATENA*, v. 37, n. 3, p. 393–399, 1999.
- ZANATTA, Felipe Augusto Scudeller; LUPINACCI, Cenira Maria; BOIN, Marcos Noberto. Dinâmica erosiva linear: estudo de caso em área rural no oeste paulista. *Geografia*, v. 44, n. 2, dez. 2019.
- ZONTA, João Henrique *et al.* *Práticas de Conservação de Solo e Água*. Campina Grande, PB: Embrapa, set. 2012.
- ZHU, X. et al. Soil splash erosion: An overlooked issue for sustainable rubber plantation in the tropical region of China. *International Soil and Water Conservation Research*, v. 11, n. 1, p. 30–42, 2023.