



I CONGRESSO PERNAMBUCANO DE RECURSOS HÍDRICOS

Água para o Desenvolvimento
Recife, 24, 25 e 26 de Março de 2026

UM NOVO MAPA DE PROPRIEDADES FÍSICAS E HIDRÁULICAS DE SOLOS PARA O ESTADO DE PERNAMBUCO BASEADO EM AQUISIÇÃO DE DADOS LEGADOS E APRENDIZADO DE MÁQUINA

Rodrigo de Q. Miranda¹, Rodolfo L. B. Nóbrega²⁻⁴, Alexandre H. C. Barros⁵, Suzana M. G. L. Montenegro⁶, Maria do S. B. Araújo⁷, and Josiclêda D. Galvêncio¹

* A lista completa de autores pode ser encontrada no link no final do artigo.

Palavras-chave: mapeamento digital de solos, aprendizado de máquina, propriedades hidráulicas, SLEEP.

INTRODUÇÃO

Mapas confiáveis de propriedades do solo são essenciais para modelagem ambiental e, em particular, para estudos hidrológicos e de recursos hídricos em bacias com forte gradiente climático e fisiográfico. Entretanto, levantamentos pedológicos convencionais são caros e demorados, sobretudo em regiões tropicais, onde a diversidade de solos é elevada e a disponibilidade de perfis é limitada (McBRATNEY et al., 2003; WADOUX et al. 2020).

Neste trabalho sintetizamos um arcabouço escalável de mapeamento digital de solos que integra o algoritmo Soil-Landscape Estimation and Evaluation Program (SLEEP; ZIADAT et al., 2020) a modelos de aprendizado de máquina do tipo Gradient Boosting, visando estimar propriedades físico-químicas do solo em múltiplas profundidades e derivar parâmetros hidráulicos por funções de pedotransferência (PTFs). O estudo foi aplicado ao estado de Pernambuco, Nordeste do Brasil (~98.000 km²), com grande contraste de relevo, clima e cobertura vegetal.

METODOLOGIA

A base de dados de solo foi composta por 223 perfis (Fig. 1) oriundos do Zoneamento Agroecológico de Pernambuco (ZAPE/Embrapa), com informação de número e profundidade de camadas (SOL_Z), textura (areia, silte e argila, incluindo areia fina e grossa), fragmentos rochosos, carbono orgânico e soma de bases (SB). Como covariáveis, foram utilizados 6 parâmetros topográficos (derivados de MDE a ~30 m), 10 variáveis climáticas (1961–2016) e 2 variáveis de vegetação (NDVI e temperatura de superfície).

¹PRODEMA, Universidade Federal de Pernambuco, Recife, Brazil

²University of Bristol, School of Geographical Sciences, University Road, Bristol BS8 1SS, UK

³Cabot Institute for the Environment, University of Bristol, Bristol, UK

⁴Imperial College London, Georgina Mace Centre for the Living Planet, Department of Life Sciences, Silwood Park Campus, Buckhurst Road, Ascot, SL5 7PY, UK

⁵Brazilian Agricultural Research Corporation – Embrapa Soils, Recife, Brazil

⁶Departamento de Engenharia Civil, Universidade Federal de Pernambuco, Recife, Brazil

⁷Departamento de Ciências Geográficas, Universidade Federal de Pernambuco, Recife, Brazil

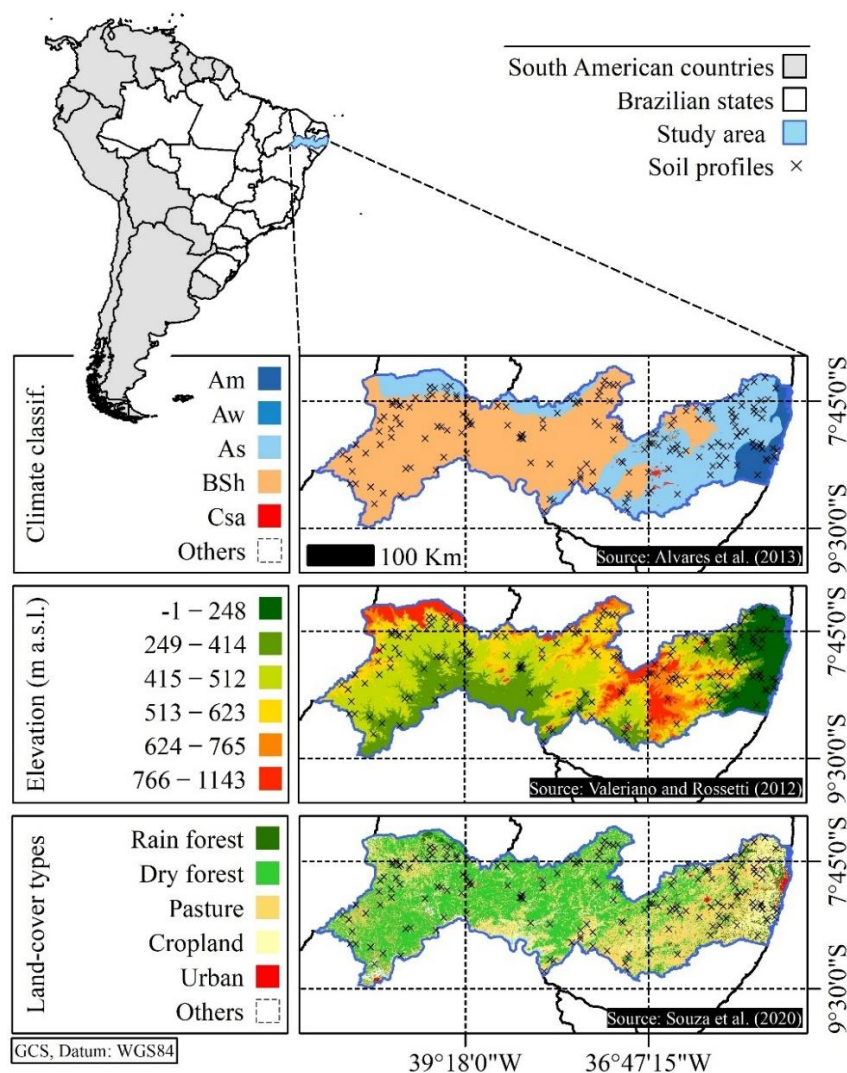


Figura 1. Distribuição espacial dos perfis de solo amostrados ao longo da área de estudo.

O SLEEP foi empregado para delinear unidades homogêneas do terreno (facetas) e agregá-las em ‘patches’ segundo similaridade de declividade, reduzindo ruído espacial (efeito ‘sal e pimenta’) e custo computacional. Para cada propriedade básica do solo, um modelo Gradient Boosting (FRIEDMAN, 2001) foi calibrado com seleção recursiva de atributos, buscando minimizar multicolinearidade e otimizar o conjunto de entradas. A validação seguiu uma estratégia de holdout (20% para teste) e calibração com validação cruzada de 2 dobras, avaliando desempenho por r^2 , RMSE e viés percentual (PBIAS).

Os mapas finais foram produzidos em resolução de 30 m e em múltiplas profundidades (camadas observadas nos perfis). Em seguida, os atributos básicos mapeados alimentaram PTFs (e.g., SAXTON et al., 2016) para estimar propriedades hidráulicas do solo, incluindo condutividade hidráulica saturada (SOL_K), densidade do solo (SOL_BD) e capacidade de água disponível (SOL_AWC), comparando diferentes formulações (globais e regionais).

RESULTADOS

O procedimento gerou 64.415 patches (área média $1,35 \pm 4,54 \text{ km}^2$) tratados como perfis virtuais. Os modelos apresentaram bom desempenho espacial, com r^2 entre 0,79 e 0,98 e PBIAS entre -1,39 e 1,14%, com menor desempenho para silte, refletindo maior complexidade e redundância de preditores (Tabela 1).

Tabela 1. Aspectos gerais dos resultados obtidos.

Item	Valor
Área de estudo	~98.000 km ² (Pernambuco)
Perfis de solo (ZAPE/Embrapa)	223
Covariáveis	6 topográficas; 10 climáticas; 2 de vegetação
Resolução espacial	~30 m
Unidades mapeadas (patches)	64.415 (área média 1,35 km ²)
Desempenho (verificação)	$r^2 = 0,79\text{--}0,98$; PBIAS = -1,39 a 1,14%

De modo geral, variáveis topográficas e climáticas dominaram a explicação da variabilidade das propriedades, enquanto variáveis de vegetação tiveram maior influência na predição das frações de areia (Fig. 2). As estimativas de propriedades hidráulicas derivadas por PTFs mostraram sensibilidade ao conjunto de PTFs empregado, indicando a necessidade de preferir funções calibradas regionalmente quando disponíveis, especialmente para aplicações hidrológicas.

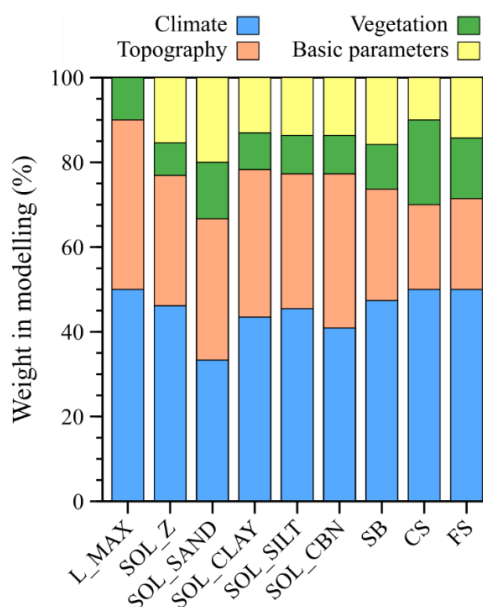


Figura 2. Pesos proporcionais das diferentes variáveis de entrada usadas para modelar cada parâmetro básico do solo.

CONCLUSÕES

Os resultados indicam que a integração SLEEP e Gradient Boosting é uma alternativa custo-efetiva para gerar mapas de propriedades do solo em regiões tropicais com escassez de dados, preservando significado pedológico ao trabalhar com profundidades reais dos perfis. O arcabouço produz mapas em alta resolução espacial (30 m) com desempenho robusto e pode ser transferido para outras regiões, desde que haja variabilidade ambiental

suficiente e um conjunto mínimo de perfis de referência. Para parâmetros hidráulicos, recomenda-se o uso de PTFs regionais e a avaliação explícita de incertezas associadas a extrapolações. O mapa de solos gerado para Pernambuco se mostra apropriado para uso direto e como entrada para outros modelos, como hidrológicos e de superfície.

A versão completa desse trabalho, juntamente com a lista de todos os autores e os dados resultantes da modelagem está publicada em: <https://doi.org/10.1016/j.soilad.2025.100064>

AGRADECIMENTOS

Agradecemos à Embrapa Solos pelo acesso aos dados do ZAPE e às agências de fomento e instituições colaboradoras que apoiaram o estudo, incluindo CAPES, FACEPE, CNPq e parceiros internacionais vinculados ao desenvolvimento metodológico.

REFERÊNCIAS

- FRIEDMAN, J. H. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, v. 29, n. 5, p. 1189-1232, 2001. DOI: 10.1214/aos/1013203451.
- McBRATNEY, A. B.; MENDONÇA-SANTOS, M. L.; MINASNY, B. On digital soil mapping. *Geoderma*, v. 117, n. 1-2, p. 3-52, 2003. DOI: 10.1016/S0016-7061(03)00223-4.
- SAXTON, K. E.; RAWLS, W. J. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, v. 70, n. 5, p. 1569-1578, 2006. DOI: 10.2136/sssaj2005.0117.
- WADOUX, A. M. J.-C.; MINASNY, B.; McBRATNEY, A. B. Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth-Science Reviews*, v. 210, 103359, 2020. DOI: 10.1016/j.earscirev.2020.103359.
- ZIADAT, F. M.; YEGANANTHAM, D.; SHOEMATE, D.; SRINIVASAN, R.; et al. Soil-Landscape Estimation and Evaluation Program (SLEEP): an approach for soil mapping in data-scarce regions. *International Journal of Agricultural and Biological Engineering*, v. 8, n. 3, p. 158-172, 2015. DOI: 10.25165/ijabe.v8i3.1270.