

# Informe Técnico-Científico de Estudos Hidrológicos e Hidrogeológicos

V. 1, N. 1 Rio de Janeiro, 2024

## Área-piloto e algoritmos de aprendizado de máquinas no mapeamento digital semidetalhado de solos da Bacia Hidrográfica Paraná 3, Região Sul do Brasil, e dificuldades inerentes

*Pilot area and machine learning algorithms for digital semi-detailed soil mapping of the Paraná 3 Basin, South of Brazil, and inherent difficulties*

Michele D. Menezes ([michele.menezes@ufla.br](mailto:michele.menezes@ufla.br))<sup>1</sup>  
Sérgio H.G. Silva ([sergio.silva@ufla.br](mailto:sergio.silva@ufla.br))<sup>1</sup>  
Samara M. Barbosa ([samara.barbosa1@ufla.br](mailto:samara.barbosa1@ufla.br))<sup>1</sup>  
Renata Andrade ([renata\\_andrade@ufla.br](mailto:renata_andrade@ufla.br))<sup>1</sup>  
André F. Rodrigues ([afrodrigues@ehr.ufmg.br](mailto:afrodrigues@ehr.ufmg.br))<sup>2</sup>

José M. Lima ([jmlima@ufla.br](mailto:jmlima@ufla.br))<sup>1</sup>  
Jhones S. Amorim ([jhones.amorim@ufrn.br](mailto:jhones.amorim@ufrn.br))<sup>3</sup>  
Edgar Shinzato ([edgar.shinzato@sbg.gov.br](mailto:edgar.shinzato@sbg.gov.br))<sup>4</sup>  
Nilton Curi ([ntcuri@gmail.com](mailto:ntcuri@gmail.com))<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras (UFLA)

<sup>2</sup> Universidade Federal de Minas Gerais (UFMG)

<sup>3</sup> Universidade Federal do Rio Grande do Norte (UFRN)

<sup>4</sup> Serviço Geológico do Brasil (SGB)

Produto do TED No 016/CPRM/2021, celebrado entre o SGB e a UFLA, tendo como unidade executora a Zetta Agência UFLA de Inovação, Geotecnologia e Sistemas Inteligentes

### Abstract

Detailed soil legacy data (soil survey at a scale of 1:10,000) was used as a pilot area for extracting soil information and mapping a surrounding area with similar soil-landscape relationships. Using a digital soil mapping framework, the algorithms C5.0, partial least square, and random forest were applied to establish quantitative relationships between soil information from the pilot area and digital terrain models. Different strategies were applied to extract soil-landscape relationships from the pilot area and create soil spatial predictive models for areas of the Paraná 3 Basin (BP3) with similar soil-landscape relationships. The accuracy of the predictive final soil map was further assessed. Digital terrain information was extracted from soil mapping units considering four different strategies: (i) removing 30 m; (ii) removing 60 m; and (iii) removing 90 m of the borders from each soil class delineation; and (iv) considering the entire polygon of the soils mapping units developed from the basaltic parent material. Random forest with the strategy of removing 60 m of the borders outperformed other algorithms, reaching a global accuracy of 0.62, resulting in a digital soil mapping with 5,028 Km<sup>2</sup> (out of 8,000 Km<sup>2</sup>) of area at a semi-detailed scale, based on soil information from a pilot area (7.57 Km<sup>2</sup>), and representing a gain of 664 times in mapped area. Considering the region is very important in terms of water and agricultural production and that the semi-detailed pedological mapping is currently being performed (PronaSolos-PR Project), it may constitute a good opportunity to check for the possibility of information complementarity from both soils mapping methodologies, which are not exclusionary.

Keywords: digital soil mapping; soil legacy; geoprocessing

### Introdução

Mesmo diante do crescimento da demanda por informações sobre solos, visando fornecer suporte ao uso da terra, manejo e políticas públicas no Brasil, ainda existe escassez de informações pedológicas, principalmente em escalas mais detalhadas (Polidoro *et al.*, 2020; Coelho; Vasques, 2022). Portanto, investir em alternativas para ganhos em área mapeada, que sejam mais rápidos, econômicos e com adequada acurácia, torna-se crucial. Um exemplo consiste no uso do conceito de área-piloto, tradicionalmente adotado em levantamentos de solos, que, segundo o IBGE (2007), prevê investigações minuciosas em áreas menores que representem feições fisiográficas definidas e possam ser extrapoladas para o restante da área de interesse. Com o avanço das técnicas de geoprocessamento e de algoritmos robustos, aliados a informações geoespaciais, com destaque para a relação solo-paisagem, o mapeamento de solos a partir de áreas-piloto ou de referência tem se popularizado (Lagacherie *et al.*, 2001; Pelegrino *et al.*, 2016; Silva *et al.*, 2016; Machado *et al.*, 2019; Mello *et al.*, 2021).

Nesse contexto, o arcabouço do mapeamento digital de solos está inserido, onde modelos preditivos espaciais quantitativos resultam em otimização e organização de um amplo conjunto de dados (McBratney *et al.*, 2003). Modelos preditivos são construídos para estabelecer relações entre informações de solos (variável-alvo) e variáveis explanatórias (variáveis preditoras digitais que representam fatores de formação do solo), a partir de técnicas mais quantitativas (McBratney *et al.*, 2003), com destaque para os algoritmos de aprendizado de máquinas. O uso do mapeamento digital de solos foi iniciado na década de 1970, com grande desenvolvimento na década de 1980, devido aos avanços tecnológicos nas áreas de tecnologia da informação, sensoriamento remoto, estatística/matemática, posicionamento global, sistemas de medição e acesso instantâneo a informações através de redes de computadores (Caten *et al.*, 2011). O uso de técnicas computacionais de manipulação, a análise e a distribuição de dados espaciais têm tornado os mapeamentos de solos mais rápidos e de menor custo

(Wolski *et al.*, 2017; Regmi; Rasmussen, 2018).

Apesar do progresso das pesquisas em mapeamento digital no Brasil, a maior parte dos mapas de solos está atualmente limitada à caracterização regional, com escassez de dados digitais em alta resolução para o território nacional, ou mesmo em nível regional. Nesse sentido, o Programa Nacional de Solos do Brasil (PronaSolos) tem a função de alavancar os levantamentos pedológicos e estabelecer um banco de dados integrado, no qual as informações são organizadas e sistematizadas para consulta pelo público em geral (Polidoro *et al.*, 2021). Com o objetivo de avançar na qualidade e na quantidade das informações sobre solos, o PronaSolos conta com a colaboração e o compartilhamento de informações locais. Considerando-se as atividades do PronaSolos na Região Sul do país, o estado do Paraná tem gerado mapas pedológicos mais detalhados, com maior potencial para apoiar estudos em mapeamento digital. Um exemplo disso consiste no levantamento detalhado de solos (escala 1:10.000) realizado em parte da Sub-bacia Hidrográfica Arroio Fundo (Gomes *et al.*, 2020), um importante legado de informações para a Mesorregião Oeste do Paraná, na qual a Bacia Hidrográfica Paraná 3, objeto deste trabalho, está inserida.

Em muitas situações, o desenvolvimento do solo ocorre em resposta à forma como a água se move na paisagem. Portanto, os atributos do terreno podem caracterizar os caminhos de fluxo de água e as interações deste com os atributos do solo, dos quais derivam as condições ambientais típicas de cada unidade de mapeamento. Atualmente, os modelos digitais de terreno atuam como variáveis predictoras do solo, descrevendo numericamente a variação espacial das características morfométricas da superfície da Terra (Sena *et al.*, 2020). Essas características, por sua vez, estão intimamente relacionadas aos processos de formação do solo. Os modelos digitais de elevação (MDE) podem ser obtidos gratuitamente e o cálculo de outras informações derivadas pode ser gerado em softwares gratuitos, ou seja, sem custos adicionais para o mapeamento digital de solos (Andrade *et al.*, 2020; Pelegrino *et al.*, 2016).

Considerando-se a hipótese de que algoritmos robustos de aprendizado de máquinas podem gerar mapas de solos a partir de informações extraídas de área-piloto, e serem reaplicadas na área de entorno (tendo características fisiográficas similares), com adequada acurácia, este trabalho teve como objetivos: (a) usar técnicas geoespaciais para identificar áreas na Bacia Hidrográfica Paraná 3, que têm feições fisiográficas semelhantes à área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo) para que nelas fossem inseridas informações de solos; e (b) criar modelos preditivos espaciais a partir dos algoritmos C5.0, partial least square e random forest, utilizando diferentes estratégias para extração de informações, e gerar o mapa preditivo de solos para as áreas da Bacia Hidrográfica Paraná 3 semelhantes à área-piloto.

## Material e Métodos

### a) Área total investigada

A caracterização e as análises geoespaciais foram realizadas, inicialmente, para toda a área da Bacia Hidrográfica Paraná 3 (BP3) (Figura 1A), para compreensão do potencial total de extrapolação das informações a partir da área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo). A BP3 situa-se na Mesorregião Oeste do Paraná, com extensão aproximada de 8.000 km.

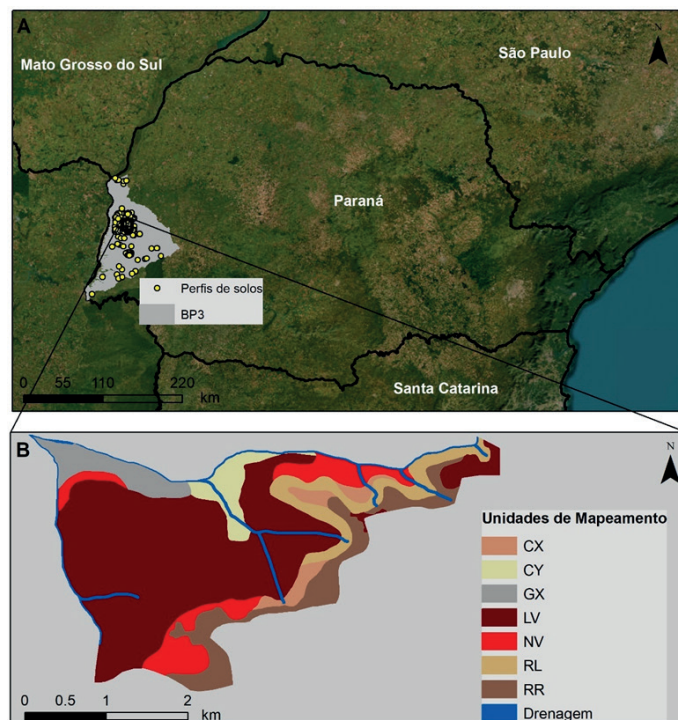


Figura 1 - Localização da Bacia Hidrográfica Paraná 3 (BP3) e perfis de solos utilizados para cálculo da acurácia do mapa final predito. Fonte: A – área total investigada e B – área-piloto (adaptado de Gomes *et al.*, 2020). CX: Cambissolo Háptico; CY: Cambissolo Flúvico; GX: Gleissolo Háptico; LV: Latossolo Vermelho; NV: Nitossolo Vermelho; RL: Neossolo Litólico; e RR: Neossolo Regolítico.

A área total investigada possui legado de mapeamento pedológico de solos em escala 1:250.000 (Gomes; Wrege, 2020). A precipitação anual varia de 1.400 mm, no noroeste, a 2.200 mm, no sudeste da região. O clima é do tipo Cfa, variando entre o clima subtropical e o subtemperado na classificação de Köppen, modificada por Maluf, para a Região Sul do Brasil (Wrege *et al.*, 2011). As elevações apresentam grande amplitude, variando de 82 a 1.008 m, o que confere diferentes componentes fisiográficos na região de estudo. O relevo tem, predominantemente, grau médio de dissecação. A maioria dos solos é originada do derrame de eruptivas básicas (basalto) pertencentes à Formação Serra Geral. Os solos restantes são derivados do Arenito Caiuá. A distribuição das áreas agrícolas encontra-se, principalmente, associada às áreas de relevo com baixa a média dissecação e solos mais profundos, localizados nos setores alto e baixo da bacia (Rocha; Bade, 2018; Gomes; Wrege, 2020).

A região é de grande importância para produção agrícola e de água, contendo bacias hidrográficas importantes, como as dos rios Paraná e Iguazu, que apresentam elevado potencial de produção de águas subterrâneas. Portanto, aumentar o nível de detalhe do mapeamento de solos da região é de grande importância para diferentes stakeholders: agricultores, hidrelétricas e tomadores de decisão, a exemplo do Serviço Geológico do Brasil (SGB), a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), o Parque Tecnológico Itaipu (PTI), as instituições ligadas à Secretaria da Agricultura do Paraná, como o Instituto de Desenvolvimento Rural (IDR), as cooperativas agrícolas, etc.

### b) Área-piloto

A área-piloto consiste em parte da Sub-bacia Hidrográfica Arroio Fundo, localizada no perímetro da mesorregião supracitada, cujo legado de solos consiste em mapeamento pedológico em escala 1:10.000 (adaptado de Gomes *et al.*, 2020). Por estar em escala detalhada, as informações das unidades de mapeamento de solos foram extraídas em sistema de informação geográfica e utilizadas para treinamento dos modelos preditivos espaciais. Os solos são derivados de eruptivas básicas (basalto). Existe o predomínio da unidade de mapeamento Latossolo Vermelho, seguido de Neossolo Regolítico, Nitossolo Vermelho, Neossolo Litólico, Gleissolo Háplico, Cambissolo Flúvico e Cambissolo Háplico (Figura 1B).

### c) Área mapeada digitalmente a partir de extrapolação das informações

Foram utilizadas técnicas geoespaciais para a caracterização da área-piloto e busca da área em seu entorno com características fisiográficas semelhantes. Como o material de origem dos solos (litologia basáltica) e o bioclima foram considerados constantes, as variações de relevo foram investigadas. Deste modo, o modelo digital de elevação (MDE) foi submetido ao algoritmo de classificação em intervalos naturais, também denominado de otimização de Jenks (Smith; Goodchild; Longley, 2021), calculado no software de sistema de informação geográfica ArcGIS, versão 10.8, da ESRI. Usando esse método de classificação, agrupamentos naturais relacionados ao relevo foram criados para a área total investigada. O algoritmo identifica quebras de classes de altitude, onde grandes contrastes ocorrem, agrupando valores semelhantes e maximizando as diferenças entre as classes. Assim, áreas geomorfologicamente contrastantes foram espacialmente delineadas. O mapa foi sobreposto àquele da área-piloto para sua caracterização e detecção espacial de condições similares na região de entorno.

### Modelo preditivo espacial

O modelo preditivo espacial seguiu o arcabouço do mapeamento digital de solos (McBratney *et al.*, 2003), conforme a Equação 1:

$$\text{Solos} = f(\text{solo}, \text{clima}, \text{organismos}, \text{relevo}, \text{material de origem}, \text{idade e localização}) + \epsilon \quad (1)$$

Onde: as informações de solos foram utilizadas como input (variável alvo) nos modelos e consistem nos dados obtidos a partir da área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo). Assim, para a extração das informações a partir dos polígonos das unidades de mapeamento de solos, tendo como base um raster com 12,5 m de resolução, foram avaliadas quatro estratégias diferentes: (i) retirada de 30 m; (ii) retirada de 60 m; (iii) retirada de 90 m das bordas de cada unidade de mapeamento de solo delineada; e (iv) considerando a área total dos polígonos das unidades de mapeamento de solos derivados de basalto. As bordas foram removidas, pois nas transições entre as unidades de mapeamento de solos ocorrem incertezas mais elevadas, que poderiam ser propagadas para os modelos preditivos espaciais. As covariáveis ambientais predictoras em formato digital representam os fatores de formação dos solos. Nesse caso, as informações de solos foram relacionadas com o relevo (variável preditora). Tendo-se em mente os fatores de formação do solo, o relevo está fortemente relacionado ao tempo (controlado pela erosão) (Resende *et al.*, 2014).

Já os demais fatores de formação dos solos foram considerados constantes nas áreas de estudo, considerando a escala de mapeamento e os objetivos aqui pretendidos. Os atributos do terreno foram calculados em sistema de informação geográfica a partir do MDE Alos-Palsar (ASF, 2023) com 12,5 m de resolução (Tabela 1). Os atributos do terreno foram extraídos da área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo), utilizando uma malha de 45.325 pontos espaçados regularmente em 12,5 metros.

Para extração das informações e treinamento dos modelos, foram selecionadas informações das classes de solos de maior expressão geográfica. Essa etapa foi realizada com base no conhecimento das relações solo-paisagem, a fim de reduzir o desbalanceamento de dados de entrada nos modelos preditivos. Devido às diferentes combinações envolvendo fatores e processos de formação dos solos, é frequente o desbalanceamento, o que pode causar superestimação ou subestimação, com consequente redução da acurácia dos modelos preditivos (Brungard *et al.*, 2015; Ma *et al.*, 2019; Shariffar *et al.*, 2019).

**TABELA 1.** Atributos do terreno, sua correspondência em inglês e descrição sucinta do algoritmo.

Atributo do terreno	Correspondência em inglês	Abreviação	Descrição sucinta
Modelo Digital de Elevação	Digital elevation model	DEM	Representa a elevação hidrologicamente consistente do terreno
Índice de resolução múltipla de nivelamento de topo de cume	Multiresolution index of ridge top flatness	MRRTF	Descreve o quão plano é o topo de um cume

CONTINUA

**TABELA 1.** Atributos do terreno, sua correspondência em inglês e descrição sucinta do algoritmo. (continuação)

Atributo do terreno	Correspondência em inglês	Abreviação	Descrição sucinta
Índice de resolução múltipla de nivelamento de fundo de vale	Multiresolution index of valley bottom flatness	MRVBF	Descreve o quão plano é o fundo do vale
Curvatura planar	Plan curvature	Plan_curv	Curvatura da superfície que é perpendicular à direção do fluxo. Valores negativos indicam fluxo divergente
Curvatura vertical	Profile curvature	Prof_curv	Forma da vertente no plano vertical (côncava, retilínea ou convexa)
Aspecto	Aspect	Aspect	Direção descendente da encosta mais íngreme
Declive	Slope	Slope	Taxa de variação da altitude
Elevação do declive	Slope height	Slope_H	Descreve a distância vertical entre a parte inferior e a superior da encosta
Elevação normalizada	Normalized height	NH	Converte os valores de elevação para o intervalo entre 0 e 1
Posição em relação ao meio da encosta	Mid-slope position	MSP	Descreve a posição da encosta entre a crista superior e o fundo do vale
Índice topográfico de umidade	Topographical Wetness Index by multiple-flow Quinn	FD8Q	Tendência do terreno de drenar ou acumular água, utilizando a metodologia de acumulação de fluxo proposta por Quinn <i>et al.</i> (1991)
Índice de umidade do Saga GIS	Saga Wetness Index	SWI	Tendência do terreno de drenar ou acumular água
Índice de potência de escoamento	Stream Power Index	SPI	Potencial erosivo do fluxo de água
Fator LS	LS Factor	LSF	Efeito do comprimento e inclinação da encosta na erosão (RUSLE)
Altura padronizada	Standardized height	SH	Distância vertical entre a base e o índice de inclinação padronizado
Profundidade do vale	Valley depth	VD	Distância vertical do nível de drenagem

As funções (f) testadas neste estudo consistem nos seguintes algoritmos de aprendizado de máquinas: (a) C5.0, que é baseado em árvores de decisão ajustadas a partir do algoritmo de Quinlan (1993); (b) partial least square (PLS), que reduz a dimensionalidade do conjunto de dados e maximiza a combinação linear de dependências com as variáveis de resposta (Kuhn, 2023); e (c) random forest, que se baseia na construção de múltiplas árvores de decisão durante o seu treinamento, as quais são posteriormente consolidadas para definir uma única predição para cada observação no conjunto de dados (Breiman, 2001). O índice mean decrease in Gini foi calculado a partir da random forest, o qual consiste em medir como cada covariável ambiental contribui para a homogeneidade dos nós e das folhas nas árvores de decisão da random forest.

Com base nas informações extraídas dos mapas e de atributos relacionados do terreno, foram gerados 12 modelos de predição, um para cada estratégia (retirando 30 m; retirando 60 m; e retirando 90 m; e considerando-se a totalidade dos

polígonos das unidades de mapeamento de solos derivados de basalto) e para cada algoritmo de aprendizado de máquina (C5.0, mínimos quadrados parciais e floresta aleatória). Os modelos foram treinados no software R (R Development Core Team, 2021), por meio do pacote caret (Kuhn, 2008). A partir dos modelos de predição treinados, foi realizada a extrapolação das informações das classes de solos de parte da Sub-bacia Hidrográfica Arroio Fundo (área-piloto) para áreas vizinhas fisiograficamente semelhantes na Bacia Hidrográfica Paraná 3 (área total investigada), sendo então inseridas as informações das classes de solos previstas por pixel.

O  $\mathcal{E}$  refere-se aos cálculos de acurácia dos mapas de solos gerados. Um conjunto de dados de validação externa, contendo perfis de solos não utilizados no modelo preditivo (dados observados), foram plotados sobre o mapa gerado (dados preditos). Duas metodologias de validação foram utilizadas:

- Validação cruzada, onde todos os pontos empregados na modelagem foram utilizados para o cálculo da acurácia; e
- Validação independente ou externa, onde 279 perfis de solos foram empregados para o cálculo da acurácia.

Os perfis de solos utilizados na Bacia Hidrográfica Paraná 3 foram obtidos de Cooper *et al.* (2005), Silva (2017) e Curcio *et al.* (2021). Tais informações foram utilizadas para cálculo de métricas de acurácia a partir da matriz de confusão (solos preditos versus solos estimados pelos modelos preditivos). A acurácia global, com valores que variam de 0 a 1, foi calculada pela soma da diagonal maior (amostras corretamente classificadas) dividida pelo número total de amostras em toda a matriz de confusão. Quanto mais próximo de 1, maior a acurácia. O coeficiente Kappa (Cohen, 1960; Sim; Wright, 2005) foi calculado a partir da seguinte equação:

$$\text{Kappa index} = (P_o - P_a) / (1 - P_a) \quad (2)$$

Onde:  $P_o$  é a proporção de amostras classificadas corretamente e  $P_a$  é a probabilidade de concordância aleatória.

### Etapas de geoprocessamento do mapa final

O mapa de maior acurácia (mapa de solos final), em formato raster, foi convertido ao formato vetorial e submetido a etapas de geoprocessamento para eliminação de polígonos espúrios decorrentes da etapa de predição espacial, sendo utilizados os filtros boundary clean e majority. Na sequência, foi aplicada a ferramenta eliminate, em que polígonos com área inferior a 10 ha foram incorporados a polígonos maiores. Todas as etapas foram calculadas no ArcGIS 10.8 da ESRI. Além disso, foi adicionado o critério fases de relevo, calculadas em SIG a partir da declividade do terreno, com os seguintes fatiamentos: relevo plano (0-3%), suave ondulado (3-8%), ondulado (8-20%), forte ondulado (20-45%), montanhoso (45-75%) e escarpado (>75%) (Santos *et al.*, 2015).

## Resultados e Discussão

### Determinação da área para extrapolação das informações

Na busca por áreas fisiograficamente semelhantes à área-piloto, inicialmente foi criado, em SIG, o perfil topográfico a partir do MDE (Figura 2A) com base na linha A-B no mapa, visando captar o máximo de variações na paisagem. O perfil topográfico (valores de altitude com relação à distância) revelou grandes contrastes geomorfológicos (Figura 2B), cujas altitudes variaram de aproximadamente 120 m, no Planalto de Foz do Iguaçu, até 800 m, no Planalto de Cascavel (Atlas [...], 2016; Gomes; Wrege, 2020). O MDE foi então reclassificado em intervalos naturais (Figura 2C). Plotando-se a área-piloto sobre o mapa (Figura 2D), foi possível observar que apenas dois grupamentos de relevos contrastantes ocorrem na área: de 82 a 318 m e de 318 a 430 m. O limite máximo de altitude foi plotado na Figura 2B (linhas vermelhas) e representa o perfil topográfico da área a ser mapeada. Os padrões de maior elevação, que se encontram fora das faixas de altitude, foram excluídos, resultando num total de 5.028 km<sup>2</sup> que são fisiograficamente semelhantes à área-piloto e que foram utilizados para extrapolação das informações via modelos preditivos apresentados a seguir.

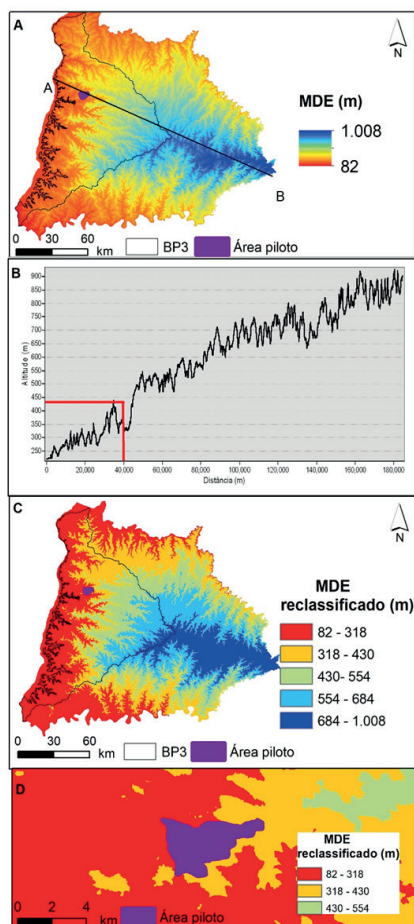


Figura 2 - Modelo digital de elevação evidenciando (A) a localização do perfil topográfico A-B que foi plotado em (B); Modelo Digital de Elevação (MDE) reclassificado a partir do algoritmo intervalos naturais de Jenks (C); e maior pormenorização do MDE reclassificado com destaque para a área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo). BP3 – Bacia do Paraná 3. Fonte: Elaborada pelos autores.

### Acurácia dos modelos preditivos espaciais

A Tabela 2 apresenta a acurácia global e o coeficiente Kappa dos modelos de predição espacial das classes de solos para os diferentes algoritmos testados e diferentes estratégias de validação. Os índices de validação cruzada, sempre com valores superiores aos obtidos via validação externa independente, denotaram adequada acurácia interna dos modelos preditivos, bem como sua estabilidade (Machado *et al.*, 2019). Os índices não refletiram variações relacionadas ao número de observações nos diferentes métodos, mas sim ao algoritmo em si, onde o PLS apresentou valores mais baixos de acurácia global e de coeficiente Kappa.

TABELA 2. Acurácia global (AG) e coeficiente Kappa (CK) dos modelos de predição espacial das classes de solos.

Algoritmo	AG	CK	AG	CK	AG	CK	AG	CK
	Estratégia							
	30 m		60 m		90 m		UM	
Validação cruzada – parte da Sub-bacia Hidrográfica Arroio Fundo								
C5.0	0,98	0,95	0,99	0,97	1,00	0,98	0,97	0,95
PLS	0,87	0,64	0,91	0,65	0,93	0,55	0,87	0,69
RF	0,97	0,93	0,98	0,95	0,99	0,97	0,97	0,94
Validação independente da área mapeada								
C5.0	0,43	0,26	0,49	0,33	0,36	0,17	0,34	0,14
PLS	0,56	0,44	0,52	0,38	0,41	0,25	0,48	0,30
RF	0,41	0,21	0,62	0,50	0,45	0,29	0,32	0,11

PLS – partial least square; RF – random forest; UM – polígono total da unidade de mapeamento de solos.

Considerando a validação independente, é possível observar que o algoritmo C5.0 apresentou o pior desempenho (valores mais baixos de acurácia global e de índice Kappa) quando comparado com PLS e random forest, independentemente da estratégia de treinamento dos modelos. Mais especificamente, as estratégias de retirada de dados próximos às bordas das unidades de mapeamento (remoção de 30 m e de 60 m) proporcionaram relativamente maior acurácia à predição espacial quando comparadas com as estratégias de remoção de 90 m e quando se considerou o polígono total das unidades de mapeamento de solos derivados de basalto.

Quanto maior a distância de remoção das bordas, menor a quantidade de pontos de aprendizado para os modelos preditivos, o que tem resultado em menor acurácia da predição espacial (Ma *et al.*, 2019; Shariffar *et al.*, 2019). Os valores mais elevados de acurácia foram obtidos a partir da random forest com remoção de 60 m das bordas, com acurácia global de 0,62 e índice Kappa de 0,50, sendo esse método utilizado para extrapolação das informações para as áreas de entorno da área-piloto. Coelho *et al.* (2021), analisando os mapeamentos digitais de classes de solos no Brasil, obtiveram um valor médio de 0,56 para a acurácia global, valor este um pouco abaixo do melhor modelo deste trabalho (0,62). A random forest tem se destacado entre os algoritmos de predição espacial das classes de solos, cujos recursos de bootstrapping e bagging apresentam elevado potencial para distinguir associações espaciais complexas (Heung *et al.*, 2016).

Diferentes abordagens sobre extrapolação de informações das classes de solos obtidas a partir de áreas-piloto (de referência) têm sido testadas nacionalmente (Crivellini *et al.*, 2009; Caten *et al.*, 2011; Bagatini *et al.*, 2016; Pelegriño *et al.*, 2016; Silva *et al.*, 2016; Wolski *et al.*, 2017; Gonçalves *et al.*, 2021) e internacionalmente (Lagacherie *et al.*, 2001; McKay *et al.*, 2010; Mello *et al.*, 2021). Na maior parte dessas pesquisas, foram obtidos índices de acurácia em torno de 0,50, devido à natureza das variações das classes de solos, fatores e processos de formação dos solos, e tamanho da área extrapolada. Os autores relatam ainda que os mapas resultantes promoveram considerável melhoria para áreas onde as informações de solos são inexistentes ou menos detalhadas. Machado *et al.* (2019), Wolski *et al.* (2017) e Silva *et al.* (2016) obtiveram acurácia de global de 0,61, 0,66 e 0,71, respectivamente, indicando que o uso de mapas legados como fonte de informação tem sido uma adequada alternativa para maior economia e rapidez no mapeamento digital de solos.

### Mapa final de solos e relação solo-paisagem

A Figura 3 mostra o mapa final de solos obtido a partir do modelo preditivo de maior acurácia. Na Tabela 3 são apresentadas as unidades de mapeamento de solos e suas respectivas expressões geográficas. Devido ao ganho de informações advindas dos modelos digitais do terreno, a área mínima mapeável elaborada em SIG foi de 10 ha. Posteriormente à esta etapa de geoprocessamento, informações referentes às fases de relevo foram adicionadas, considerando-se a importância do relevo nas tomadas de decisão em várias situações, com destaque para a mecanização agrícola, a vulnerabilidade à erosão hídrica, a dinâmica da água no solo e a recarga potencial de aquíferos, além do ambiente em geral.

Na busca pelo refinamento da relação solo-paisagem, foi calculado o índice mean decrease in Gini, cujos valores para cada covariável ambiental podem ser visualizados na Figura 4.

Quanto maior o índice, maior é a importância da variável no modelo preditivo, e, desse modo, destacaram-se neste estudo, principalmente as covariáveis ambientais: modelo digital de elevação (DEM), relacionado à altitude, e dois índices relacionados às distâncias entre as partes mais baixas e as mais elevadas das encostas, que são a elevação do declive (Slope\_H) e a altura padronizada (SH).

A Figura 5 destaca duas paisagens contrastantes na área mapeada. A Figura 5B destaca a região representativa de paisagens cuja amplitude entre as distâncias dos fundos do vale e dos topos são maiores, consistindo nas paisagens mais suavizadas (convexas). Em alguns locais, esses padrões podem ser também representados por menores variações de altitude em longas distâncias. Tais paisagens favorecem a ocorrência de solos profundos, representados pelos Latossolos, os quais são predominantes na porção oeste da bacia, são mais aptos para produção sustentada de grãos, são resistentes à erosão laminar (mas não em sulcos) e têm elevado potencial de recarga de aquíferos. Já a Figura 5C destaca a região representativa das paisagens mais a leste da bacia, cujo relevo caracteriza-se por ser mais dissecado (patamarizado). As amplitudes entre o nível mais baixo da drenagem até o topo das encostas são menores (vertentes mais curtas). Nessas porções predominam as unidades de mapeamento representadas por solos rasos e pedregosos, os quais são mais aptos para utilização com cultivos menos intensivos, ou mesmo, em casos extremos, devem compor áreas de reserva de flora e fauna. Naturalmente, essas áreas são muito vulneráveis à erosão, têm baixíssimo potencial de recarga de aquíferos e requerem cuidados especiais no seu manejo para não comprometer o ambiente.

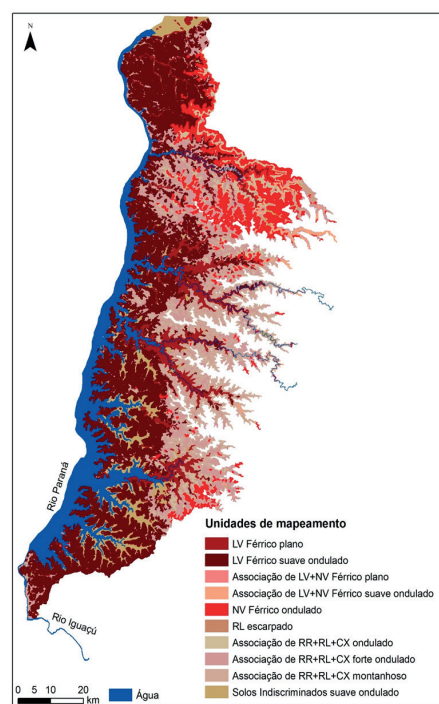


Figura 3 - Mapa digital final de solos de parte (5.028 Km<sup>2</sup>) da Bacia Hidrográfica Paraná 3 em escala semidetalhada. LV – Latossolo Vermelho; NV – Nitossolo Vermelho; RL – Neossolo Litólico; RR – Neossolo Regolítico; e CX – Cambissolo Háplico. Fonte: Elaborada pelos autores.

**TABELA 3.** Unidades de mapeamento de solos e suas respectivas expressões geográficas de parte (5.028 km<sup>2</sup>) da Bacia Hidrográfica Paraná 3.

Siglas	Unidades de mapeamento	Área (km <sup>2</sup> )	Área (%)
LV1	LV Férrico plano	466,13	9,27
LV2	LV Férrico suave ondulado	1.811,53	36,03
LV3	Associação de LV+NV Férrico plano	0,24	0,00
LV4	Associação de LV+NV Férrico suave ondulado	38,33	0,76
NV	NV Férrico ondulado	396,99	7,90
RL	RL escarpado	0,24	0,00
RR1	Associação de RR+RL+CX ondulado	204,90	4,08
RR2	Associação de RR+RL+CX forte ondulado	1.277,26	25,40
RR3	Associação de RR+RL+CX montanhoso	461,87	9,19
SI	Solos Indiscriminados suave ondulado	370,79	7,37
Total		5.028,28	100,00

## Considerações finais

As dificuldades relacionadas ao mapeamento digital de solos na Bacia Hidrográfica Paraná 3 envolvem padrões intrincados de distribuição das classes de solos nas encostas suavizadas e extensas mais próximas ao divisor das Bacias Hidrográficas Paraná 3 e Piquiri, nas encostas situadas mais próximas ao reservatório de Itaipu, nas encostas situadas em altitudes superiores a 430 m, nas encostas dobradas e de topos mais curtos em altitudes intermediárias, nas encostas de transição de eruptivas básicas-arenitos e nas encostas de arenitos propriamente ditas. A última dificuldade é explicada porque na área-piloto não há registro de ocorrência de arenitos (Gomes *et al.*, 2020).

Vale a pena ressaltar que o mapeamento pedológico semidetalhado da região Oeste do Paraná, na qual a Bacia Hidrográfica Paraná 3 está incluída (Curcio; Debrino; Ferlin, 2021), está em curso e constitui parte importante do Projeto PronaSolos PR.

## Conclusões

A partir de informação legada de solos, sem custo adicional de campanhas de campo, foi possível mapear digitalmente solos com adequada acurácia a partir do algoritmo random forest em escala semidetalhada, extraindo-se informações da área-piloto (parte da Sub-bacia Hidrográfica Arroio Fundo – 7,57 km<sup>2</sup>), relacionando-se classes de solos com modelos digitais do terreno e extrapolando as informações para 5.028 km<sup>2</sup> (de um total de 8.000 km<sup>2</sup>) da Bacia Hidrográfica Paraná 3, resultando em um ganho de 664 vezes em área mapeada.

Considerando que a Bacia Hidrográfica Paraná 3 é muito importante em termos de produção agrícola e de água e que o mapeamento pedológico semidetalhado está em curso (Projeto PronaSolos PR), isso pode constituir uma boa oportunidade para se checar sobre a possibilidade de complementariedade de informações geradas por essas duas metodologias de mapeamento de solos, as quais não são excludentes.

## Agradecimentos

Ao SGB, CNPq, Capes e Fapemig pelo suporte financeiro. MM, SHGS, SMB, RA, AFR e NC agradecem ao SGB pelas bolsas de estudo concedidas. À UFLA, UFMG, UFRB, UFRN e SGB pelo apoio. Ao pesquisador da Embrapa Florestas, João Bosco Vasconcellos Gomes, pela valiosa ajuda em várias partes deste trabalho.

## Referências

ALASKA Satellite Facility - ASF. Alos Palsar. Alaska, [2023]. Disponível em: <https://search.asf.alaska.edu/#/>. Acesso em: 12 set. 2023.

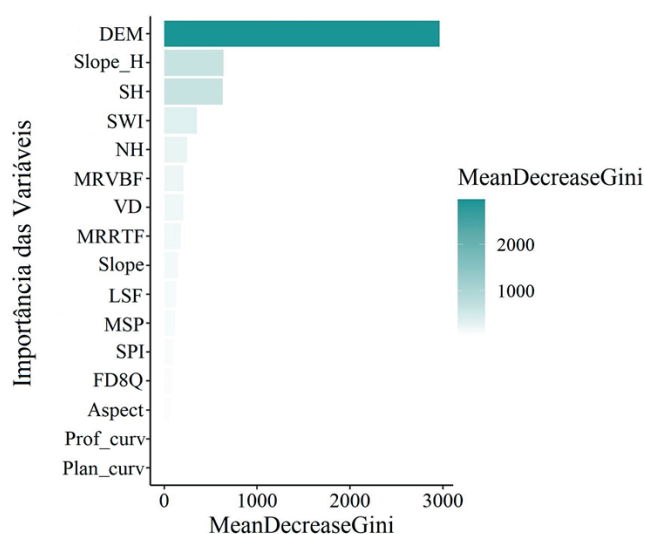


Figura 4. Valores do Mean Decrease Gini correspondentes às covariáveis ambientais calculados a partir do algoritmo random forest com a estratégia de extração de 60m das bordas das unidades de mapeamento de solos. DEM – digital elevation model; Slope\_H – slope height; SH - Standardized height; SWI - Saga Wetness Index; NH - Normalized height; MRVBF - Multiresolution index of valley bottom flatness; VD - Valley depth; MRRTF - Multiresolution index of ridge top flatness; LSF – LS factor; MSP - Mid-slope position; SPI - Stream Power Index; FD8Q - Topographical Wetness Index by multiple-flow Quinn; Prof\_curv - Profile curvature; e Plan\_curv - Plan curvature. A correspondência em inglês das abreviações pode ser vista na Tabela 1. Fonte: Elaborada pelos autores.

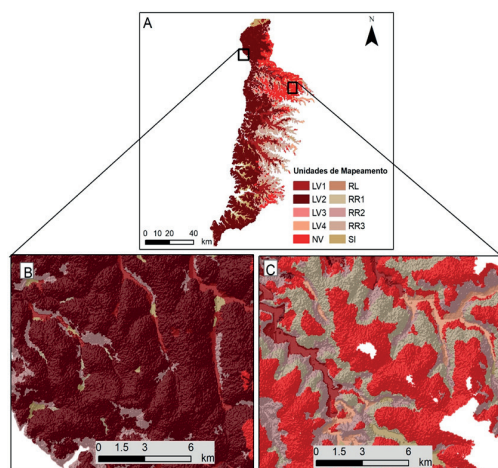


Figura 5 - Mapa de solos (A) com destaque para os dois tipos de paisagens contrastantes, onde (B) é a superfície mais suavizada (convexada) e (C) a mais dissecada (patamarizada). LV – Latossolo Vermelho; RL – Neossolo Litólico; RR – Neossolo Regolítico; NV – Nitossolo Vermelho; e SI – Solos Indiscriminados. Fonte: Elaborada pelos autores.

- ANDRADE, R.; SILVA, S. H. G.; FARIA, W. M.; POGGERE, G. C.; BARBOSA, J. Z.; GUILHERME, L. R. G.; CURI, N. Proximal sensing applied to soil texture prediction and mapping in Brazil. **Geoderma Regional**, Amsterdam, v. 23, e00321, 2020. DOI: <https://doi.org/10.1016/j.geodrs.2020.e00321>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S2352009420300705?via%3Dihub>. Acesso em: 31 out. 2023.
- ATLAS geomorfológico do estado do Paraná: escala base 1:250.000; modelos reduzidos 1:500.000. Curitiba: Mineropar; Universidade Federal do Paraná, 2016.
- BAGATINI, T.; GIASSON, E.; TESKE, R. Expansão de mapas pedológicos para áreas fisiograficamente semelhantes por meio de mapeamento digital de solos. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 51, p. 1317-1325, 2016.
- BREIMAN, L. Random forest. **Machine Learning**, v. 45, p. 5-32, 2001. DOI: <https://doi.org/10.1023/A:1010933404324>. Disponível em: <https://link.springer.com/article/10.1023/A:1010933404324>. Acesso em: 31 out. 2023.
- BRUNGARD, C. W.; BOETTINGER, J. L.; DUNIWAY, M. C.; WILLS, S. A.; EDWARDS, T. C. Machine learning for predicting soil classes in three semi-arid landscapes. **Geoderma**, Amsterdam, v. 239-240, p. 68-83, 2015. DOI: <https://doi.org/10.1016/j.geoderma.2014.09.019>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0016706114003516?via%3Dihub>. Acesso em: 31 out. 2023.
- CATEN, A. ten; DALMOLIN, R. S. D.; PEDRON, F. D. A.; SANTOS, M. D. L. M. Extrapolação das relações solo-paisagem a partir de uma área de referência. **Ciência Rural**, Santa Maria, RS, v. 41, p. 812-816, 2011.
- COELHO, F. F.; GIASSON, E.; CAMPOS, A. R.; TIECHER, T.; COSTA, J. J. F.; COBLINSKI, J. A. Digital soil class mapping in Brazil: a systematic review. **Scientia Agricola**, Piracicaba, SP, v. 78, n. 5, e20190227, 2021. DOI: <https://doi.org/10.1590/1678-992X-2019-0227>. Disponível em: <https://www.scielo.br/j/sa/a/YfN4xN33zCj8SBGw4TrwD3s/?lang=en>. Acesso em: 31 out. 2023.
- COELHO, M. R.; VASQUES, G. M. **Importância e potencial de mapas de atributos do solo em apoio à formulação e implementação de políticas públicas no Brasil: uma revisão**. Rio de Janeiro: Embrapa Solos, 2022. E-book. (Embrapa Solos. Documentos, 234). Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/242264/1/CNPS-DOC-234-2022.epub>. Acesso em: 10 set. 2023.
- COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, Los Angeles, v. 20, n. 1, p. 37-46, 1960. DOI: <https://doi.org/10.1177/001316446002000104>. Disponível em: <https://journals.sagepub.com/doi/10.1177/001316446002000104>. Acesso em: 31 out. 2023.
- COOPER, M.; MENDES, L. M. S.; SILVA, W. L. C.; SPAROVEK, G. A national soil profile database for Brazil available to international scientists. **Soil Science Society of America Journal**, New York, v. 69, p. 649-652, 2005. DOI: <https://doi.org/10.2136/sssaj2004.0140>. Disponível em: <https://access.onlinelibrary.wiley.com/doi/10.2136/sssaj2004.0140>. Acesso em: 31 out. 2023.
- CRIVELENTI, R. C.; COELHO, R. M.; ADAMI, S. B. F.; OLIVEIRA, S. R. D. M. Mineração de dados para inferência de relações solo-paisagem em mapeamentos digitais de solo. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 44, p.1707-1715, 2009.
- CURCIO, G. R.; BONNET, A.; CAVIGLIONE, J. H.; KACHAROUSKI, M.; FERLIN, D.; FERRIANI, L. M.; OLIVEIRA, D. P.; WIGO, M.; DEBRINO, M. A.; PETRY, A. L.; BRUSTOLON, R. **Levantamento semidetalhado e aptidão agrícola dos solos dos aldeamentos indígenas Tekoha Añetete e Tekoha Itamarã, Diamante D'Oeste, Paraná**. Colombo, PR: Embrapa Florestas, 2021.
- CURCIO, G. R.; DEBRINO, M. A.; FERLIN, D. Produtos do Projeto PronaSolos PR – Bloco 1. Curitiba: PronaSolos Paraná, 2021. Disponível em: <https://www.agricultura.pr.gov.br/PronasolosPR/Notícia/Produtos-do-Projeto-Pronasolos-PR-Bloco-1>. Acesso em: 14 out. 2023.
- GOMES, J. B. V.; WREGGE, M. S. **Municípios formadores da Bacia do Paraná 3 e Palotina: estudos de clima, solos e aptidão das terras para o cultivo do eucalipto**. Brasília, DF: Embrapa, 2020.
- GOMES, J. B. V.; BOGNOLA, I. A.; MATSUSHITA, M. S.; HOLLER, W.; WREGGE, M. S. **Levantamento detalhado de solos de uma área-piloto em parte da Microbacia do Arroio Fundo, município de Marechal Cândido Rondon, PR**. [Brasília]: Embrapa, 2020. GeoInfo Data Set. Disponível em: [https://geoinfo.cnpf.embrapa.br/layers/geonode%3Asolos\\_arroiofundo\\_wgs84\\_1](https://geoinfo.cnpf.embrapa.br/layers/geonode%3Asolos_arroiofundo_wgs84_1). Acesso em: 14 out. 2023.
- GONÇALVES, T. G.; PONS, N. A. D.; MELLONI, E. G. P.; MANCINI, M.; CURI, N. Digital soil mapping: predicting soil classes distribution in large areas based on existing soil maps from similar small areas. **Ciência e Agrotecnologia**, Lavras, MG, v. 45, e007921, 2021. DOI: <http://dx.doi.org/10.1590/1413-7054202145007921>. Disponível em: <https://www.scielo.br/j/cagro/a/H4Mj3zQNVpt-DdnTTvRC>



FP8w/?lang=en. Acesso em: 31 out. 2023.

HEUNG, B.; HO, H. C.; ZHANG, J.; KNUDBY, A.; BULMER, C. E.; SCHMIDT, M. G. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. **Geoderma**, Amsterdam, v. 265, p. 62-77, 2016. DOI: <https://doi.org/10.1016/j.geoderma.2015.11.014>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0016706115301300?via%3Dihub>. Acesso em: 31 out. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Manual técnico de Pedologia**. 2. ed. Rio de Janeiro: IBGE, 2007. Disponível em <https://biblioteca.ibge.gov.br/visualizacao/livros/liv37318.pdf>.

KUHN, M. Building predictive models in R using the caret package. **Journal of Statistical Software**, Zurich, v. 28, n. 5, p. 1-26, 2008. DOI: <https://doi.org/10.18637/jss.v028.i05>. Disponível em: 31 out. 2023.

KUHN, M. **C5.0 decision trees and rule-based models**: version 0.1.8. [S.l.]: CRAN UTC, 2023. Disponível em <https://cran.r-project.org/web/packages/C50/C50.pdf>. Acesso em: 9 set. 2023.

LAGACHERIE, P.; ROBBEZ-MASSON, J. M.; NGUYEN-THE, N.; BARTHÈS, J. P. Mapping of reference area representativity using a mathematical soilscape distance. **Geoderma**, Amsterdam, v. 101, p.105-118, 2001.

MA, Y.; MINASNY, B.; MALONE, B. P.; MCBRATNEY, A. B. Pedology and digital soil mapping (DSM). **European Journal of Soil Science**, London, v. 70, n. 2, p. 216-235, 2019. DOI: <https://doi.org/10.1111/ejss.12790>. Disponível em: <https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12790>. Acesso em: 31 out. 2023.

MACHADO, D. F. T.; SILVA, S. H. G.; CURTI, N.; MENEZES, M. D. Soil type spatial prediction from Random Forest: different training datasets, transferability, accuracy and uncertainty assessment. **Scientia Agricola**, Piracicaba, SP, v. 76, n. 3, p. 243-254, 2019. DOI: <https://doi.org/10.1590/1678-992X-2017-0300>. Disponível em: <https://www.scielo.br/j/sa/a/NLTch6nvK4Gx8hMgdJwzvtz/?lang=en>. Acesso em: 31 out. 2023.

MCBRATNEY, A. B.; MENDONÇA-SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, Amsterdam, v.17, p. 3-52, 2003.

MCKAY, J.; GRUNWALD, S.; SHI, X.; LONG, R. F. Evaluation of the transferability of a knowledge-based soil-landscape model. *In*: BOETTINGER, J. L.; HOWELL, D. W.; MOO-

RE, A. M.; HARTEMINK, A. E.; KIENAST-BROWN, S. (ed.). **Digital soil mapping**: bridging research, environmental application, and operation. London: Springer, 2010, p. 165-177.

MELLO, F. A.; DEMATTÊ, J. A.; RIZZO, R.; DOTTO, A. C.; POPPIEL, R. R.; MENDES, W. D. S.; GUIMARÃES, C. C. Expert-based maps and highly detailed surface drainage models to support digital soil mapping. **Geoderma**, Amsterdam, v. 384, e114779, 2021. DOI: <https://doi.org/10.1016/j.geoderma.2020.114779>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0016706120325349?via%3Dihub>. Acesso em: 31 out. 2023.

PELEGRINO, M. H. P.; SILVA, S. H. G.; MENEZES, M. D.; SILVA, E.; OWENS, P. R., CURTI, N. Mapping soils in two watersheds using legacy data and extrapolation for similar surrounding areas. **Ciência e Agrotecnologia**, Lavras, MG, v. 40, p. 534-546, 2016.

POLIDORO, J. C.; COELHO, M. R.; CARVALHO FILHO, A.; LUMBRERAS, J. F.; OLIVEIRA, A. P.; VASQUES, G.; MACARIO, C. G. N.; VICTORIA, D. C.; BHERING, S. B.; FREITAS, P. L.; QUARTAROLI, C. F.; MENDONÇA-SANTOS, M. L. (ed.) **Programa Nacional de Levantamento e Interpretação de Solos do Brasil (PronaSolos)**: diretrizes para implementação. Rio de Janeiro: Embrapa Solos, 2021. E-book. (Embrapa Solos. Documentos, 225). Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/226841/1/CNPS-DOC-225-2021.epub>. Acesso em: 18 dez. 2022.

QUINLAN, R. **C4.5**: programs for machine learning. San Mateo, US: Morgan Kaufmann Publishers, 1993.

QUINN, P.; BEVEN, K.; CHEVALLIER, P.; PLANCHON, O. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. **Hydrological Processes**, New York, v. 5, n. 1, p. 59-79, 1991. DOI: <https://doi.org/10.1002/hyp.3360050106>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/hyp.3360050106>. Acesso em: 31 out. 2023.

REGMI, N. R.; RASMUSSEN, C. Predictive mapping of soil-landscape relationships in the arid Southwest United States. **Catena**, Amsterdam, v.165, p. 473-486, 2018.

RESENDE, M.; CURTI, N.; REZENDE, S. B.; CORRÊA, G. F.; KER, J. C. **Pedologia**: base para distinção de ambientes. 6. ed. Lavras, MG: Editora UFLA, 2014.

ROCHA, A. S.; BADE, M. R. **Geografia da Bacia Hidrográfica do Paraná 3**: fragilidade e potencialidades socioambientais. Jundiaí, SP: Editora In House, 2018.

SANTOS, R. D.; SANTOS, H. G.; KER, J. C.; ANJOS, L. H. C.; SHIMIZU, S. H.; **Manual de descrição e coleta de solos no campo**. 7. ed. Viçosa, MG: SBCS, 2015.

SENA, N. C.; VELOSO, G. V.; FERNANDES-FILHO, E. I.; FRANCELINO, M. R.; SCHAEFER, C. E. G. R. Analysis of terrain attributes in different spatial resolutions for digital soil mapping application in Southeastern Brazil. **Geoderma Regional**, Amsterdam, v. 21, e00268, 2020. DOI: <https://doi.org/10.1016/j.geodrs.2020.e00268>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S2352009420300171?via%3Dihub>. Acesso em: 31 out. 2023.

SHARIFIFAR, A.; SARMADIAN, F.; MALONE, B. P.; MINASNY, B. Addressing the issue of digital mapping of soil classes with imbalanced class observations. **Geoderma**, Amsterdam, v. 350, p. 84-92, 2019. DOI: <https://doi.org/10.1016/j.geoderma.2019.05.016>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0016706118319499?via%3Dihub>. Acesso em: 31 out. 2023.

SILVA, B. A. **Mapeamento convencional e digital de solos na folha topográfica de Marechal Cândido Rondon-PR-BR**. 2017. 92 f. Dissertação (Mestrado em Geografia) - Universidade Estadual do Oeste do Paraná, Cascavel, PR, 2017.

SILVA, S. H. G.; ANDRADE, R.; ROSADO, L. R.; CURCIO, G. R.; GOMES, J. B. V.; DEBRINO, M. A.; KACHAROUSKI, M.; FERLIN JÚNIOR, D. P.; GUILHERME, L. R. G.; CURI, N. **Boletim Técnico**: predição de atributos de solos das Bacias Hidrográficas Paraná 3 e parte da Piquiri, de maneira ágil, acurada, com baixo custo e sem geração de efluentes químicos, com apoio da fluorescência de raios-X portátil e de variáveis obtidas em campo. [Curitiba]: Embrapa Florestas; Pronasolos PR; Itaipu; UFLA, [2023]. Disponível em: [https://www.agricultura.pr.gov.br/sites/default/arquivos\\_restritos/files/documento/2023-02/i\\_boletim\\_tecnico\\_pti\\_0.pdf](https://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/documento/2023-02/i_boletim_tecnico_pti_0.pdf) Acesso em: 17 out. 2023.

SILVA, S. H. G.; MENEZES, M. D.; OWENS, P. R.; CURI, N. Retrieving pedologist's mental model from existing soil map and comparing data mining tools for refining a larger area map under similar environmental conditions in Southeastern Brazil. **Geoderma**, Amsterdam, v. 267, p. 65-77, 2016.

SIM, J.; WRIGHT, C. C. The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. **Physical Therapy Rehabilitation Journal**, Oxford, v. 85, n. 3, p. 257-268, 2005. DOI: <https://doi.org/10.1093/ptj/85.3.257>. Disponível em: <https://academic.oup.com/ptj/article/85/3/257/2805022>. Acesso em: 31 out. 2023.

SMITH, M. J.; GOODCHILD, M. F.; LONGLEY, P. A.. **Geospatial analysis: a comprehensive guide**. 6. ed. [S. l.]: Spatial Analysis Online, 2021. Disponível em <https://www.spatialanalysisonline.com/HTML/index.html>. Acesso em: 7 de setembro de 2023.

WOLSKI, M. S.; DALMOLIN, R. S. D.; FLORES, C. A.; MOURA-BUENO, J. M.; CATEN, A. ten; KAISER, D. R. Digital soil mapping and its implications in the extrapolation of soil-landscape relationships in detailed scale. **Pesquisa Agropecuária Brasileira**, Brasília, v. 52, p. 633-642, 2017.

WREGE, M. S.; STEINMETZ, S.; REISSER JUNIOR, C.; ALMEIDA, I. R. **Atlas climático da Região Sul do Brasil**: estados do Paraná, Santa Catarina e Rio Grande do Sul. Pelotas, RS: Embrapa Clima Temperado; Colombo, PR: Embrapa Florestas, 2011.



### INFORME TÉCNICO BACIA DO PARANÁ 3 V.1, N.1, 2024

Publicação on-line seriada do Serviço Geológico do Brasil – SGB  
Diretoria de Hidrologia e Gestão Territorial  
Departamento de Gestão Territorial – DEGET

Disponível em: [rigeo.cprm.gov.br](http://rigeo.cprm.gov.br)

Serviço Geológico do Brasil – SGB  
Av. Pasteur, 404 Urca - Rio de Janeiro - RJ - BRASIL  
CEP: 22.290-255  
Telefone:(21) 2295-0032  
Contatos: [seus@cprm.gov.br](mailto:seus@cprm.gov.br) /

### CRÉDITOS DA BACIA DO PARANÁ 3

**Diretora de Hidrologia e Gestão Territorial:** Alice Silva de Castilho

**Departamento de Hidrologia:** Andrea de Oliveira Germano

**Editor:** Marta Vasconcelos Ottoni

**Corpo Editorial:** Nilton Curi

**Corpo de revisores:** João Bosco Vasconcellos Gomes

**Revisão de texto:** Irinéa Barbosa da Silva

**Normalização bibliográfica:** Rede de Bibliotecas Ametista Editoração eletrônica: Divisão de Editoração Geral – DIEDIG