



Foto: Arquivo do Grupo de Pesquisa da UFSM.

## DADOS PARA APLICAÇÕES PEDOMÉTRICAS EM LARGA ESCALA NO BRASIL

*Alessandro Samuel-Rosa  
Gustavo M. Vasques*

**A** pedometria é a disciplina da Ciência do Solo dedicada ao estudo da variação espacial e temporal do solo e, assim da pedogênese, lançando mão de tecnologias da informação (TIs) para a coleta, armazenamento, manipulação, modelagem e distribuição de dados do solo. O insumo básico de uma aplicação pedométrica são os dados do solo, que adquirem o papel de variáveis de resposta, ou

dependentes. As variáveis de resposta são aquelas variáveis que se deseja compreender ou estimar, ou seja, as propriedades, características, atributos do solo, como o conteúdo de argila e a classe taxonômica.

O mapeamento digital do solo, por exemplo, é um método pedométrico que pressupõe que o solo e suas características são resultado de

processos controlados pelos componentes ambientais que moldam a paisagem. Nesse caso, dados representando os componentes ambientais na paisagem são usados como variáveis explicativas ou independentes para prever variáveis de resposta do solo. Os dados do solo também podem servir como variáveis explicativas, como no caso das chamadas funções de pedotransferência, em que uma

variável do solo - como o conteúdo de matéria orgânica - pode ser estimada a partir de outra variável, como a sua cor, por exemplo.

Historicamente, a maneira mais comum de obtenção de dados do solo é a sua amostragem, usando métodos e ferramentas para a sua coleta e descrição no campo, seguidos pela análise das amostras no laboratório. Essa foi a maneira com a qual inúmeros projetos de levantamento e de pesquisa de solos geraram, ao longo de décadas, um grande volume de dados legados, literalmente 'deixados para as próximas gerações. Muitos desses dados estão compilados em repositórios institucionais, com abrangências nacional, regional e global (Tabela 1), porém uma boa parte deles encontra-se dispersa e sob o risco de não ser reaproveitada ou mesmo ser perdida para sempre.

No Brasil, a principal fonte de dados compilados é o Sistema de Informação de Solos Brasileiros ([https://www.bdsolos.cnptia.embrapa.br/consulta\\_publica.html](https://www.bdsolos.cnptia.embrapa.br/consulta_publica.html)), da Embrapa, que conta com dados de cerca de 9.000 perfis de solos. O IBGE também disponibiliza dados, principalmente os do Projeto Radambrasil, em sua página (<http://www.downloads.ibge.gov.br>). Juntos, os dados sob salvaguarda da Embrapa e do IBGE somam cerca de 10 mil observações (pontos amostrais, perfis) do solo que estão sendo usados para construir o futuro Repositório Brasileiro Livre para Dados Abertos do Solo (<http://coral.ufsm.br/febr/>) (Figura 1).

Além dos métodos convencionais de amostragem de solos (por exemplo, locais de amostragem são escolhidos a dedo e amostras de terra são coletadas para análise

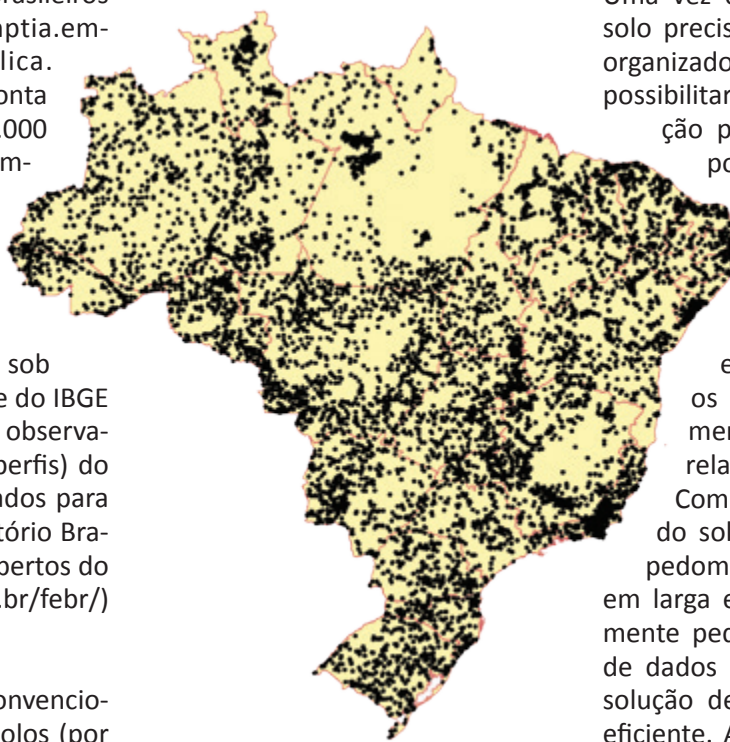
**Tabela 1** Bases de dados de solos e estimativa do número de observações que contém

Base de dados	Endereço	Observações
Africa Soil Profiles Database	<a href="http://africasoils.net/services/data/soil-databases">http://africasoils.net/services/data/soil-databases</a>	18.500
Archivo Digital de Perfiles de Suelo de México	<a href="http://www.inegi.org.mx/geo/contenidos/recnat/edafologia/PerfilesSuelo.aspx">http://www.inegi.org.mx/geo/contenidos/recnat/edafologia/PerfilesSuelo.aspx</a>	10.000
BrazilSoilDB	<a href="http://www.esalq.usp.br/gerd">http://www.esalq.usp.br/gerd</a>	5.500
Sistema de Informação de Solos Brasileiros	<a href="https://www.bdsolos.cnptia.embrapa.br/consulta_publica.html">https://www.bdsolos.cnptia.embrapa.br/consulta_publica.html</a>	9.000
Sistema de Información de Suelos del INTA	<a href="http://sisinta.inta.gob.ar">http://sisinta.inta.gob.ar</a>	3.000
Soil and Landscape Grid of Australia	<a href="http://www.clw.csiro.au/aclep/soilandlandscapegrid">http://www.clw.csiro.au/aclep/soilandlandscapegrid</a>	280.000
USA National Cooperative Soil Survey	<a href="https://sdmdataaccess.nrcs.usda.gov">https://sdmdataaccess.nrcs.usda.gov</a>	50.000
World Soil Information Service	<a href="http://www.isric.org/explore/wosis">http://www.isric.org/explore/wosis</a>	100.000

se em laboratório), existem hoje métodos alternativos tanto para definir os locais de amostragem segundo critérios objetivos, usando técnicas de inteligência artificial<sup>1</sup>, quanto para analisar as amostras de terra de maneira mais limpa e não-destrutiva, ao invés de via análise química, como a espectro-

metria de infravermelho ou a análise da susceptibilidade magnética do solo. Tais métodos, chamados em conjunto de sensoriamento proximal do solo, podem ser aplicados *in situ* para obter dados do solo sem a necessidade de coletar amostras de terra (Figura 2).

Uma vez coletados, os dados do solo precisam ser armazenados e organizados apropriadamente para possibilitar uma eficiente distribuição para uso atual e futuro por qualquer pessoa interessada. Existem inúmeras alternativas tecnológicas para o armazenamento dos dados, desde planilhas eletrônicas (CSV) até os sistemas de gerenciamento de bases de dados relacionais (PostgreSQL). Como o volume de dados do solo usado em aplicações pedométricas, mesmo quando em larga escala, ainda é relativamente pequeno (<< 1 Gb), bases de dados relacionais ainda são a solução de armazenamento mais eficiente. Além disso, elas são capazes de acomodar dados de dife-



**Figura 1.** Localização das cerca de 10 mil observações do solo disponíveis para uso nas plataformas on-line da Embrapa e do IBGE. Fonte: Acervo do Grupo de Pesquisa da UFSM.

<sup>1</sup> Boletim Informativo SBCS V 42.N 1, 2016.





**Figura 2.** Coleta de dados do solo utilizando sensores proximais (espectrômetro de raios gama à esquerda e medidor de condutividade elétrica e susceptibilidade magnética à direita).

rentes tipos, georreferenciados ou não, com diferentes conteúdos e formatos: desde observações pontuais contendo poucos (somente análises de rotina do solo) a milhares de dados (análises de rotina + curva espectral com ~ 2000 bandas), além de permitirem consultas complexas, com resultados produzidos em frações de segundo.

Considerando as tecnologias disponíveis, o sistema de gerenciamento de dados do solo deve fornecer as características mínimas necessárias para as etapas posteriores de modelagem dos dados. No caso da pedometria, envolvem a análise dos dados do solo e da sua relação com os componentes da paisagem para produzir modelos de representação do comportamento espacial e/ou temporal do solo; modelos de predição dos seus atributos; modelos das relações solo-paisagem, enfim, para permitir o entendimento da formação e distribuição do solo e dos seus atributos. Entre essas características, idealmente, o sistema teria: (a) possibilidade de acesso on-line gratuito via navegador ou serviços web; (b) ferramentas básicas de visualização dos dados e geopro-

cessamento, como zoom, inspeção de atributos (botão info), buffering, spatial join; (c) ferramentas de consulta de dados espacial e/ou por valor de atributos; (d) possibilidade de exportação dos dados em formatos abertos, como CSV e KML; e (e) ferramentas de comunicação e colaborativas para a manutenção, conferência e atualização dos dados pela comunidade de mantenedores e usuários do sistema. Infelizmente, essa não é a realidade das bases de dados de solos.

A geração de informações atualizadas do solo para todo o Brasil requer a manipulação de um grande volume de dados, os quais idealmente deveriam estar organizados e disponíveis para acesso em bases de dados públicas e/ou privadas abertas. No entanto, estima-se que boa parte (talvez a maior parte) dos dados legados estão dispersos em repositórios institucionais de acesso restrito, em computadores pessoais e/ou em meios analógicos (impresos em papel). Caberia aqui um esforço conjunto da comunidade científica para recuperar esses dados e torná-los disponíveis para uso, bem como para revisar a qualidade dos

dados já disponíveis, a exemplo do Repositório Brasileiro de Dados de Ferro do Solo ([www.ufsm.br/febr](http://www.ufsm.br/febr)), da Universidade Federal de Santa Maria, e da Biblioteca Espectral de Solos do Brasil (<http://bibliotecaespectral.wixsite.com/esalq>), da Universidade de São Paulo.

Plataformas multiusuário, sistemas de gerenciamento e compartilhamento de dados on-line, ferramentas robustas de análise de Big Data, e processos de geração contínua de informações do solo, em múltiplas escalas, se tornarão cada vez mais comuns e devem ser invariavelmente assimilados pela comunidade de cientistas do solo e traduzidos para a comunidade de usuários de informações. A pedometria pode ajudar sobremaneira para esse fim, fornecendo ferramentas para a coleta, organização, análise e distribuição dos dados legados, atuais e futuros de solos.

**Alessandro Samuel-Rosa** Universidade Federal de Santa Maria. E-mail: [alessandrosamuel@mail.ufsm.br](mailto:alessandrosamuel@mail.ufsm.br)  
**Gustavo M. Vasques** Embrapa Solos. E-mail: [gustavo.vasques@embrapa.br](mailto:gustavo.vasques@embrapa.br).