



## NOVIDADES E DESAFIOS DO MAPEAMENTO DIGITAL DO SOLO DE ALTA RESOLUÇÃO

Alessandro Samuel-Rosa<sup>(1)</sup>, Lucas Rios do Amaral<sup>(2)</sup>

**RESUMO:** O mapeamento digital do solo (MDS) ganhou popularidade nos últimos anos. O principal motivo é a facilidade e rapidez com que permite produzir informações do solo. Neste artigo nós descrevemos algumas das mais importantes novidades e alguns dos maiores desafios do MDS de alta no Brasil. Para isso, o MDS é analisado em função das suas aplicações em grandes e pequenas extensões territoriais. No primeiro caso, a grande novidade é o rápido desenvolvimento da capacidade computacional instalada, principalmente via computação na nuvem. Sem a computação de alto desempenho não seria possível realizar MDS de alta resolução em grandes extensões. Da mesma maneira como vem sendo fundamental a disponibilidade cada vez maior de dados espaciais de alta resolução do sensoriamento remoto. Mas para que as informações do solo produzidas em escala global ou nacional tenham utilidade local, é necessário vencer o desafio de aumentar a disponibilidade de dados de observações do solo. O compartilhamento aberto dos dados existentes é fundamental para isso, mas nada substituirá a coleta de novos dados do solo no campo. Já nas aplicações do MDS de alta resolução em pequenas extensões, sendo a agricultura de precisão o principal exemplo, a disponibilidade de dados não é um problema. De fato, nessas aplicações é mais o alto custo com a obtenção periódica de dados que mais preocupa. Por isso, ganha cada vez mais espaço o uso de sensores proximais para obtenção de dados indiretos sobre as propriedades do solo. Mas outras possibilidades também precisam ser exploradas. Uma delas são as configurações amostrais otimizadas via algoritmos inteligentes. Outra são informações espaciais auxiliares sobre o terreno e as práticas de uso e manejo das lavouras ao longo do tempo. Inclusive, essas informações, quando disponíveis em alto nível de detalhe espacial e temporal, poderiam ser usadas para construir modelos mecanísticos do funcionamento espaço-temporal das lavouras e fazendas. Isso permitiria produzir informações relevantes sobre o solo com menor demanda de dados e, portanto, custos mais baixos.

**Palavras-chave:** Modelagem espacial preditiva, Agricultura de precisão, Geoestatística, Bases de dados, Covariáveis ambientais.

<sup>(1)</sup> Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil. E-mail: [alessandrorosa@utfpr.edu.br](mailto:alessandrorosa@utfpr.edu.br).

<sup>(2)</sup> Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola, Campinas, São Paulo, Brasil. E-mail: [lucas.amaral@feagri.unicamp.br](mailto:lucas.amaral@feagri.unicamp.br).

## **NOVELTIES AND CHALLENGES OF HIGH RESOLUTION DIGITAL SOIL MAPPING**

**ABSTRACT:** Digital soil mapping (DSM) has gained popularity in recent years. The main reason is the ease and speed with which it allows to produce soil information. In this article we describe some of the most important novelties and some of the biggest challenges of high resolution DSM in Brazil. DSM is analyzed according to its applications in large and small territorial extensions. In the first case, the great novelty is the rapid development of the installed computational power, mainly through cloud computing. Without high performance computing it would not be possible to perform high resolution DSM over large areas. In the same way, the increasing availability of high resolution remote sensing data has been fundamental. But for soil information produced on a global or national scale to have local utility, the challenge of increasing the availability of soil observation data needs to be addressed. Openly sharing the existing soil data is critical to this, but it will not replace the sampling of new soil data in the field. In the case of small-scale high-resolution DSM applications, precision agriculture being the main example, data availability is not a problem. In fact, in these applications it is more the high costs with the periodical collection of data that worries most. Therefore, the use of proximal sensors to obtain indirect data on soil properties is gaining more and more space. But other possibilities also need to be explored. One of these are sample configurations optimized via intelligent algorithms. Another is ancillary spatial information about the terrain and land use and management practices over time. These information, when available in a high level of spatial and temporal detail, could be used to construct mechanistic models of the spatio-temporal functioning of fields and farms. This would make it possible to produce relevant soil information with lower data demand and, therefore, lower costs.

**Key words:** Predictive spatial modeling, Precision agriculture, Geostatistics, Databases, Environmental covariates

O mapeamento digital do solo (MDS) é uma técnica que ganhou bastante popularidade nos últimos anos. O principal motivo é a facilidade e rapidez com que o MDS permite produzir informações espaciais do solo. Não é à toa que o MDS está se tornando a técnica oficial de produção de informação espacial do solo tanto para instituições nacionais (Polidoro et al., 2016) como internacionais (Hengl et al., 2014). Mas assim como qualquer tecnologia, apesar da sua popularidade, o MDS ainda enfrenta alguns desafios.

## **DEFININDO O MAPEAMENTO DIGITAL DO SOLO**

Um dos desafios enfrentados pelo MDS no Brasil é a percepção equivocada de que o mesmo consiste numa técnica usada apenas por pedólogos – cientistas do solo especializados em gênese, morfologia, classificação e levantamento do solo – para produzir mapas de tipos de solo em áreas de média a grande extensão (municípios, estados, países). Talvez a principal razão para o surgimento dessa percepção é o fato de que foram os pedólogos, abordando o mapeamento de tipos de solo via métodos numéricos – regressão logística, árvore de decisão, redes neurais –, que primeiro usaram a expressão *digital soil mapping* em publicações (Giasson et al., 2006; Figueiredo et al., 2008; ten Caten et al., 2012). Uma característica interessante dessas primeiras publicações é a tradução da expressão *digital soil mapping* como *mapeamento digital de solos*, com o termo *solos*, no plural. Para pedólogos, mapear *solos* significa estabelecer limites abruptos arbitrários para separar áreas homogêneas, sendo cada uma delas considerada *um solo*, o qual recebe um nome conforme a classe de solo mais comum. Publicada concomitantemente aos primeiros trabalhos dos pedólogos brasileiros, a definição formal do MDS mostrava que o mesmo trata o solo como um todo contínuo, sem limites abruptos arbitrários, abrangendo muito mais do que o mapeamento de classes de solo:

*O mapeamento digital do solo consiste na [...] criação e população de sistemas de informação espacial do solo usando modelos numéricos que inferem as variações espacial e temporal dos tipos e propriedades do solo a partir de observações e conhecimento do solo e de variáveis ambientais relacionadas.* (McBratney et al., 2003; Lagacherie e McBratney, 2007; Ma et al., 2019)

Se revisarmos o uso de métodos numéricos e observações do solo coletadas no campo para produção de informação espacial do solo, veremos que os primeiros trabalhos de MDS no Brasil foram realizados, pelo menos, já na década de 1980 (Tabela 1). Trata-se de trabalhos conduzidos principalmente por edafólogos – cientistas do solo especializados em química, física e biologia do solo com enfoque na produção vegetal –, que aplicaram o método numérico da geoestatística para estudar a variação espacial de propriedades do solo em campos

experimentais e lavouras. Mais tarde, tais trabalhos serviram de base para os primeiros estudos de agricultura de precisão no Brasil, a qual é um sistema de gerenciamento agrícola que leva em conta o conhecimento e manejo ultradetalhado da variação espacial e temporal das lavouras (Molin et al., 2015), conforme revelado pelo uso de dados empíricos e métodos numéricos. Infelizmente, muitos dos cientistas do solo que trabalham com agricultura de precisão, ao mapearem as propriedades do solo usando métodos numéricos como a geoestatística, ou mesmo interpoladores determinísticos, não percebem que o que estão praticando é MDS.

**Tabela 1.** Alguns dos trabalhos pioneiros de mapeamento digital do solo conduzidos no Brasil utilizando o método numérico da krigagem ordinária para predição espacial.

Ano	Trabalho
1987	PREVEDELLO, B. M. S. Variabilidade espacial de parâmetros de solo e planta. Piracicaba, SP, 1987. 166 p. Tese (Doutorado em Solos e Nutrição de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 1987.
1988	SILVA, A. P. Variabilidade espacial de atributos físicos do solo. Piracicaba, SP, 1988. 98 p. Tese (Doutorado em Solos e Nutrição de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 1988.
1989	DOURADO NETO, D. Variabilidade espacial das alturas de chuva e irrigação e de potenciais da solução do solo. Piracicaba, SP, 1989. 180 p. Dissertação (Mestrado em Irrigação e Drenagem) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 1989.
1990	TURATTI, A. L. Variabilidade espacial do armazenamento de água em Terra Roxa Estruturada. Piracicaba, SP, 1990. 88 p. Tese (Doutorado em Solos e Nutrição de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 1990.
1991	CARVALHO, T. M. Variabilidade espacial de propriedades físico-hídricas de um Latossolo Vermelho Amarelo através da geoestatística. Lavras, MG, 1991. 84 p. Dissertação (Mestrado em Solos e Nutrição de Plantas) - Escola Superior de Agricultura de Lavras, Lavras, MG, 1991.
1993	GUIMARÃES, E. C. Variabilidade espacial da umidade e da densidade do solo em um Latossolo Roxo. Campinas, SP, 1993. 138 p. Dissertação (Mestrado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, SP, 1993.

Há muito tempo, edafólogos e pedólogos têm alguma dificuldade para trabalhar juntos e aprender uns com os outros. Uma das principais razões para isso é a maneira como cada um desses especialistas aborda o estudo da variação espacial e temporal do solo. Ao mesmo tempo em que cada especialidade se aprofunda, afunila o seu conhecimento, reduzem-se as oportunidades de contato com outras áreas do conhecimento, as quais nem sempre são tão distantes quanto se poderia imaginar. Como era de se esperar, essa dificuldade trouxe mais malefícios do que benefícios à ciência do solo e à agricultura brasileiras. Ambos os especialistas, edafólogos e pedólogos, deixaram de acompanhar e compartilhar importantes

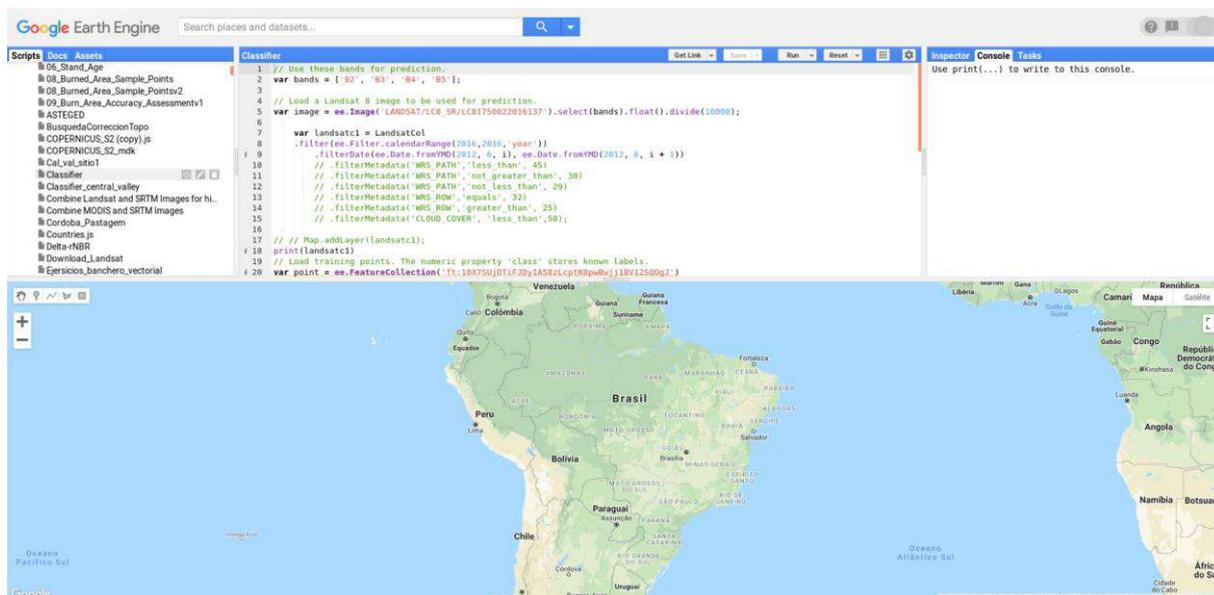
evoluções tecnológicas que estavam ocorrendo no início da década de 1990. Entre elas, a combinação de dados espaciais auxiliares e métodos numéricos (Heuvelink e Bierkens, 1992; Moore et al., 1993; Odeh et al., 1995) e os novos métodos numéricos de mineração de dados e aprendizado de máquina (Wolpert, 1992; Breiman, 1996; Schumacher et al., 1996). Demoraram quase duas décadas para que os mapeadores do solo, atuando em pequenas, médias ou grandes extensões, começassem a atualizar e compartilhar suas caixas de ferramentas.

## **MDS DE ALTA RESOLUÇÃO**

Conforme dito acima, o objetivo do MDS é produzir mapas, tanto de classes, como de propriedades do solo. Para isso, são usados métodos numéricos – geostatística, mineração de dados, aprendizado de máquina – e observações do solo coletadas no campo. Por muito tempo, a produção dessas informações do solo em alta resolução esteve restrita às áreas de pequena extensão (lavouras e campos experimentais). Contudo, o acelerado desenvolvimento da capacidade computacional instalada observada nos últimos anos fez cair por terra muitas das limitações impostas pela extensão da área mapeada. Amplamente customizáveis, os serviços de computação de alto desempenho na nuvem são capazes de atender a praticamente qualquer necessidade que por ventura possa surgir em um trabalho de MDS. Eles permitem utilizar volumes imensos de dados para realizar MDS de alta resolução – por exemplo, 5 m de resolução espacial – para qualquer extensão espacial, desde o nível de campo experimental até todo o planeta, inclusive considerando as dimensões vertical (profundidade) e temporal.

Dentre os serviços de computação de alto desempenho na nuvem, o mais popular deles é o [Google Earth Engine](#) (GEE) (Gorelick et al., 2017), sendo utilizado por iniciativas como o [MapBiomass](#), projeto multi-institucional de mapeamento anual da cobertura e uso do solo no Brasil que cobre o período de 1985 a 2017. Além de gratuito, o GEE dispõe de uma longa lista de métodos de mineração de dados e um imenso catálogo de dados do sensoriamento remoto, incluindo os programas Landsat, Sentinel, MODIS, ASTER, SRTM, WorldClim, WorldPop, entre outras. A Figura 2 mostra a interface do GEE, que pode ser acessada utilizando qualquer navegador da Internet. A interface é composta por quatro janelas, sendo a inferior aquela onde são mostrados os resultados dos processamentos. Na parte superior, ao centro, localiza-se a aba para entrada dos comandos para processamento dos dados, utilizando a linguagem de programação [JavaScript](#). Para a maioria dos cientistas do solo interessados em utilizar o GEE, a necessidade de aprender a linguagem JavaScript será o maior desafio a ser enfrentado. Entretanto, trata-se de desafio que não pode ser encarado como impeditivo. A atividade científica depende cada vez mais da computação e o cientista do solo moderno, assim como os

demais cientistas, precisa conhecer, pelo menos, uma linguagem de programação. Do contrário, será incapaz de acompanhar a evolução da ciência.



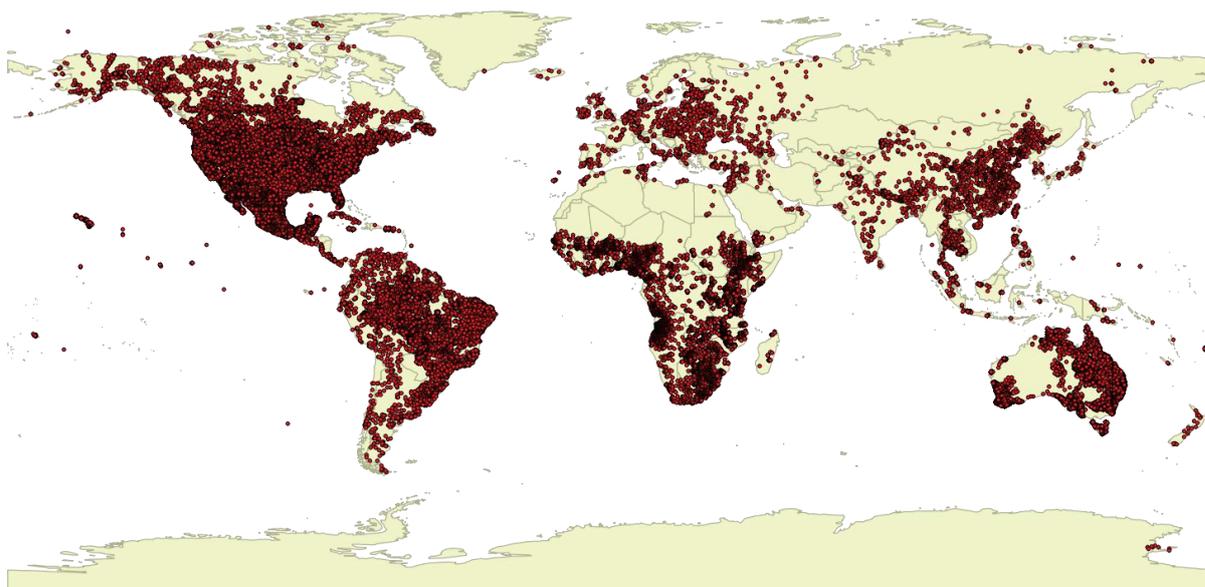
**Figura 2.** Visão do ambiente de desenvolvimento interativo do Google Earth Engine.

## DADOS DE OBSERVAÇÕES DO SOLO

Os resultados do uso dos serviços de computação de alto desempenho no MDS aplicado em grandes extensões já são conhecidos há alguns anos. O principal desses resultados é o sistema de mapeamento global [SoilGrids](#), desenvolvido e mantido pelo [ISRIC World Soil Information](#). O SoilGrids já produziu versões de mapas de propriedades e classes de solo, com cobertura global e em várias profundidades, com resoluções espaciais de 1 km (Hengl et al., 2014) e 250 m (Hengl et al., 2017). Em breve, versões com resolução espacial de 90-100 m deverão ser disponibilizadas – preparadas a partir do trabalho conjunto com os projetos globais [Global Soil Partnership](#) (GSP) e [GlobalSoilMap](#).

Garantida a capacidade computacional de alto desempenho e a disponibilidade de métodos numéricos de predição espacial, o maior desafio do MDS de alta resolução aplicado a grandes extensões territoriais é conseguir um maior volume de dados de observações do solo (Mulder et al., 2016; Somarathna et al., 2017). Quando maior o volume de dados, maior é a capacidade dos métodos numéricos capturarem estruturas de mais curto alcance de variação espacial e temporal do solo, ou seja, produzir informação mais detalhada. A Figura 3 mostra a distribuição espacial das cerca de 150 000 observações do solo, abertas para uso, disponíveis no [World Soil Information Service](#) (WoSIS), sistema que alimenta o SoilGrids. Além do pequeno número, uma característica marcante da Figura 3 é a heterogeneidade da distribuição

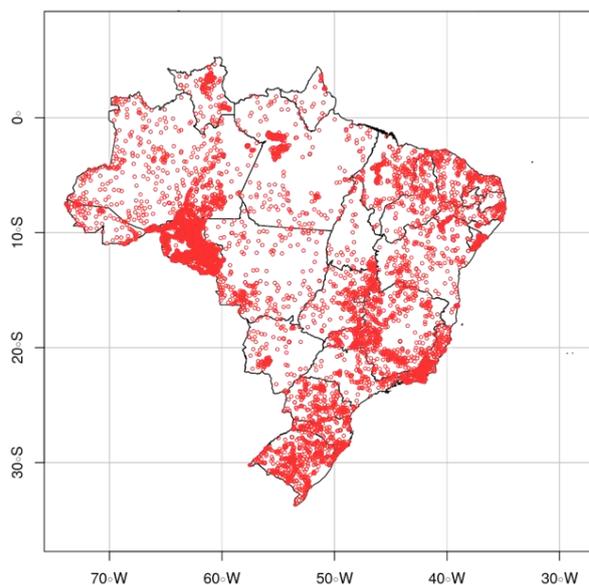
espacial das observações do solo: poucos países, entre eles o Brasil, concentram a maior parte delas. Isso significa que, além do baixo nível de detalhe, há grande chance de que os mapas resultantes sejam incapazes de representar a variação espacial do solo daqueles países em que este é menos investigado ou cujos dados sejam menos comumente compartilhados de maneira aberta como, por exemplo, a Mongólia, o Paraguai, a Arábia Saudita e a Suécia.



**Figura 3.** Localização das observações do solo utilizadas para produção dos mapas de propriedades e classes de solo pelo sistema SoilGrids.

A Figura 3 fornece uma visão relativamente otimista da disponibilidade de dados de observações do solo abertas para uso no Brasil. Um olhar mais aproximado revela que o cenário não é assim tão positivo. A Figura 4 mostra as cerca de 15 mil observações do solo disponíveis no [Repositório Brasileiro Livre para Dados Abertos do Solo](#) (febr). Trata-se da maior coleção pública de dados abertos do solo do Brasil, prontamente disponíveis para trabalhos de MDS. Assim como na escala global, a distribuição espacial das observações do solo é bastante concentrada em alguns estados brasileiros – como Rondônia, Rio Grande do Sul e Rio de Janeiro. Na iminência do início do Programa Nacional de Levantamento e Interpretação de Solos do Brasil ([PronaSolos](#)), a maior preocupação deve ser a obtenção de maior volume de dados de observações do solo que fique aberto para uso. Contudo, o processo de coleta de novos dados no campo de maneira a cobrir todo o território nacional em alta densidade deverá demorar, pelo menos, algumas décadas. A solução para o momento seria resgatar e compartilhar, abertamente, os dados de trabalhos preexistentes de mapeamento do solo, de pequenas a grandes extensões, e iniciar a produção de versões preliminares dos mapas de

propriedades e classes de solo. Isso possibilitaria indicar as áreas prioritárias para as futuras campanhas de coleta de dados no campo e paulatinamente aumentar o nível de detalhe das informações do solo em escala nacional.

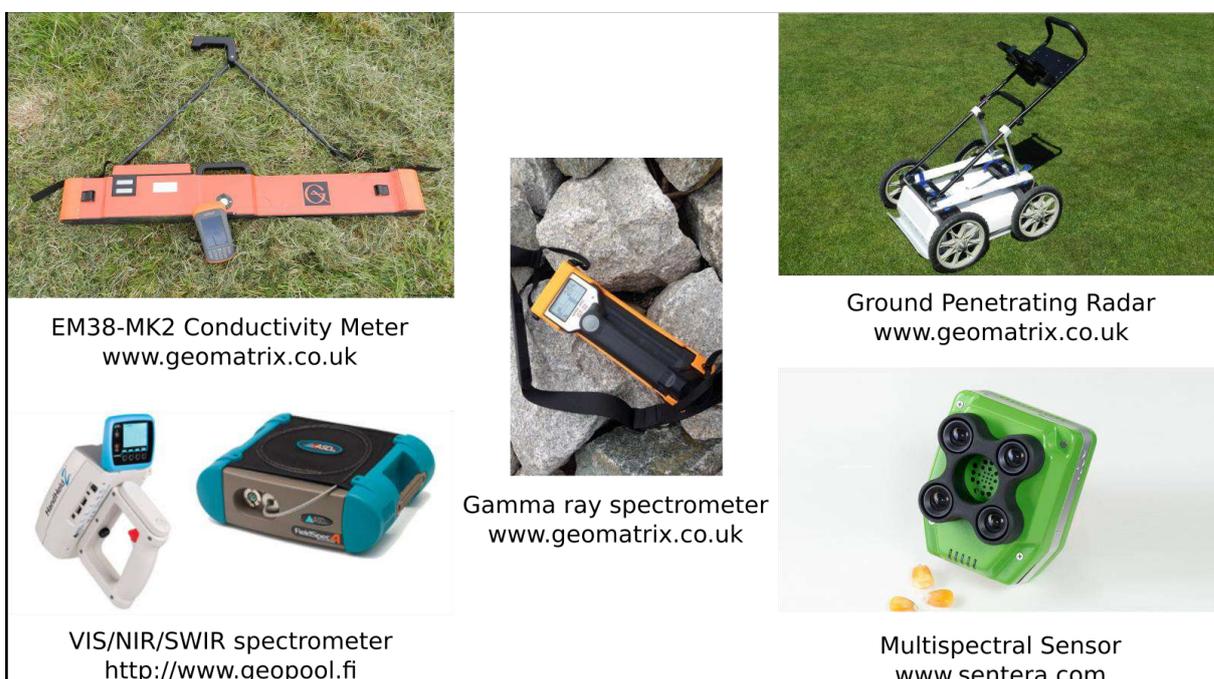


**Figura 4.** Localização das observações do solo abertas para uso disponíveis no Repositório Brasileiro Livre para Dados Abertos do Solo.

### **SENSORIAMENTO PROXIMAL E INFORMAÇÕES ESPACIAIS AUXILIARES**

A disponibilidade de dados do solo não costuma ser um problema no mapeamento de campos experimentais, sítios arqueológicos e lavouras sob agricultura de precisão. Isso ocorre porque a pequena extensão territorial da área mapeada exige menor volume de dados, mesmo que em alta densidade amostral. Na agricultura de precisão, por exemplo, ao invés da coleta de uma única amostra composta do solo para representar a fertilidade de todo um talhão, são coletadas dezenas a centenas de amostras individuais. O lado negativo dessa prática é o aumento do custo para mapeamento do solo: quanto maior o nível de detalhe espacial e/ou temporal desejado, maior costuma ser o número de amostras necessárias. Assim, novas estratégias de obtenção de dados vêm ganhando espaço. Cada vez mais se tem pesquisado o uso de sensores no campo, os quais utilizam-se de princípios físicos para obtenção de informações indiretas sobre as propriedades do solo (Figura 5). Essas informações podem então ser utilizadas no auxílio à predição das propriedades do solo por métodos numéricos de correlação empírica (Heil e Schmidhalter, 2017). Diversas são as tecnologias para essa finalidade, indo de imagens diárias coletadas por satélites de alta resolução (sensoriamento remoto) até sensores e câmeras

embarcadas em máquinas agrícolas e drones (sensoriamento proximal) (Heege, 2013).



**Figura 5.** Sensores proximais utilizados no mapeamento digital do solo de alta resolução.

Dois equipamentos sensores têm se destacado no auxílio ao mapeamento de propriedades do solo: os sensores de condutividade elétrica aparente do solo (CEa) e os espectroradiômetros. O primeiro permite que, de forma simples e rápida, obtenha-se dados de solo com alta densidade espacial (Corwin e Lesch, 2005). Diversas são as propriedades do solo que influenciam as leituras deste tipo de equipamento, sendo que uma das mais importantes é a umidade do solo, a qual geralmente está intimamente relacionada à granulometria do solo. Isso permite que a variação espacial da granulometria do solo seja predita usando dados de CEa. Já o segundo equipamento permite que diversas propriedades do solo, especialmente a granulometria e o conteúdo de matéria orgânica – e propriedades correlatas –, sejam preditas simultaneamente com base no comportamento espectral do solo (Soriano-Disla et al., 2013). A principal limitação do uso da técnica de espectroscopia do solo é que muitas das propriedades de interesse agrícola, como a disponibilidade de nutrientes e o pH, não têm influência direta no comportamento espectral, o que torna a criação de bancos de dados e modelos preditivos um processo altamente complexo e desafiador. Muitas iniciativas para resolver essa limitação existem ao redor do mundo, mas ainda nenhuma alcançou a eficiência preditiva necessária para a agricultura de precisão.

Além das técnicas de coleta de dados por sensores remotos e proximais, outros tipos de

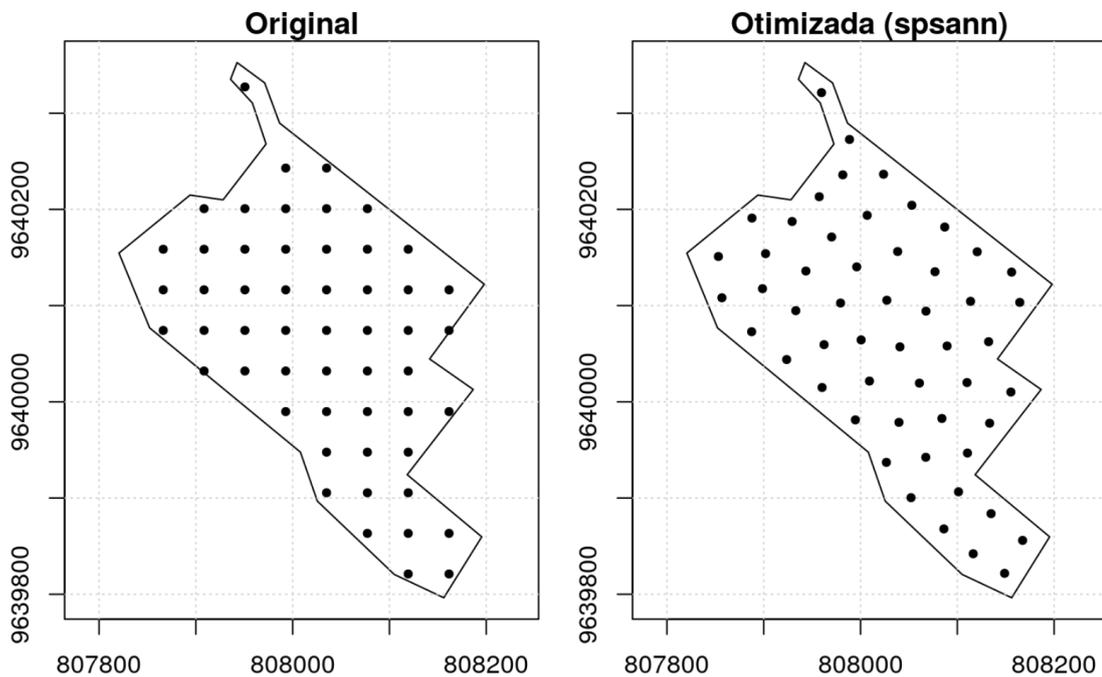
informações espaciais auxiliares também podem auxiliar na predição das propriedades do solo em pequenas áreas – assim como já é feito no MDS aplicado a grandes extensões (Samuel-Rosa et al., 2015a). Novamente, o principal interesse no uso de tais variáveis espaciais auxiliares é a redução da dependência de amostras do solo coletadas diretamente no campo. No longo prazo, uma menor dependência de observações diretas do solo deve reduzir os custos de produção de informação do solo na agricultura de precisão.

Um tipo de informação auxiliar com grande potencial são os atributos de terreno, como declividade, índice de umidade topográfica e curvatura, derivados de modelos digitais de elevação – como aquele do projeto [TOPODATA](#), com 30 m de resolução espacial (Valeriano e Rossetti, 2012). Isso porque o relevo costuma apresentar marcada influência na formação do solo e, assim nas suas propriedades, via redistribuição das águas e sedimentos.

Outro exemplo de variável espacial auxiliar em áreas de produção agrícola, pecuária ou florestal, é o histórico das práticas de manejo e conservação do solo. Por exemplo, numa área subdividida em talhões delimitados por terraços, em geral, é de se esperar que as propriedades do solo variem de talhão para talhão; alguma variação também é esperada entre os piquetes de uma área de produção de gado de corte. Já numa área de integração lavoura-pecuária-floresta, será a distância das linhas de cultivo da espécie florestal que produzirá variação nas propriedades químicas e na disponibilidade de água no solo. Curiosamente, apesar do seu grande potencial e facilidade de obtenção, a utilização desse tipo de informação espacial auxiliar ainda é pouco usada no MDS na agricultura de precisão.

## **MODELOS GEOESTATÍSTICOS E OTIMIZAÇÃO DE CONFIGURAÇÕES AMOSTRAIS**

Apesar do grande potencial de uso das informações espaciais auxiliares, o MDS sempre dependerá dos dados de observações do solo obtidas no campo. No Brasil, o MDS aplicado a pequenas extensões costuma utilizar, principalmente, configurações amostrais regulares retangulares para obter os dados do solo no campo (Figura 6, painel esquerdo). Contudo, já se sabe há bastante tempo que, quando o interesse é a interpolação espacial (krigagem), as configurações amostrais regulares retangulares (também chamadas de grades ou malhas amostrais) somente são eficientes se a área for perfeitamente retangular (van Groenigen, 1999) – na verdade, as melhores configurações amostrais para esse caso são as configurações regulares triangular ou hexagonal (Yfantis et al., 1987). A popularidade da configuração amostral regular retangular está associada ao ainda comum limitado planejamento amostral em trabalhos de mapeamento do solo.

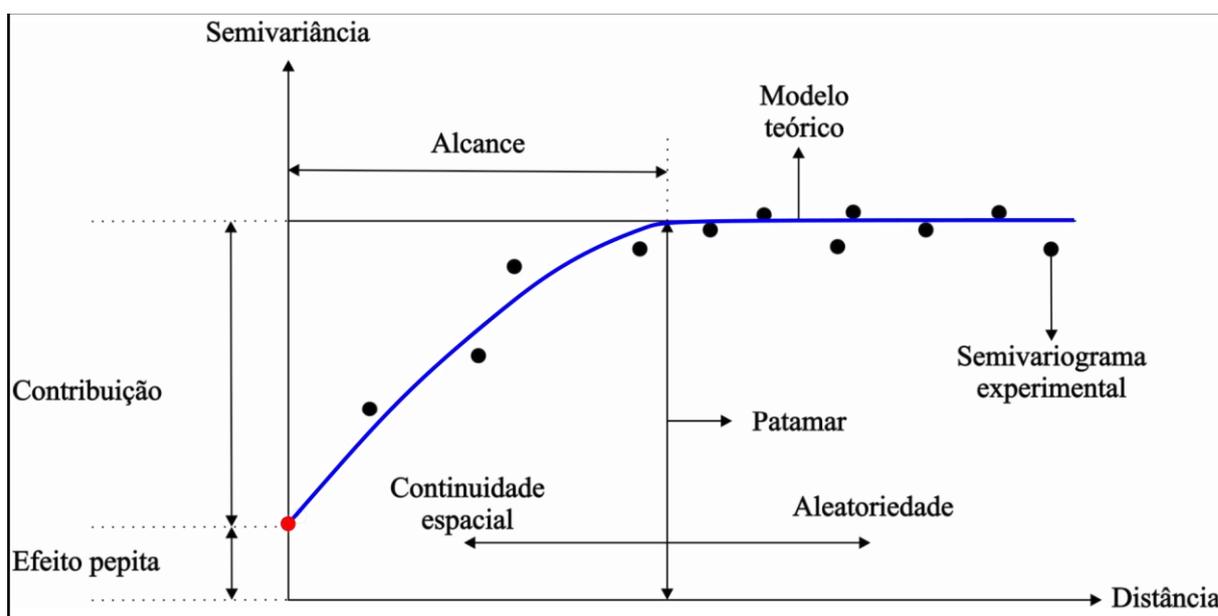


**Figura 6.** Comparação entre uma configuração amostral regular retangular (esquerda) e uma configuração amostral otimizada usando o pacote *spsann* em ambiente R (Samuel-Rosa et al., 2015b) para maximizar a cobertura espacial (direita).

Quando o formato da área em questão é irregular, apresentando áreas alongadas e estreitas, ou mesmo regiões não mapeáveis em seu interior, uma configuração amostral regular retangular geralmente apresenta amplos espaços não amostrados. Além disso, muitas amostras acabam sendo posicionadas muito próximas das bordas – ou mesmo sobre as bordas. Isso prejudica a capacidade dos dados capturarem a estrutura da variação espacial do solo, resultando em mapas de menor qualidade, o que significa uma má aplicação dos recursos disponíveis. Para obter uma distribuição espacial ótima dos pontos amostrais, em função do método numérico a ser utilizado, deve-se lançar mão de algoritmos inteligentes (Minasny e McBratney, 2006; Brus e Heuvelink, 2007; Walvoort et al., 2010), já comumente usados no MDS aplicado a grandes extensões. Com esses algoritmos, uma configuração amostral otimizada para a interpolação espacial vai adquirir um padrão praticamente irregular, mas mantendo distâncias aproximadamente regulares entre os pontos vizinhos, de maneira a reduzir ao máximo o tamanho dos espaços vazios (subamostrados) (Figura 6, painel direito).

Outro elemento raramente encontrado em trabalhos de MDS no Brasil que utilizam métodos de interpolação espacial como a krigagem é a quantificação explícita do chamado

efeito pepita do modelo do variograma (Figura 7), a porção da variância dos dados devida a erros aleatórios. Para determinar tais erros basta selecionar, aleatoriamente, algumas observações e determinar as propriedades do solo com, pelo menos, uma repetição laboratorial. Feito isso, a média da variância dessas repetições pode ser utilizada como estimativa do efeito pepita do modelo do variograma (Kuensch et al., 2011). Na prática, o efeito pepita costuma ser estimado de maneira indireta, extrapolando o modelo do variograma até a origem. O problema dessa estratégia é que ela pode sub ou superestimar o efeito pepita. A subestimativa poderia mostrar dependência espacial quando não há nenhuma; já a superestimativa poderia incorretamente indicar ausência de dependência espacial nos dados. Em ambos os casos, o prejuízo aparece nos mapas gerados, que invariavelmente fornecerão indicações imprecisas para recomendação de fertilizantes.



**Figura 7.** Representação esquemática do variograma experimental (pontos pretos) e respectivo modelo teórico ajustado (linha azul). O ponto vermelho representa o efeito pepita, estimado usando repetições de laboratório (distância de separação igual a zero).

### MODELOS PREDITIVOS “MECANÍSTICOS”

Conforme visto acima, a agricultura de precisão é um sistema que utiliza e produz um volume muito grande de informações espaciais e temporais, em alta resolução, como imagens orbitais e mapas de produtividade. Com esse conhecimento ultradetalhado, deveria ser possível aventurar-se mais frequentemente no uso de modelos preditivos de cunho mais mecanístico. Em vez de depender apenas das correlações empíricas entre propriedades do solo e informações

espaciais auxiliares, esses modelos mecânicos seriam construídos a partir do conhecimento das práticas de manejo do solo, executadas ao longo das safras, utilizando relações algébricas e físicas baseadas, por exemplo, nos princípios da conservação da matéria e da energia. Considere, por exemplo, o seguinte modelo mecânico de explicação do conteúdo de potássio ( $K^+$ ) no solo:

$$K^+ = K_{i=0}^+ + K_{i=1}^+ + K_{i=2}^+ + K_{i=3}^+ - K_{i=4}^+ - K_{i=5}^+ - K_{i=6}^+$$

Segundo o modelo acima, o conteúdo atual de  $K^+$  no solo depende, em primeiro lugar, do conteúdo inicial de  $K^+$  ( $i = 0$ ) no solo, por exemplo, antes da semeadura da última cultura agrícola. Essa informação costuma estar disponível na forma de um mapa do conteúdo de  $K^+$ . O conteúdo atual de  $K^+$  também depende da adubação potássica daquela cultura ( $i = 1$ ) – geralmente disponível na forma de um mapa de aplicação de  $K^+$  – e dos ganhos de  $K^+$  por mineralização da matéria orgânica ( $i = 2$ ) – que pode ser estimada a partir de um mapa de cobertura do solo com palha – e deposição de sedimentos vindos das partes mais altas do terreno ( $i = 3$ ) – calculada em função do modelo digital de elevação. O conteúdo atual de  $K^+$  também depende das perdas, por exemplo, por remoção de sedimentos para as partes mais baixas do terreno ( $i = 4$ ) ou por lixiviação ( $i = 5$ ) e pela exportação da cultura anterior ( $i = 6$ ) – disponível na forma de mapa de produtividade. Outros fatores podem interferir no conteúdo de  $K^+$  no solo, mas é provável que a solução de uma equação com apenas essas informações já produzia uma estimativa, pelo menos, bastante aproximada do conteúdo de  $K^+$  no solo. É provável que novas observações do solo sejam necessárias para corrigir erros sistemáticos, mas certamente em número muito inferior ao requerido tradicionalmente em métodos empíricos de interpolação espacial com a krigagem.

## CONSIDERAÇÕES FINAIS

O está se tornando a técnica oficial de produção de informação espacial do solo em todo o mundo. Sua aplicação em grandes extensões, inclusive em alta resolução, se tornou uma realidade a partir do momento em que os cientistas do solo começaram a ter maior acesso à computação de alto desempenho na nuvem. Com uma grande disponibilidade de informações espaciais auxiliares, produzidas por um número cada vez maior de sistemas sensores orbitais, e uma ampla gama de avançados métodos numéricos de mineração de dados e aprendizado de máquina, os cientistas do solo enfrentam dois grandes desafios. O primeiro é aprender uma linguagem de programação para que possam desfrutar dos avanços computacionais. O segundo é obter mais dados do solo, em quantidade e qualidade, no espaço e no tempo, para que seja possível produzir informação do solo em escala global e com nível de detalhe local. A sugestão

é que compartilhem mais os dados existentes para construir e compartilhar uma grande base de dados do solo.

Na agricultura de precisão, o uso de métodos numéricos para a produção de informação do solo já é uma realidade há muito mais tempo. Da mesma maneira que o uso de técnicas de sensoriamento remoto de alta resolução e proximal para obter dados indiretos do solo. Hoje o maior desafio é utilizar novas estratégias para reduzir a dependência dos dados do solo coletados no campo. Dentre elas estão a otimização das configurações amostrais usando algoritmos inteligentes e o uso de outras informações espaciais auxiliares, como aquelas relacionadas ao terreno e ao uso e manejo das áreas de produção. Nas áreas onde a agricultura de precisão é praticada há mais tempo, o desafio é transformar o grande volume de informações espaçotemporais disponíveis para construir modelos mecanísticos. Modelos baseados no mecanismo de funcionamento da unidade de produção, que permitam fazer previsões espaciais e temporais com menor demanda de dados e, portanto, menor custo para os agricultores.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BREIMAN, L. Bagging predictors. *Mach Learn*, 24:123-140, 1996.
- BRUS, D.J.; HEUVELINK, G.B.M. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, 138:86-95, 2007.
- TEN CATEN, A.; DALMOLIN, R.S.D.; MENDONÇA-SANTOS, M. D.L.; GIASSON, E. Digital soil mapping: characteristics of the Brazilian approach. *Cienc Rural*, 42:1989-1997, 2012.
- FIGUEIREDO, S.R.; GIASSON, E.; TORNQUIST, C.G.; NASCIMENTO, P.C. D. Multiple logistic regression applied to soil survey in Rio Grande do Sul state, Brazil. *Rev. Bras. Cienc. Solo*, 32:2779-2785, 2008.
- GIASSON, E.; CLARKE, R.T.; INDA JUNIOR, A.V.; MERTEN, G.H.; TORNQUIST, C.G. Digital soil mapping using multiple logistic regression on terrain parameters in Southern Brazil. *Sci Agric*, 63:262-268, 2006.
- GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens Environ*, 202:18-27, 2017.
- VAN GROENIGEN, J.-W. Constrained optimisation of spatial sampling: a geostatistical approach. Wageningen, Wageningen University, 1999. 148p. (Doctorate Thesis)
- HEIL, K.; SCHMIDHALTER, U. The application of EM38: determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. *P Soc Photo-opt Ins*, 17:2540, 2017.
- HENGL, T.; JESUS, J.M.; HEUVELINK, G.B.M.; GONZALEZ, M.R.; KILIBARDA, M.; BLAGOTIĆ, A.; SHANGGUAN, W.; WRIGHT, M.N.; GENG, X.; BAUER-MARSCHALLINGER, B.; GUEVARA, M.A.; VARGAS, R.; MACMILLAN, R.A.; BATJES, N.H.; LEENAARS, J.G.B.; RIBEIRO, E.; WHEELER, I.; MANTEL, S.; KEMPEN, B. SoilGrids250m: Global Gridded Soil Information Based on Machine Learning. *PLoS One*, 12:e0169748, 2017.
- HENGL, T.; JESUS, J.M. D.; MACMILLAN, R.A.; BATJES, N.H.; HEUVELINK, G.B.M.; RIBEIRO, E.; SAMUEL-ROSA, A.; KEMPEN, B.; LEENAARS, J.G.B.; WALSH, M.G.; GONZALEZ, M.R. SoilGrids1km—global soil information based on automated mapping. *PLoS One*, 9:e105992,

- 2014.
- HEUVELINK, G.B.M.; BIERKENS, M.F.P. Combining soil maps with interpolations from point observations to predict quantitative soil properties. *Geoderma*, 55:1-15, 1992.
- KUENSCH, H.R.; PAPRITZ, A.; SCHWIERZ, C.; STAHEL, W.A. Robust estimation of the external drift and the variogram of spatial data. In: PROCEEDINGS OF THE ISI 58TH WORLD STATISTICS CONGRESS OF THE INTERNATIONAL STATISTICAL INSTITUTE, 2011. Anais. 2011. p.1-8.
- LAGACHERIE, P.; MCBRATNEY, A.B. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: P. LAGACHERIE, A.M.; VOLTZ, M. (eds.). *Digital soil mapping – an introductory perspective*. Elsevier, 2007. 3 - 22p.
- MA, Y.; MINASNY, B.; MALONE, B.P.; MCBRATNEY, A.B. Pedology and digital soil mapping (DSM). *Eur J Soil Sci*, 70:216-235, 2019.
- MCBRATNEY, A.B.; MENDONÇA-SANTOS, M.L.; MINASNY, B. On digital soil mapping. *Geoderma*, 117:3-52, 2003.
- MINASNY, B.; MCBRATNEY, A.B. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput Geosci*, 32:1378-1388, 2006.
- MOLIN, J.P.; DO AMARAL, L.R.; COLAÇO, A.F. *Agricultura de precisão. Oficina de Textos*, 2015. 224p.
- MOORE, I.D.; GESSLER, P.E.; NIELSEN, G.A.; PETERSON, G.A. Soil attribute prediction using terrain analysis. *Soil Sci. Soc. Am. J.*, 57:443-452, 1993.
- MULDER, V.L.; LACOSTE, M.; DE-FORGES, A.C.R.; MARTIN, M.P.; ARROUAYS, D. National versus global modelling the 3D distribution of soil organic carbon in mainland France. *Geoderma*, 263:16-34, 2016.
- ODEH, I.O.A.; MCBRATNEY, A.B.; CHITTLEBOROUGH, D.J. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma*, 67:215-226, 1995.
- POLIDORO, J.C.; MENDONÇA-SANTOS, M.L.; LUMBRERAS, J.F.; COELHO, M.R.; CARVALHO FILHO, A.; MOTTA, P.E.F.; CARVALHO JUNIOR, W.; ARAÚJO FILHO, J.C.; CURCIO, G.R.; CORREIA, J.R.; MARTINS, É.S.; SPERA, S.T.; OLIVEIRA, S.R.M.; BOLFE, E.L.; MANZATTO, C.V.; TÔSTO, S.G.; VENTURIERI, A.; SÁ, I.B.; OLIVEIRA, V.Á.; SHINZATO, E.; ANJOS, L.H.C.; VALLADARES, G.S.; RIBEIRO, J.L.; MEDEIROS, P.S.C.; MOREIRA, F.M.S.; SILVA, L.S.L.; SEQUINATTO, L.; AGLIO, M.L.D.; DART, R.O. *Programa Nacional de Solos do Brasil (PronaSolos)*. Rio de Janeiro, 2016. 53p.
- SAMUEL-ROSA, A.; HEUVELINK, G.B.M.; VASQUES, G.M.; ANJOS, L.H.C. Do more detailed environmental covariates deliver more accurate soil maps? *Geoderma*, 243–244:214-227, 2015a.
- SAMUEL-ROSA, A.; HEUVELINK, G.B.M.; VASQUES, G.M.; ANJOS, L.H.C. spsann – optimization of sample patterns using spatial simulated annealing. In: *GEOPHYSICAL RESEARCH ABSTRACTS – EGU GENERAL ASSEMBLY 2015*, 2015b. Anais. Copernicus, 2015b. p. EGU2015-7780.
- SCHUMACHER, M.; ROSSNER, R.; VACH, W. Neural networks and logistic regression: Part I. *Comput. Stat. Data Anal.*, 21:661-682, 1996.
- SOMARATHNA, P.D.S.N.; MINASNY, B.; MALONE, B.P. More data or a better model? Figuring out what matters most for the spatial prediction of soil carbon. *Soil Sci. Soc. Am. J.*, 81:1413, 2017.
- SORIANO-DISLA, J.M.; JANIK, L.J.; ROSSEL, R.A.V.; MACDONALD, L.M.; MCLAUGHLIN, M.J. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Appl Spectrosc Rev*, 49:139-186, 2013.
- VALERIANO, M.M.; ROSSETTI, D.F. Topodata: Brazilian full coverage refinement of SRTM data. *Appl Geogr*, 32:300-309, 2012.
- WALVOORT, D.J.J.; BRUS, D.J.; DE GRUIJTER, J.J. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. *Comput Geosci*, 36:1261-1267, 2010.
- WOLPERT, D.H. Stacked generalization. *Neural Networks*, 5:241-259, 1992.

YFANTIS, E.A.; FLATMAN, G.T.; BEHAR, J.V. Efficiency of kriging estimation for square, triangular, and hexagonal grids. *Math Geol*, 19:183-205, 1987.