

## Predição de classes de solos com dados coletados em pixels delimitados por *buffers* em perfis de solo georreferenciados

Alcinei Ribeiro Campos<sup>1</sup>, Elvio Giasson<sup>1</sup>, José Janderson Ferreira Costa<sup>1</sup>, Fabrício Fernandes Coelho<sup>1</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil. E-mail: alcineicampos@gmail.com (ORCID: 0000-0003-0070-2101); giasson@ufrgs.br (ORCID: 0000-0003-3659-6873); janderson.ferreirac@gmail.com (ORCID: 0000-0002-3891-8111); fabricio.coelho@ufrgs.br (ORCID: 0000-0001-7589-6014)

**RESUMO:** Os estudos de mapeamento digital de classes de solos (MDS) têm utilizado mapas legados como principal fonte de informação para calibração dos modelos preditores. Entretanto, são necessárias novas abordagens com técnicas que permitam o uso de informações contidas em perfis de solos georreferenciados, permitindo a aplicação do MDS em áreas amostradas que não disponibilizem de mapas convencionais de solos. O objetivo deste estudo foi avaliar o desempenho na predição de ocorrência de solos de amostras coletadas em pixels de perfis de solos georreferenciados e em pixels coletados em *buffers* com raio de 50, 100, 150, 200 e 250 m dos perfis de solos nas bacias dos rios Lajeado Grande e Santo Cristo. As duas áreas possuem dados de levantamento de solos publicados em escala 1:50.000. Para predição da ocorrência das classes de solos foram utilizadas dez variáveis preditoras geradas a partir de um modelo digital de elevação com resolução espacial de 30 m. Para a predição foi utilizado o algoritmo Random Forest. Os mapas preditos foram avaliados quanto à exatidão em relação aos perfis de solos e quanto a reprodutibilidade dos mapas convencionais. A utilização dos pixels amostrais coletados nos *buffers* não alterou de forma expressiva a acurácia geral dos mapas preditos na bacia do rio Lajeado Grande, mas permitiu um ganho de 15,6% de exatidão na bacia do rio Santo Cristo. O uso apenas dos perfis georreferenciados resultou em mapas preditos com exatidão superior a 75% e concordância de reprodutibilidade igual ou superior a 67% em relação ao mapa convencional.

**Palavras-chave:** mapeamento digital de solos; Random Forest; técnicas pedométricas

## Prediction of soil classes with data collected in buffers delimited pixels in georeferenced soil profiles

**ABSTRACT:** Digital soil mapping studies (MDS) have used legacy maps as the main source of information for calibration of predictor models. However new approaches are needed with techniques that allow the use of information contained in georeferenced soil profiles, allowing the application of MDS in sampled areas that do not provide conventional soil maps. The objective of this study was to evaluate the performance in the prediction of soil occurrence of samples collected in pixels of profiles of georeferenced soils and in pixels collected in buffers with radius of 50, 100, 150, 200 and 250 m of soil profiles in the Lajeado Grande and Santo Cristo Rivers Watersheds. Two areas with availability of soil survey data published at scale 1: 50,000 were used for the study. For prediction of the occurrence of soil classes, ten predictive variables were generated from a digital elevation model with spatial resolution of 30 m. For the prediction, Random Forest algorithm was used. The predicted maps were evaluated for accuracy in relation to soil profiles and the reproducibility of conventional maps. The use of the sampling pixels collected in the buffers did not significantly alter the overall accuracy of the predicted maps in the Lajeado Grande river watershed, but allowed a 15.6% the gain at overall accuracy the Santo Cristo river watershed. The use of georeferenced profiles resulted in predicted maps with overall accuracy greater than 75% and agreement of reproducibility equal to or high than 67% in relation to the conventional map.

**Key words:** digital mapping of soils; Random Forest; pedometer techniques

## Introdução

No mapeamento digital de classes de solos, diversos estudos têm testado as mais variadas abordagens buscando, principalmente, a reprodução de mapas legados objetivando calibrar modelos matemáticos capazes de reproduzir o modelo criado pelo pedólogo na delimitação das classes de solos (Giasson et al., 2006; Behrens et al., 2010; Arruda et al., 2013; Dias et al., 2016). No entanto, essa abordagem exige a existência de mapas legados para calibração dos modelos de predição, sendo inviável para regiões onde não há disponibilidade dessas informações. Uma alternativa para a ausência de mapas legados é a predição da ocorrência dos solos a partir de informações coletadas em perfis de solos georreferenciados (Hengl et al., 2007; Alves et al., 2015; Arruda et al., 2016). Dessa forma, pode-se realizar o mapeamento de áreas que já foram amostradas, além da confecção de bancos de dados estruturados e a extrapolação para áreas fisiograficamente semelhantes.

Para uma boa calibração dos modelos preditores é necessária uma densidade de amostras superior à utilizada pelo método convencional de levantamento de solos (ten Caten et al., 2013; Bagatini et al., 2015), o que restringe a aplicação do MDS (mapeamento digital de solos) com essa abordagem em áreas com baixa disponibilidade de perfis georreferenciados. Nesse contexto, o baixo número de amostras pode ser contornado com a utilização de estratégias de amostragem mais representativas, permitindo a coleta de pontos amostrais suficientes para o treinamento dos modelos preditivos. Dentre as alternativas de amostragem pode ser utilizado a delimitação de *buffers* em torno dos perfis de solos georreferenciados, permitindo assim obter um maior número de amostras nas áreas vizinhas aos perfis (Chagas et al., 2010; Dias et al., 2016).

O uso do *buffer* permite a extrapolação da classificação do perfil de solo para as áreas adjacentes a este, permitindo a coleta de um maior número de pontos amostrais, o que pode levar a uma maior eficiência na calibração dos modelos preditores aplicados para predição de ocorrência das classes de solos. O objetivo deste estudo foi avaliar o desempenho na predição de ocorrência de solos de amostras coletadas em pixels de perfis de solos georreferenciados e em pixels coletados em *buffers* com raio de 50, 100, 150, 200 e 250 metros dos perfis de solos nas bacias dos rios Lajeado Grande e Santo Cristo, localizadas na região noroeste do estado do Rio Grande do Sul, Brasil.

## Material e Métodos

Foram utilizados dados de solos, hidrográficos e do modelo digital de elevação (MDE) das Bacias dos rios Santo Cristo (SC) e Lajeado Grande (LG), localizadas na região noroeste do estado do Rio Grande do Sul, com a bacia SC limitada pelas coordenadas 54°44' e 54°21'W e 27°34' e 28°1'S e a LG pelas coordenadas 54°13' e 53°51'W e 27°24' e 27°40'S (Figura 1). A geologia das duas áreas é caracterizada por

derrames basálticos da formação Serra Geral (Freitas et al., 2012). As áreas das duas bacias são 900,7 km<sup>2</sup> e 533,3 km<sup>2</sup>, respectivamente nas bacias SC e LG, e solos estão mapeados na escala 1:50.000 (Kämpf et al., 2004a, 2004b) conforme unidades de mapeamento apresentadas na Tabela 1.

O MDE utilizado foi produzido a partir de dados do sensor *Aster/GDEM v2 (Global Digital Elevation Models)*, com resolução espacial de 30 metros, obtidos no Serviço Geológico Americano (Tachikawa et al., 2011), a partir do qual foram derivadas oito variáveis preditoras (nível da base da rede de drenagem, orientação de vertentes, insolação difusa, abertura positiva do terreno, convexidade, declividade, índice de convergência e posição média da declividade). Além destas, foram derivadas as variáveis índice de densidade de drenagem (Otto et al., 2017) e distância euclidiana dos rios a partir da rede hidrográfica na escala 1:50.000 (Hasenack & Weber, 2010), totalizando dez variáveis utilizadas para predição da ocorrência das UM. Para derivar as variáveis foi utilizado o software SAGA GIS v.2.2.2 integrado ao *software R* (R Core Team, 2018).

Para coleta dos dados amostrais foram utilizados pontos de observação georreferenciados (perfis de solos e amostras extras) e pixels delimitados em torno de cada perfil de solos (*buffers*). Foram testados cinco *buffers* com raios de 50, 100, 150, 200, 250 m, em torno dos perfis georreferenciados (Figura 1C), além dos pontos sem *buffer* (BF00), totalizando seis bases de pontos amostrais (BF00, BF50, BF100, BF150, BF200 e BF250).

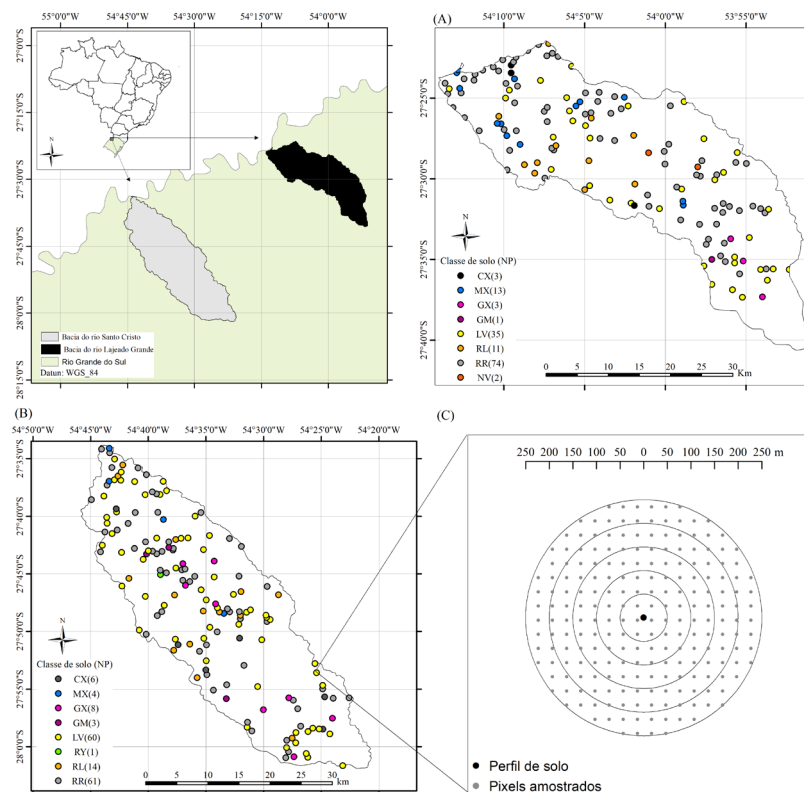
Foram utilizados perfis de solos classificados no nível de subordem (Santos et al., 2018), sendo 169 na bacia LG e 196 na bacia SC. Para coleta dos dados para treinamento foram eliminados pontos vizinhos com distância inferior a 500 m, distancia equivalente ao diâmetro do maior *buffer* testado, de forma a selecionar amostras para todas as classes descritas nos levantamentos de solos, restando 142 pontos na LG e 157 na SC, distribuídos em oito classes taxonômicas de solos (Figura 1). Para ambas as áreas de estudo as classes Latossolos Vermelhos e Neossolos Regolíticos juntos correspondem a mais de 75% dos perfis de solos descritos.

Na amostragem das áreas delimitadas pelos *buffers*, foram considerados todos os pixels dentro do raio de alcance dos respectivos *buffers* (Figura 1). Com esses dados foram preditos mapas de ocorrência de classes taxonômicas de solos (CT00, CT50, CT100, CT150, CT200 e CT250) a partir da classificação dos perfis de solos no nível de subordem (Santos et al., 2018). Para a predição dos mapas de solo foi utilizando o método *Random Forest* (RF) (Breiman, 2001). As RFs foram implementadas no pacote *RandomForest* (Breiman et al., 2018) no *software R* (R Core Team, 2018). Para utilização da RF foram utilizados os seguintes parâmetros: 200 árvores na floresta (*n*tree), sendo esse valor definido por meio de testes que variaram de 100 a 500, a quantidade mínima de dados por nó terminal (*nodesize*) o padrão cinco para cada nó terminal e o número de variáveis utilizadas por árvore (*m*try) a raiz quadrada do número total de variáveis preditoras (Breiman, 2001).

**Tabela 1.** Unidades de mapeamento de solos (UM) que ocorrem nas áreas das bacias dos rios Lajeado Grande e Santo Cristo localizados no noroeste do Rio Grande do Sul (Brasil).

UM	Composição taxonômica	Proporção (%)	Inclusões	Área (%)
<b>Bacia hidrográfica do rio Lajeado Grande (LG)</b>				
G	Gleissolos			1,61
LV1	Latossolo Vermelho distroférico		RR, CX	27,21
LV2	Latossolo Vermelho + Neossolo Regolítico *	60/40	CX	5,84
M1	Chernossolo		RR, RF	0,67
M2	Chernossolo + Neossolo Regolítico*	60/40	CX, RF	0,73
NV1	Nitossolo Vermelho			0,45
NV2	Nitossolo Vermelho + Neossolo Regolítico*	60/40		1,68
RL	Neossolo Litólico + Neossolo Regolítico*		AR	0,41
RR1	Neossolo Regolítico		AR, RL, CX, M, LV	30,35
RR2	Neossolo Regolítico + Neossolo Litólico, relevo forte ondulado*	60/40	AR, CX	23,91
RR3	Neossolo Regolítico + Latossolo Vermelho*	60/40	CX	5,26
RR4	Neossolo Regolítico + Chernossolo*	60/40	CX, RF	0,86
RR5	Neossolo Regolítico + Cambissolo + Nitossolo Vermelho*	50/30/20		0,52
RR6	Neossolo Regolítico + AR *	70/30		0,49
<b>Área total (ha)</b>				<b>53.388</b>
<b>Bacia hidrográfica do rio Santo Cristo (SC)</b>				
G1	Gleissolos			2,46
LV1	Latossolo Vermelho distroférico		RR, CX	38,13
LV2	Latossolo Vermelho + Neossolo Regolítico*	60/40	CX	7,92
M1	Chernossolo Háplico		CX	0,20
RL	Neossolo Litólico			0,01
RR1	Neossolo Regolítico + Cambissolo Háplico*	60/40	AR, RL, CX	34,79
RR2	Neossolo Regolítico + Neossolo Litólico, relevo forte ondulado**	50/50	AR, CX	8,24
RR3	Neossolo Regolítico + Latossolo Vermelho*	60/40	CX	7,84
RR4	Neossolo Regolítico + Neossolo Litólico*	70/30		0,03
RR5	Neossolos Regolíticos + Cambissolo Háplico + Latossolo Vermelho*	50/30/20		0,38
<b>Área total (ha)</b>				<b>90.073</b>

\*Associações; \*\*Complexos; AR = afloramento de rocha; CX = Cambissolo; LV = Latossolo Vermelho; M = Chernossolo; RY = Neossolo Flúvico; RL = Neossolo Litólico; RL = Neossolo Litólico; RR = Neossolo Regolítico



**Figura 1.** Mapa com o número (NP) e distribuição dos perfis de solos georreferenciados das bacias Lajeado Grande (A) e Santo Cristo (B), e o esquema dos buffers (C) utilizados para coleta dos pixels amostrais.

Nos mapas preditos foram removidos todos os agrupamentos de pixels com menos de cinco hectares. Para validação foram avaliadas as concordâncias dos mapas preditos com a exatidão em relação as classes descritas nos perfis de solos, sendo que para essa avaliação foram considerados todos os perfis disponíveis para as respectivas áreas de estudo (169 perfis na bacia LG e 196 perfis na bacia SC). Também foi avaliada a reprodutibilidade dos mapas convencionais das áreas de estudo, por meio da comparação dos mapas preditos com os respectivos mapas convencionais, para essa avaliação foi utilizado com referência a concordância com as classes taxonômicas (CT) que compõe as respectivas unidades de mapeamento de solo, como pode ser visto na Tabela 2. De acordo com as formas de avaliação descritas acima foram geradas matrizes de erro, das quais foram calculados os valores de acurácia geral (AG) e acurácia do mapeador (AM) (Congalton, 1991).

## Resultados e Discussão

O uso do *buffer* permitiu aumentar o número de pixels coletados por perfil de solo georreferenciado (amostras), no

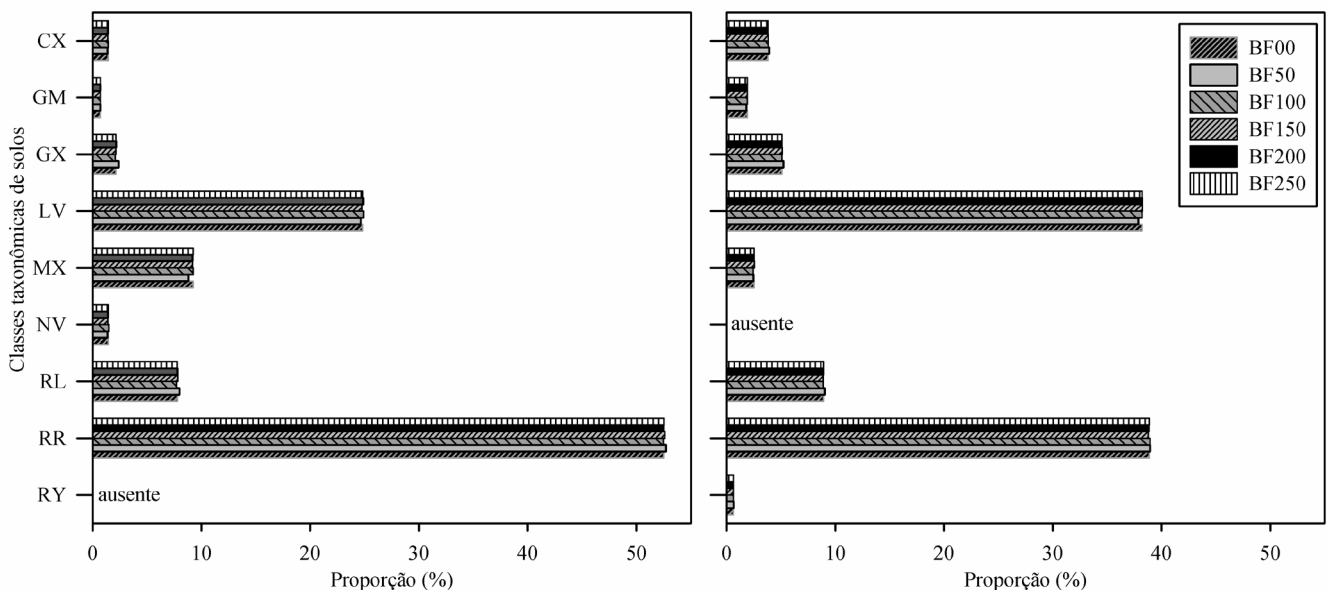
entanto, não alterou a proporcionalidade de amostras entre as classes taxonômicas de solos, com os Latossolos Vermelhos e Neossolos Regolíticos, que juntos correspondem mais de 75% das amostras coletadas em todos os tamanhos de *buffer* (Figura 2).

Na bacia do rio Lajeado Grande os maiores valores de AG (91%) foram obtidos nos mapas CT100, CT150, CT200 e CT250 (Tabela 3). Na bacia Santo Cristo os valores de AG variaram de 77% a 89% (Tabela 3), representado um ganho de 15,6% entre o mapa de menor concordância (CT00) e os de maiores concordâncias (CT100 e CT150). Para essa área todas as classes apresentaram acurácia do mapeador igual ou superior a 64%, sendo que as duas com maior número de perfis (LV e RR) apresentaram concordância superior a 90%. Na validação com os perfis de solos para ambas às áreas de estudo, os valores de AG obtidos podem ser considerados satisfatórios, uma vez que estão de acordo com outros estudos que utilizaram perfis de solos para predição de ocorrência dos solos no MDS (Håring et al., 2012; Bagheri Bodaghabadi et al., 2015; Vasques et al., 2015; Dias et al., 2016).

**Tabela 2.** Classes taxonômicas consideradas como corretas para avaliação da reprodutibilidade dos mapas de unidades de mapeamento de solos com os perfis de solos.

CT	Unidades de mapeamento de solos													
	G	LV1	LV2	M1	M2	NV1	NV2	RL	RR1	RR2	RR3	RR4	RR5	RR6
CX														X
MX				X	X									
GX	X													
GM	X													
LV		X	X								X		X	
RY														
RL								X	X <sup>(1)</sup>	X			X	
RR			X		X		X	X	X	X	X	X	X	X
NV						X	X							

CT – Classe taxonômica de solo, CX = Cambissolos, LV = Latossolo Vermelho, MX = Chernossolo Háptico, RY= Neossolo Flúvico, RL = Neossolo Litólico, RR = Neossolo Regolítico, NV – Nitossolo Vermelho, <sup>(1)</sup> associação ocorre apenas no mapa da bacia do rio Santo Cristo.



CX – Cambissolo Háptico, MX - Chernossolo Háptico, GX – Gleissolo Háptico, GM – Gleissolo Melânico, LV – Latossolo Vermelho, RL – Neossolo Litólico, RR - Neossolo Regolítico e NV – Nitossolo Vermelho.

**Figura 2.** Proporção de pixels coletados em cada buffer (BF) nas unidades de mapeamento e classes taxonômicas de solos nas bacias do Lajeado Grande (A e C) e Santo Cristo (B e D).



**Tabela 3.** Exatidão (%) dos mapas preditos de classes taxonômicas de solos com os perfis de solos nas bacias dos rios Lajeado Grande e Santo Cristo.

Mapas	Unidades de mapeamento de solos									Acurácia geral
	CX	MX	GX	GM	LV	RY	RL	RR	NV	
Bacia hidrográfica do Lajeado Grande										
CT00	0	94	100	100	82	-	24	100	0	82
CT50	33	88	100	100	95	-	59	99	20	89
CT100	67	88	100	100	97	-	65	98	40	91
CT150	67	88	100	100	100	-	65	98	40	91
CT200	67	88	100	100	100	-	65	96	40	91
CT250	67	88	100	100	100	-	65	96	40	91
Nº de Perfis	3	16	3	2	38	0	17	85	5	168
Bacia hidrográfica do Santo Cristo										
CT00	14	67	100	50	95	0	10	91	-	77
CT50	64	67	100	50	93	100	55	95	-	86
CT100	64	89	100	100	96	100	70	92	-	89
CT150	64	89	100	100	93	100	70	94	-	89
CT200	64	89	100	100	93	100	70	91	-	88
CT250	64	89	100	100	91	100	70	90	-	86
Nº de Perfis	14	9	9	4	66	1	19	74	-	196

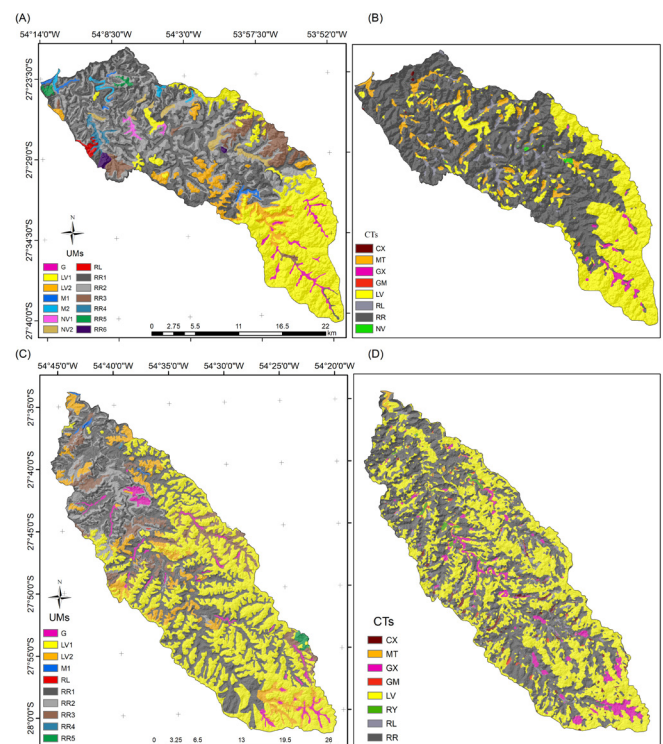
CX – Cambissolo Háplico, MT – Chernossolo Argilúvico, GX – Gleissolo Háplico, GM – Gleissolo Melânico, LV – Latossolo Vermelho, RL – Neossolo Litólico, RR – Neossolo Regolítico e NV – Nitossolo Vermelho.

O aumento no número de pixels coletados nos *buffers* promoveu ganho de acurácia do mapeador para a maioria das classes, entretanto, ocorreu uma perda de aproximadamente 4% nas classes LV e RR (Tabela 3). Essa redução de concordância pode ser atribuída a melhor predição das classes que disponibilizavam de menor número de perfis de solos e que obtiveram aumento no número de pontos amostrais com a utilização do *buffer* e inserção de erro nas classes com maior proporção de perfis georreferenciados. Esses resultados concordam com os observados na predição de solos com amostragem estratificada, que favorece classes de menor extensão e acarretam em erros nas classes de maior extensão (Teske et al., 2015).

Como pode ser visto na Figura 3, os mapas preditos com os dados coletados nos *buffers* de 150 m, para as bacias dos rios Lajeado Grande (Figura 3A) e Santo Cristo (Figura 3B), apresentam grande concordância de ocorrência espacial dos Latossolos e Neossolos com as unidades LV e RR1 do mapa convencional de solos (Figura 3A e 3C), sendo que essas duas UM correspondem a aproximadamente 70% dos solos em ambas as áreas de estudo.

De modo geral, a reprodutibilidade do mapa convencional de solos (Tabela 4) com os dados coletados apenas nos pixels dos perfis de solos alcançou valores próximos ao valor médio de 60% observado na literatura em estudos com outras estratégias de amostragem utilizada no mapeamento digital de solos (Höfig et al., 2014; Giasson et al., 2015; Bagatini et al., 2016; Dias et al., 2016; Pelegrino et al., 2016). Vale destacar que em ambas as áreas de estudo apenas duas unidades representam mais de 75% da composição dos respectivos mapas de solos, sendo estas as mais importantes na predição e as quais apresentaram maior concordância com os mapas legados utilizados na validação.

Na avaliação de concordância dos mapas preditos de classes taxonômicas de solo com os mapas de unidades de



UM – Unidades de mapeamento, CTs – Classes taxonômicas de solos, CX – Cambissolo Háplico, MT – Chernossolo Argilúvico, GX – Gleissolo Háplico, GM – Gleissolo Melânico, LV – Latossolo Vermelho, RL – Neossolo Litólico, RR – Neossolo Regolítico e NV – Nitossolo Vermelho.

**Figura 3.** Mapa convencional de solos da bacia do rio Lajeado Grande (A) e Santo Cristo (C) e mapa predito de classes taxonômicas de solos com buffer de 150 m na bacia Lajeado Grande (B) e Santo Cristo (D).

mapeamento de solo constatou-se uma ligeira redução nos valores de AG (Tabela 4). Para a bacia do Lajeado Grande, o maior valor de AG (86%) foi obtido no mapa CT00, predito com os dados coletados nos pixels dos perfis de solo. Na bacia Santo Cristo, o maior valor de AG (68%) foi obtido no mapa

**Tabela 4.** Concordância de reprodutibilidade dos mapas preditos de classes taxonômicas de solos com os mapas convencionais de solos das bacias dos rios Lajeado Grande e Santo Cristo.

Mapas	Unidades de mapeamento de solo														
	G	LV1	LV2	M1	M2	NV1	NV2	RL	RR1	RR2	RR3	RR4	RR5	RR6	AG
Bacia do rio Lajeado Grande															
CT00	32	72	98	43	98	0	73	100	88	97	98	99	94	100	86
CT50	45	72	98	51	98	0	62	100	83	94	96	97	91	96	83
CT100	42	72	99	47	97	0	53	100	80	91	95	96	78	95	81
CT150	37	73	98	47	97	0	55	100	77	89	95	95	75	93	80
CT200	32	73	97	47	96	0	54	100	75	88	95	95	74	92	79
CT250	28	73	97	46	96	0	57	100	75	87	93	94	72	91	79
Área (%)	1,61	27,21	5,84	0,67	0,73	0,45	1,68	0,41	30,35	23,91	5,26	0,86	0,52	0,49	533,3 (km <sup>2</sup> )
Bacia do rio Santo Cristo															
CT00	40	62	91	54	-	-	-	0	61	81	89	100	45	-	67
CT50	28	62	94	27	-	-	-	57	62	82	92	100	30	-	68
CT100	27	64	94	29	-	-	-	60	59	78	92	100	28	-	67
CT150	29	60	91	25	-	-	-	59	59	78	89	100	16	-	65
CT200	31	59	90	28	-	-	-	100	56	74	88	90	15	-	63
CT250	31	57	90	31	-	-	-	100	56	71	87	93	15	-	62
Área (%)	2,46	38,13	7,92	0,20	-	-	-	0,01	34,79	8,24	7,84	0,03	0,38	-	900,7 (km <sup>2</sup> )

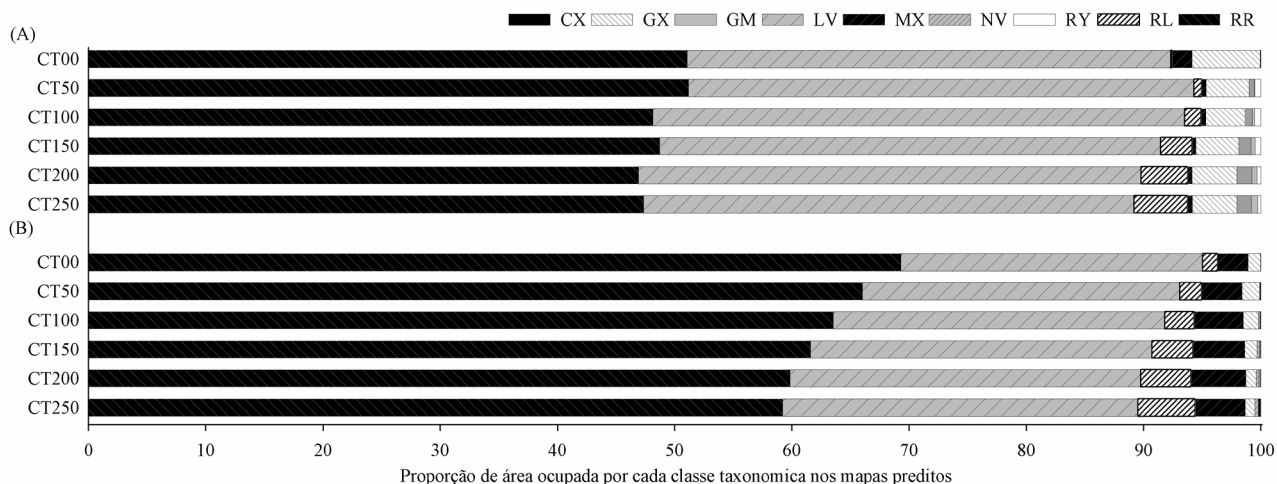
AG - Acurácia Geral. Área – áreas em porcentagem correspondente a cada unidade de mapeamento no mapa convencional de solos, (-) Unidade de mapeamento ausente no mapa convencional.

CT50, sendo que nestes mapas também são observados os maiores valores de acurácia do mapeador nas UM de maiores extensões de área (Tabela 4). Com exceção das unidades LV1 e RR1, que são unidades simples, os valores observados nas outras UM de maiores extensões são explicados por estas serem compostas por duas ou mais classes taxonômicas de solos (CT), para as quais são considerados corretos os acertos em qualquer uma das CTs que compõe as unidades de mapeamento de solo.

As unidades de mapeamento LV1, LV2, RR1 e RR2 são compostas por Latossolos Vermelhos (LV) e Neossolos Regolíticos (RR) ou associação destes, e ocorrem predominantemente em ambas às áreas de estudo, classes estas que também possuem o maior número de perfis de solos e, conseqüentemente, maior concordância de acertos em todas as avaliações realizadas. As classes taxonômicas LV e RR

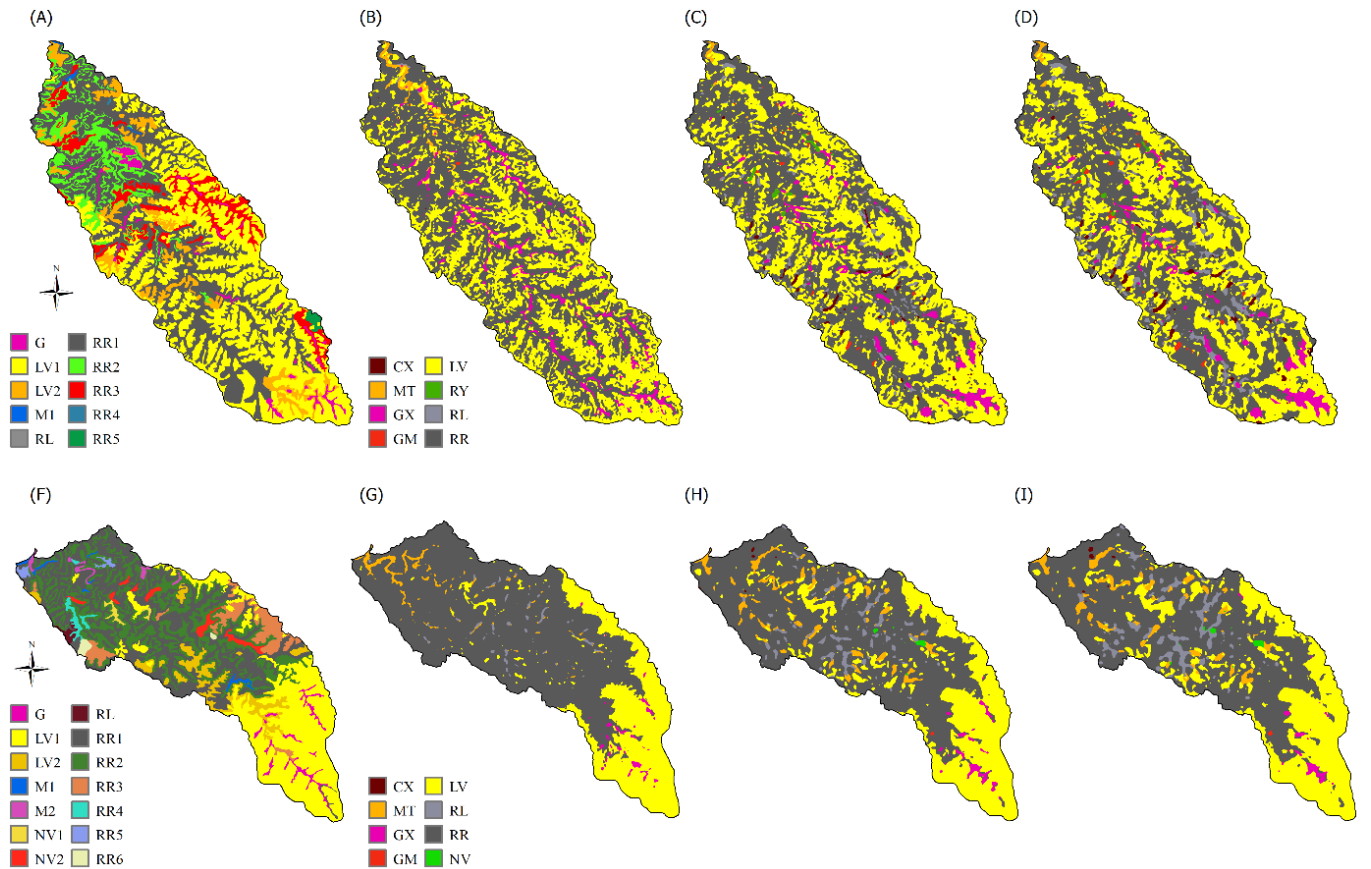
representam mais de 85% da composição dos mapas preditos (Figura 4), concordando com os mapas convencionais de solos, uma vez que estas CTs estão presentes na composição da maioria das UM.

De modo geral, o aumento no número de pixels amostrados promovido pelo *buffer* levou à redução na proporção das CTs predominantes e a um ligeiro aumento na predição das classes com menor proporção, indicando redução na superestimação das classes de maior extensão, como pode ser observado nos tratamentos CT00 e CT250 (Figura 4). Quando foram utilizados os pixels coletados nos *buffers*, houve redução na superestimação das UM de maiores extensões, com a composição do mapa CT250 ficando próximo do mapa convencional, sendo que esse efeito foi mais expressivo na bacia do Lajeado Grande e, principalmente, nas unidades LV2 e RR2 (Figura 4).



CX – Cambissolo Háplico, MX - Chernossolo Háplico, GX – Gleissolo Háplico, GM – Gleissolo Melânico, LV – Latossolo Vermelho, RL – Neossolo Litólico, RR - Neossolo Regolítico e NV – Nitossolo Vermelho.

**Figura 4.** Proporção de área de cada classe taxonômica nos mapas preditos das bacias dos rios Santo Cristo (A) e Lajeado Grande (B).



CX – Cambissolo Háplico, MT – Chernossolo Argilúvico, GX – Gleissolo Háplico, GM – Gleissolo Melânico, LV – Latossolo Vermelho, RL – Neossolo Litólico, RR – Neossolo Regolítico e NV – Nitossolo Vermelho.

**Figura 5.** Mapas convencional de unidades de mapeamento solos das bacias dos rios Santo Cristo (A) e Lajeado Grande (F) e mapas de classes taxonômicas preditos com os *buffers* de 50, 150 e 250m das bacias Santo Cristo (B, C e D) e Lajeado Grande (G, H e I).

As alterações nos mapas preditos, promovidas pela utilização de pixels amostrais coletados nos *buffers* também podem ser constatadas através da análise visual dos mapas preditos. As principais alterações ocorreram entre os mapas CT00, CT150 e CT250 para as bacias Santo Cristo (Figuras 5B, 5C e 5D) e Lajeado Grande (Figuras 5G, 5H e 5I). Nestes mapas é possível visualizar a evolução na individualização de áreas correspondentes as classes de menor ocorrência, como a RL, GX e CX.

As classes Latossolos e Neossolos ocupam posição bem definidas na paisagem, com os Latossolos ocorrendo principalmente nas posições superiores e planas e os Neossolos ocorrendo nos vales encaixados, características estas que permitem um treinamento mais eficiente do classificador para as respectivas CTs e, conseqüentemente, resultando nos maiores valores de concordância constatados. Para ambas às áreas de estudos, as classes predominantes foram preditas de forma satisfatória em todos os mapas, estes resultados corroboram outros estudos (Höfig et al., 2014; ten Caten et al., 2011) e é atribuído principalmente a forte influencia que o relevo exerce na formação destas classes de solos, as quais ocorre predominantemente em condições distintas do relevo.

A utilização de amostras coletadas apenas nos pixels dos perfis de solos foi eficiente na predição da ocorrência

das classes de solos predominantes, já a utilização do *buffer* para coleta de amostras nos locais com disponibilidades de perfis georreferenciados permite melhorar a predição das classes com menor proporção de amostras. Estes resultados demonstram que o uso desta abordagem no MDS resulta em mapas semelhantes aos obtidos no levantamento tradicional e permite reduzir ou eliminar o delineamento manual dos polígonos de solos exigido pelo mapeamento tradicional o que leva a economia na cartografia dos solos e principalmente no tempo requerido para produções e disponibilização destas informações (Demattê et al., 2015; Teske et al., 2015).

## Conclusões

A utilização dos pixels amostrais coletados nos *buffers* não alterou de forma expressiva a acurácia geral dos mapas preditos na bacia do rio Lajeado Grande, mas permitiu um ganho de 15,6% de concordância na bacia do rio Santo Cristo na validação com os perfis de solos.

Nos mapas preditos, as classes Latossolos Vermelhos e Neossolos Regolíticos foram preditas como predominantes, concordando com as proporções observadas nos perfis de solos e nos mapas convencionais disponíveis para ambas às áreas de estudo.



O uso apenas dos perfis georreferenciados resultou em mapas preditos com concordância superior a 75% de exatidão com os perfis de solos e concordância igual ou superior a 67% em relação a reprodutibilidade do mapa convencional para as duas áreas de estudo.

## Literatura Citada

- Alves, M.R.; Demattê, J.A.M.; Barros, P.P.S. Multiple geotechnological tools applied to digital mapping of tropical soils. *Revista Brasileira de Ciência do Solo*, v. 39, n. 5, p. 1261–1274, 2015. <https://doi.org/10.1590/01000683rbc20140410>.
- Arruda G.P.; Demattê, J.A.M.; Chagas, C. S.; Fiorio, P.R.; Souza, A.B.; Fongaro, C.T. Digital soil mapping using reference area and artificial neural network. *Scientia Agricola*, v. 73, n. 3, p. 266–273, 2016. <https://doi.org/10.1590/0103-9016-2015-0131>.
- Arruda, G.P.; Demattê, J.A.M.; Chagas, C.S.S. Mapeamento Digital de solos por redes neurais artificiais com base na relação solo-paisagem. *Revista Brasileira de Ciência do Solo*, v. 37, n. 1, p. 327–338, 2013. <https://doi.org/10.1590/S0100-06832013000200004>.
- Bagatini, T.; Giasson, E.; Teske, R. Expansão de mapas pedológicos para áreas fisiograficamente semelhantes por meio de mapeamento digital de solos. *Pesquisa Agropecuária Brasileira*, v. 51, n. 9, p. 1317–1325, 2016. <https://doi.org/10.1590/01000683rbc20140289>.
- Bagatini, T.; Giasson, E.; Teske, R. Seleção de densidade de amostragem com base em dados de áreas já mapeadas para treinamento de modelos de árvore de decisão no mapeamento digital de solos. *Revista Brasileira de Ciência do Solo*, v. 39, n. 4, p. 960–967, 2015. <https://doi.org/10.1590/S0100-204X2016000900031>.
- Bagheri Bodaghabadi, M.; Martínez-Casasnovas, J.A.; Salehi, M.H.; Mohammadi, J.; Esfandiarpour Borujeni, I.; Toomanian, N.; Gandomkar, A. Digital soil mapping using artificial neural networks and terrain-related attributes. *Pedosphere*, v. 25, n. 4, p. 580–591, 2015. [https://doi.org/10.1016/S1002-0160\(15\)30038-2](https://doi.org/10.1016/S1002-0160(15)30038-2).
- Behrens, T.; Zhu, A.X.; Schmidt, K.; Scholten, T. Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma*, v. 155, n. 3–4, p. 175–185, 2010. <https://doi.org/10.1016/j.geoderma.2009.07.010>.
- Breiman, L.; Cutler, A.; Liaw, A.; Wiener M. Breiman and Cutler's Random Forests for classification and regression. Vienna: The R Project for Statistical Computing; CRAN, 2018. <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>. 10 Jul. 2018.
- Breiman, L. Random Forest. *Machine Learning*, v. 45, n.1, p. 5–32, 2001. <https://doi.org/10.1023/A:1010933404324>.
- Chagas, C.S.; Fernandes Filho, E.I.; Vieira, C.A.O.; Schaefer, C.E.G.R.; Carvalho Júnior, W. Atributos topográficos e dados do Landsat7 no mapeamento digital de solos com uso de redes neurais. *Pesquisa Agropecuária Brasileira*, v. 45, n. 5, p. 497–507, 2010. <https://doi.org/10.1590/S0100-204X2010000500009>.
- Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, v. 37, n. 1, p. 35–46, 1991. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).
- Demattê, J.A.M.; Rizzo, R.; Botteon, V.W. Pedological mapping through integration of digital terrain models spectral sensing and photopedology. *Revista Ciencia Agronomica*, v. 46, n. 4, p. 669–678, 2015. <https://doi.org/10.5935/1806-6690.20150053>.
- Dias, L.M.S.; Coelho, R.M.; Valladares, G.S.; Assis, A.C.C.; Ferreira, E.P.; Silva, R.C. Predição de classes de solo por mineração de dados em área da bacia sedimentar do São Francisco. *Pesquisa Agropecuária Brasileira*, v. 51, n. 9, p. 1396–1404, 2016. <https://doi.org/10.1590/S0100-204X2016000900038>.
- Freitas, M.; Binotto, R.; Nanni, A.; Rodrigues, A.; Bortoli, C. Avaliação do potencial hidrogeológico, vulnerabilidade intrínseca e hidroquímica do sistema aquífero Serra Geral no noroeste do estado do Rio Grande do Sul. *Revista Brasileira de Recursos Hídricos*, v. 17, n. 2, p. 31–41, 2012. <https://doi.org/10.21168/rbrh.v17n2.p31-41>.
- Giasson, E.; Clarke, R.T.; Vasconcellos, A.; Junior, I.; Henrique, G.; Tornquist, C.G. Digital soil mapping using multiple logistic regression on terrain parameters in southern Brazil. *Scientia Agricola*, v. 63, n. 3, p. 262–268, 2006. <https://doi.org/10.1590/S0103-90162006000300008>.
- Giasson, E.; ten Caten, A.; Bagatini, T.; Bonfatti, B. Instance selection in digital soil mapping: a study case in Rio Grande do Sul, Brazil. *Ciência Rural*, v. 45, n. 9, p. 1592–1598, 2015. <https://doi.org/10.1590/0103-8478cr20140694>.
- Häring, T.; Dietz, E.; Osenstetter, S.; Koschitzki, T.; Schröder, B. Spatial disaggregation of complex soil map units: A decision-tree based approach in Bavarian forest soils. *Geoderma*, v. 185–186, p. 37–47, 2012. <https://doi.org/10.1016/j.geoderma.2012.04.001>.
- Hasenack, H.; Weber, E. Base cartográfica vetorial contínua do Rio Grande do Sul - escala 1:50.000 Porto Alegre: Centro de Ecologia; UFRGS, 2010. CD Rom.
- Hengl, T.; Toomanian, N.; Reuter, H.I.; Malakouti, M.J. Methods to interpolate soil categorical variables from profile observations: Lessons from Iran. *Geoderma*, v. 140, n. 4, p. 417–427, 2007. <https://doi.org/10.1016/j.geoderma.2007.04.022>.
- Höfig, P.; Giasson, E.; Vendrame, P.R.S. Mapeamento digital de solos com base na extrapolação de mapas entre áreas fisiograficamente semelhantes. *Pesquisa Agropecuária Brasileira*, v. 49, n. 12, p. 958–966, 2014. <https://doi.org/10.1590/S0100-204X2014001200006>.
- Kämpf, N.; Giasson, E.; Streck, E.V. Levantamento pedológico e análise qualitativa do potencial de uso dos solos para o descarte de dejetos suínos da microbacia do Rio Lageado Grande. Porto Alegre: Secretaria do Meio Ambiente do Rio Grande do Sul, 2004a. 27p. (Relatório).
- Kämpf, N.; Giasson, E.; Streck, E.V. Levantamento semidetalhado dos solos da microbacia do rio Santo Cristo. Porto Alegre: Secretaria do Meio Ambiente do Rio Grande do Sul; 2004b. (Relatório).
- Otto, J.-C.; Prasicek, G.; Blöthe, J.; Schrott, L. GIS Applications in geomorphology. In: Cova, T.J.; Tsou, M.-H.; Bareth, G.; Song, C.; Song, Y.; Cao, K.; Silva, E.A., (Eds.). *Comprehensive geographic information systems*. Amsterdam: Elsevier, 2017. p. 81–111.
- Pelegrino, M.H.P.; Silva, S.H.G.; Menezes, M.D.; Silva, E.; Owens, P.R.; Curi, N. Mapping soils in two watersheds using legacy data and extrapolation for similar surrounding areas. *Ciência e Agrotecnologia*, v. 40, n. 5, p. 534–546, 2016. <https://doi.org/10.1590/1413-70542016405011416>.



- R Core Team. R: a language and environment for statistical computing. Vienna: The R Project for Statistical Computing, 2018. <https://www.r-project.org>. 10 Jul. 2018.
- Santos, H.G.; Jacomine, P.K.T.; Anjos, L.H.C.; Oliveira, V.A.; Lumberras, J.F.; Coelho, M.R.; Almeida, J.A.; Araújo Filho, J.C.; Oliveira, J.B.; Cunha, T.J.F. Sistema brasileiro de classificação de solos. 5.ed. Rio de Janeiro: Embrapa Solos, 2018. 356p.
- Tachikawa, T.; Kaku, M.; Iwasaki, A.; Gesch, D.B.; Oimoen, M.J.; Zhang, Z.; Danielson, J.J.; Krieger, T.; Curtis, B.; Haase, J.; Abrams, M.; Carabajal, C. ASTER global digital elevation model version 2 - summary of validation results. Washington: NASA, 2011. 27p. <http://pubs.er.usgs.gov/publication/70005960>. 10 Jul. 2018.
- ten Caten, A.; Dalmolin, R.S.D.; Pedron, F.A.; Mendonça, S.M.L. Extrapolação das relações solo-paisagem a partir de uma área de referência. *Ciência Rural*, v.41, p.812-816, 2011. <https://doi.org/10.1590/S0103-84782011000500012>.
- ten Caten, A.; Simão, R.; Dalmolin, D.; Pedron, F.D.A. An appropriate Data Set Size for Digital Soil. *Revista Brasileira de Ciência do Solo*, v. 37, n. 1, p. 359–366, 2013. <https://doi.org/10.1590/S0100-06832013000200007>.
- Teske, R.; Giasson, E.; Bagatini, T. Produção de um mapa pedológico associando técnicas comuns aos mapeamentos digitais de solos com delineamento manual de unidades de mapeamento. *Revista Brasileira de Ciência do Solo*, v. 39, n. 4, p. 950–959, 2015. <https://doi.org/10.1590/01000683rbcs20140285>.
- Vasques, G.M.; Demattê, J.A.M.; Viscarra, R.R.A.; Ramírez, L.L.; Terra, F.S.; Rizzo, R.; Souza Filho, C.R.; Filho, C.R.S. Integrating geospatial and multi-depth laboratory spectral data for mapping soil classes in a geologically complex area in southeastern Brazil. *European Journal of Soil Science*, v. 66, n. 4, p. 767–779, 2015. <https://doi.org/10.1111/ejss.12255>.