

## Identificação de padrões de erros em estimativa de densidade do solo

### Identification of error patterns in soil bulk density estimation

Gabriela Pereira da Silva Souza<sup>1\*</sup>, Raquel Stucchi Boschi<sup>2</sup>, Maria Leonor Ribeiro Casimiro Lopes-Assad<sup>3</sup>

<sup>1</sup> Centro de Ciências Agrárias, Universidade Federal de São Carlos – UFSCar, Araras, SP, Brasil. Autor para correspondência: gabrielasouza\_12@hotmail.com

<sup>2</sup> Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo – USP, Piracicaba, SP, Brasil.

<sup>3</sup> Departamento de Recursos Naturais e Proteção Ambiental, Centro de Ciências Agrárias, Universidade Federal de São Carlos – UFSCar, Araras, SP, Brasil.

---

### RESUMO

O objetivo deste estudo foi analisar e interpretar os padrões nos erros de quatro funções de pedotransferência (PTF), desenvolvidas para estimar a densidade do solo (DS), utilizando a técnica de indução de árvore de decisão (AD). O conjunto de dados usado era composto por 974 amostras de solos de diversos biomas do Brasil e continha informações sobre granulometria (areia, silte e argila), teor de carbono (C) e pH. Os dados de granulometria foram reclassificados em classes de textura. A primeira PTF foi desenvolvida no conjunto de dados utilizado neste trabalho e as três outras, selecionadas na literatura, foram geradas para diferentes locais. A seleção das três PTFs foi feita com base nos atributos preditores utilizados, sendo necessário que todos estivessem presentes no banco de dados utilizado. Na avaliação, foram adotados limiares de 0,08 g cm<sup>-3</sup>, 0,1 g cm<sup>-3</sup> e 0,12 g cm<sup>-3</sup> para definição de erros aceitáveis e erros não aceitáveis. Quatro valores de poda (mínimo de objetos nas folhas de 5%, 10%, 15% e 20%) foram testados buscando obter melhor desempenho das ADs. As árvores geradas foram consistentes para os diferentes valores de limiar de erro e coeficiente de poda. As ADs foram semelhantes, para uma mesma PTF, independente do limiar utilizado. A PTF Benites apresentou a maior taxa de sensibilidade (83,3%), com a maior taxa de erros classificados como aceitáveis e que realmente eram aceitáveis. As PTFs apresentaram padrão nos erros e foi possível detectá-los utilizando AD.

**Palavras-chave:** árvore de decisão, funções de pedotransferência, análise de erros.

---

### ABSTRACT

The objective of this study was to analyze patterns in the errors of four pedotransfer functions (PTF), developed for estimate bulk density (DS), by using a decision tree (AD) technique. The dataset used was composed of 974 soil samples from several Brazilian biomes and contained information of particle size distribution (sand, silt, and clay), carbon content (C) and pH. The granulometry data were reclassified into texture classes. The first PTF was developed in the dataset used in this study and the three others, selected in the literature, were generated for different sites. In the evaluation, thresholds of 0.08 g cm<sup>-3</sup>, 0.1 g cm<sup>-3</sup> and 0.12 g cm<sup>-3</sup> were used for the definition of acceptable and non-acceptable errors. Four values of pruning (minimum number of objects on the 5%, 10%, 15% and 20%) were tested to obtain better performance of the ADs. The trees obtained were consistent for different thresholds and pruning values. The Benites PTF performed the highest sensitivity rate (83.3%) presenting the highest errors classified as acceptable that really were acceptable. The PTFs presented patterns in the errors, and it was possible to detect them using AD.

**Keywords:** decision tree, pedotransfer functions, error analysis.



## INTRODUÇÃO

A densidade do solo (DS) é um atributo físico dos mais importantes, pois, a partir de sua determinação, é possível avaliar diversas propriedades do solo. Compactação, drenagem e armazenamento de água são alguns dos processos influenciados pela DS.

O método mais utilizado e considerado padrão para determinação de DS envolve a coleta de amostras de solo com estrutura indeformada em anéis metálicos de bordas cortantes e volume interno conhecido. Essa prática é trabalhosa, demorada e muitas vezes impraticável (Benites et al., 2007; Heuscher et al., 2005). Existem vários outros métodos para determinação de DS, como torrão parafinado, tomografia computadorizada, atenuação gama e sonda de superfície nêutron-gama. Mas a definição de qual método é o mais adequado para medidas de DS é praticamente impossível uma vez que depende de uma série de parâmetros tais como disponibilidade de equipamentos (geralmente caros para as técnicas nucleares), análise em laboratório ou diretamente no campo, tamanho de amostra a ser analisada e tipo de estudo a ser conduzido (Pires et al., 2011). Por isso, é muito comum a falta desta informação em estudos de caracterização de solos.

Como alternativa, são usados modelos matemáticos para estimativa da DS. As funções de pedotransferência (PTFs) (Bouma, 1989) têm sido amplamente utilizadas para estimar DS, destacando-se aqui as propostas por Manrique e Jones (1991), Tomasella e Hodnett (1998), Kaur et al. (2002), Benites et al. (2007), Nemes et al. (2010) e Nanko et al. (2014). As PTFs, em geral, utilizam atributos facilmente determináveis e que fazem parte de análises de rotina dos laboratórios ou de levantamentos de solos. Essas funções têm tido uso crescente face à atual facilidade de geração de dados quantitativos e ao desenvolvimento de tecnologias de informação e computação de dados (McBratney et al., 2011).

Os métodos mais usados para gerar PTFs têm sido regressão estatística, por meio de modelos lineares ou não lineares. O desempenho das PTFs pode ser avaliado por meio de índices como ME (erro médio), RMSE (raiz quadrada do erro quadrático médio) e  $R^2$  (coeficiente de determinação). Porém, o desempenho das PTFs em conjuntos de dados diferentes do utilizado na sua geração

é incerto e pode não ser satisfatório (De Vos et al., 2005; Benites et al., 2007).

A técnica de AD pode ser empregada para avaliação dos erros de PTFs (Boschi et al., 2015b) pois possibilitam a identificação e o entendimento de padrões nos erros. Por serem representações simples, as AD agregam clareza às análises, facilitando sua utilização tanto por um especialista da área, como por um iniciante (Han et al., 2011).

O objetivo deste estudo foi analisar e interpretar os padrões nos erros de quatro PTFs, desenvolvidas para estimar a densidade do solo, utilizando árvore de decisão.

## MATERIAL E MÉTODOS

### Conjunto de Dados Utilizado

Foram utilizados dados de 974 amostras de perfis do solo, de diversos biomas do Brasil, com informações de granulometria (teores de areia, silte e argila), teor de carbono (C) e pH, para os quais foram desenvolvidas 12 PTFs (Boschi et al., 2015a). Os dados de granulometria foram reclassificados em cinco grupos de classes de textura (arenosa, siltosa, média, argilosa e muito argilosa), conforme Santos et al. (2005).

### Funções de Pedotransferência (PTFs) Utilizadas

Foram selecionadas quatro PTFs: a PTF proposta por Boschi et al. (2015a), aqui designada por PTF Boschi; a PTF C proposta por Benites et al. (2007), aqui designada por PTF Benites; a PTF proposta por Tomasella e Hodnett (1998), aqui designada por PTF T&H; e a PTF proposta por Manrique e Jones (1991), aqui designada PTF M&J.

As PTFs selecionadas e suas respectivas características são:

**a) PTF Boschi** (Equação 1): foi desenvolvida no conjunto de dados provenientes de diferentes biomas do Brasil, utilizado neste trabalho, e que apresentou o melhor desempenho ( $R^2=0,59$ ) dentre as 12 testadas (Boschi et al., 2015a).

$$DS=1,112+(0,0002913Areia)-(0,007817C)-(0,0002217Argila)+(0,06125pH) \quad (1)$$

**b) PTF Benites** (Equação 2): PTF gerada em um conjunto de 1.396 dados de solos, provenientes de

diversos biomas do Brasil (Benites et al., 2007), que, no conjunto de origem apresentou  $R^2 = 0,42$ .

$$DS = 1,5224 - (0,0005 \text{ Argila}) \quad (2)$$

- c) **PTF T&H** (Equação 3): PTF gerada em um conjunto com 396 dados de solos coletados na Amazônia brasileira (Tomasella e Hodnett, 1998), que no conjunto de origem apresentou  $R^2 = 0,6$ .

$$DS = 1,578 - (0,054C) - (0,006 \text{ Silte}) - (0,004 \text{ Argila}) \quad (3)$$

- d) **PTF M&J** (Equação 4): PTF gerada em um conjunto com 19.651 dados de solos, provenientes dos Estados Unidos, incluindo o Havaí, Porto Rico e outros países (Manrique e Jones, 1991), que no conjunto de origem apresentou  $R^2 = 0,41$ .

$$DS = 1,66 - (0,318C^{0,5}) \quad (4)$$

Nas quais:

DS = densidade do solo, expressa em  $\text{g cm}^{-3}$ ;

Areia = fração da terra fina com tamanho entre 2 mm e 0,05 mm, expressa em  $\text{g kg}^{-1}$ ;

C = teor de carbono orgânico, expresso em porcentagem;

Argila = fração da terra fina com tamanho menor que 0,002 mm, expressa em  $\text{g kg}^{-1}$ ;

pH = potencial hidrogeniônico;

Silte = fração da terra fina com tamanho entre 0,05 mm e 2  $\mu\text{m}$ , expressa em  $\text{g kg}^{-1}$ .

#### Modelagem

Na classificação dos erros das PTFs foi utilizada a técnica de AD. Os atributos preditores utilizados foram teores de argila, silte e areia, classe textural, C e pH. O atributo meta foi definido com base na categorização dos erros da estimativa da DS pelas PTFs. Na categorização dos erros considerou-se o limiar de  $0,1 \text{ g cm}^{-3}$  e foram adotadas duas classes: erro aceitável (A) e erro inaceitável (I). Este limiar foi definido considerando que o valor significativo da DS apresenta variação decimal. O valor do limiar foi alterado em 20% para testar a sensibilidade do modelo, sendo assim foram utilizados os limiares de  $0,08 \text{ g cm}^{-3}$  e  $0,12 \text{ g cm}^{-3}$ .

O programa computacional Weka 3.6 (Witten et al., 2011) foi utilizado para geração das ADs. O algoritmo utilizado para a indução das ADs foi o J48, que é uma implementação do algoritmo C4.5 no Weka. Para geração

da AD, este algoritmo subdivide o conjunto de dados em subconjuntos homogêneos e caracterizados por sua classe. Neste processo, o atributo com o maior ganho de informação é selecionado a cada subdivisão. Na divisão do conjunto de dados em conjunto de treinamento e conjunto de teste, foi utilizada a técnica de validação cruzada (*cross-validation 10 - folds*) e foram testadas pré-podas de no mínimo 5%, 10%, 15% e 20% objetos nas folhas.

A acurácia das AD foi utilizada como métrica para a seleção dos melhores modelos. A acurácia foi calculada com base nos valores de falso positivo (FP), falso negativo (FN) e número total de amostras:  $1 - (\text{FP} + \text{FN}/n)$ .

Visando o entendimento dos padrões dos erros, foi feita a análise descritiva dos exemplos pertencentes às classes A (erro aceitável) e I (erro inaceitável).

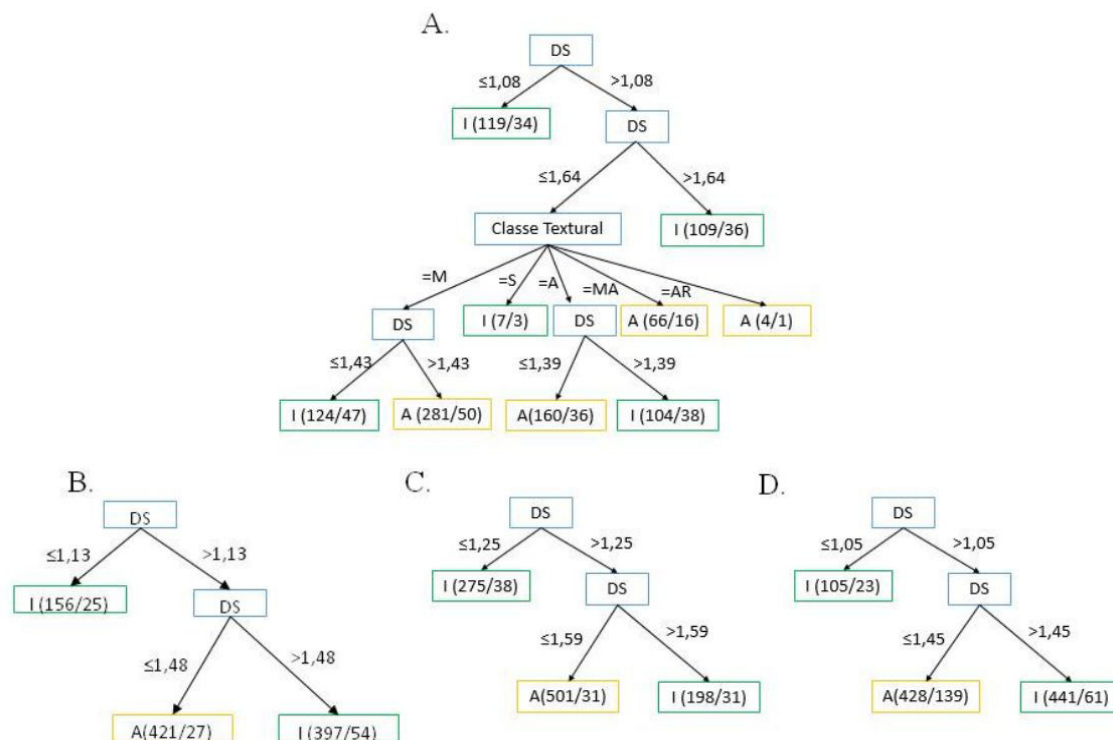
## RESULTADOS E DISCUSSÃO

As ADs foram semelhantes, para uma mesma PTF, independente do limiar utilizado ( $0,1 \text{ g cm}^{-3}$ ,  $0,08 \text{ g cm}^{-3}$  e  $0,12 \text{ g cm}^{-3}$ ). Assim, será apresentada uma árvore para cada modelo (Figuras 1A-D).

A melhor árvore para a PTF Boschi foi obtida com poda de 10% e limiar de  $0,12 \text{ g cm}^{-3}$  e com acurácia de 73,2%. A melhor árvore para a PTF Benites foi obtida com limiar de  $0,1 \text{ g cm}^{-3}$  e poda de 15%, com acurácia de 89,1%. No caso da PTF T&H, a melhor árvore foi obtida quando o limiar foi estabelecido em  $0,12 \text{ g cm}^{-3}$ , houve poda de 20%, com acurácia de 85,1%. Já para a PTF M&J, a melhor árvore foi obtida quando o limiar foi estabelecido em  $0,12 \text{ g cm}^{-3}$  e houve poda de 10% com acurácia de 77,1%.

As ADs para a PTF Boschi apresentaram nove regras e dois atributos preditivos, sendo a DS o atributo de maior relevância, seguido pela classe textural. A regra que apresentou maior número de exemplos foi considerada como regra principal. No caso da PTF Boschi foi obtida para a classe A e englobou solos com DS entre  $1,08$  e  $1,64 \text{ g cm}^{-3}$  e textura média (Figura 1A).

As ADs para as PTFs Benites, T&H e M&J apresentaram três regras, DS como atributo preditivo e a principal regra foi obtida para a classe A. Na PTF Benites essa regra englobou solos com DS entre  $1,13$  e  $1,48 \text{ g cm}^{-3}$  (Figura 1B), na PTF T&H solos com



**Figura 1.** Árvore de decisão obtida na avaliação dos erros na estimativa da DS usando: (A) PTF Boschi, proposta por Boschi et al. (2015a), com limiar de  $0,12 \text{ g cm}^{-3}$  e poda de 10%. (B) PTF Benites, proposta por Benites et al. (2007), com limiar de  $0,1 \text{ g cm}^{-3}$  e poda de 15%. (C) PTF T&H, proposta por Tomasella e Hodnett (1998), com limiar de  $0,12 \text{ g cm}^{-3}$  e poda de 20%. (D) PTF M&J, proposta por Manrique e Jones (1991), com limiar de  $0,12 \text{ g cm}^{-3}$  e poda de 10%. DS: densidade do solo; M: textura média; S: textura silteosa; A: textura argilosa; MA: textura muito argilosa; AR: textura arenosa; A: erros aceitáveis; I: erros inaceitáveis.

DS entre  $1,25$  e  $1,59 \text{ g cm}^{-3}$  (Figura 1C) e na PTF M&J solos com DS entre  $1,05$  e  $1,45 \text{ g cm}^{-3}$  (Figura 1D).

A partir da regra principal para classe A de cada PTF foram avaliados os exemplos do conjunto de dados que atendiam essa regra. Para a PTF Boschi, os exemplos englobados pela principal regra ( $SE 1,08 < DS < 1,64$  E textura média ENTÃO Classe A) apresentaram em média  $690,4 \text{ g kg}^{-1}$  ( $\pm 134,2 \text{ g kg}^{-1}$ ) de areia,  $119 \text{ g kg}^{-1}$  ( $\pm 89 \text{ g kg}^{-1}$ ) de silte,  $190,6 \text{ g kg}^{-1}$  ( $\pm 89,7 \text{ g kg}^{-1}$ ) de argila e DS de  $1,5 \text{ g cm}^{-3}$  (Tabela 1) e 72,1% dos erros aceitáveis foram classificados corretamente. Nessa regra foram enquadradas amostras com classe textural média, argilosa e muito argilosa. Todos os exemplos do conjunto de dados enquadrados na classe I apresentaram, em média,  $566,2 \text{ g kg}^{-1}$  de argila ( $\pm 183,5 \text{ g kg}^{-1}$ ),  $217,2 \text{ g kg}^{-1}$  ( $\pm 146,7 \text{ g kg}^{-1}$ ) de silte,  $216,6 \text{ g kg}^{-1}$  ( $\pm 176 \text{ g kg}^{-1}$ ) de areia e DS de  $0,9 \text{ g cm}^{-3}$  (Tabela 1). Amostras com textura argilosa, muito argilosa ou silteosa foram encontradas na classe I da PTF Boschi.

Para a PTF Benites, os exemplos do conjunto de dados que atendiam à principal regra ( $SE 1,13 < DS \leq 1,48$

ENTÃO Classe A) apresentaram em média  $377,3 \text{ g kg}^{-1}$  ( $\pm 179,5 \text{ g kg}^{-1}$ ) de argila,  $178,4 \text{ g kg}^{-1}$  ( $\pm 127,3 \text{ g kg}^{-1}$ ) de silte,  $444 \text{ g kg}^{-1}$  ( $\pm 231,7 \text{ g kg}^{-1}$ ) de areia e DS de  $1,3 \text{ g cm}^{-3}$  (Tabela 2) e 83,3% dos erros aceitáveis foram classificados corretamente. Como observado com a PTF Boschi, nessa regra foram enquadradas amostras das cinco classes texturais observadas no conjunto de dados. Para a classe I, os exemplos do conjunto de dados apresentaram como característica teores médios de  $571,2 \text{ g kg}^{-1}$  ( $\pm 185,4 \text{ g kg}^{-1}$ ) de argila,  $210,2 \text{ g kg}^{-1}$  ( $\pm 139,7 \text{ g kg}^{-1}$ ) de silte,  $218,6 \text{ g kg}^{-1}$  ( $\pm 177,3 \text{ g kg}^{-1}$ ) de areia e DS de  $0,96 \text{ g cm}^{-3}$  (Tabela 2). Na classe I da PTF Benites apenas amostras com textura arenosa não foram englobadas.

Para a PTF T&H, os exemplos do conjunto de dados que atendiam à principal regra ( $SE 1,25 < DS \leq 1,59$  ENTÃO Classe A) apresentaram, em média,  $446 \text{ g kg}^{-1}$  ( $\pm 67,1 \text{ g kg}^{-1}$ ) de argila,  $169,3 \text{ g kg}^{-1}$  ( $\pm 103,3 \text{ g kg}^{-1}$ ) de silte,  $384,6 \text{ g kg}^{-1}$  ( $\pm 139,2 \text{ g kg}^{-1}$ ) de areia e DS de  $1,42 \text{ g cm}^{-3}$  (Tabela 3) e 70,7% dos erros aceitáveis foram classificados corretamente. Para a classe I, os exemplos no

conjunto de dados apresentaram, em média, 537,7 g kg<sup>-1</sup> ( $\pm 179,3$  g kg<sup>-1</sup>) de argila, 208,5 g kg<sup>-1</sup> ( $\pm 130,3$  g kg<sup>-1</sup>) de silte, 253,8 g kg<sup>-1</sup> ( $\pm 176,7$  g kg<sup>-1</sup>) de areia e DS de 1,06 g cm<sup>-3</sup> (Tabela 3). Nas Classes A e I da PTF T&H foram encontradas amostras de todas as classes texturais.

Para a PTF M&J, os exemplos do conjunto de dados que atendiam à principal regra ( $SE 1,05 < DS \leq 1,45$

ENTÃO Classe A) apresentaram, em média, 409,9 g kg<sup>-1</sup> ( $\pm 235,1$  g kg<sup>-1</sup>) de argila, 183,4 g kg<sup>-1</sup> ( $\pm 127$  g kg<sup>-1</sup>) de silte, 208,5 g kg<sup>-1</sup> ( $\pm 130,3$  g kg<sup>-1</sup>) de areia e DS de 1,06 g cm<sup>-3</sup> (Tabela 3). Nas Classes A e I da PTF T&H foram encontradas amostras de todas as classes texturais. Para a classe ENTÃO Classe A) apresentaram, em média, 409,9 g kg<sup>-1</sup> ( $\pm 235,1$  g kg<sup>-1</sup>) de argila, 183,4 g kg<sup>-1</sup> ( $\pm 127$  g kg<sup>-1</sup>) de silte, 208,5 g kg<sup>-1</sup> ( $\pm 130,3$  g kg<sup>-1</sup>) de areia e DS de 1,06 g cm<sup>-3</sup> (Tabela 3) e 77,5% dos erros aceitáveis foram classificados corretamente. No conjunto de dados, as amostras que englobaram esses exemplos apresentaram todas as cinco classes de textura observadas. Para a classe

**Tabela 1.** Estatística descritiva dos exemplos enquadrados na principal regra para as classes A (erros aceitáveis) e I (erros inaceitáveis), das árvores geradas para avaliar os erros na estimativa da DS pela PTF desenvolvida por Boschi et al. (2015a) (PTF Boschi), nos limiares de 0,08 g cm<sup>-3</sup>, 0,1 g cm<sup>-3</sup> e 0,12 g cm<sup>-3</sup>.

Atributos	Classe A (erros aceitáveis)				Classe I (erros inaceitáveis)			
	Areia	Silte	Argila	DS	Areia	Silte	Argila	DS
	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>
Média	690,4	119,0	190,6	1,5	216,6	217,2	566,2	0,9
DP	134,2	89,0	89,7	0,1	176,0	146,7	183,5	0,1
Mínimo	180	10	2	1,44	15,7	21	2	0,45
Máximo	966	590	346	1,64	929	696	910	1,08
Mediana	690	98	196	1,56	150	150	600	0,98

DP: desvio padrão.

**Tabela 2.** Estatística descritiva dos exemplos enquadrados na principal regra para as classes A (erros aceitáveis) e I (erros inaceitáveis) das árvores geradas para avaliar os erros na estimativa da DS pela PTF Benites, proposta por Benites et al. (2007).

Atributos	Classe A (erros aceitáveis)				Classe I (erros inaceitáveis)			
	Areia	Silte	Argila	DS	Areia	Silte	Argila	DS
	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>
Média	444,3	178,4	377,3	1,3	218,6	210,2	571,2	0,96
DP	231,7	127,3	179,5	0,1	177,3	139,7	185,4	0,15
Mínimo	30	20	2	1,14	14,9	21	1	0,45
Máximo	951	680	880	1,48	932	696	920	1,13
Mediana	460	146,2	370	1,35	152,5	150	600	1

DP: desvio padrão.

**Tabela 3.** Estatística descritiva dos exemplos enquadrados na principal regra para as classes A (erros aceitáveis) e I (erros inaceitáveis), das árvores geradas para avaliar os erros na estimativa da DS pela PTF T&H, proposta por Tomasella e Hodnett (1998).

Atributos	Classe A (erros aceitáveis)				Classe I (erros inaceitáveis)			
	Areia	Silte	Argila	DS	Areia	Silte	Argila	DS
	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>
Média	384,6	169,3	446,1	1,42	253,8	208,5	537,7	1,06
DP	139,2	103,3	67,1	0,09	176,7	130,3	179,3	0,16
Mínimo	70	30	350	1,26	14,9	21	1	0,45
Máximo	600	430	587	1,59	932	696	920	1,25
Mediana	430	140	440	1,40	223	165	560	1,11

DP: desvio padrão.

**Tabela 4.** Estatística descritiva dos exemplos enquadrados na principal regra para as classes A (erros aceitáveis) e I (erros inaceitáveis), das árvores geradas para avaliar os erros na estimativa da DS pela PTF M&J, proposta por Manrique e Jones (1991).

Atributo	Classe A (erros aceitáveis)				Classe I (erros inaceitáveis)			
	Areia	Silte	Argila	DS	Areia	Silte	Argila	DS
	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>	-----g kg <sup>-1</sup> -----			g cm <sup>-3</sup>
Média	406,6	183,4	409,9	1,3	208,1	224,7	567,2	0,90
DP	235,1	127,0	192,9	0,1	173,6	152,5	187,0	0,14
Mínimo	14,9	20	1	1,06	15,7	21	2	0,45
Máximo	951	680	920	1,45	929	696	910	1,05
Mediana	389,5	150	420	1,3	147	150	600	0,93

DP: desvio padrão.

I, os exemplos do conjunto de dados apresentaram, em média, 567 g kg<sup>-1</sup> ( $\pm$  187 g kg<sup>-1</sup>) de argila, 224,7 g kg<sup>-1</sup> ( $\pm$  152,5 g kg<sup>-1</sup>) de silte, 208 g kg<sup>-1</sup> ( $\pm$  173,6 g kg<sup>-1</sup>) de areia e DS de 0,90 g cm<sup>-3</sup> (Tabela 4). Amostras com textura argilosa, muito argilosa, siltosa ou média foram encontradas na classe I da PTF M&J.

O valor médio de DS observado nos exemplos que enquadram a classe A, tanto na PTF Boschi como nas PTFs Benites e T&H, foi próximo e situado em torno de 1,5 g cm<sup>-3</sup>. Isto pode ser explicado pelo fato das PTFs terem sido desenvolvidas para solos brasileiros, embora com conjuntos de dados diferentes. Foi possível observar uma diferença entre as médias da DS para a classe A e para a classe I, que ficou em torno de 0,9 g cm<sup>-3</sup> para todas as PTFs, mostrando que independente da PTF, os erros foram classificados como inaceitáveis para valores extremos de DS.

As árvores para as PTFs Benites, T&H e M&J apresentaram o mesmo atributo preditor e regras semelhantes, entretanto, as características texturais do solo foram diferentes nas três árvores.

## CONCLUSÃO

- 1) As PTFs para estimativa da densidade do solo apresentam baixo desempenho na estimativa de valores extremos.
- 2) Os erros na estimativa de DS com PTFs apresentam padrões.

## AGRADECIMENTOS

Ao CNPq pela Bolsa concedida à primeira autora no PIBIC/CNPq – UFSCar, e à FAPESP (Processo 2015/06804-0), pela Bolsa de Pós-Doutorado concedida à segunda autora.

## REFERÊNCIAS

BENITES, V.M., MACHADO, P.L.O.A., FIDALGO, E.C.C., COELHO, M.R. & MADARI, B.E., 2007. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. *Geoderma*, vol. 139, no. 1-2, pp. 90-97. <http://dx.doi.org/10.1016/j.geoderma.2007.01.005>.

BOSCHI, R.S., LOPES-ASSAD, M.L.R.C., SOUZA, G.P.S., CIDIN, A.C.M. & ASSAD, E.D., 2015a. Efeito

de estimativa de densidade do solo no cálculo de estoque de carbono. *Ciência, Tecnologia & Ambiente*, vol. 2, no. 1, pp. 8-16.

BOSCHI, R.S., RODRIGUES, L.H.A. & LOPES-ASSAD, M.L., 2015b. Analysis of Patterns of Pedotransfer Function Estimates: An Approach Based on Classification Trees. *Soil Science Society of America Journal*, vol. 0, no. 3, pp. 720-729. <http://dx.doi.org/10.2136/sssaj2014.11.0452>.

BOUMA, J., 1989. Using soil survey data for quantitative land evaluation. *Advances in Soil Science*, vol. 9, pp. 177-213. [http://dx.doi.org/10.1007/978-1-4612-3532-3\\_4](http://dx.doi.org/10.1007/978-1-4612-3532-3_4).

DE VOS, B., VAN MEIRVENNE, M.V., QUATAERT, P., DECKERS, J. & MUYS, B., 2005. Predictive quality of pedotransfer functions for estimating bulk density of forest soils. *Soil Science Society of America Journal*, vol. 69, no. 2, pp. 500-510. <http://dx.doi.org/10.2136/sssaj2005.0500>.

HAN, J., KAMBER, M. & PEI, J., 2011. *Data mining: concepts and techniques*. 3. ed. San Francisco: Morgan Kaufmann Publishers Inc. 740 p.

HEUSCHER, S.A., BRANDT, C.C. & JARDINE, P.M., 2005. Using soil physical and chemical properties to estimate bulk density. *Soil Science Society of America Journal*, vol. 69, pp. 51-56. <http://dx.doi.org/10.2136/sssaj1991.03615995005500020030x>.

KAUR, R., KUMAR, S. & GURUNG, H.P., 2002. A pedo-transfer function soil data and its comparison with existing PTFs. *Australian Journal of Soil Research*, vol. 40, no. 5, pp. 847-857. <http://dx.doi.org/10.1071/SR01023>.

MANRIQUE, L.A. & JONES, C.A., 1991. Bulk-density of soils in relation to soil physical and chemical properties. *Soil Science Society of America Journal*, vol. 55, no. 2, pp. 476-481. <http://dx.doi.org/10.2136/sssaj1991.03615995005500020030x>.

MCBRATNEY, A.B., MINASNY, B. & TRANTER, G., 2011. Necessary meta-data for pedotransfer functions. *Geoderma*, vol. 160, no. 3-4, pp. 627-629. <http://dx.doi.org/10.1016/j.geoderma.2010.09.023>.

NANKO, K., UGAWA, S., HASHIMOTO, S., IMAHA, A., KOBAYASHI, M., SAKAI, H., ISHIZUKA, S., MIURA, S., TANAKA, N., TAKAHASHI, M. & KANEKO, S., 2014. A pedotransfer function for estimating bulk density of forest soil in Japan affected by volcanic ash.

*Geoderma*, vol. 213, pp. 36-45. <http://dx.doi.org/10.1016/j.geoderma.2013.07.025>.

NEMES, A., QUEBEDEAUX, B. & TIMLIN, D.J., 2010. Ensemble approach to provide uncertainty estimates of soil bulk density. *Soil Science Society of America Journal*, vol. 74, no. 6, pp. 1938-1945. <http://dx.doi.org/10.2136/sssaj2009.0370>.

PIRES, L.F., ROSA, J.A. & TIMM, L.C., 2011. Comparação de métodos de medida da densidade do solo. *Acta Scientiarum. Agronomy*, vol. 33, no. 1, pp. 161-170. <http://dx.doi.org/10.4025/actasciagron.v33i1.5507>.

SANTOS, R.D., DE LEMOS, R.C., SANTOS, H.G., KER, J.C. & ANJOS, L.H., 2005. *Manual de descrição e coleta de solo no campo*. 5. ed. Viçosa: Sociedade Brasileira de Ciência de Solo. 100 p.

TOMASELLA, J. & HODNETT, M.G., 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Science*, vol. 163, no. 3, pp. 190-202. <http://dx.doi.org/10.1097/00010694-199803000-00003>.

WITTEN, I., FRANK, E. & HALL, M.A., 2011. *Data mining: practical machine learning tools and techniques*. 3. ed. Burlington: Morgan Kaufmann. 664 p.