



## FUNÇÃO DE PEDOTRANSFERÊNCIA BASEADA EM ENSEMBLE DE ROTAÇÃO PARA ESTIMATIVA DA RETENÇÃO DE ÁGUA NO SOLO

E.J. Ferreira<sup>1</sup>, A.A. Netto<sup>1</sup>, S. Crestana<sup>1</sup>

(1) Embrapa Instrumentação, Rua XV de novembro, 1452, 13560-970, São Carlos, SP, ednaldo.ferreira@embrapa.br, nettoangelotti@yahoo.com.br, silvio.crestana@embrapa.br

**Resumo:** A estimativa da capacidade de retenção de água de um solo é uma propriedade importante à atividade agrícola por estar intrinsecamente relacionada com a disponibilidade de água às plantas, estresse hídrico, condutividade hidráulica, manejo de irrigação, entre outras. Os métodos experimentais para levantamento das curvas de retenção de água são relativamente precisos, mas demandam tempo, equipamentos de custos elevados e aderência à realidade de campo. O desenvolvimento de *Funções de Pedotransferência* (FPTs) para estimar a retenção da água a partir propriedades básicas (trivialmente adquiridas) tem sido uma área de pesquisa intensa nas últimas décadas. Nesse sentido, os métodos para construção de *ensemble de Redes Neurais Artificiais* (RNAs) merecem destaque pelo potencial de ampliar a acurácia e a precisão das FPTs induzidas. Resultados promissores para aplicações de diversas naturezas têm sido obtidos por *ensembles* construídos por um método denominado “baseado em rotação” (RB: *Rotation-Based*). Nesse contexto, o propósito deste trabalho foi avaliar o desempenho do método RB para construção de *ensemble* de RNAs visando estimar a retenção de água em solos da microbacia hidrográfica do ribeirão Canchim (São Paulo). Os resultados de validação mostraram que o RB obteve erro absoluto significativamente inferior ( $p$ -valor  $< 0,05$ ) ao de uma RNA e a outro método de *ensemble* (*Bagging*). A variância também foi reduzida pelo método RB e o coeficiente de determinação médio foi superior a 0,9. Esses resultados indicam que o RB é um método promissor e adequado para aumento da exatidão e precisão das FPTs.

**Palavras-chave:** pedotransferência, *ensembles*, redes neurais, retenção de água, *rotation-based*.

### PEDOTRANSFER FUNCTION BASED ON ROTATION-BASED ENSEMBLE FOR ESTIMATION OF SOIL WATER RETENTION

**Abstract:** The water holding capacity of a soil is an important property for agricultural activity due to its intrinsic relation with water availability to plants, water stress, hydraulic conductivity, irrigation management, among others. Experimental methods for fitting curves of water retention are relatively accurate but they demand time, costs with equipment and adherence to field. The development of Pedotransfer Functions (PTFs) for estimation of water retention from basic properties (trivially acquired) has been an intensive research area in recent decades. In this sense, methods for building an ensemble of Artificial Neural Networks (ANN) are highlighted because of the potential for improving the accuracy and precision of PTFs. Promising results for several applications have been obtained by using ensembles built by a method called *Rotation-based* (RB). In this context, the purpose of this study was the performance evaluation of the RB method for building an ensemble of ANNs aiming to estimate water retention in soils from the watershed Canchim (São Paulo). The results obtained from validation experiments showed that RB method obtained an absolute error significantly smaller ( $p$ -value  $< 0.05$ ) compared to an ANN and also to another ensemble method (*Bagging*). The variance was reduced by using RB and the coefficient of determination was greater than 0.9. These results indicate that RB method is suitable for increasing the accuracy and precision of PTFs.

**Keywords:** pedotransfer, ensembles, neural networks, retention of water, rotation-based.

#### 1. Introdução

A premente expansão da produção agrícola mundial tem produzido impactos importantes nos solos, especialmente nos países como o Brasil, que tem o negócio da agricultura como motor imane da economia. A intensificação da produção de alimentos está diretamente relacionada com uso da água. O produtor rural expande a produção agrícola por meio, dentre outras, do manejo de irrigação. Estima-se que aproximadamente 64% da água doce utilizada pelo homem são destinadas à irrigação (GUERRA et al., 2004).

Os levantamentos realizados em solos brasileiros exibem um quantitativo relativamente abundante de dados químicos e morfológicos, mas ainda tímidos em termos das propriedades hídricas. A capacidade de retenção de água em potenciais mátricos específicos e a condutividade hidráulica são propriedades importantes por permitir a mensuração do fluxo de água e dos solutos na região não saturada de um solo. A estimativa da capacidade de

retenção de água de um solo é deveras crucial aos estudos de disponibilidade de água às plantas, delineamento dos projetos de irrigação e drenagem, infiltração, condutividade hidráulica, estresse hídrico e movimento de solutos. As curvas de retenção de água do solo descrevem a relação entre conteúdo de água e o potencial matricial, possibilitando a obtenção de outras propriedades inerentes. Por outro lado, os métodos experimentais para elaboração dessas curvas geralmente demandam tempo considerável, requerem equipamentos onerosos e/ou, ainda, apresentam ligeira inadequação à realidade de campo (WILLIAMS et al., 1992; OLIVEIRA e SARVASI, 1997). Nesse sentido, o desenvolvimento de modelos de regressão capazes de estimar a retenção de água a partir de propriedades básicas do solo, de obtenção trivial e de baixo custo, tem sido uma área de interesse que Bouma (1989) denominou como *Funções de Pedotransferência* (FPTs). Diversas abordagens têm sido utilizadas para tratar as FPTs, nem sempre gerando os mesmos resultados. Tomasella e colaboradores (2003), utilizando-se de uma ampla base de dados de retenção de água de solos brasileiros, desenvolveram FPTs comparando-se dois métodos donde um deles mostrou-se superior.

No bojo das pesquisas com FPTs, modelos de regressão baseados em Redes Neurais Artificiais (RNAs) têm recebido destaque por demonstrarem desempenhos superiores nas estimativas da retenção de água (SCHAAP e BOUTEN, 1996; PACHEPSKY et al., 1996). As RNAs são modelos computacionais inspirados na estrutura e funcionamentos dos neurônios biológicos. Uma fase de treinamento confere à RNA a capacidade de reconhecer padrões e classificá-los. O poder de indução da natureza matemática inerente aos dados torna as RNAs capazes de ajustar desde funções matemáticas simples a funções de alta complexidade, espectro que inclui qualquer FPT. Um dos modelos de RNAs utilizado como FPT tem sido a MLP (do inglês: *MultiLayer Perceptron*) treinada pelo algoritmo *backpropagation* (RUMELHART et al., 1986). As MLPs aplicadas à retenção de água têm mostrado resultados superiores (NETTO, 2007).

Outra maneira de aumentar a exatidão e a precisão de FPTs é por meio dos denominados *ensembles de modelos*. Um *ensemble* é um modelo composto por versões acuradas e diversificadas de FPTs (ao invés de uma única função). A média das respostas providas por todos os modelos componentes é, tipicamente, a resultante provida pelo *ensemble*. Em geral, a vantagem de adotar *ensembles* é o aumento da acurácia e/ou da precisão das estimativas. Netto (2007) avaliou uma FPT baseada em um *ensemble* de MLPs, construído pelo método clássico conhecido como *Bagging* (BREIMAN, 1996). Corroborando a teoria, os resultados por ele obtidos foram ligeiramente melhores do que com apenas uma MLP.

Recentemente, um novo método para construção de *ensembles*, denominado “baseado em rotação” (RB, do inglês *Rotation-Based*), tem demonstrado desempenho superior a outros clássicos (RODRÍGUEZ; ALONSO, 2004), incluindo *Bagging*. O método RB produz resultados promissores por meio da diversificação de rotações proporcionadas por transformações baseadas na extração de componentes principais. Ante a esse potencial de desempenho superior, o propósito deste trabalho foi avaliar o método RB na tarefa de ajustar uma FPT baseada em um *ensemble* de MLPs para estimar a retenção de água em solos de uma microbacia de São Paulo.

## 2. Materiais e Métodos

Os solos utilizados foram amostrados da microbacia hidrográfica do ribeirão Canchim sob três diferentes condições de manejo: sem cobertura vegetal; com cobertura; e de mata. As classes de solos foram: Latossolo Vermelho Eutrófico e Distroférico, Nitossolo Vermelho Eutrófico e Latossolo Vermelho-Amarelo Distroférico. Para cada classe-manejo, uma trincheira foi cavada para coleta de amostras indeformadas em cinco profundidades: 0,1, 0,2, 0,3, 0,4 e 0,6m. A determinação da retenção de água de cada tripla *classe-manejo-profundidade* dos solos foi realizada pelo método padrão da câmara de Richards, submetendo-os às tensões de 0,001; 0,033; 0,10; 0,30 e 1,5 MPa.

As determinações das frações de argila, silte e areia para cada *classe-manejo-profundidade* foram realizadas com auxílio de um analisador granulométrico (VAZ et al., 1999). A densidade do solo foi obtida por meio dos procedimentos propostos por Lemos e Santos (2002). Dessa forma, a base de dados utilizada neste trabalho foi constituída por 275 amostras, instanciadas para cada *classe-manejo-profundidade* para as variáveis independentes (propriedades básicas) de profundidade, tensão ( $\theta$ ), densidade, areia, silte e argila; assim como para a variável dependente da retenção de água, obtida pelo método padrão. A referida base de dados foi a mesma utilizada por Netto (2007), excluindo o atributo de resistência à penetração do solo.

Para análises e comparações foi utilizada a ferramenta Weka (HALL et al., 2009). O procedimento de validação cruzada em 10 *folds* foi repetido 10 vezes para obtenção das estimativas de desempenho em múltiplas aleatorizações de particionamento. Os desempenhos obtidos pelo *Backpropagation* (BP-MLP), *Bagging* e RB foram comparados estatisticamente usando o teste *t-pareado modificado* com nível de significância de 5%. A métrica de desempenho adotada foi o Erro Absoluto Médio (EAM). O coeficiente de determinação ( $r^2$ ) médio foi calculado para elucidar a força da relação entre valores medidos e preditos pelos modelos.

Os parâmetros dos métodos *Bagging* e RB foram os predefinidos no Weka. Os tamanhos dos *ensembles* foram de 50 MLPs. As topologias das MLPs foram as mesmas para todos os algoritmos avaliados: 6 elementos na camada de entrada; 3 neurônios na camada intermediária e 1 neurônio na camada de saída. Foram utilizadas 10000 épocas no treinamento. Os demais parâmetros do algoritmo *Backpropagation* foram aqueles predefinidos na ferramenta.

### 3. Resultados e Discussão

Os EAMs e seus respectivos desvios das estimativas de  $\theta$ , obtidos para os três métodos (algoritmos), são apresentados na Figura 1. Os métodos BP-MLP, *Bagging* e RB foram comparados, par-a-par, para diferenças estatisticamente significativas ( $\alpha=0,05$ ). Os círculos internos às barras indicam que o EAM da respectiva barra do método foi significativamente menor do que os EAMs dos métodos equivalentes às cores de preenchimento dos círculos.

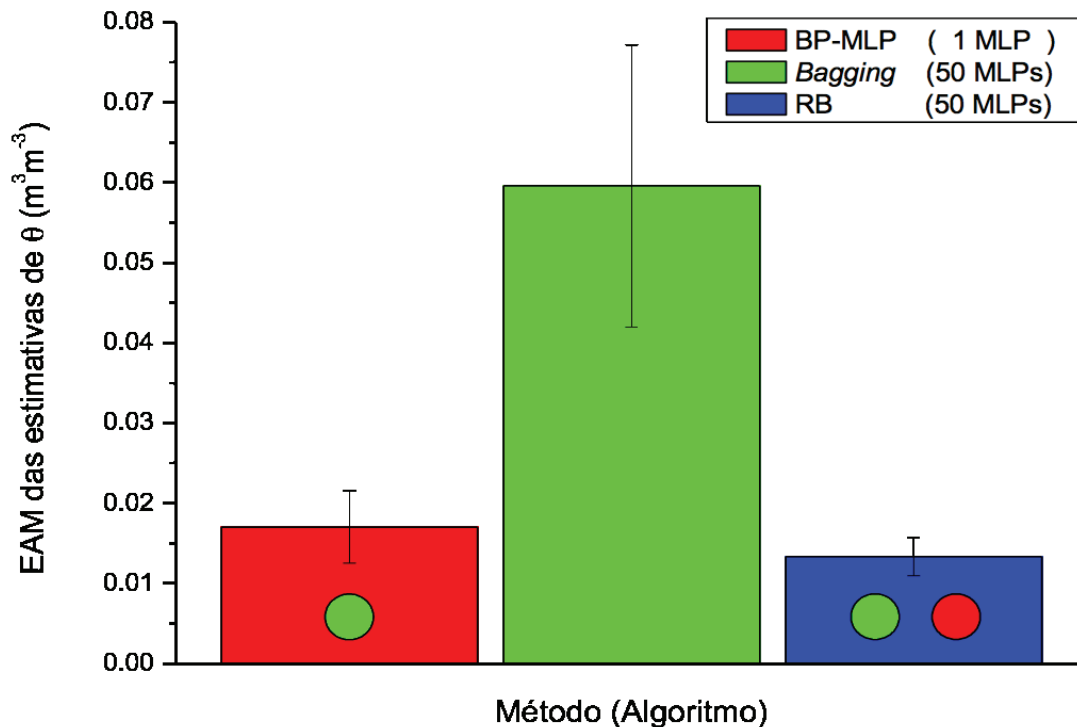


Figura 1. Desempenho em validação cruzada dos métodos na estimativa de retenção de água.

Diferentemente dos resultados obtidos por Netto (2007), *Bagging* apresentou o pior desempenho na comparação com BP-MLP e, ainda mais acentuado, com RB, exibindo EAM (0,059) significativamente superior aos demais nas comparativas resultantes do teste *t*. Por outro lado, o método RB apresentou um EAM (0,013) ligeira, mas significativamente menor na comparação com BP-MLP e, de maneira ainda mais discrepante, com respeito a *Bagging*; corroborando a superioridade esperada em relação aos demais.

Os desvios (desvio padrão) exibidos no topo das barras do gráfico da Figura 1 também indicam que o método RB apresenta uma variância (0,0024) ligeiramente inferior ao de BP-MLP (0,0045) e consideravelmente menor em relação a *Bagging* (0,0176). Adicionalmente, os coeficientes de determinação médios de BP-MLP, *Bagging* e RB foram, respectivamente, 0,88; 0,17 e 0,91. Esses achados indicam que o método RB, além da maior exatidão, contribui para maior precisão (menor variabilidade) das estimativas da retenção de água no solo.

### 4. Conclusões

O método RB tem apresentado resultados promissores em aplicações de diversas naturezas. Neste trabalho o RB foi avaliado no desempenho de uma FPT com dados de uma microbacia hidrográfica de São Paulo. Os resultados obtidos são animadores. O EAM de validação estimado foi inferior a  $0,015 \text{m}^3 \text{m}^{-3}$ , valor significativamente menor do que os demais métodos considerados nas comparativas. A variância também sofreu redução importante e o coeficiente de determinação médio (0,91) foi o maior em relação aos demais. Os resultados obtidos induzem duas asserções: (1) o método RB é promissor para construção de FPTs baseadas em *ensembles* de maiores acurácia e precisão; (2) a FPT baseada no *ensemble* construído pelo RB permite uma estimativa relativamente acurada e precisa da retenção de água dos solos similares aos obtidos na microbacia hidrográfica do Canchim e, consequentemente, a elaboração das curvas de retenção.

### Agradecimentos

Ao CNPq pelo apoio financeiro que viabilizou as coletas em campo.

**Referências**

- BOUMA, J. Using soil survey data for quantitative land evaluation. *Advances in Soil Science*, v. 9, p. 177-213, 1989.
- BREIMAN L. Bagging predictors. *Machine Learning*, v. 24, p. 123-140, 1996.
- GUERRA, A.F.; LIMA, J.E.F.W.; RODRIGUES, L.N. Água e irrigação. In: *Agroanalysis: Embrapa: excelência no agronegócio brasileiro*. FGV, ed. especial, v. 24, p. 12-13, 2004.
- HALL, M. et al. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, p. 10-18, November 2009. ISSN 1931-0145.
- LEMOS, R.C.; SANTOS, R.D. Manual de descrição e coleta de solo no campo. 4ed. Viçosa: Sociedade Brasileira de Ciência do Solo, 2002. 83p.
- NETTO, A. A. Estimativa da retenção de água no solo a partir do uso de equipamentos não convencionais, redes neurais artificiais e funções de pedotransferência. 2007. (Tese de Doutorado), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.
- OLIVEIRA, L.F.C.; SARVASI, F.O.C. Determinação da curva de retenção de um latossolo vermelho-escuro da região de Alfenas por tensiometria. *Ciência e Agrotecnologia*, v. 21, p. 91-95, 1997.
- PACHEPSKY, Y. A.; TIMLIN, AD.; VARALLYAY, G. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Science Society of America Journal*, v. 60: 727-733, 1996.
- RODRIGUEZ, J. J.; ALONSO, C. J. Rotation-based ensembles. In: *Current Topics in Artificial Intelligence: 10th Conference of the Spanish Association for Artificial Intelligence*. [S.l.]: Springer, 2004. p. 498-506.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representation by back-propagating errors. *Nature*, London, v. 323, p. 533-536, 1986.
- SCHAAP, M.; BOUTEN, W. Modeling water retention curves of sandy soils using neural networks. *Water Research Resource*, v. 32, p. 3033-3040, 1996.
- TOMASELLA et al.,. Comparison of two techniques to develop pedotransfer functions for water retention. *Soil Sci. Am. J.*, p. 1085-1092, v. 67, 2003.
- VAZ, C.M.P.; NAIME, J.M.; MACEDO, A. Soil particle size fractions determined by gamma-ray attenuation. *Soil Science*, v. 164, p. 403-410, 1999.
- WILLIAMS, R.D.; AHUJA, L.R.; NANEY, J.W. Comparison of methods to estimate soil water characteristics from soil texture, bulk density, and limited data. *Soil Science*, v.153, p.172-184, 1992.