

Estimação da curva de retenção de água no solo considerando erros heteroscedásticos

Carolina Costa Mota Paraíba - DEs, UFSCar¹²

Aline de Holanda Nunes Maia - Embrapa Meio Ambiente

Carlos Alberto Ribeiro Diniz - DEs, UFSCar

Lineu Neiva Rodrigues - Embrapa Cerrados

Resumo: *A curva de retenção de água no solo (CRA) é uma representação gráfica da relação funcional entre o conteúdo de água (umidade) no solo e seu potencial matricial. Em agricultura irrigada, a CRA tem sido identificada como uma informação chave do solo, sendo necessária para a interpretação de propriedades do solo, tal como a condutividade hidráulica. O método mais utilizado para estimação dos parâmetros da CRA é o método de quadrados mínimos (MQM) não lineares, considerando-se que os erros experimentais são independentes e normalmente distribuídos com média zero e variância constante. Porém, a estimação da CRA é geralmente feita sem que nenhum estudo diagnóstico do modelo estimado seja conduzido para verificar sua adequabilidade. A violação dos referidos pressupostos compromete a quantificação de incertezas das estimativas e testes de hipóteses sobre os parâmetros da CRA, uma vez que as inferências estatísticas são construídas com base na especificação de uma função de distribuição de probabilidade para os erros. Além disso, em teoria estatística, sabe-se que os estimadores de MQM serão apropriados e ótimos apenas se os erros de mensuração são não correlacionados com variância constante. No presente trabalho apresentamos uma proposta alternativa para o ajuste do modelo de van Genuchten (1980) quando os erros experimentais afastam-se da suposição de homoscedasticidade.*

Palavras-chave: *modelos não lineares, heterocedasticidade, modelo de van Genuchten.*

1 Introdução

A CRA é de fundamental importância no estudo das relações entre água e solo, fenômeno físico que afeta o uso de solos nos mais diferentes propósitos, funcionando como uma ferramenta de interpretação e avaliação do comportamento dos solos não saturados de acordo com as mudanças na sucção, estando disponíveis diversas relações teóricas e empíricas para a modelagem das propriedades dos solos não saturados. As curvas de retenção podem ser estimadas através do ajuste de modelos não lineares a conjuntos de dados experimentais ou a partir de funções de pedotransferência, que são modelos de regressão múltipla que relacionam os parâmetros das CRA a atributos físicos do solo.

¹Doutoranda do PPG - DEs, UFSCar

²Contato: carolina_paraiba@yahoo.com.br

Quando construída a partir de dados experimentais, a CRA é elaborada utilizando um conjunto de pares (θ_{ij}, ψ_{ij}) , $i = 1, \dots, k$ tensões e $j = 1, \dots, J$ replicações, obtidos aplicando-se diferentes tensões (ψ_i , sucções) a amostras de solo, e observando-se a umidade (θ_{ij}) que permanece na amostra após a aplicação de cada uma das tensões nas n unidades experimentais. Desta forma, a CRA relaciona a variável resposta θ com a variável regressora ψ .

2 Estimação do modelo de van Genuchten (1980) para a representação da CRA

A seguir, apresentamos o modelo de VAN GENUCHTEN (1980), comumente utilizado na modelagem da CRA por sua característica de adequar-se bem a diferentes tipos de solo.

Sejam ψ_{ij} o potencial matricial e $\Theta_{ij} = \frac{\theta_{ij} - \theta_r}{\theta_s - \theta_r}$ a umidade volumétrica normalizada, sendo que θ_s é a umidade saturada, θ_r é a umidade residual e θ_{ij} é a umidade da amostra. O modelo de van Genuchten (1980) é dado pela expressão:

$$\theta_{ij} = \theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_{ij})^n]^m}; \psi_{ij} \geq 0, \quad (1)$$

onde α é um parâmetro relacionado a ψ_{ij}^{-1} e n e m são parâmetros relacionados à inclinação da curva. O número de parâmetros do modelo de van Genuchten (1980) pode ser reduzido usando-se a relação $m = 1 - 1/n$ do modelo de Mualem, ou $m = 1 - 2/n$ do modelo de Burdine (VAN GENUCHTEN, 1980).

Em VAN GENUCHTEN (1980), o autor destaca que a umidade saturada, θ_s , é de fácil obtenção experimental, estando disponível na maioria das vezes. CORNELIS et al. (2005) destacam que θ_r é geralmente definido como o conteúdo de água do solo para o qual a movimentação da água cessa, como o conteúdo de água do solo à tensão de $-15atm$, ou ainda como um parâmetro de ajuste igual ao conteúdo de água onde a primeira derivada de θ em relação a ψ $d\theta/d\psi$ se anula.

Em estudos de determinação da CRA, a expressão analítica descrita em (1) é ajustada considerando-se um problema de regressão não linear do tipo

$$\theta = f(\psi; \theta_s, \theta_r, \beta) + \varepsilon; \varepsilon \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (2)$$

onde β é o vetor de parâmetros de ajuste da curva e ε são os erros aleatórios (erros experimentais) normalmente distribuídos com média zero e variância constante σ^2 . Como θ_s e θ_r são medidos experimentalmente tem-se a estimação de $\beta = (\alpha, n, m)$ ou de $\beta = (\alpha, n)$ usando-se a restrição de Mualem ou de Burdine.

Os problemas de modelagem como em (2) possuem uma vasta literatura, tendo sido amplamente estudados. Exemplos podem ser encontrados em BATES E WATTS (1988), DRAPER E SMITH (1998), GALLANT (1987) e SEBER E WILD (1989), entre outros. Além disso, para estes casos, os estimadores de máxima verossimilhança (EMV) dos parâmetros são iguais aos encontrados pelo método de quadrados mínimos.

No caso do ajuste da CRA, o método mais utilizado para estimação dos parâmetros é o método de quadrados mínimos (MQM) não lineares (DOURADO-NETO et al. (2000), CHAO

et al. (2008), CORNELIS et al. (2005) e SILVA et al., (2006) e YATES et al. (1992)).

Uma das vantagens do uso do MQM é que ele é de fácil aplicação, além de produzir boas estimativas mesmo quando o conjunto de dados disponível é relativamente pequeno, além de possuir uma teoria bem desenvolvida para a computação de intervalos de confiança e predição. Porém, o MQM possui sérias desvantagens que podem comprometer a qualidade da estimação da CRA e das predições feitas a partir da mesma: i) grande sensibilidade à presença de outliers; ii) apresentam propriedades ótimas apenas se os erros de mensuração são não correlacionados com variância constante. Usualmente, a estimação da CRA é apresentada sem nenhum estudo diagnóstico do modelo ajustado, o que pode comprometer sua aplicabilidade uma vez que a estimação dos parâmetros e inferências estatísticas são construídas com base na especificação da estrutura dos erros, ε .

3 Discussões

O uso de modelos na agricultura trouxe consigo uma necessidade de obtenção de informações sobre propriedades físicas do solo, como a condutividade hidráulica. Além disso, há um interesse crescente na busca de técnicas de estimação de características do solo que melhorem o desenvolvimento de modelos ambientais, principalmente modelos de produtividade agrícola. Através de uma análise diagnóstica de curvas de retenção ajustadas a conjuntos de dados reais, nota-se que, nem sempre a suposição de erros experimentais constante é adequada. Neste cenário, o MQM não fornecerá estimadores ótimos, comprometendo a estimação da condutividade hidráulica a ser utilizada no propósito de modelagem ambiental em agricultura irrigada.

O presente trabalho propõe a estimação dos parâmetros da CRA através do método de máxima verossimilhança, levando em conta erros experimentais heteroscedásticos.

4 Resultados

Inicialmente, propomos o uso de procedimentos diagnósticos para verificar se as suposições de independência e homoscedasticidade são válidas. A seguir, apresentamos exemplos de estudos diagnósticos de CRAs ajustadas pelo MQM para dois perfis de solo na profundidade de 0 – 5cm (Figura 1). Os perfis foram selecionados de uma base de dados coletados na área da Bacia Hidrográfica do Rio Buriti Vermelho, localizada na parte leste de Distrito Federal.

Observa-se que há evidência de violação da suposição de homoscedasticidade dos resíduos (Figura 1c e 1d). Desta forma, é razoável considerar estruturas alternativas para a distribuição dos erros na construção dos estimadores dos parâmetros do modelo, como uma tentativa de melhor quantificar as incertezas das estimativas. Por esta razão, o presente trabalho propõe o estudo de métodos alternativos para a estimação da CRA supondo $\varepsilon \sim N_n(\mathbf{0}, \Sigma)$ com as seguintes estruturas para a matriz de covariâncias: $\Sigma = \begin{bmatrix} \sigma_1^2 & \dots & \sigma_1^2 & \sigma_2^2 & \dots & \sigma_2^2 \end{bmatrix} \mathbf{I}_n$, $\Sigma = \sigma^2 \psi^\lambda \mathbf{I}_n$.

Nas Tabelas 1 e 2 apresentamos as estimativas dos parâmetros da CRA obtidos pelo MQM. Os resultados foram obtidos utilizando o *procedure* NLIN do software estatístico SAS.

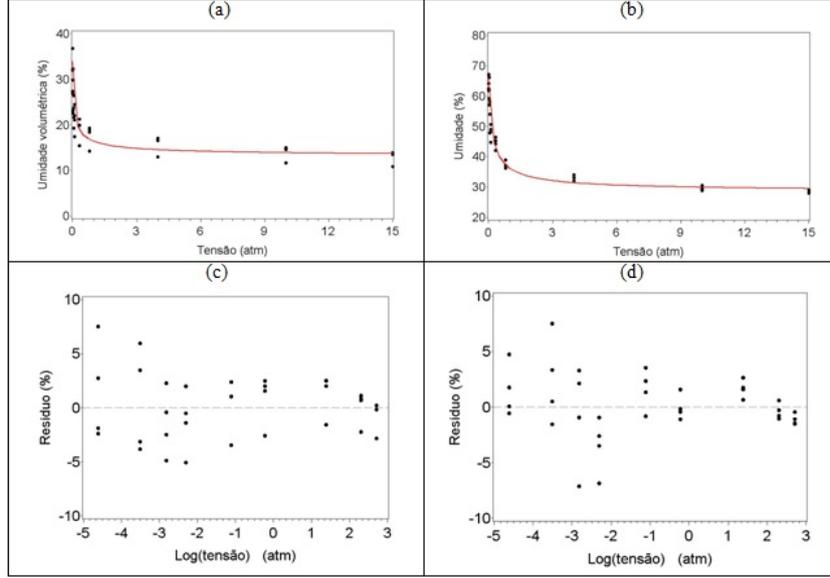


Figura 1: Curvas de retenção ajustadas ((a) e (b)) para dois perfis de solo da Bacia Hidrográfica do Rio Buriti Vermelho, DF e respectivos gráficos de dispersão dos resíduos ((c) e (d))

Tabela 1: Estimativa dos parâmetros via MQM para o perfil (a)

Parâmetros	Estimativa	Erro padrão aproximado	Intervalo de confiança 95% aproximado	
a	654.988	139.085	372.335	937.640
n	16.364	0.0509	15.329	17.398

Tabela 2: Estimativa dos parâmetros via MQM para o perfil (b)

Parâmetros	Estimativa	Erro padrão aproximado	Intervalo de confiança 95% aproximado	
a	202.281	37.833	125.395	279.167
n	17.946	0.052	16.890	19.002

A função de log-verossimilhança considerando $\varepsilon \sim N_n(\mathbf{0}, \Sigma)$, onde $\Sigma = \sigma^2 \psi^\lambda \mathbf{I}_n$ é dada por:

$$l(\alpha, n, \sigma^2, \lambda, \theta, \psi) = \frac{n}{2} \log \frac{1}{2\pi} + \frac{1}{2\sigma^2} \sum_{i=1}^N \log \frac{1}{\psi_i^\lambda} - \frac{1}{2\sigma^2} \sum_{i=1}^N \frac{1}{\psi_i^\lambda} \left[\theta_i - \left(\theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_i)^n]^{(1-1/n)}} \right) \right]^2. \quad (3)$$

Os EMV dos parâmetros são obtidos resolvendo-se numericamente o sistema de equações abaixo e os intervalos de confiança dos parâmetros são calculados assintoticamente.

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \alpha} l(\alpha, n, \sigma^2, \lambda, \theta, \psi) = -\frac{1}{2\sigma^2} \sum_{i=1}^N \frac{2n(\theta_s - \theta_r)(1 - \frac{1}{n})(\alpha\psi_i)^n \left[\theta_i - \left(\theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_i)^n]^{(1-1/n)}} \right) \right]}{\alpha\psi_i^\lambda [1 + (\alpha\psi_i)^n]^{(2-1/n)}} = 0 \\ \frac{\partial}{\partial n} l(\alpha, n, \sigma^2, \lambda, \theta, \psi) = -\frac{1}{2\sigma^2} \sum_{i=1}^N \frac{2(\theta_s - \theta_r) \left[\theta_i - \left(\theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_i)^n]^{(1-1/n)}} \right) \right] \left[\frac{\ln[1 + (\alpha\psi_i)^n]}{n^2} + \frac{(1 - \frac{1}{n})(\alpha\psi_i)^n \ln(\alpha\psi_i)}{1 + (\alpha\psi_i)^n} \right]}{\psi_i^\lambda [1 + (\alpha\psi_i)^n]^{(1-1/n)}} = 0 \\ \frac{\partial}{\partial \sigma^2} l(\alpha, n, \sigma^2, \lambda, \theta, \psi) = -\frac{1}{2(\sigma^2)^2} \sum_{i=1}^N \ln \frac{1}{\psi_i^\lambda} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^N \frac{1}{\psi_i^\lambda} \left[\theta_i - \left(\theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_i)^n]^{(1-1/n)}} \right) \right]^2 = 0 \\ \frac{\partial}{\partial \lambda} l(\alpha, n, \sigma^2, \lambda, \theta, \psi) = -\frac{1}{2\sigma^2} \sum_{i=1}^N \ln \psi_i + \frac{1}{2\sigma^2} \sum_{i=1}^N \frac{\ln \psi_i \left[\theta_i - \left(\theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha\psi_i)^n]^{(1-1/n)}} \right) \right]^2}{\psi_i^\lambda} = 0 \end{array} \right.$$

Referências

- [1] BATES, D.M.; WATTS, D.G. Nonlinear Regression Analysis and Its Applications, 3rd ed., New York: Wiley, 1998.
- [2] CHAO, K.C.; NELSON, J. D.; OVERTON, D. D.; CUMBERS, J. M. Soil water retention curves for remolded expansive soils. *Unsaturated Soils: Advances in Geo-Engineering. Proceedings of the 1st European Conference, E-UNSAT 2008, Durham, United Kingdom, 2-4 July 2008.*
- [3] CORNELIS, W. M.; KHLOSI, M.; HARTMANN, R.; VAN MEIRVENNE, M.; DE VOS, B.; Comparison of unimodal analytical expressions for the soil-water retention curve. *Soil Science Society of America Journal*, vol. 69, p. 1902-1911, 2005.
- [4] DRAPER, N. R.; SMITH, H. Applied regression analysis, 3rd ed., New York: Wiley Series in Probability in Statistics, 1998.
- [5] DOURADO-NETO, D.; NIELSEN, D. R.; HOPMANS, J. W.; REICHARDT, K.; BACCHI, O. O. S. Software to model soil water retention curves (SWRC, version 2.00). *Scientia Agricola*, vol. 57, nº 1, 2000.
- [6] GALLANT, A. R. Nonlinear Statistical Inference, New York: John Wiley and Sons, 1987.
- [7] SEBER, G.A.F.; WILD, C.J. Nonlinear Regression, Wiley: New York, 1989.
- [8] SILVA, E. M.; LIMA, E. F. W.; AZEVEDO, J. A.; RODRIGUES, L. N. Valores de tensão na determinação da curva de retenção de água de solos do Cerrado. *Pesq. Agropec. Bras.*, vol. 41, nº 2, p. 323-330, 2006.
- [9] VAN GENUCHTEN, M.T. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, vol. 44, p. 892-898, 1980.
- [10] YATES, S. R.; VAN GENUCHTEN, M. T.; WARRICK, A. W.; LEIJ, F. J. Analysis of measured, predicted, and estimated hydraulic conductivity using the RETC computer program. *Soil Science Society of America Journal*, vol. 56, p. 347-354, 1992.