

Programa SAS® para ajuste de curvas de retenção

ALINE DE HOLANDA NUNES MAIA⁽¹⁾, LINEU NEIVA RODRIGUES⁽²⁾ DENÍLSON PEREIRA PASSO⁽³⁾

RESUMO – Neste trabalho apresentamos um programa SAS para ajuste de curvas de retenção, informação essencial para simulações da dinâmica da água no solo e seus impactos no rendimento agrícola de culturas. Os parâmetros são estimados utilizando métodos de quadrados mínimos não lineares, via processos iterativos implementados no procedimento NLIN (Proc NLIN) do software estatístico SAS/STAT® do SAS System. Uma das vantagens do uso desse programa, é a possibilidade de quantificar incertezas associadas às estimativas dos parâmetros da curva de retenção e aos valores preditos da umidade volumétrica para cada nível específico de tensão da água no solo. Podem ser ajustadas curvas para um grande número de estratos (ex. classes de solo, profundidades) com geração de arquivos em diversos formatos (ex. planilha Excel), contendo as informações sobre qualidade de ajuste dos modelos, estimativas de parâmetros e suas incertezas para cada estrato.

Palavras-Chave: Van Genuchten, modelos não lineares, métodos computacionais.

Introdução

Curvas de retenção de água no solo se constituem numa informação essencial para manejo eficiente da água em agricultura irrigada e para simulações da dinâmica da água e sua influência na produtividade das culturas. São representadas por modelos não lineares que relacionam potencial matricial com umidade volumétrica, a partir de estimativas obtidas em laboratório. Diversos modelos têm sido propostos para este fim, entre eles os modelos de Rogowski [1] e Van Genuchten [2]. O ajuste de curvas de retenção requer o uso algoritmos de minimização de quadrados mínimos não lineares (Hartley [3]) que envolvem processos iterativos para otimização das estimativas dos parâmetros. Por esse motivo, é necessário o uso de softwares específicos para esse tipo de estimação. Neste trabalho, apresentamos um programa SAS® ajuste do modelo de Van Genuchten e quantificação de incertezas associadas às estimativas dos parâmetros e valores preditos.

Material e Métodos

A. Dados

Como exemplo, foi utilizado um conjunto de dados provenientes de dois perfis de solo (1 e 2), em duas profundidades (0 e 60 cm), coletados na área da Bacia

do Buriti Vermelho, em Brasília, DF. O solo nos locais de coleta é classificado como Latossolo Vermelho.

As variáveis utilizadas na análise foram umidade volumétrica (θ , m³/m³), potencial matricial (Ψ , atm), com três repetições para cada nível de Ψ (0,01; 0,03; 0,06; 0,10; 0,33; 0,80, 4; 10 e 15). Os parâmetros θ_R e θ_S do modelo de Van Genuchten foram considerados fixos e informados para cada repetição. A forma de organização dos dados numa planilha Excel® é apresentada na Tabela 1.

B. Modelo de Van Genuchten

Para representar a relação não linear entre umidade volumétrica e potencial matricial, foi ajustado o modelo de Van Genuchten:

$$\theta = \theta_R + \theta_S + (\theta_S - \theta_R) \left[\frac{1}{1 + (\alpha \Psi_m)^n} \right]^m$$

com $m = 1 - 1/n$ (Mualen [4]), sendo $\theta(\Psi_m)$ a umidade volumétrica correspondente para cada o potencial matricial (Ψ_m); θ_R e θ_S a umidade volumétrica residual e de saturação, respectivamente, e n e α parâmetros de ajuste.

Os parâmetros α e n foram estimados utilizando o método de quadrados mínimos não lineares, via procedimento iterativo de Marquadt [5], implementado no programa SAS descrito na próxima seção.

C. Programa SAS®

O programa CRVG.sas (curva de retenção Van Genuchten) foi desenvolvido em linguagem de programação SAS, utilizando os procedimentos do software estatístico SAS/STAT® (SAS [6]) descritos a seguir:

Importação de dados: para importação dos dados é necessário informar o diretório (caminho), o tipo de arquivo (dbms) e a opção de usar ou não o nome das variáveis do cabeçalho (getnames).

```
proc import out=d1
datafile="C:\caminho\dados.xls"
dbms=excel replace;
getnames=yes;
```

```
run;
```

Impressão do arquivo de dados:

```
proc print data=d1 noobs;
var Perfil Prof Rep Tensao TetaS TetaR
Umidade;
run;
```

⁽¹⁾ Pesquisadora da Embrapa Meio Ambiente, Rodovia Campinas-Mogi Mirim, Km 127,5, Caixa Postal 69, Jaguariúna, SP, CEP 13083-00.

E-mail: ahmaia@cnpma.embrapa.br.

⁽²⁾ Pesquisador da Embrapa Cerrados, BR 020 Km 18, Planaltina, DF, Caixa Postal: 08223, CEP 73310-970. E-mail: Lineu@cpac.embrapa.br

⁽³⁾ Bolsista CNPq, Embrapa Cerrados, BR 020 Km 18, Planaltina, DF, Caixa Postal: 08223, CEP 73310-970. E-mail: geodenilson@gmail.com

Construção de gráficos exploratórios: é útil para identificação preliminar de valores discrepantes ('outliers') e visualização do forma das curvas. As opções vpercent e hpercent controlam o tamanho dos eixos.

```
proc plot data=d1 hpercent=50
vpercent=80;
by perfil prof;
plot umidade*tensao;
run;
```

Ajuste do Modelo: nesse módulo é utilizado o procedimento NLIN para ajuste do modelo de Van Genuchten. Para ajuste de um outro modelo, a equação informada no 'model' deve ser substituída. Em 'parms', são informados valores iniciais para α e n .

```
proc nlin data=d2 ;
by perfil prof;
parms alfa=0.1 n=1;
model Umidade=TetaR + ((TetaS-TetaR)/
(1+(alfa*((abs(tensao)**n))**
(1-(1/n)))));
```

Geração de arquivos com resultados do ajuste do modelo: o código abaixo é utilizado para gerar arquivos contendo valores preditos (pred), resíduos (res) e limites do intervalo de confiança de 95% (LI, LS) para os valores preditos (PREDRES), status de convergência (convergencia) e estimativas dos parâmetros (parâmetros):

```
output out=PREDRES p=pred r=res
stdp=ep_pred l95=LI u95=LS;
ods output
ConvergenceStatus=CONVERGENCIA;

ods output ParameterEstimates=PARAMETROS;
run;
```

Impressão dos arquivos: impressão dos arquivos gerados na janela 'output' do SAS.

Valores observados, preditos e resíduos:

```
proc print data=PREDRES;
run;
```

Status de convergência:

```
proc print data=CONVERGENCIA;
run;
```

Estimativas dos parâmetros e intervalos de confiança:

```
proc print data=PARAMETROS;
run;
```

Exportação de dados: os arquivos gerados no módulo anterior são exportados para planilhas Excel®.

```
proc export data=PREDRES
outfile="C:\caminho\CONVERGENCIA.xls"
DBMS=Excel replace;
run;
```

```
proc export data=convergencia
outfile="C:\caminho\CONVERGENCIA.xls"
DBMS=Excel replace;
run;
```

```
proc export data=PARAMETROS
outfile="C:\caminho\CONVERGENCIA.xls"
DBMS=Excel replace;
run;
```

Construção de gráfico da curva ajustada: utilizando o PROC GPLOT do software gráfico SAS/GRAPH (SAS, [7]), são gerados gráficos das curvas ajustadas com intervalos de confiança para cada perfil e profundidade. Para construção dos gráficos, inicialmente são excluídos os valores preditos das repetições 2 e 3, que são similares aos da repetição 1.

```
data d2;
set predres;
if rep>1 then pred=".";
if rep>1 then li=".";
if rep>1 then ls=".";
```

A seguir, são informadas as opções do gráfico, tais como tipo de linha e símbolo, legendas, tipo de interpolação, entre outras. Para maiores detalhes, ver o manual do PROC GPLOT (SAS [6]).

Opções gerais:

```
goptions ftext='Arial' htext=4 gunit=pct;
```

Eixo x:

```
axis1 label=(a=90 'Umidade volumétrica
(m3/m3)') order=0 to 0.6 by 0.1 minor=none;
```

Eixo y:

```
axis2 label=('Potencial matricial (atm)')
offset=(1,1)order=0 to 15 by 5 minor=none
length=65;
```

Tabela 2. Estimativas dos parâmetros n e α do modelo, com respectivos limites de confiança (LI, LS) para cada perfil e profundidade do solo.

Perfil	Profundidade (cm)	Parametro	Estimativa	Erro Padrão	LI	LS
1	0	alfa	86,6068	16,6202	52,3771	120,8364
1	0	n	1,7154	0,0501	1,6121	1,8186
1	60	alfa	50,6947	7,8554	34,5164	66,8730
1	60	n	1,8499	0,0448	1,7576	1,9422
2	0	alfa	55,5906	7,6074	39,9230	71,2583
2	0	n	1,7103	0,0351	1,6380	1,7826
2	60	alfa	51,1304	13,1368	24,0749	78,1859
2	60	n	1,8864	0,0764	1,7291	2,0438

Tabela 3. Valores observados e preditos da umidade volumétrica para cada nível de tensão com respectivos limites de confiança (LI, LS) para o perfil 1 e profundidade zero.

Perfil	Prof	Rep	Tensao	TetaS	TetaR	Umidade	preditos	LI	LS	res
1	0	1	0,01	0,5417	0,2701	0,5190	0,4894	0,4570	0,5218	0,0296
1	0	1	0,03	0,5417	0,2701	0,4421	0,4484	0,4162	0,4806	-0,0064
1	0	1	0,06	0,5417	0,2701	0,4021	0,4162	0,3843	0,4481	-0,0141
1	0	1	0,1	0,5417	0,2701	0,3836	0,3915	0,3596	0,4233	-0,0079
1	0	1	0,33	0,5417	0,2701	0,3528	0,3396	0,3074	0,3718	0,0132
1	0	1	0,8	0,5417	0,2701	0,3313	0,3120	0,2800	0,3440	0,0193
1	0	1	4	0,5417	0,2701	0,3026	0,2849	0,2538	0,3161	0,0177
1	0	1	10	0,5417	0,2701	0,2872	0,2780	0,2472	0,3088	0,0092
1	0	1	15	0,5417	0,2701	0,2749	0,2760	0,2453	0,3068	-0,0012
1	0	2	0,01	0,5417	0,2701	0,4855	0,4894	0,4570	0,5218	-0,0039
1	0	2	0,03	0,5417	0,2701	0,4509	0,4484	0,4162	0,4806	0,0025
1	0	2	0,06	0,5417	0,2701	0,4032	0,4162	0,3843	0,4481	-0,0130
1	0	2	0,1	0,5417	0,2701	0,3732	0,3915	0,3596	0,4233	-0,0182
1	0	2	0,33	0,5417	0,2701	0,3377	0,3396	0,3074	0,3718	-0,0019
1	0	2	0,8	0,5417	0,2701	0,3181	0,3120	0,2800	0,3440	0,0060
1	0	2	4	0,5417	0,2701	0,2900	0,2849	0,2538	0,3161	0,0051
1	0	2	10	0,5417	0,2701	0,2788	0,2780	0,2472	0,3088	0,0008
1	0	2	15	0,5417	0,2701	0,2685	0,2760	0,2453	0,3068	-0,0076
1	0	3	0,01	0,5417	0,2701	0,5271	0,4894	0,4570	0,5218	0,0377
1	0	3	0,03	0,5417	0,2701	0,4632	0,4484	0,4162	0,4806	0,0148
1	0	3	0,06	0,5417	0,2701	0,3878	0,4162	0,3843	0,4481	-0,0284
1	0	3	0,1	0,5417	0,2701	0,3859	0,3915	0,3596	0,4233	-0,0056
1	0	3	0,33	0,5417	0,2701	0,3453	0,3396	0,3074	0,3718	0,0057
1	0	3	0,8	0,5417	0,2701	0,3221	0,3120	0,2800	0,3440	0,0100
1	0	3	4	0,5417	0,2701	0,2921	0,2849	0,2538	0,3161	0,0072
1	0	3	10	0,5417	0,2701	0,2785	0,2780	0,2472	0,3088	0,0005
1	0	3	15	0,5417	0,2701	0,2669	0,2760	0,2453	0,3068	-0,0091

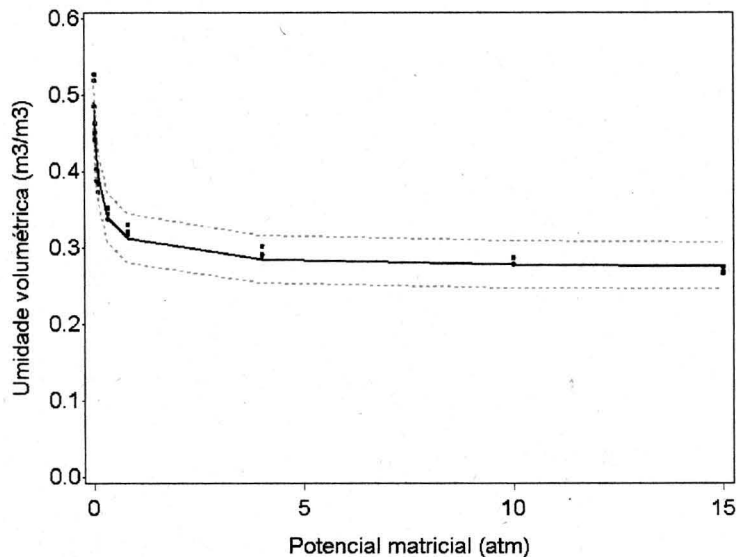


Figura 1. Modelo ajustado para o perfil=1 e profundidade=0, com intervalo de confiança de 95% para os valores preditos da umidade volumétrica.

Tabela 1. Organização dos dados em planilha Excel para uso no ajuste das curvas de retenção.

Perfil	Prof	Rep	Tensao	TetaS	TetaR	Umidade
1	0	1	0,01	0,541689	0,270091	0,5190
1	0	1	0,03	0,541689	0,270091	0,4421
1	0	1	0,06	0,541689	0,270091	0,4021
1	0	1	0,10	0,541689	0,270091	0,3836
1	0	1	0,33	0,541689	0,270091	0,3528
1	0	1	0,80	0,541689	0,270091	0,3313
1	0	1	4,00	0,541689	0,270091	0,3026
1	0	1	10,00	0,541689	0,270091	0,2872
1	0	1	15,00	0,541689	0,270091	0,2749
1	0	2	0,01	0,541689	0,270091	0,4855
1	0	2	0,03	0,541689	0,270091	0,4509
1	0	2	0,06	0,541689	0,270091	0,4032
1	0	2	0,10	0,541689	0,270091	0,3732
1	0	2	0,33	0,541689	0,270091	0,3377
1	0	2	0,80	0,541689	0,270091	0,3181
1	0	2	4,00	0,541689	0,270091	0,2900
1	0	2	10,00	0,541689	0,270091	0,2788
1	0	2	15,00	0,541689	0,270091	0,2685
1	0	3	0,01	0,541689	0,270091	0,5271
1	0	3	0,03	0,541689	0,270091	0,4632
1	0	3	0,06	0,541689	0,270091	0,3878
1	0	3	0,10	0,541689	0,270091	0,3859
1	0	3	0,33	0,541689	0,270091	0,3453
1	0	3	0,80	0,541689	0,270091	0,3221
1	0	3	4,00	0,541689	0,270091	0,2921
1	0	3	10,00	0,541689	0,270091	0,2785

Tipos de símbolos, cores, linhas e interpolações:

Dados observados:

```
symbol1 interpol=none v=dot h=1 w=2  
c=black;
```

Modelo ajustado:

```
symbol2 interpol=line v=none line=1 h=2  
v=none w=2 c=black;
```

Limite inferior do intervalo de confiança:

```
symbol3 interpol=line v=none line=2 h=2  
v=none w=2 c=grey;
```

Limite superior do intervalo de confiança:

```
symbol4 interpol=line v=none line=2 h=2  
v=none w=2 c=grey;
```

Com os comandos abaixo, o arquivo é ordenado para geração dos gráficos por perfil e profundidade:

```
proc sort data=d2;  
by perfil prof tensao;  
run;
```

Utilizando o Proc GPLOT, os gráficos são gerados de acordo com as especificações anteriores:

```
proc gplot data=d2;  
by perfil prof;  
plot (umidade pred li ls)*tensao/overlay  
vaxis=axis1 haxis=axis2;  
run;
```

Resultados

Na Tabela 2, são apresentadas as estimativas dos parâmetros n e α do modelo, com respectivos intervalos de confiança, para cada perfil e profundidade, geradas pelo *Output Delivery System*. Uma planilha com os valores observados e preditos da umidade volumétrica para cada nível de tensão é apresentada na Tabela 3. Na Figura 1, observa-se a curvas de retenção ajustadas para o perfil 1 e profundidade zero, gerado pelo procedimento GPLOT.

Discussão

A principal vantagem do uso do programa CRVG.sas é a quantificação das incertezas associadas às estimativas dos parâmetros da curva de retenção e às estimativas das umidades volumétricas correspondentes a valores fixos de potencial matricial. As incertezas dependem da variabilidade entre os valores de θ de amostras submetidas à mesma tensão (repetições) e da qualidade de ajuste do modelo, expresso em termos do

quadrado da correlação entre valores observados e preditos.

O ajuste simultâneo de modelos para os diferentes estratos via comando BY é um procedimento rápido e eficiente quando se deseja estimar curvas um grande número de locais e profundidades. O sistema para seleção e exportação da informação gerada pelo PROC NLIN (*Output Delivery System*) para arquivos externos de diferentes formatos (ex. planilha Excel, arquivo tipo txt ou dat) facilita seu uso como entrada para modelos de simulação.

Outra função útil do programa CRVG.sas é a geração de um arquivo com informação sobre o *status* de convergência de cada modelo, útil para identificar prováveis erros no registro dos dados ou indicar modelos alternativos, via posterior análise gráfica dos casos onde não houve do algoritmo de estimação dos parâmetros. O programa pode ser facilmente generalizado para ajuste de outros modelos de curvas de retenção.

Conclusões

O programa apresentado é útil para obtenção de estimativas dos parâmetros do modelo de Van Genuchten e valores preditos com incertezas associadas, informação importante, por exemplo, para uso em modelos de simulação que incorporam incertezas dos dados de entrada, para geração de previsões probabilísticas.

Agradecimentos

Ao *Challenge Program on Water and Food*, CNPq e Embrapa pelo apoio financeiro aos projetos de pesquisa.

Referências

- [1] Rogowski, A.S. Watershed physics: model of soil moisture characteristics. *Water Resources Research*, v.7, p.1575-1582. 1971.
- [2] van Genuchten, M.T. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44:892-898, 1980.
- [3] Hartley, H.O. The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares. *Technometrics*, 3: 269-280, 1961
- [4] Mualem, Y. A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resource Research*, 12:513-522, 1976.
- [5] Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters." *Journal for the Society of Industrial and Applied Mathematics*, 11:431-441, 1963.
- [7] SAS Institute Inc, *SAS/STAT® User's Guide*, Version 9.1, v.1-7. SAS Institute Inc., Cary, NC, 2004.
- [8] SAS institute Inc, *SAS/GRAPH Software: Reference*, Version 8/. SAS Institute Inc., Cary, NC, 2004.