

## PREVISÃO DA TRANSMISSIVIDADE HIDRÁULICA USANDO ALGORITMO FLORESTA ALEATÓRIA

Paula Josyane dos Santos Francisco<sup>1</sup>; Juliana Pertille da Silva<sup>2</sup>; Annelise de Almeida Vetromile Lapuente dos Santos<sup>3</sup>; Henrique Sanchez Franz<sup>4</sup>; Uuelle San Martins<sup>5</sup>; Janice Bilhalva de Bilhalva<sup>6</sup>

<sup>1</sup>*Universidade Federal de Pelotas – paula.josyane.eng@gmail.com*

<sup>2</sup>*Universidade Federal de Pelotas – juliana.pertill@gmail.com*

<sup>3</sup>*Universidade Federal de Pelotas – annelise.lapuente@gmail.com*

<sup>4</sup>*Universidade Federal de Pelotas – franzhenrique@yahoo.com.br*

<sup>5</sup>*Universidade Federal de Pelotas – uielsm@gmail.com*

<sup>6</sup>*Universidade Federal de Pelotas – evandro\_gesso@hotmail.com*

### 1. INTRODUÇÃO

As águas subterrâneas possuem um papel fundamental no abastecimento das cidades brasileiras. Segundo a ANA (2021), 40% das sedes urbanas brasileiras dependem exclusivamente de águas subterrâneas para seu abastecimento. Nesse sentido, é de extrema importância caracterizar os sistemas aquíferos explotados a fim de entender e quantificar o volume de água neles contidos.

Os poços tubulares são obras hidráulicas, que além de permitirem a retirada de água dos aquíferos, permitem a observação e compreensão da geologia na qual esse poço está inserido e mensurar a produtividade de um determinado aquífero através do teste de produção, que também pode ser chamado de teste de bombeamento (FEITOSA, 2008). A partir de metodologias de análise de fluxo, é possível estimar características como transmissividade e coeficiente de armazenamento. Os métodos de análise podem variar principalmente de acordo com a condição de confinamento e o tipo de regime de escoamento. Esses métodos de análise unidos a técnicas de previsão, vinculadas a inteligência artificial, podem garantir a espacialização das características de Transmissividade e Coeficiente de armazenamento.

Nesse trabalho foram utilizados o método de Theis (THEIS, 1935) para obtenção da transmissividade e o algoritmo *RandomForestRegressor* (BREIMAN, 2001) para previsão da transmissividade em 12 diferentes poços no perímetro urbano da cidade de Alegrete, no estado do Rio Grande do Sul, compreendendo a porção confinada do Sistema Aquífero Guarani.

### 2. METODOLOGIA

A área de estudo, na qual foi aplicada essa metodologia, está localizada na porção confinada gaúcha do Sistema Aquífero Guarani. Esse sistema aquífero se estende por quatro países (Argentina, Brasil, Paraguai e Uruguai) e no território brasileiro ocupa oito estados (Mato Grosso, Mato Grosso do Sul, Goiás, Minas Gerais, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul).

A primeira etapa envolveu a obtenção dos dados dos testes de bombeamento poços do perímetro urbano de Alegrete. A segunda e a terceira etapa foram realizadas utilizando exclusivamente a linguagem de programação *Python* e a IDE (*Integrated Development Environment*) *Visual Studio Code*.

A segunda etapa incluiu o ajuste da curva padrão do método de Theis aos dados de campo, que se referem aos tempos, em min, e os rebaixamentos, em m.

Alguns dos poços já possuíam as curvas e outros não, dessa forma optou-se por ajustar a curva para todos os poços. Na Figura 1 é apresentado o gráfico referente ao ajuste do poço1 (COR ALG 02). As variáveis utilizadas baseiam-se nas variáveis utilizadas por Dashti et al. (2023). A Tabela 1 apresenta o conjunto de variáveis preditoras utilizadas para prever a transmissividade e medidas estatísticas calculadas. Todas as variáveis contidas na Tabela 1 foram obtidas através dos testes de bombeamento e Si, Sm e Sf são as declividades obtidas através da divisão dos testes de bombeamento em três períodos: inicial (i), médio (m) e final (f). As outras variáveis preditoras se referem a características de construção dos poços ou do aquífero explotado, são elas: Latitude, Longitude, Vazão ( $m^3/h$ ), espessura do aquífero explotado (m) e profundidade (m) do poço. As variáveis preditoras  $t_{\text{init}}$ ,  $t_{\text{end}}$ ,  $s_{\text{ini}}$  e  $s_{\text{end}}$  referem-se ao início e ao fim do teste de bombeamento. A terceira etapa envolveu a aplicação do algoritmo *RandomForestRegressor* (ou Floresta Aleatória) da biblioteca *pandas* para previsão da transmissividade. Ainda foram calculadas as métricas  $R^2$  e a importância das variáveis, essa última indica quais variáveis preditoras obtiveram o maior impacto na previsão da transmissividade.

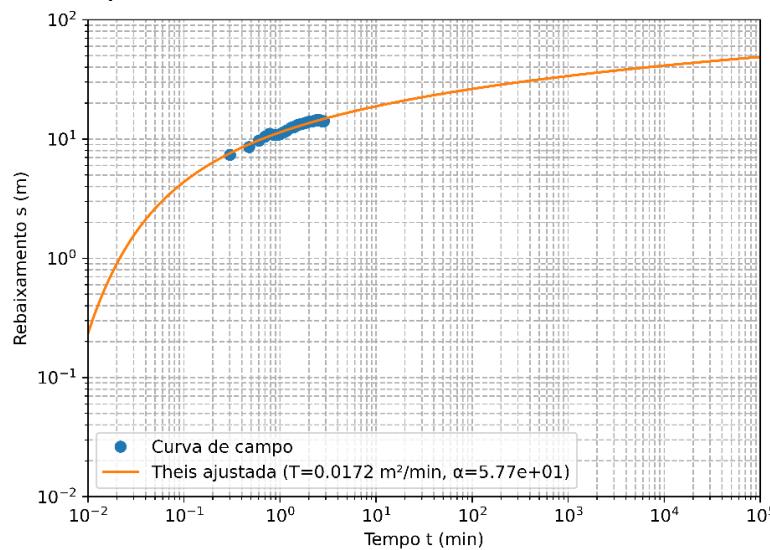


Figura 1 – Curva padrão de Theis e curva de campo do poço COR ALG 02.

Tabela 1 – Variáveis preditoras e transmissividade calculada.

Poço	X	Y	Q ( $m^3/h$ )	Esp. (m)	Prof. (m)	Si	Sm	Sf	$t_{\text{ini}}$	$t_{\text{end}}$	$s_{\text{ini}}$ (m)	$s_{\text{end}}$ (m)	$T_c$ ( $m^2/\text{min}$ )
Poço 1 (C.A. 02)	670175,2	614986,1	42,3	20,0	98,0	7,7	7,8	1,7	1,0	720,0	4,9	14,1	1,72E-02
Poço 2 (C.A. 07)	670175,2	614995,8	26,6	16,0	118,0	12,3	1,2	-1,6	1,0	720,0	7,3	24,0	5,28E-03
Poço 3 (C.A. 09)	670382,2	616264,2	128,5	2,0	137,0	5,3	3,5	1,0	1,0	720,0	7,9	16,1	5,26E-02
Poço 5 (C.A. 11C)	670250,8	615796,2	360,0	84,0	360,0	2,3	2,2	1,6	1,0	1440,0	20,2	26,4	1,73E-01
Poço 6 (C.A. 14A)	670375,8	616002,5	50,0	16,0	179,0	7,9	2,3	0,1	1,0	720,0	6,1	14,9	1,50E-02
Poço 7 (C.A. 15B)	670317,3	619998,4	40,0	43,0	195,0	23,0	2,0	0,8	1,0	720,0	3,8	18,4	2,31E-02
Poço 8 (C.A. 20A)	670473,0	617414,5	370,0	234,0	354,0	51,9	221,9	546,4	1,0	1440,0	14,0	582,7	2,99E-04
Poço 9 (C.A. 22)	67024,0	621440,0	30,0	2,0	126,0	15,5	3,0	1,0	1,0	720,0	6,6	23,8	9,18E-03
Poço 10 (C.A. 23)	670175,8	620659,8	34,0	6,0	135,0	15,8	2,9	0,5	1,0	720,0	7,5	16,9	3,14E-02
Poço 11 (C.A. 24)	670105,8	61457,3	10,0	6,0	66,0	6,1	0,8	-0,7	1,0	720,0	4,5	9,0	1,73E-02
Poço 12 (C.A. 25)	670239,8	618917,1	48,0	51,0	156,0	4,7	1,4	-0,8	1,0	720,0	18,1	23,2	3,26E-02

Poço 13 (C.A. 28)	670593,9	617349,9	90,0	58,0	264,0	16,1	12,1	4,2	1,0	720,0	47,6	68,2	1,46E-02
Máximo	670593,9	621440,0	370,0	234,0	360,0	51,9	221,9	546,4	1,0	1440,0	47,6	582,7	1,73E-01
Mínimo	67024,0	61457,3	10,0	2,0	66,0	2,3	0,8	-1,6	1,0	720,0	3,8	9,0	2,99E-04
Média	620024,1	571273,5	102,5	44,8	182,3	14,1	21,8	46,2	1,0	840,0	12,4	69,8	3,26E-02
Desvio Padrão	166735,8	153729,7	121,2	62,3	91,7	12,8	60,4	150,8	0,0	268,3	11,8	155,3	4,44E-02
CV (%)	26,9	26,9	118,3	138,8	50,3	91,1	277,7	326,5	0,0	31,9	95,4	222,5	1,36E+02
Assimetria	-3,5	-3,5	1,8	2,6	1,0	2,3	3,5	3,5	-	2,1	2,5	3,4	2,94E+00

Notas: C. A. – COR ALG; Esp. – Espessura; Prof. – Profundidade; Tc – Transmissividade calculada.

### 3. RESULTADOS E DISCUSSÃO

O primeiro resultado refere-se ao ajuste da curva padrão a curva de campo, obtendo assim o valor de transmissividade calculada contidos na Tabela 1. Os valores de transmissividade prevista estão contidas na Tabela 2, exibindo também os resíduos. Os valores previstos, pelo uso do algoritmo Floresta aleatória, apresentam uma sutil diminuição em relação ao valor calculado, sendo 0,40 vezes menor no Poço 12 (COR ALG 25), 0,86 vezes menor no Poço 11 (COR ALG 24), 0,86 vezes menor no Poço 1 (COR ALG 02, 0,31 vezes menor no Poço 10 (COR ALG 23), 0,42 vezes menor no Poço 7 (COR ALG 15B) e 0,28 vezes menor no Poço 3 (COR ALG 09), do que o valor calculado ( $T_{\text{teste}}$ ).

O coeficiente de determinação ( $R^2$ ) para o conjunto de dados testados foi igual a -1,77, provando um desbalanceamento entre o erro quadrático médio e a variância dos dados testados. Além dessas métricas, o tamanho do conjunto de dados impacta negativamente no  $R^2$ , o que faz com que essa métrica considere apenas seis valores de transmissividade para o treinamento e seis para o teste.

Tabela 2 – Dados de transmissividade no teste, na previsão e resíduos.

Poço	$T_{\text{teste}}$ (m <sup>2</sup> /min)	$T_{\text{pred}}$ (m <sup>2</sup> /min)	Resíduo (m <sup>2</sup> /min)
Poço 12 (COR ALG 25)	0,0326	0,0130	-0,0195
Poço 11 (COR ALG 24)	0,0173	0,0150	-0,0023
Poço 1 (COR ALG 02)	0,0172	0,0149	-0,0024
Poço 10 (COR ALG 23)	0,0314	0,0098	-0,0216
Poço 7 (COR ALG 15B)	0,0231	0,0098	-0,0133
Poço 3 (COR ALG 09)	0,0526	0,0150	-0,0376

Com relação a importância das variáveis, a declividade do trecho inicial do teste de bombeamento e a profundidade do poço são as variáveis preditoras que mais impactam a previsão da transmissividade. Parte desse valor de importância de profundidade, deve-se a determinação analítica, sem considerar métodos de análise de fluxo, multiplicando a condutividade hidráulica pela espessura do aquífero (FETTER, 2021).

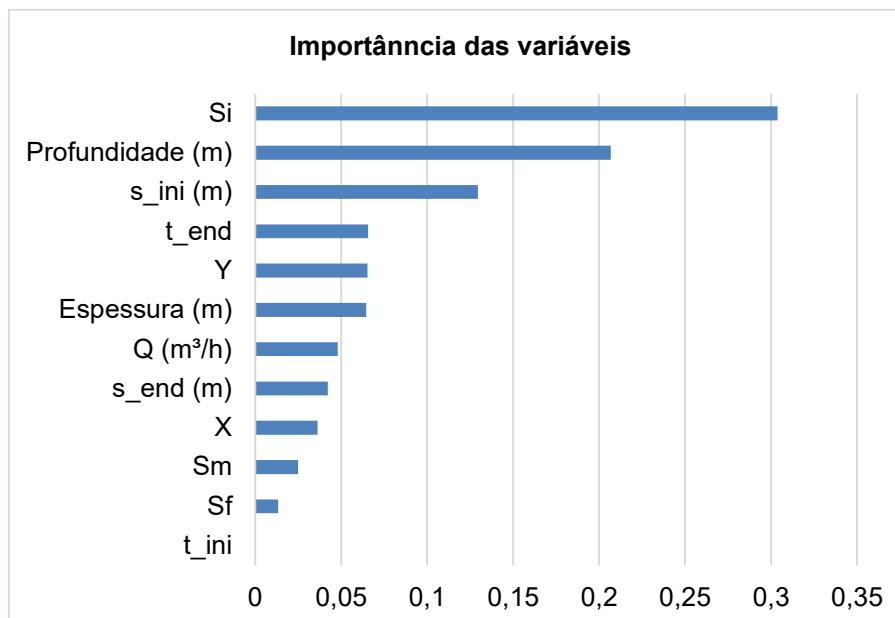


Figura 2 – Importância das variáveis do algoritmo floresta aleatória.

#### 4. CONCLUSÕES

O algoritmo floresta aleatória, nesse caso não executou boas previsões, considerando as variáveis preditoras e o tamanho do conjunto de dados. Em previsões espaciais, que consideram as coordenadas, com o objetivo de espacialização da transmissividade, não se obteve um valor aceitável de coeficiente de determinação. Sendo assim, são recomendados a utilização de algoritmos simplificados, mais poços e consequentemente mais teste de bombeamento para melhor prever a transmissividade.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

ANA (Agência Nacional de Águas e Saneamento Básico). **Atlas águas: segurança hídrica do abastecimento urbano**. Brasília: ANA, 2021. 332 p.

BREIMAN, L. (2001). Random forests. **Machine Learning**, 45, 5–32. Kluwer Academic Publishers. <https://doi.org/10.1023/A:1010933404324>

Dashti, Z., Nakhaei, M., Vadiati, M., Karami, G. H., & Kisi, O. (2023). Estimation of unconfined Aquifer Transmissivity using a comparative study of machine learning models. **Water Resources Management**, 37(12), 4909-4931.

FEITOSA, Fernando A. Carneiro, et al. **Hidrogeologia: conceitos e aplicações**. CPRM, 2008.

FETTER, Charles Willard; KREAMER, David. **Applied hydrogeology**. Waveland Press, 2021.

THEIS, Charles V. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage. **Eos, Transactions American Geophysical Union**, v. 16, n. 2, p. 519-524, 1935.