

Um modelo de regressão linear multivariado para a permeabilidade na região da Oceania

Alesson Mansur Torres Kopp¹

Programa de Pós-Graduação em Ciências Computacionais, UERJ, Rio de Janeiro, RJ

Alisson da Silva Pinto²

Programa de Pós-Graduação em Ciências Computacionais, UERJ, Rio de Janeiro, RJ

Antônio José Boness dos Santos³

Departamento de computação científica, UFPB, João Pessoa, PB

Cristiane Oliveira de Faria⁴

Instituto de Matemática e Estatística, UERJ, Rio de Janeiro, RJ

Jhoab Pessoa de Negreiros⁵

Programa de Pós-Graduação em Ciências Computacionais, UERJ, Rio de Janeiro, RJ

João Paulo Alves Barros⁶

UERJ, Rio de Janeiro, RJ

Resumo. Na avaliação de reservatórios de exploração de petróleo, as propriedades permeabilidade e porosidade exercem uma forte influência na caracterização e predição dos poços petrolíferos. Por isso, determinar esta relação entre as variáveis requer a utilização de técnicas preditivas robustas. Neste trabalho é proposto uma nova abordagem para a obtenção de valores da permeabilidade através de modelos regressivos multivariados. A permeabilidade é estimada em função de medidas da porosidade e da região onde estão localizados os poços. Esta abordagem foi construída e testada com dados disponibilizados no relatório petrolífero da região da Oceania. Para validação da função encontrada para a permeabilidade, esta foi aplicada no problema de Darcy, e a pressão obtida numericamente pelo método de Galerkin.

Palavras-chave. Regressão linear, Meios porosos altamente heterogêneos, Elementos Finitos, Problema de Darcy, Estudo da Permeabilidade

1 Introdução

Em aplicações de meios porosos altamente heterogêneos, permeabilidade e porosidade são propriedades cruciais para o planejamento de transporte e armazenamento em reservatórios de petróleo [1]. Uma estimativa refinada para a permeabilidade é fundamental para aprimorar o processo de extração de petróleo e captura de CO_2 , a seleção do método de recuperação com melhor custo-benefício e escolha da localização dos poços, dentre outras finalidades. Uma medida adequada para a permeabilidade pode ser obtida experimentalmente através de registros coletados de amostras de poços, sendo fundamental a utilização de um modelo eficaz para caracterizar medidas

¹alesson.mansur@pos.ime.uerj.br

²alisson.pinto@pos.ime.uerj.br

³ajboness@gmail.com

⁴cofaria@ime.uerj.br

⁵jhoab.negreiros@pos.ime.uerj.br

⁶jpab_7@hotmail.com

de permeabilidade na avaliação de reservatórios. Diversas propostas eficientes para previsão de permeabilidade e porosidade foram desenvolvidas para avaliação em meios porosos homogêneos, como em [1, 3, 5, 6]. E devido todos os recentes desafios na indústria petrolífera para estimar a permeabilidade em meios altamente heterogêneos, novas vertentes estão sendo desenvolvidas, destacando-se as técnicas de regressão multivariada e os métodos baseados no conceito de *Machine Learning* [1, 6]. Os recentes avanços computacionais, como o poder de processamento dos computadores e a capacidade de tratar um grande volumes de dados, viabilizam que estas novas abordagem sejam exploradas. No entanto, em particular, surgem desafios de lidar com as técnicas tradicionais e realizar modificações necessárias para o tratamento de novos problemas [9].

Neste trabalho, é proposta uma metodologia baseada no método de Regressão Linear para realizar estimativas para a permeabilidade em função da porosidade e das coordenadas dos poços. Inicialmente, esta nova estratégia para obtenção de medidas de permeabilidade é utilizada para obter a porosidade em função da localização e, em seguida, os valores para a permeabilidade. Esta abordagem foi construída e testada com dados petrolíferos da região da Oceania coletados e disponibilizados no relatório elaborado pelo serviço geológico e de recursos minerais australiano [2], e para validação do modelo para a permeabilidade, a partir do sistema de Darcy, a pressão hidráulica p foi encontrada utilizando o método de Galerkin.

Este trabalho é estruturado como segue: na Seção 2 é apresentado o problema de Darcy e definidas as variáveis. A Seção 3 é dedicada aos modelos de regressão utilizados para a obtenção de valores para a permeabilidade K . Os experimentos numéricos para validação na nova metodologia para obtenção da permeabilidade está na Seção 4 e por fim a Conclusão encontrada com estes testes.

2 Problema modelo

Nesta seção formula-se o problema de Darcy que modela o escoamento de um fluido em um meio poroso baseado em [4, 8].

Seja o domínio $\Omega \subset \mathbb{R}^2$ e cuja fronteira $d\Omega = \Gamma_D \cup \Gamma_N$ é dada por Γ_D (condição de contorno de Dirichlet) e por Γ_N (condição de contorno de Neumann). Conhecidas as funções f , \bar{h} e o tensor de permeabilidade \mathbf{K} , desprezando os efeitos gravitacionais, a pressão hidráulica $p : \Omega \rightarrow \mathbb{R}$ e o campo de velocidades $\mathbf{u} : \Omega \rightarrow \mathbb{R}^2$ são dados por

$$\mathbf{u} = -\mathbf{K}\nabla p \quad \text{em } \Omega, \tag{1}$$

$$\text{div } \mathbf{u} = f \quad \text{em } \Omega, \tag{2}$$

com as condições de contorno $\mathbf{u} \cdot \mathbf{n} = \bar{h}$ em Γ_N , $p = \bar{p}$ em Γ_D .

O problema de Darcy pode ser representado através de um problema de Poisson substituindo (1) na equação de conservação da massa (2), obtendo uma equação apenas para o campo da pressão:

$$-\text{div}(\mathbf{K}\nabla p) = f \quad \text{em } \Omega, \tag{3}$$

com condições de contorno $-\mathbf{K}\nabla p \cdot \mathbf{n} = \bar{h}$ sobre Γ_h , $p = \bar{p}$ sobre Γ_p . Neste caso, a velocidade é calculada em um segundo momento com a (2) e a solução p encontrada.

3 Um Modelo de Regressão para a Permeabilidade

3.1 Regressão Linear Multivariável

Em muitos casos, a contribuição de uma única variável independente não é suficiente para explicar a variável dependente Y . Quando isso acontece, é possível executar uma regressão linear

multivariável (ou multivariada ou múltipla) para estudar os efeitos das múltiplas variáveis na variável dependente [9]. No modelo de regressão multivariada, a variável dependente Y é descrita como uma função linear das variáveis independentes x_i , na forma

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n. \tag{4}$$

O modelo permite que os coeficientes de regressão b_i sejam calculados para cada variável independente X_i , e como no caso de uma variável, o coeficiente de determinação \mathbf{R}^2 descreve a relação geral entre as variáveis dependentes e independentes [9]. De acordo com Montgomery, em [7], é mais conveniente representar os modelos de regressão múltipla em forma matricial, de maneira que a equação (4) é reescrita como um sistema visto na equação (5):

$$\mathbf{Y} = \mathbf{XB}, \tag{5}$$

onde:

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \text{ e } \mathbf{B} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix}.$$

A solução do sistema (5) retorna os coeficientes e a constante da regressão. Maiores detalhes sobre o modelo podem ser obtidos em [7].

3.2 Metodologia para obtenção da permeabilidade

O relatório técnico apresentado em [2] possui dados de 551 poços na região da Austrália e Papua-Nova Guiné (Oceania). A partir destes dados é proposto construir uma relação matemática adequada que represente a permeabilidade em todo o domínio definido.

Isso conduz a escrever uma regressão linear para a permeabilidade em termos das medidas Longitude, Latitude e Porosidade encontradas em [2]. No entanto, a porosidade também é uma função que depende da posição geométrica, o que leva a propor uma regressão linear para a porosidade em termos das medidas Longitude e Latitude. Dado que para a Permeabilidade e a Porosidade foram encontradas ambas por regressão linear, combinando as duas funções encontra-se uma função para a permeabilidade que depende apenas das coordenadas.

Assim, a metodologia implementada para a obtenção da função da permeabilidade é descrita como segue:

- 1 Obter os coeficientes de regressão linear para a Porosidade em termos da Longitude e da Latitude, resultando uma relação linear

$$\alpha = a.Longitude + b.Latitude + c_1$$

- 2 Obter os coeficientes de regressão linear para a Permeabilidade em termos da Porosidade, Longitude e da Latitude, resultando em

$$K = d.Longitude + e.Latitude + f.Porosidade + c_2$$

- 3 Substituir a regressão linear da Porosidade na função da Porosidade, resultando em uma relação que dependerá apenas da Latitude e da Longitude.

$$K = A.Longitude + B.Latitude + C$$

onde $A = (d + fa)$, $B = (e + fb)$ e $C = (c_2 + fc_1)$.

Observação: A vantagem de se obter uma função que depende apenas das coordenadas, permitindo encaixá-la facilmente no código implementado de elementos finitos para o problema de Darcy. Se a permeabilidade fosse considerada dependente da porosidade, seria necessário alterar os dados de entrada e armazenar estes valores.

4 Experimentos Numéricos

4.1 Tratamento dos dados

A definição da região de aplicação para a obtenção da permeabilidade e da pressão hidráulica seguiu as coordenadas disponibilizadas em [2]. Na Figura 1 é possível ver um mapa de onde as leituras foram realizadas e que estão sendo representadas pelo símbolo '+/.

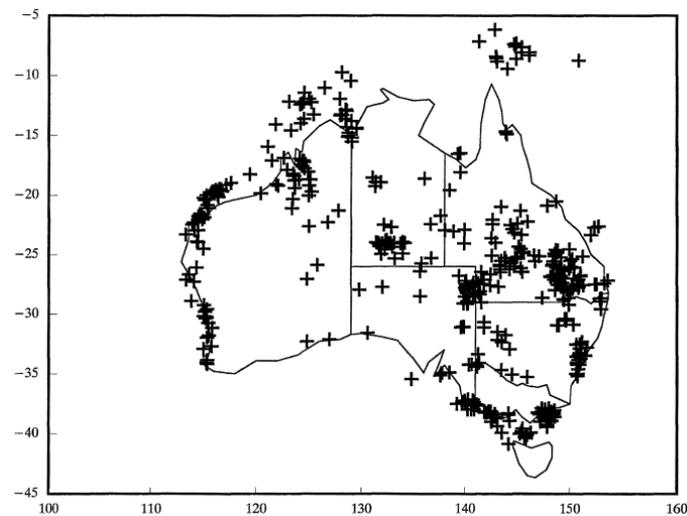


Figura 1: Imagem da região de aplicação do modelo. Retirado de [2]

Em [2] os dados são apresentados de forma tabular em um arquivo .pdf. Por essa razão a coleta e organização deles aconteceu manualmente, com auxílio de um editor de texto. E, para melhor visualização das informações, elas foram guardadas num banco de dados relacional MySQL.

Foram definidas as seguintes regras de exclusão: poços sem nenhuma informação, não constar longitude e latitude e não conter profundidade e porosidade. Com isso, dos 6278 dados disponíveis, restaram 2060 após a aplicação dos critérios. As regressões foram implementadas em linguagem Python com a biblioteca *statsmodels*⁷.

4.2 Regressões Lineares para Permeabilidade

Os coeficientes obtidos para as regressões lineares da porosidade e da permeabilidade são mostrados na Tabela 1. A regressão para a porosidade apresentou $R^2 = 0.018$, e a da permeabilidade $R^2 = 0.098$. Esses valores estão dentro da nossa expectativa, uma vez que não existe, de fato, uma relação linear entre as variáveis independentes e dependentes.

Na Figura 2 é apresentado o gráfico da regressão linear para a Porosidade e para a Permeabilidade em função da Latitude e Longitude. Os pontos vermelhos representam a nuvem de dispersão

⁷<https://www.statsmodels.org/>

Tabela 1: Coeficientes de Regressão Linear.

	Constante	Longitude	Latitude	Porosidade
Porosidade	16.5009	-0.0334	-0.1403	—
Permeabilidade	67.5209	-2.3663	2.3307	36.6775

de todos os dados utilizado na análise e os pontos verdes representam o plano da regressão linear encontrado para ambas variáveis. Na Figura 2(b) é possível observar que a nossa estratégia de regressão para a permeabilidade aproximou os resultados obtidos dos valores reais.

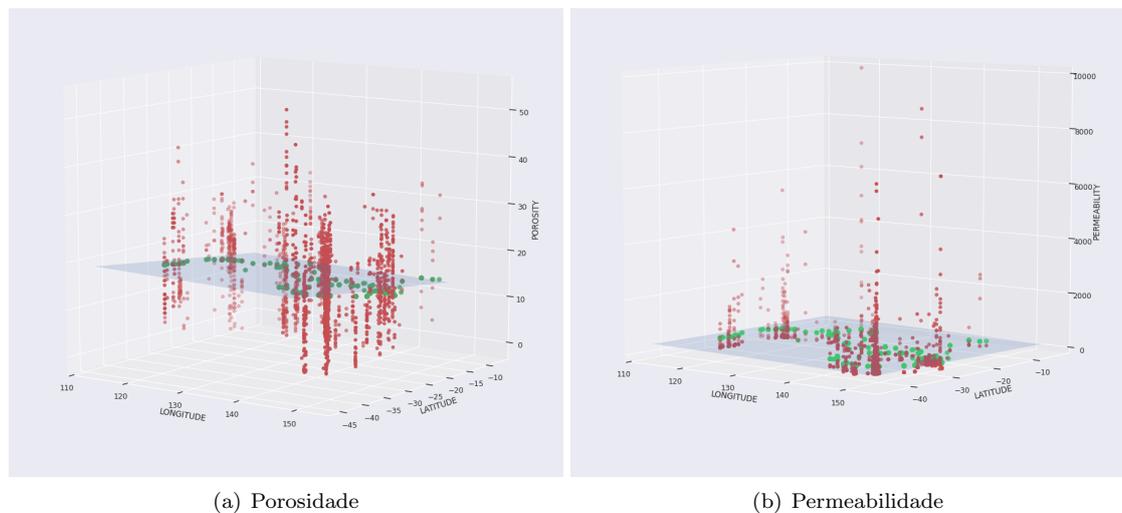


Figura 2: Representação da nuvem de dispersão dos dados utilizados em vermelho, e os planos em azul representam a Regressão Linear para (a) porosidade e (b) a permeabilidade.

4.3 Influência da Permeabilidade na Pressão

Com o intuito de validar a função de regressão linear para a permeabilidade, nesta etapa deve-se encontrar a pressão hidráulica a partir da formulação de Galerkin aplicada ao problema de Poisson (3). O domínio da região foi definido na região $[112, 154] \times [-45, -3]$ e a discretização contou com 129×129 nós em cada uma das direções. Foi utilizado elementos quadriláteros e interpolação lagrangeana de ordem igual a 1, resultando em 16384 elementos. O termo fonte é: $f(x, y) = 2\pi^2 \sin(\pi x) \sin(\pi y)$.

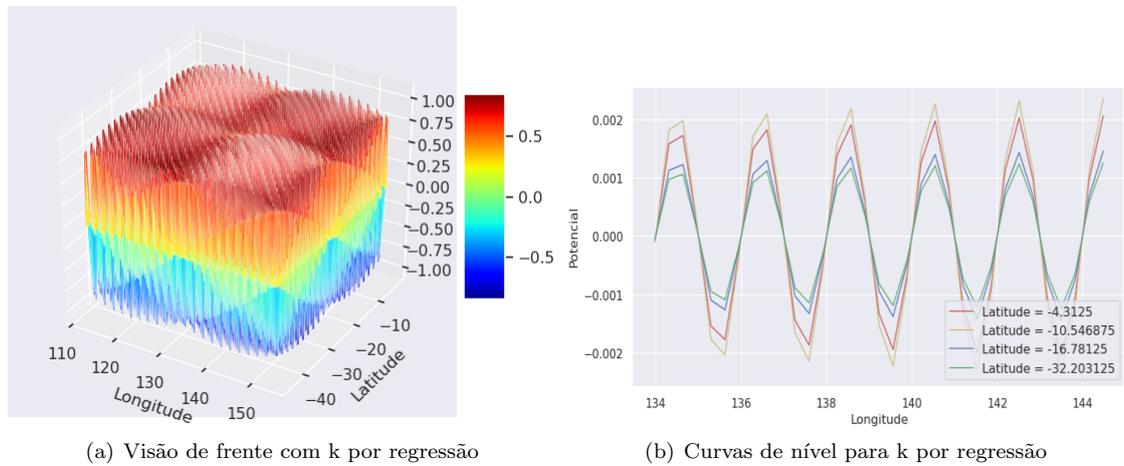


Figura 3: (a) Perfil da variação da pressão hidráulica na região da Oceania para permeabilidade obtida pela regressão. (b) Curvas de nível para permeabilidade .

Na Figura 3(a) é mostrado o perfil da distribuição de pressão p obtido pelo método de Galerkin no domínio definido. O primeiro resultado a ser observado é a influência da função que descreve a permeabilidade na solução. É possível verificar graficamente que a pressão depende da permeabilidade, condizendo com o que é esperado para um meio altamente heterôgeneo. Sendo o termo fonte uma função oscilatória a solução também apresenta estas oscilações e vemos que os picos não possuem a mesma altura como no caso de um meio homogêneo (permeabilidade constante). Já a Figura 4 mostra a distribuição padronizada da projeção 2D da pressão hidráulica.

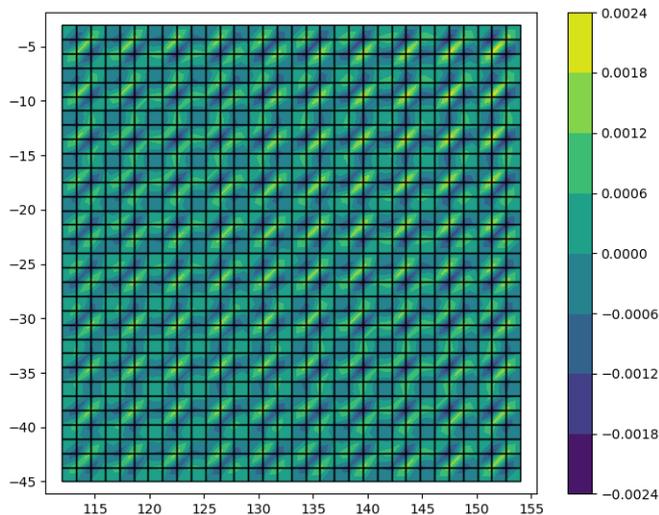


Figura 4: Projeção 2D.

5 Conclusões

A utilização de modelos regressivos, em detrimento a um formalismo determinístico rígido, tem como objetivo prever uma expectativa de permeabilidade, correspondente a um dado conjunto de parâmetros. A vantagem do uso desta metodologia nesta aplicação é que com uma pequena quantidade de dados ligados aos poços foi possível estimar uma função para a permeabilidade que é válida para qualquer ponto do domínio da região. E assim, é possível estimar a pressão hidráulica e o campo de velocidades resolvendo o sistema de Darcy em toda a região.

O cálculo da pressão foi feito utilizando o método de Galerkin na formulação do potencial. Mesmo sabendo que o método de Galerkin não é o mais adequado para a obtenção da solução numérica, já foi possível visualizar o padrão da solução e as diferentes alturas nos picos da solução, comprovando a influência da permeabilidade. Para trabalhos futuros pretende-se aplicar métodos mais estáveis, como o método misto híbrido estabilizado de Galerkin descontínuo (SDHM).

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

- [1] Ahmadi M.A. e Chen, Z. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs. *Petroleum*, v. 5, n. 3, p. 271-284, 2019.
- [2] Australia. Bureau of Mineral Resources, Geology and Geophysics (1990). Program summary. Bureau of Mineral Resources, Geology & Geophysics, Canberra, 1990.
- [3] Balan B., Mohaghegh, S., Ameri, S. State-of-the-art in permeability determination from well log data: Part 1-A comparative study, model development, *SPE Eastern Regional Meeting*. Society of Petroleum Engineers, 1995.
- [4] Correa, M. R. Métodos de Elementos Finitos Estabilizados para Escoamentos de Darcy e de Stokes-Darcy Acoplados, Tese de Doutorado, LNCC, 2006.
- [5] Correa, M. R. e Borges, M. R. A. A semi-discrete central scheme for scalar hyperbolic conservation laws with heterogeneous storage coefficient and its application to porous media flow. *International Journal for Numerical Methods in Fluids*, v. 73, n. 3, p. 205-224, 2013.
- [6] Male, F. e Duncan, I. J., Lessons for machine learning from the analysis of porosity-permeability transforms for carbonate reservoirs. *Journal of Petroleum Science and Engineering*, v. 187, p. 106825, 2020.
- [7] Montgomery, D. C., Peck, E. A. and Vining, G. G. *Introduction to Linear Regression Analysis*, 5th ed. Wiley, New Jersey, USA, 2012.
- [8] Rodríguez Núñez, Y. Métodos de Elementos finitos híbridos aplicados a escoamentos miscíveis em meios porosos heterogêneos, Tese de Doutorado, LNCC, 2014.
- [9] Schneider, A., Hommel, G. and Blettner, M. Linear regression analysis, *part 14 of a series on evaluation of scientific publications*, Dtsch Arztebl Int, 107(44): 776-82, 2010. DOI: 10.3238/arztebl.2010.0776