

10^o pdpetro

Realização



10º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS

TÍTULO DO TRABALHO:

CARACTERIZAÇÃO DE REDE DE POROS EM ROCHAS RESERVATÓRIO POR SEGMENTAÇÃO E QUANTIFICAÇÃO DE FATORES DE FORMA

AUTORES:

Rafael Guntzel Arenhart ¹
Anderson Camargo Moreira ¹
Odair José Bellini ¹
Sandra Nelis Tonietto ²
Celso Peres Fernandes ¹

INSTITUIÇÃO:

¹ Laboratório de Meios Porosos e Propriedades Termofísicas – LMPT, Departamento de Engenharia Mecânica, Universidade Federal de Santa Catarina - UFSC

² CENPES - Centro de Pesquisas e Desenvolvimento Leopoldo Américo Miguez de Mello - PETROBRAS

CARACTERIZAÇÃO DE REDE DE POROS EM ROCHAS RESERVATÓRIO POR SEGMENTAÇÃO E QUANTIFICAÇÃO DE FATORES DE FORMA

Abstract

The pore system in reservoir rocks is usually a complex structure, where a quantitative characterization is not a trivial process, sometimes requiring experimental procedures with high costs and long execution times. A solution for the evaluation of such systems that is rising in use is that of Digital Rocks, defined as three-dimensional images of the rocks, frequently acquired with x-ray microtomography, that allows for a wide range of characterizations of the porous media properties. These characterizations may be obtained either directly from the images, or by physical simulations on the images. One of the more significant properties of porous media is the absolute permeability. In simpler rock structure, such as sandstones, it is possible to establish a direct correlation between porosity and permeability. However, in more complex rocks, such as carbonates, it is not possible to observe a clear correlation between porosity and permeability. The insufficiency of the porosity to completely describe the porous system leads to the need of development of new metrics for this description. A group of metrics used in the quantification of geometric parameters is that of the shape factors. These factors are numeric values that represent one or more characteristics of the geometry of an object. The use of formation factors is promising, since there is a large variety of shape factors available for use, they can be calculated quickly (when compared to alternative methods such as three point correlation) and they can be used in Machine Learning algorithms as features in supervised methods for classification and regression of permeability. Shape factors are calculated individually for each object in the target image. For a percolating pore network, it is not possible to use the common technique of labeling connected voxels, since the pores form interconnected cluster. To circumvent this limitation, it is possible to use a segmenting technique for labeling, such as a watershed algorithm. In the present work, the Imago3d software was used. Pore network partitioning was done using a watershed algorithm, and four shape factors were used to characterize the pores. The image samples analyzed were obtained from testimony samples from a reservoir in the gulf of Mexico. The calculated shape factors were then used to estimate the permeability of the samples using a support vector machine algorithm.

Introdução

As técnicas de rochas digitais tem visto um aumento na sua aplicação industriais e de engenharia na última década. Esse método consiste na obtenção, digitalização e tratamento de imagens volumétricas de meios porosos, e então na simulação de propriedades físicas desse meio (Andra, 2013a). A simulação pode se dar por meios analíticos, numéricos (ex.: volumes finitos), por autômatos (ex.: Lattice Boltzmann), por abstração da rede (ex.: rede de poros e ligações) ou por correlação. Entre as propriedades simuladas no meio poroso estão a permeabilidade, as condutividades térmica e elétrica e o módulo de elasticidade. As imagens, por sua vez, são usualmente obtidas por microtomografia de raio-x ou de elétrons.

Os meios porosos, e portanto as propriedades de rochas digitais, são encontrados em vários campos de atividade de engenharia, como por exemplo atividades de exploração de petróleo e gás, avaliação de aquíferos, dispersão de contaminantes em água, e secagem e aquecimento industrial

(Sartoni, 2017). O uso de rochas digitais é um avanço considerável em relação aos métodos experimentais, tanto por questões econômicas quanto de amplitude de informações. Economicamente, a simulação em rochas digitais costuma ter custo e tempo de execução inferiores aos dos meios experimentais, como ocorre por exemplo com a simulação da permeabilidade e a avaliação experimental por intrusão de mercúrio. Em questão da precisão dos valores obtidos, Andras *et al*, (2013b) fazem uma análise ampla comparando resultados experimentais e simulados com rochas digitais para permeabilidade, condutividade e elasticidade, encontrando valores simulados que divergem entre 10% e 40% dos valores experimentais.

Especificamente na indústria de petróleo e gás, a obtenção da permeabilidade de meios porosos por rochas digitais é utilizada para mapear a qualidade de reservatórios, na simulação numérica de reservatórios em escala geológica, na estimativa de reservas, no planejamento de perfuração e em cálculos de produção (Anazi, 2010).

Além dos métodos de simulação tradicionais já mencionados, nos últimos anos o uso de Aprendizagem de Máquina (*Machine Learning* - ML) tem sido aplicado em rochas digitais. O uso de ML é particularmente adequado para lidar com a grande disponibilidade de amostras, que se amplia conforme técnicas de microtomografia se tornam mais acessíveis, e também para solucionar o aumento de custo computacional e uso de memória dinâmica das imagens de maior resolução. Algumas das aplicações de ML na área de petróleo e gás encontradas na literatura são na classificação de curvas de ressonância magnética (Silva, 2015), classificação de estruturas em registros de poço (Anazi, 2010), segmentação de fases (Chauhan, 2016), e identificação de estruturas litográficas (Xie, 2018).

Das aplicações mencionadas, duas são efetivamente de ML em rochas digitais, e ambas destacam a técnica de Support Vector Machine (SVM), que apesar de ser mais lenta do que outras técnicas de ML avaliadas, minimiza o limite superior de variância das previsões e tem melhor desempenho na identificação de padrões. Adicionalmente, apesar dos trabalhos mencionados utilizarem SVM para classificação (separar as amostras em grupos discretos), o método pode também ser utilizado para regressão (estimar variáveis contínuas), e por isso ele foi utilizado neste trabalho.

Os trabalhos encontrados também comparam processos de ML supervisionados e não-supervisionados. Embora ambos apresentem resultados positivos nas técnicas de classificação, apenas os métodos supervisionados se aplicam para regressão.

O objetivo do presente trabalho é avaliar a aplicação do método de ML de regressão supervisionada por SVM na determinação da permeabilidade de rochas digitais, comparando o resultado com a regressão linear simples da permeabilidade em função da porosidade. Os processos de binarização, segmentação e cálculos de fatores de forma foram feitos com o software Imago3d, as operações de ML foram programadas em Python utilizando a biblioteca scikit-learn (Pedregosa, 2011), e os gráficos foram criados também em Python utilizando a biblioteca matplotlib (Hunter, 2007).

Metodologia

As amostras de rocha utilizadas são de uma formação de carbonato em um reservatório na baía do golfo do México. As amostras foram extraídas por Tonietto (2014), enquanto que a tomografia e binarização das imagens foi realizada por Bellini (2019). Um total de 6 amostras do conjunto foram selecionadas.

Para atender a demanda de um número elevado de amostras para os algoritmos de ML, cada

uma das amostras, originalmente com tamanho de 500^3 voxels, foi dividida em 125 amostras de tamanho 100^3 . Das amostras resultantes, foram removidas as amostras que não apresentavam ao menos um agregado percolante, utilizando o método de Hoshen-Kopelman (Hoshen, 1976). Após a seleção, restaram 292 amostras.

Cada uma das rochas digitais teve a sua permeabilidade calculada pela aproximação de Arns-Adler (Arns, 2018). A aproximação de Arns-Adler aproxima o problema não-linear do escoamento viscoso para um problema linear, permitindo a obtenção de valores de permeabilidade de forma mais rápida e com menos uso de memória dinâmica. O princípio do método envolve aproximar para cada elemento de poro uma permeabilidade equivalente igual ao quadrado da distância do voxel até a interface com a rocha, e então resolver o sistema laplaciano de equações de Darcy e conservação. Do resultado desse sistema de equações, foi obtido o gradiente de velocidade a permeabilidade efetiva do meio, sendo a permeabilidade a variável alvo do algoritmo de ML.

As imagens de rocha binarizadas foram segmentadas utilizando a técnica de *watershed* descrita por Gostick (2017). O método de watershed é composto por duas etapas: a criação de marcadores nos máximos locais do mapa de distância, e então a inundação dos marcadores para a identificação das linhas divisoras entre eles. A abordagem de Gostick adiciona duas etapas ao processo de criação de marcadores, uma é a utilização de um filtro de média gaussiana no mapa de distância para remover as irregularidades, e outra é a aglutinação de marcadores muito próximos. Essas duas operações reduzem a ocorrência de super segmentação por aparecimento de múltiplos máximos locais em partículas muito irregulares.

Os fatores de forma utilizados foram volume, superfície, raio hidráulico, e irregularidade. Volume e superfície foram obtidos pela técnica de Marching Cubes (Lewiner, 2003), enquanto raio hidráulico e irregularidade são calculados a partir dessas medidas. As estatísticas utilizadas para descrever esses parâmetros foram média, mediana, desvio padrão, quartil inferior e quartil superior, medidos tanto por número de objetos quanto ponderado pelo volume do objeto. Cada um desses valores foi utilizado como um parâmetro no algoritmo de ML, além dos valores de porosidade e número de partículas.

Foram determinados um total de 29 parâmetros na etapa anterior. É importante evitar parâmetros excessivos para que não ocorra o efeito de *overfitting*, no qual os coeficientes de ML se ajustam tão bem a uma amostra que eles perdem capacidade de generalização. Para reduzir o número de parâmetros, foi utilizado o método KBest com a métrica de regressão de informação mútua. Esse método ordena os parâmetros em relação a quais são mais dependentes da variável alvo, e então escolhe os K parâmetros mais dependentes. A forma de escolher K foi identificando o escore de cada valor K e então utilizar o K com maior escore. A definição de escore será feita a seguir.

O método de ML escolhido foi o Linear Support Vector Regression (LSVR), que é uma aplicação de SVM para regressão linear dos parâmetros. O princípio básico desse algoritmo é a determinação de que as amostras estão espalhadas em um espaço n-dimensional, onde n é o número de parâmetros, e esse espaço é dividido por hiperplanos, sendo que então a predição da variável alvo de qualquer ponto nesse espaço é uma função da distância do ponto a cada um dos hiperplanos.

Para avaliar a efetividade do LSVR, foi utilizada a técnica de validação cruzada com KFold embaralhados e R^2 como métrica. O principal objetivo desse método é obter a métrica com predições feitas sempre dividindo as amostras em dois grupos não-sobrepostos, um de treinamento e outro de teste. O KFold embaralhado divide a amostra em K grupos (sendo $k = 20$ no presente trabalho), e para cada grupo o LSVR é treinado com os outros e então o escore R^2 é calculado com o grupo remanescente. Isso garante que todos os pontos sejam testados apenas uma vez, mas nunca estando no

próprio grupo de treinamento. Adicionalmente, a variância das predições também foi determinada para gerar uma avaliação visual em um gráfico do tipo *boxplot*. Como linha base de comparação do LSVR, o mesmo procedimento de validação foi feito com uma regressão simples da permeabilidade em função da porosidade.

Resultados e Discussão

Os resultados de R^2 para a seleção de parâmetros por *KBest* estão demonstrados na Figura 1.

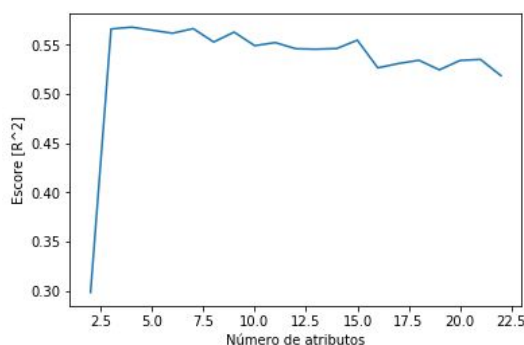


Figura 1 – Valores de R^2 para diferentes números de parâmetros.

O maior valor de R^2 obtido foi para a seleção de 4 parâmetros (porosidade, superfície média em número, irregularidade média em número, e o quartil superior da irregularidade em número), com um valor de $0,5615 \pm 0.1411$. Para comparação, o valor de R^2 para a regressão linear da porosidade foi de $0,3826 \pm 0.1394$. A Figura 2 apresenta a representação gráfica de dispersão da variância das predições para os dois métodos.

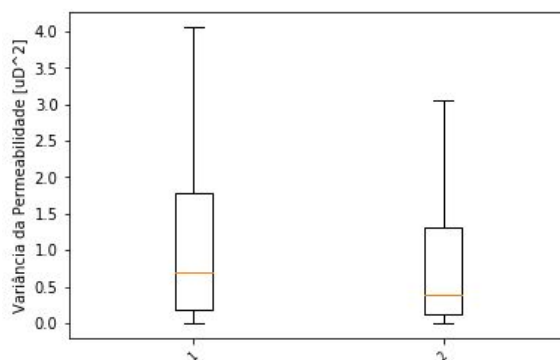


Figura 2 – Dispersão das variâncias das amostras. 1) regressão linear simples, 2) LSVR

Conclusões

O método de ML para regressão da permeabilidade foi implementado e testado. O desempenho foi superior ao da regressão linear simples, porém o intervalo de confiança indica baixa significância estatística. Considerando que dos 29 parâmetros medidos, apenas 4 foram selecionados, o mais provável é que seja necessário adicionar mais fatores de forma, e com medidas mais independentes, para que o modelo seja aprimorado. Porém, a eficiência em termos de tempo computacional do método se comprova, enquanto que para cada volume eram necessários entre 15 e 25 minutos para simular a permeabilidade, todo o processo de ML, incluindo a segmentação por

watershed e a medida dos fatores de forma demorou aproximadamente 30 minutos para as 292 amostras. Considerando que a simulação da permeabilidade tem um comportamento exponencial, essa diferença deve se acentuar ainda mais para imagens de maior resolução.

Agradecimentos

Esse projeto foi financiado pela Petrobras e pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis.

Referências Bibliográficas

AL-ANAZI, A., GATES I.D., A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs, **Engineering Geology**, 114, 2010.

ANDRA, H. *et al*, Digital rock physics benchmarks — Part I : Imaging and segmentation, **Computers & Geosciences**, 50, 2013a.

ANDRA, H. *et al*, Digital rock physics benchmarks — Part II : Computing effective properties, **Computers & Geosciences**, 50, 2013b.

ARNS, C. H., ADLER, P. M., Fast Laplace Solver for Pore-Scale Permeability, **Physical Review E**, 97, 2018.

BELLINI, O. J., Caracterização Tridimensional Do Sistema Poroso E Categorização De Rochas Carbonáticas Com Aprendizado De Máquina, Tese de Doutorado, **Universidade Federal de Santa Catarina**, 2019.

CHAUHAN, S. *et al*, Processing of rock core microtomography images: Using seven different machine learning algorithms, **Computers & Geoscience**, 86, 2016.

HOSHEN, J., KOPELMAN, R., Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm, **Phys. Rev. B** 14, p. 3438, 1976.

HUNTER, J. D., Matplotlib: A 2D Graphics Environment. **Computing in Science & Engineering**, 9, pp. 90-95, 2007.

LEWINER, T., LOPES H., VIEIRA, A. W., TAVARES, G. Efficient implementation of Marching Cubes cases with topological guarantees. **Journal of Graphics Tools** 8(2) pp. 1-15, 2003.

PEDREGOSA, F., *et al*, Scikit-learn: Machine Learning in Python, **JMLR** 12, pp. 2825-2830, 2011.

SILVA, P. N. *et al*, Automatic classification of carbonate rocks permeability from ¹H NMR relaxation data, **Expert Systems with Applications**, 42, 2015.

STARNONI, M., POKRAJAC, D., NEILSON, J., D., Computation of fluid flow and pore-space properties estimation on micro-CT images of rock samples, **Computers & Geosciences**, 106, 2017.

TONIETTO, S. N., Pore Characterization And Classification In Carbonate Reservoirs And The Influence Of Diagenesis On The Pore System. Case Study: Thrombolite And Grainstone Units Of The Upper Jurassic Smackover Formation, Gulf Of Mexico, Tese de Doutorado, **Texas A&M University**, 2014.

XIE, Y. *et al*, Evaluation of machine learning methods for formation lithology identification: A comparison of tuning processes and model performances, **Journal of Petroleum Science and**

