

9º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



9º CONGRESSO
Brasileiro de P&D em
PETRÓLEO E GÁS

Maceió, AL
de 09 a 11 de novembro
2017

Realização:



ABPG
ASSOCIAÇÃO BRASILEIRA DE P&D EM
PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

Utilização de Inferência Bayesiana para Estimativa de Primeira Quebra em Ondas P

AUTORES:

Leonardo de Assis Pinto¹, Marco Antonio Rodrigues de Ceia¹, Irineu de Azevedo Lima Neto¹,
Roseane Marchezi Missagia¹, Fernando Vizeu Santos²

INSTITUIÇÃO:

¹Laboratório de Engenharia e Exploração de Petróleo - LENEP, Campus Macaé, Universidade Estadual do Norte Fluminense Darcy Ribeiro – UENF; ²Departamento de Geologia, Instituto de Geociências, Campus Gragoatá, Universidade Federal Fluminense

Este Trabalho foi preparado para apresentação no 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás - 9º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás - ABPG, no período de 09 a 11 de novembro de 2017, em Maceió/AL. Esse Trabalho foi selecionado pelo Comitê Científico do evento para apresentação, seguindo as informações contidas no documento submetido pelo(s) autor(es). O conteúdo do Trabalho, como apresentado, não foi revisado pela ABPG. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões da Associação Brasileira de P&D em Petróleo e Gás. O(s) autor(es) tem conhecimento e aprovação de que este Trabalho seja publicado nos Anais do 9º PDPETRO.

UTILIZAÇÃO DE INFERÊNCIA BAYESIANA PARA ESTIMATIVA DE PRIMEIRA QUEBRA EM ONDAS P

Abstract

The estimate of acoustic velocities is an important information which could be extracted from several type data, like: exploration seismic, well logging (acoustic logging) and laboratory experiments. To realize this estimate, a technique called first break picking is used, which consist in an analysis of the amplitude of a measured signal which has a variation over time to identify the time of the wave arrival. The complexity of this analysis is in the fact of the data, generally is contaminated with noise and this one could be either inherent to the measuring equipment or other factors, that would make a simple visual analysis impossible. Considering this problem, the development of techniques which would be capable to identify the wave arrival automatic, accurate and faster, justify the study of advanced methods to accomplish this task.

Introdução

Medidas de ondas acústicas podem ser obtidas a partir da sísmica, perfilagem de poços e medidas laboratoriais. O dado obtido é uma informação da distribuição da amplitude do sinal acústico pelo tempo, onde este tempo, pode ser entendido como sendo o tempo que a onda levou para sair da sua fonte até chegar ao receptor e a amplitude como sendo a intensidade que esse sinal chega até o receptor.

Contudo, tais leituras comumente estão contaminadas com ruído, o qual pode ser inerente ao equipamento ou originado de outros fatores externos. Uma importante informação a ser obtida a partir desses dados é a velocidade de propagação da onda acústica, a qual pode-se conseguir através de uma técnica chamada primeira quebra, esta consiste em identificar o momento correto da chegada da onda ao receptor, contudo a presença do ruído, como citado anteriormente, pode tornar essa tarefa mais complexa que o esperado.

Algumas metodologias para identificação da primeira quebra automática podem ser encontradas na literatura, como indicado em Akram (2011), como por exemplo: baseada em propriedades de correlação (Gelchinsky & Shtivelman, 1983); dimensões fractais e/ou entropia (Jiao & Moon, 2000, Sabbione & Velis, 2010); redes neurais (McCormack et al., 1993; Dai and MacBeth, 1995); e critério estatístico (Akram, 2011).

Além desses métodos já testados, outros métodos estatísticos, como o uso de Inferência Bayesiana, se apresentam como possível solução. Em Davidson-Pilon (2015) é apresentado um problema onde um conjunto de dados é distribuído com relação ao tempo e é utilizada Inferência Bayesiana para identificação de uma variação nessa distribuição, algo que se assemelha ao problema da primeira quebra em ondas acústicas.

Metodologia

A técnica apresentada neste trabalho foi desenvolvida assumindo que os dados medidos poderiam ser divididos em duas partes diferentes, uma consistindo basicamente em ruído, que tem seu fim na chegada da onda, e a segunda, que começa na chegada da onda, a qual consiste na medida da onda e também ruído, mas para ambas as partes dos dados, foi considerada uma distribuição normal com média zero e variância σ_1^2 e σ_2^2 .

Essas considerações foram feitas para aplicar a Inferência Bayesiana, que é o método estatístico utilizado para obter o parâmetro baseado no teorema de Bayes, que poderia ser expresso como na equação abaixo:

$$P(\tau, \sigma_1^2, \sigma_2^2 | x) \propto \frac{e^{-\frac{1}{2} \left(\sum_{i=1}^{n_1} \frac{x_i^2}{\sigma_1^2} + \sum_{i=n_1+1}^N \frac{x_i^2}{\sigma_2^2} \right)}}{\sigma_1^{n_1} \times \sigma_2^{n_2}} \times P(\tau, \sigma_1^2, \sigma_2^2)$$

Onde x é o dado, x_i são amostras do dado, n_1 é o número de amostras antes de $t = \tau$, τ é o tempo de chegada, n_2 é o número de amostras após $t = \tau$, N é o número de amostras do dado e $P(\tau, \sigma_1^2, \sigma_2^2)$ é a distribuição *a priori* dos parâmetros.

A Figura 1 apresenta os dados de modelagem, o modelo proposto no presente trabalho considera todas as amostras de dados independentes um do outro.

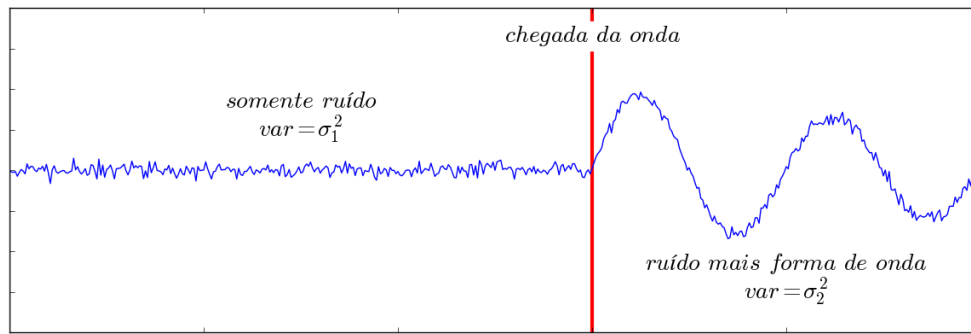


Figura 1 - Modelo de dado utilizado, com considerações adotadas para resolução do problema proposto.

Para a distribuição *a priori* do tempo da primeira quebra foi usada uma distribuição uniforme em todos os tempos medidos para o tempo de chegada τ , uma vez que τ é o principal parâmetro de interesse e não há informações sobre o valor real. Para as variâncias, foi selecionada uma distribuição gama inversa, porque é a conjugada *a priori* para a variância de uma distribuição normal com média conhecida.

Para a distribuição de gama é necessária, determinar os parâmetros α e β que poderiam ser feitos com poucas amostras de dados. Foram utilizados o primeiro 1% do dado, onde sabemos que há apenas ruído para o σ_1^2 e foi usado o último 1% do dado, onde sabemos que há sinal e ruído para o σ_2^2 e dessas amostras, α foi definido como metade do número de amostras utilizadas e β como metade da soma do quadrado dos valores de dados.

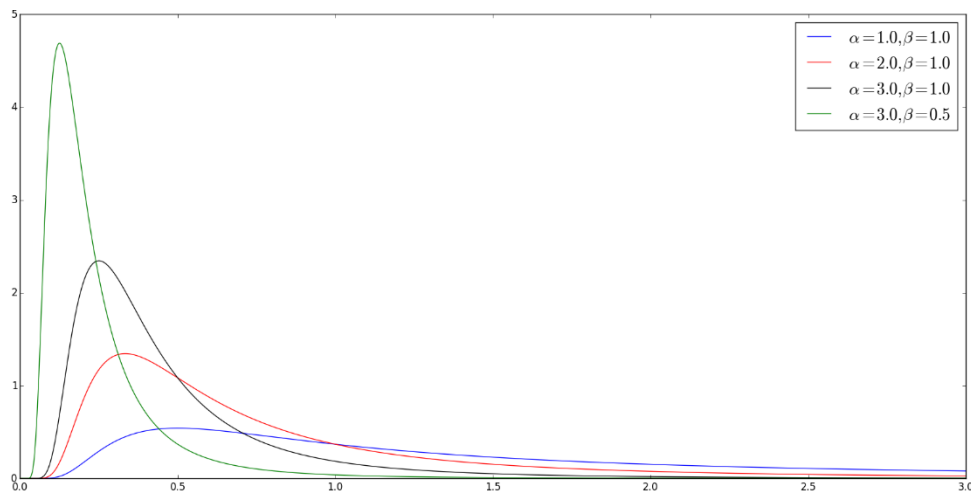


Figura 2 - Distribuição gama inversa para vários valores de alfa e beta.

Para obter esses três parâmetros, realizou-se a simulação Monte Carlo (MCMC) da cadeia Markov, utilizando o algoritmo de Metropolis-Hastings. A implementação deste algoritmo foi com biblioteca Python, chamada PyMC (Patil et al, 2010). A distribuição a *posteriori* foi amostrada 50000 vezes descartando o primeiro 10000 para a cadeia de Markov convergir para o estado estacionário e a média a *posteriori* de τ foi utilizada para estimar o tempo de chegada da onda.

Resultados e Discussão

Para realizar a simulação da técnica proposta, foram utilizadas duas medidas. A primeira consiste numa medida sem a presença da amostra com o emissor e receptor em contato, tal medida é conhecida como “face-to-face”. A segunda será realizada com uma amostra padrão de alumínio com 50 mm de comprimento, durante o experimento a amostra esteve submetida a uma pressão axial de 3 toneladas. Todo experimento foi realizado utilizando o Sistema Triaxial de Física de Rochas pertencente ao LENEP/UENF. O método da primeira quebra em ambas as medidas foi realizado utilizando a técnica proposta e um segundo método para efeito de comparação, que no caso será o método apresentado em Akram (2011).

Na figura 3, é exemplificado o que ocorre durante o processo para determinação dos parâmetro do nosso problema em que ocorre a amostragem do dado em 50000 vezes e que a parte vermelha seriam as 10000 primeiras a serem descartadas, que seria o local antes da convergência da cadeia de Markov. É observado que após o trecho em vermelho os dados da nossa amostragem em azul passam a apresentar uma estabilidade nos valores para as variâncias e o parâmetro de interesse τ , porém não um valor fixo, sendo então necessário estimar qual valor melhor se enquadra entre os possíveis.

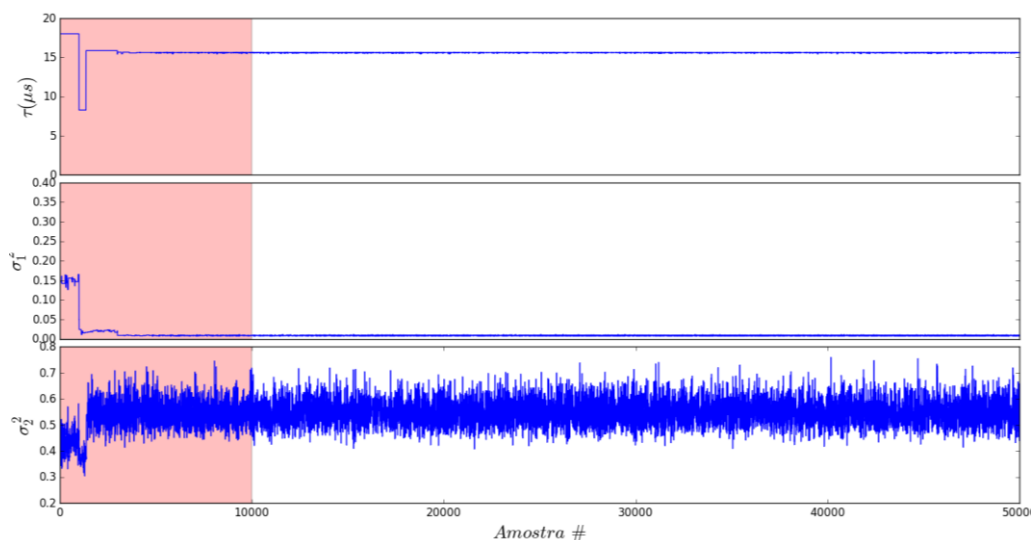


Figura 3 - Obtenção dos parâmetros da cadeia de Markov.

Após essa etapa, determina-se os valores para cada parâmetro através de um histograma e assume-se como valor a ser utilizado a média das amostras da cadeia, a qual assume-se sendo o valor que converge para a média da distribuição a *posteriori* como pode ser visto na Figura 4 a seguir.

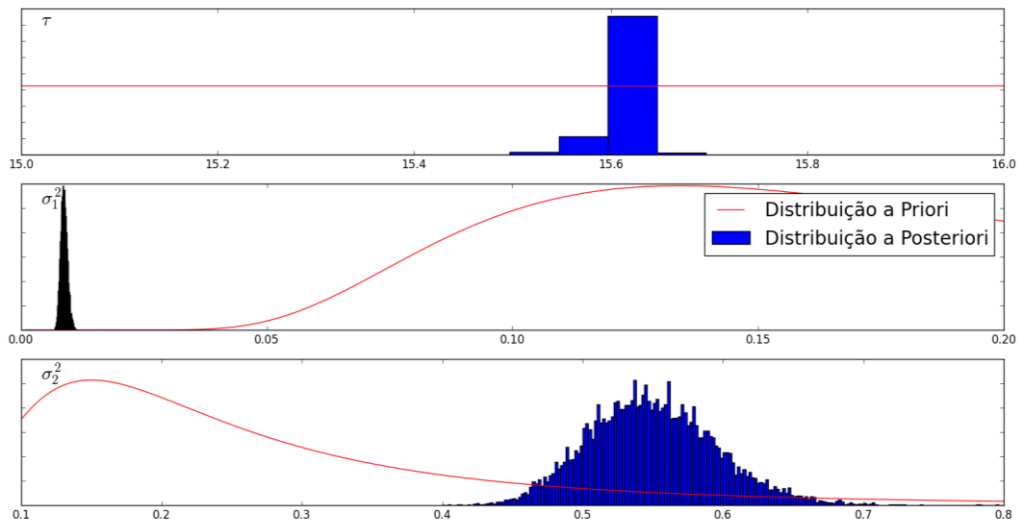


Figura 4 - Histograma da distribuição *a posteriori* e curva da distribuição *a priori*.

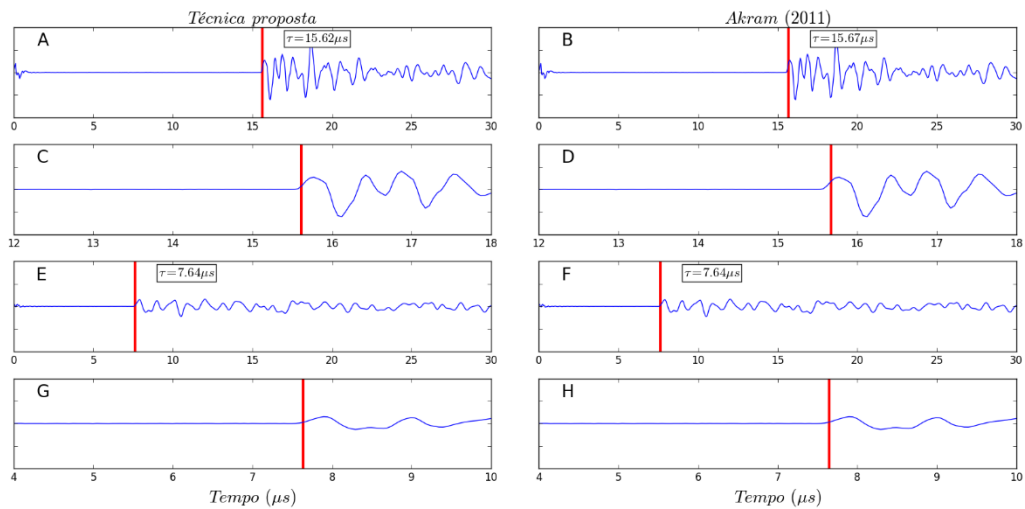


Figura 5 - Primeira quebra automática usando o método proposto e o de Akram (2011). A linha vermelha vertical marca a primeira quebra obtida. A coluna esquerda (A, C, E e G) mostra os resultados do nosso método e a coluna direita (B, D, F e H) para o método de Akram. A, B, C e D referem-se à experiência da amostra de alumínio de 50 mm. E, F, G e H referem-se ao experimento face-a-face. C, D, G e H são versões ampliadas das figuras mostradas acima.

É possível observar na Figura 5 que os dois métodos são semelhantes, mas o método proposto pode ser considerado mais preciso, o que pode ser melhor observado nas figuras 1C, 1D, 1G e 1H (figuras ampliadas). No entanto, o cálculo da velocidade da estimativa usando cada tempo é de 6.265,66 m / s para o método proposto e 6.226,65 m / s para o método de Akram, uma diferença de apenas 0,6%.

Em ambos os dados, a presença de ruído não foi considerável, portanto, para testar a sensibilidade dos métodos a diferentes níveis de ruído, todos os dados da amostra de alumínio foram contaminados com diferentes níveis de ruído gaussiano, usando uma função aleatória. Ambos os métodos processaram novamente os dados e os resultados são mostrados na Figura 6.

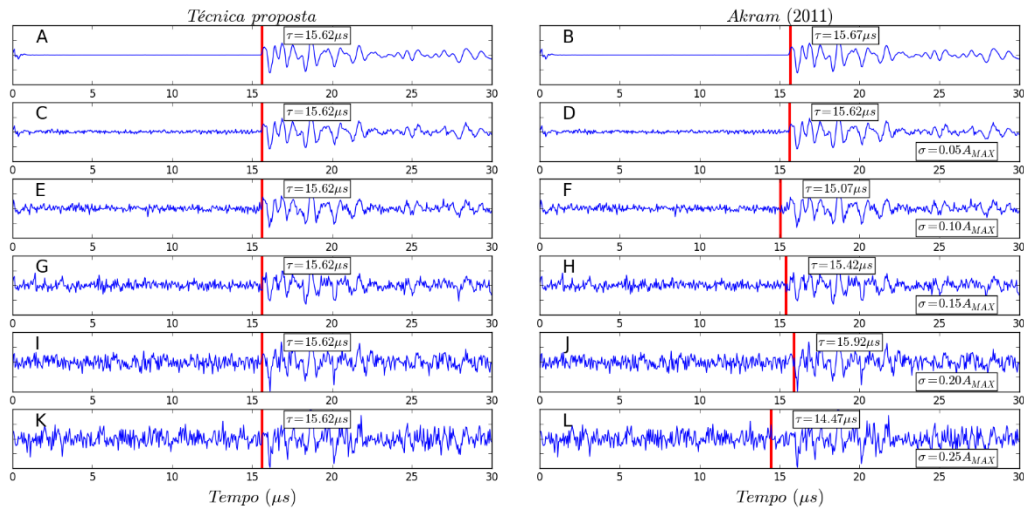


Figura 6 – Primeira quebra automática em diferentes condições de ruído. A linha vermelha vertical marca a primeira quebra obtida. A coluna da esquerda (A, C, E, G, I, K) refere-se ao método proposto e a coluna direita (B, D, F, H, J, L) ao método de Akram (2011). O nível de ruído aumenta de cima para baixo da figura. O desvio padrão de ruído foi ajustado para 5% da amplitude máxima do sinal em C e D, 10% em E e F, 15% em G e H, 20% em I e J e 25% em K e L e A e B não possuem aumento de ruído, ambos são os mesmos dados da Figura 5.

É possível observar que o método proposto conseguiu manter o mesmo resultado mesmo com a variação do ruído, enquanto o método de Akram (2011) não conseguiu, apesar da variação no resultado ter sido de no máximo, aproximadamente 7,66%. Contudo, este resultado pode não se repetir visto que o método proposto é estocástico.

Conclusões

A necessidade de novas técnicas para analisar os dados das medidas acústicas pode ser observada ao analisarmos como os apresentados nos gráficos das Figuras 2I, 2J, 2K e 2L. Já que poderia ser considerado quase impossível dar uma resposta precisa com uma análise visual, confirmando essa necessidade.

Sendo assim, o presente trabalho apresenta uma técnica que se mostra como uma solução robusta ao problema da primeira quebra automática de sinais acústicos, onde apresentou-se com considerável precisão para diversos tipos de sinais, inclusive mantendo estável o resultado.

No entanto, o custo computacional se apresentou como a maior desvantagem. Outra desvantagem que é necessário mencionar é a natureza estocástica do método usado, o que pode levar a resultados ligeiramente diferentes em cada corrida.

Agradecimentos

Os autores gostariam de agradecer a LENEP/CCT/UENF pela estrutura fornecida e PRH-226/PETROBRAS e a Capes pelo suporte financeiro.

Referências Bibliográficas

Akram, J., 2011. Automatic P-wave Arrival Time Picking Method for Seismic and Microseismic Data. CSPG CSEG CWLS Convention.

Coppens, F., 1985, First arrivals picking on common-offset trace collections for automatic estimation of static corrections: Geophysical Prospecting, 33, 1212–1231.

Davidson-Pilon, C., 2015. Bayesian Methods for Hackers. Addilson-Wesley Professional; 1 edition.

Gelchinsky B. and Shtivelman V., 1983. Automatic picking of first arrivals and parameterization of traveltimes curves. Geophysical Prospecting, 31, 915-928.

Jiao, L. and Moon, W.M., 2000. Detection of seismic refraction signals using a variance fractal dimension technique. Geophysics, 65, 286-292.

McCormack, M., Zaucha, D. and Dushek, D., 1993. First-break refraction event picking and seismic data trace editing using neural networks. Geophysics, 58, 67-78.

Patil, A., Huard, D., Fonnesbeck, C. J., 2010. PyMC: Bayesian Stochastic Modelling in Python. J Stat Softw, 35(4), 1-81.

Sabbione, J. I. and Velis, D., 2010. Automatic first-breaks picking: New strategies and algorithms. Geophysics, 75, V67 – V76.