

5º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

Análise de Agrupamentos dos Dados de DFA Oriundos de
Perfis Elétricos de Indução de Poços de Petróleo

AUTORES:

Maria das Vitórias Medeiros da Mata¹, Umberto Laino Fulco^{1,2}, Gilberto Corso^{1,2,3}

INSTITUIÇÃO:

¹ Universidade Federal do Rio Grande do Norte. Programa de Pós-Graduação em Ciência e Engenharia de Petróleo. Campus UFRN, Natal-RN. 59072-970

vitoriafisica@uol.com.br

² Departamento de Biofísica e Farmacologia, Centro de Biociências. Campus UFRN – Natal/RN

³ Departamento de Física Teórica e Experimental, Campus UFRN – Natal/RN

Este Trabalho foi preparado para apresentação no 5º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás- 5º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás-ABPG, no período de 15 a 22 de outubro de 2009, em Fortaleza-CE. 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 5ºPDPETRO.

Análise de Agrupamentos dos Dados de DFA Oriundos de Perfis Elétricos de Indução de Poços de Petróleo

Abstract

We apply the Detrended Fluctuation Analysis (DFA) in well data. The main objective of this study is to characterize spatially oil fields. After obtained the DFA values for all wells, we perform group analysis with DFA exponents. We used the non-hierarchical method called K-means. The K-means consists in dividing the elements of a data matrix in k groups so that the similarities among elements belonging to the same groups are the smallest possible. In order to certify if a dataset generated by the K-means method form spatial patterns, we created the parameter Ω (index of neighborhood). High values of Ω ; involve more aggregated data and low values of Ω ; in scattered data or without spatial correlation. We compare the DFA exponents of the well grouped by the clustering analysis with randomly generated data. We concluded that data from the DFA of 54 wells are grouped and can be used to characterize spatial fields.

Keywords: Detrended Fluctuation Analysis (DFA); analysis of groups, K-means; oil fields; Index of neighborhood (Ω).

Introdução

Hoje sabemos dos diversos e contraditórios desafios existentes na indústria petrolífera e o quanto é importante o estudo das incertezas existentes nas atividades de exploração e produção de petróleo. Deste modo, o uso de técnicas e *softwares* empregados de forma sistemática no desenvolvimento e avaliação dos campos são vistos como potenciais ferramentas para esse propósito, facilitando assim o processo de tomada de decisão que é bastante complexo e nem sempre de fácil solução.

No presente trabalho aplicamos o Detrended Fluctuation Analysis (DFA) à dados de perfis elétricos de indução de poços de petróleo. O objetivo deste trabalho é estudar a viabilidade de se usar o DFA na construção espacial de um campo petrolífero. Como caso de estudo, utilizamos dados de perfis elétricos de indução (ILD) de 54 poços de petróleo do Campo Escola de Namorado, Bacia de Campos – RJ.

Metodologia

De início transportamos os dados referentes ao perfil de indução (ILD) de cada poço para o *software* Matlab R2006a, onde aplicamos a técnica de DFA. Depois transferimos os dados obtidos para outro *software*, o OriginPro 7.5, onde foi possível construir os gráficos referentes a cada poço, além de obtermos a equação da reta, determinando assim, o índice do DFA. Por fim, os dados foram transferidos para o *software* Statistica onde fizemos a Análise de Agrupamentos.

Com o objetivo de verificar se um conjunto de dados gerados pelo DFA e agrupados pelo método do k-média forma padrões espaciais criamos o parâmetro Ω (índice de vizinhança). Altos valores de Ω implicam em dados mais agregados, baixos valores de Ω em dados dispersos ou sem correlação espacial. Estes valores foram comparados com dados gerados aleatoriamente.

DFA

A técnica do DFA é um método de dimensionamento comumente utilizado para detecção de correlações de longo alcance em séries temporais não-estacionárias pela lei de potência. Este método é baseado na teoria do passeio aleatório (IVANOVA & AUSLOOS, 1999).

A técnica de DFA aplicada consiste em dividir uma sequência $y(t)$ em intervalos não sobrepostos de tempo iguais I_n de tamanho s , onde $n = 1, \dots, N$. Introduzimos a função tendência local $Y_s(t)$ definida por $Y_s = a_n + b_n t$ para $t \in I_n$, onde os coeficientes a_n e b_n representam o ajuste linear dentro de cada janela (caixa). Cada série de pontos foi dividida em oito caixas. Calculamos a função de flutuação $F(n)$ definida como:

$$F(n) = \sqrt{\frac{1}{N} \sum_{i=1}^N [Y(i) - Y_s(i)]^2} \quad (1)$$

Por fim, plotamos o gráfico $\log [F(n)]$ por $\log(n)$, como mostrado abaixo (Figura 1) e ajustamos uma reta aos pontos, que gerou a equação da reta e a declividade da reta α é o valor do DFA para cada série de pontos.

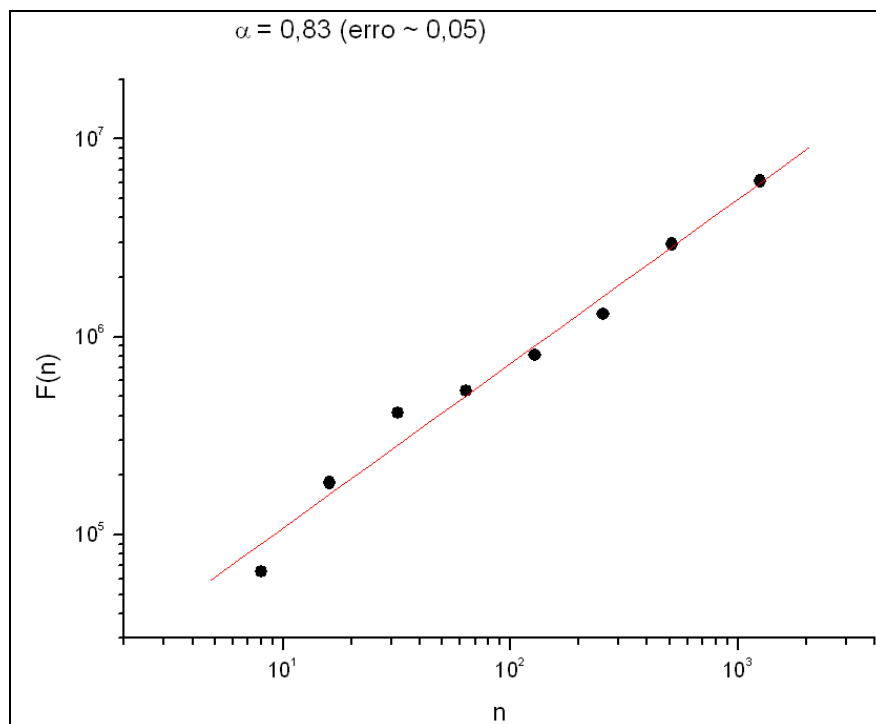


Figura 1: Equação da reta a partir de uma série de pontos do poço 01, localizado no Campo Escola de Namorado.

Desta forma $F(n)$ representa uma média da flutuação para cada segmento e n representa um comprimento de escala.

Análise de Agrupamentos

A Análise de Agrupamentos é uma técnica que procura organizar um conjunto de dados de uma matriz n em grupos mais ou menos homogêneos levando em consideração apenas as similaridades ou dissimilaridades existentes entre os elementos da matriz (Mingoti, 2007). Para este

estudo fizemos uma análise de agrupamento com base nas características de DFA do perfil de indução calculadas para cada poço. Para isto, fizemos uso do método de agrupamento por K- média baseado na distância euclidiana e que tem como critério de agregação o método do centróide.

O Método K-média (*K-means*)

O agrupamento por K- média é um método não-hierárquico que consiste em dividir os dados de uma matriz n em k grupos mais ou menos homogêneos, cada grupo constituindo uma população bem definida, ao mesmo tempo em que as semelhanças entre os elementos de grupos distintos sejam as menores possíveis (Lourenço & Matias, 2000).

O método é composto pelos seguintes passos:

1. Escolha do número k de grupos iniciais.
2. Continuar o processo através da lista de dados, alocando cada elemento ao grupo cujo centróide é mais próximo do vetor de médias amostrais, geralmente através da distância euclidiana.
3. Recalculam-se os valores dos centróides para cada novo grupo formado até que os grupos estejam bem definidos.
4. Repetir das etapas 2 e 3 até que o reagrupamento dos itens não seja mais necessário.

Método Monte Carlo

Com o método Monte Carlo utilizamos dois algoritmos para entrarmos uma distribuição de valores de Ω para um conjunto aleatório de dados que podem ser descritos da seguinte forma:

1º caso: Algoritmo para achar Ω

A- Dado o problema (caso concreto) entramos com os dados e colocamos nas K classes, neste caso já sabíamos o número de classes e o número de elementos por classe, que é o dado de origem.

a) Estrutura dos grupos

$K =$ grupos, com n_i $i = 1 \dots k$; onde i é o índice dentro do grupo.

b) Estrutura dos poços

$j = 1, \dots, N$, com $(d_{x,y})_j$ onde j é o índice dos poços e N é número de poços.

B- Definimos B (tamanho da bola aberta). Para cada elemento l dentro de uma classe procuramos um vizinho dentro da classe:

Se entre dois poços $|(d_{x,y})_{l_1} - (d_{x,y})_{l_2}| < B$, então l_1 e l_2 são vizinhos.

- Se tem vizinho $V(l_1, l_2) = 1$;

- Se não tem vizinho $V(l_1, l_2) = 0$.

C- Entramos com o valor de P_i (índice de vizinhança dentro do grupo) da seguinte forma

$$P_i = \sum_{l_1=1}^{n_1} \sum_{l_2=1}^{n_2} V(l_1, l_2); \quad (2)$$

onde $V(l_1, l_2)$ é o número total de vizinhos de cada grupo.

D- Por fim encontramos o índice de vizinhança global Ω utilizando a seguinte equação:

$$\Omega = \frac{1}{K} \sum_{i=1}^K P_i \quad (3)$$

2º caso: Algoritmo tipo Monte Carlo para entrarmos com a distribuição de valores de Ω para um conjunto aleatório de dados

A- Dados de entrada:

K - número de grupos e n_i - elementos para cada grupo.

B- Geração de NT amostras aleatórias. Neste passo embaralhamos os valores dos poços e distribuímos dentro dos K grupos, de forma que

$r = 1, \dots, NT_{total}$; onde NT_{total} é o número total de pontos.

Feito isto, calculamos Ω_r e plotamos os gráficos para análise.

C- Por fim calculamos a média e o desvio padrão dos dados através das seguintes equações, respectivamente:

$$\bar{\Omega} = \frac{1}{NT_{total}} (\sum \Omega) \quad (4)$$

e

$$\Delta\Omega = \frac{1}{NT_{total}} \sum (\bar{\Omega} - \Omega)^2 \quad (5)$$

Resultados e Discussão

Utilizando o K-média, foram feitas análises para $K = 5, 6$ e 7 . A Figura 2 mostra o resultado obtido para $K = 7$. Observamos nos mapas que à medida que era feito um recálculo do centróide, os elementos se realocavam nos grupos até o ponto que os grupos permaneceram bem definidos com relação às características de DFA.

Percebemos que o melhor resultado se deu com sete grupos; isto ficou claro depois que traçamos as curvas de níveis (Figura 3) dos dados de DFA do perfil de indução dos 54 poços estudados e comparamos estes dados com os obtidos do K-média.

Com auxílio do índice de vizinhança e do método Monte Carlo (que chamamos de **MC**) verificamos que os dados agrupados aleatoriamente apresentam uma distribuição mais baixa de Ω do

que os obtidos dos dados empíricos e agrupados pelo método do k-média. Para o programa **MC**, estimamos 1000 vezes Ω para $k = 5, 6$ e 7 conjuntos gerados aleatoriamente. Integrado no programa **MC** está o tamanho da bola b que define a vizinhança entre os elementos em análise. Para $k = 7$ e $b = 0,5$ obtivemos os seguintes resultados para Ω , $\overline{\Omega}$ e $\Delta\Omega$ (Tabela 1):

Tabela 1: Resultados obtidos do programa **MC** para $K = 7$ e $b = 0,5$.

Ω (Índice de vizinhança)	$\overline{\Omega}$ (médio)	$\Delta\Omega$ (desvio padrão)
0,76	0,19	0,03

A tabela 2 mostra os resultados da análise visual dos elementos (poços) agrupados com relação as suas semelhanças, gerados através da análise de curvas de níveis comparados com os resultados obtidos do método K-média para $k = 7$.

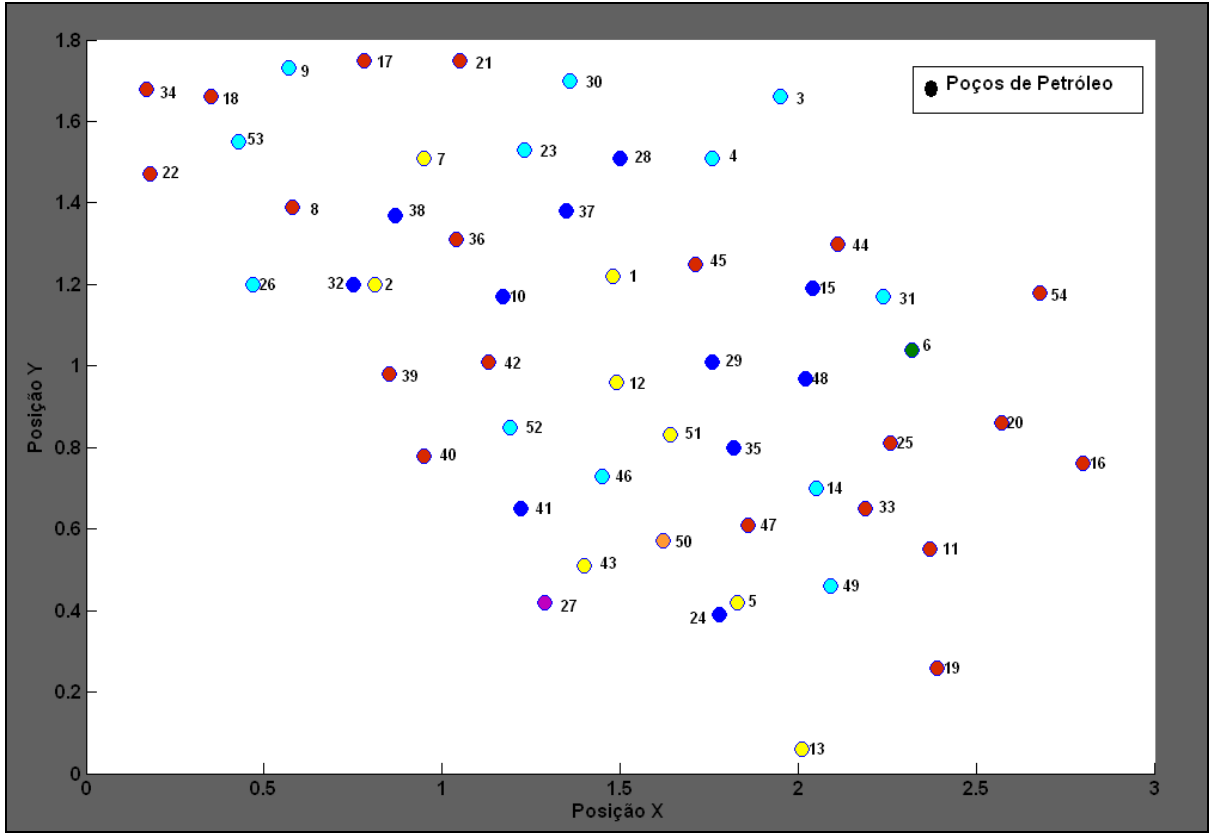


Figura 2: Para $K = 7$, agrupamentos utilizando o método K- Média dos valores de DFA. As unidades de distância são arbitrárias. (Figura do Autor).

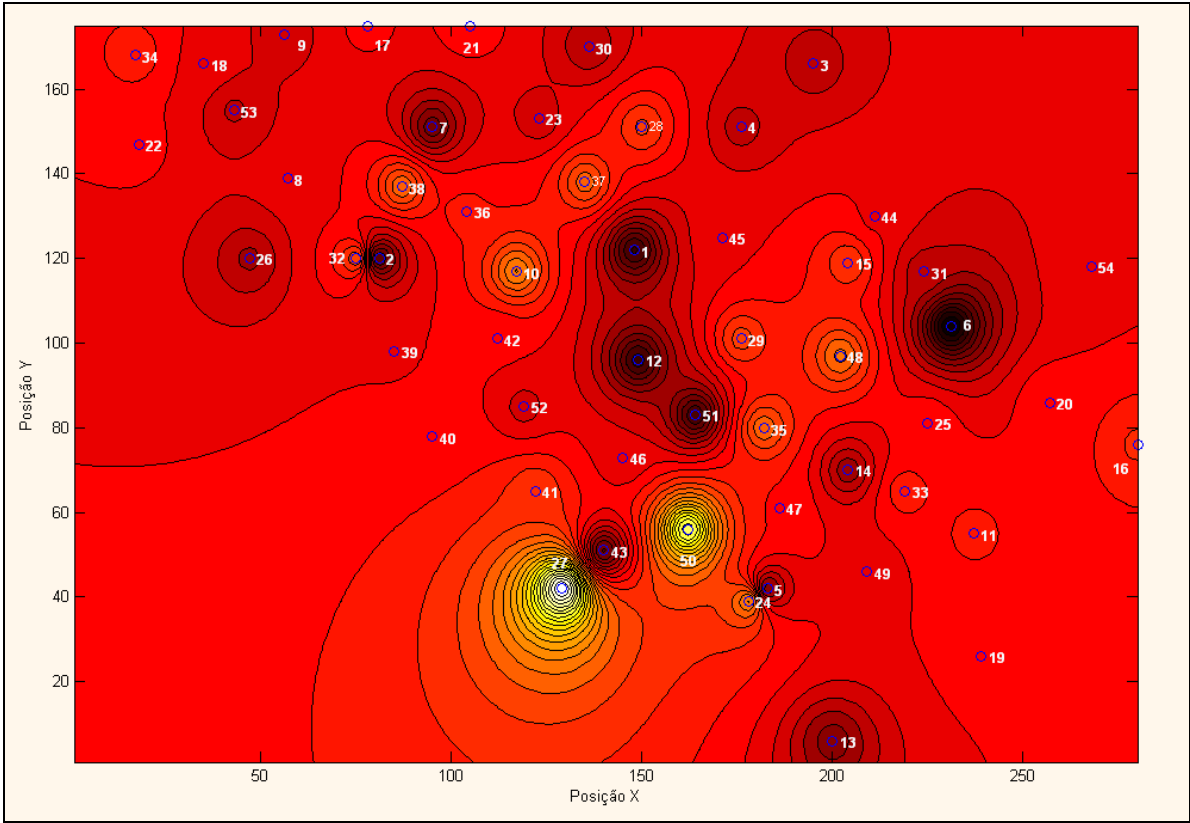


Figura 3: Mapa de distribuição das características de DFA para a curva ILD (perfil de indução) na área em estudo. Intervalos de contorno = 30. (Figura do Autor).

Tabela 2 – Comparação entre as ilhas formadas no agrupamento $k = 7$ e nas CN (curvas de níveis).

Poços no agrupamento $k=7$ pelo K-média	Poços no agrupamento com Curvas de Níveis (CN)	Poços comuns entre em $k=7$ e CN	Nº de poços em $k=7$	Nº de poços em CN
3, 4, 9, 14, 23, 26, 30, 31, 46, 49, 52, 53	3, 4, 9, 23, 30, 31, 52, 53	3, 4, 9, 23, 30, 31, 52, 53	12	8
6	6	6	1	1
50	50	50	1	1
8, 11, 16, 17, 18, 19, 20, 21, 22, 25, 33, 34, 36, 39, 40, 42, 44, 45, 47, 54	8, 11, 15, 16, 17, 18, 19, 20, 21, 22, 25, 28, 29, 32, 33, 34, 36, 38, 39, 40, 41, 42, 44, 45, 46, 47, 49, 54	8, 11, 16, 17, 18, 19, 20, 21, 22, 25, 33, 34, 36, 39, 40, 42, 44, 45, 47, 54	20	28
1, 2, 5, 7, 12, 13, 43, 51	1, 2, 5, 7, 12, 13, 14, 26, 43, 51	1, 2, 5, 7, 12, 13, 43, 51	8	10
10, 15, 24, 28, 29, 32, 35, 37, 38, 41, 48	10, 24, 35, 37, 48	10, 24, 35, 37, 48	11	5
27	27	27	1	1

Conclusões

Através da criação do parâmetro Ω (índice de vizinhança) e da simulação Monte Carlo, foi possível observar que um conjunto de dados gerados pelo método do k-média forma padrões espaciais.

Destarte, concluímos que os dados de DFA obtidos nos 54 poços estão agrupados e podem ser usados na caracterização espacial de campos.

Todavia, para uma possível tomada de decisão é preciso uma análise conjunta dos vários perfis elétricos de poços (perfis de raios gama (GR), resistividade ou indução (ILD), porosidade neutrônica (NPHI) e densidade (RHOB)).

Agradecimentos

Este trabalho teve apoio financeiro da FAPERN (Fundação de Apoio a Pesquisa do Estado do Rio Grande do Norte) e a base de dados foi fornecida pela Agência Nacional do Petróleo (ANP) e contém dados de 54 poços pertencentes ao Campo Escola de Namorado, Bacia de Campos, Rio de Janeiro.

Referências Bibliográficas

IVANOVA, K., AUSLOOS, M. Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking. Phys. Rev. A. 274 (1999) 349-354.

JOHNSON, Richard A.; WICHERN, Dean W.. Applied Multivariate Statistical Analysis. 3rd ed. Prentice-Hall, 1992.

KANTELHARDT *et al.* *Detecting Long-range Correlations with Detrended Fluctuation Analysis*. Artigo científico, 10f. 2001.

LIMA, K. T. P., 2006. *Utilização de Métodos Sísmicos, Perfilagem e Testemunhos de Poços para Caracterização dos Turbiditos da Formação Urucutuca na Bacia de Almada, BA*. Universidade Estadual do Norte Fluminense Darcy Ribeiro / Laboratório de Engenharia e Exploração de Petróleo, Macaé.

LOURENÇO, Alexandre; MATIAS, Rui P.. *Estatística Multivariada*. Instituto Superior de Engenharia do Porto, 2000.

MENEZES, S. X. & ADAMS, T. 1990. Ocorrência de resistividades anômalas no Campo de Namorado, Bacia de Campos. Boletim de Geociências da Petrobrás, 4: 183-188.

MINGOTI, S.A.. *Análise de Dados Através de Método de Estatística Multivariada*, editora UFMG, Belo Horizonte – MG. 2007. p.155-211.

ROCHA, A.C. B. & SOUZA, F. A. M. *Aplicação do K- Means Cluster a Dados de Perfis de Poços Petrolíferos*. 2º Congresso Brasileiro de P & D em Petróleo e Gás. Rio de Janeiro, 2003.

THOMAS, J. E. (organizador) *Fundamentos de Engenharia de Petróleo*. Rio de Janeiro: 2ª ed. Interciência, 2004.