



Aplicação da regra fuzzy para reconhecimento de padrões em perfis de poços

Ana Carolina G. T. Magalhães(*), Jadir da Conceição da Silva, Universidade Federal do Rio de Janeiro (ANP-PRH18).

Copyright 2003, SBGf - Sociedade Brasileira de Geofísica

This paper was prepared for presentation at the 8th International Congress of The Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 14-18 September 2003.

Contents of this paper were reviewed by The Technical Committee of The 8th International Congress of The Brazilian Geophysical Society and does not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction, or storage of any part of this paper for commercial purposes without the written consent of The Brazilian Geophysical Society is prohibited.

Abstract

The fuzzy pattern recognition method has been used for the interpretation of some geophysical exploration data. Amongst the geophysical exploration methodologies, borehole logging has an effective role in the location and evaluation of geological resources. The great challenge for the scientists is to segment these logs in areas that identify some oil and gas reservoir's patterns. The main objective of this project is to use fuzzy recognition method in Namorado oil field, Campos' Basin well loggings. In the first part of this project, the fuzzy method is applied using the MATLAB software, to identify areas with shale's pattern. Used the borehole synthetic and real data are GR, PHIN, ILD, RHOB and DT; in them defined the shale standard feature. Obtained results are consonant with previous lithologic interpretation. For the sequence of this project, the Fuzzy method will again be used two times to segment and identify, in well logging, oil-saturated rock.

Introdução

A atividade de exploração de petróleo demanda altos riscos e grandes investimentos. Quanto mais ferramenta dispuser para minimizar os riscos sem adicionar custos, melhor para a empresa interessada na exploração. A metodologia proposta neste trabalho é apresentada como uma ferramenta adicional que, juntamente com outras disponíveis, aumenta a confiabilidade da avaliação de reservatórios potenciais.

Os perfis de poços têm grande aplicação na localização e avaliação de recursos minerais, dentre eles, o petróleo. A segmentação destes perfis em zonas capazes de identificar certos padrões encontrados nos reservatórios de petróleo e gás pode ser feita por análise discriminante, redes neurais, análise de "clusters". A metodologia fuzzy, utilizada neste trabalho, foi originalmente proposta por Bezdek (1980), como aperfeiçoamento dos métodos de agrupamento já existentes. Ela é uma técnica numérica de seleção e agrupamento de dados que apresentam certo grau de similaridade entre si. No caso das geociências, dizemos que cada dado no espaço n -dimensional pertence a uma fácies, cujo grau de similaridade é especificado por uma matriz de partição ou "membership grade". O agrupamento se processa a partir de um número ótimo de grupos (fácies), criteriosamente selecionado. Foi adotado o critério de mínimo valor da

medida de "Validade Fuzzy", proposta por Fukuyama & Sugeno (1989). Os perfis utilizados foram: Raios Gama Naturais (GR), Resistividade (ILD), Porosidade Neutrônica (PHIN), Densidade (RHOB) e Tempo de Trânsito (DT) do perfil sônico convencional (Ellis, 1987), primeiramente em perfis sintéticos e depois em dados reais, provenientes de um poço da Bacia de Campos.

A regra, em si, começa com a seleção dos valores centrais dos grupos, para cada um dos perfis. Para isso, gerou-se uma matriz de partição fuzzy aleatória, geradora dos valores iniciais dos centros dos grupos, obviamente incorretos. Com repetidas atualizações da matriz de partição fuzzy, a regra move iterativamente os centros dos grupos para os valores corretos. Esta iteração é baseada na minimização de uma função objetiva que relaciona a distância Euclidiana de cada ponto com os centros dos grupos, ponderados por cada elemento da matriz de partição. Após concluir a seleção dos centros de cada grupo, são propostas variáveis lingüísticas para caracterizá-los. Então, o padrão a ser identificado é comparado, ponto a ponto, com estes centros. As variáveis lingüísticas para os centros dos cinco grupos são: MB (Muito Baixo), B (Baixo), M (Médio), A (alto) e MA (Muito Alto).

Na primeira etapa deste trabalho, já concluída, definiu-se o "padrão ou fácies folhelho", a fim de identificar os segmentos dos perfis que apresentem este padrão. Com isto, estas regiões, consideradas como sendo folhelho, podem ser descartadas, restando então os possíveis reservatórios. Foram formulados perfis sintéticos, com litologias pré-definidas, para desenvolvimento da metodologia de identificação de folhelhos. A técnica elaborada foi então aplicada em dados reais de um poço. Nas etapas seguintes do projeto, pretende-se utilizar a regra fuzzy, separando hidrocarboneto de água e, por fim, fracionando óleo e gás, identificando assim, o padrão rocha saturada em óleo, que representa o objetivo principal deste trabalho.

A regra fuzzy

A regra fuzzy parte do princípio de que se um grupo de dados x_k tem vários padrões dentro de si, sendo que cada grupo de padrão individual pode ser agrupado em torno de um centro. Desta forma, espera-se que a distância dos pontos naquele grupo possa ser correlacionada através de uma seqüência fuzzy a qual descreve uma matriz de partição U ou "membership grade", satisfazendo as seguintes condições:

$$\mu_{ik} \in [0,1], \text{ onde } 1 \leq i \leq c \text{ e } 1 \leq k \leq n; \quad (1)$$

$$\sum_{i=1}^c \mu_{ik} = 1; \quad (2)$$

$$0 \leq \sum_{k=1}^n \mu_{ik} \leq n, \quad (3)$$

onde c é o número de grupos ou fácies selecionado, e n é o número de dados. A condição dada pela equação (1) significa que os dados pertencem a vários grupos a diferentes graus. Já as condições dadas pelas equações (2) e (3) requerem apenas que o grau total da matriz de partição para cada dado seja normalizado a 1, e não podem pertencer a mais fácies do que o número existente.

A localização de um grupo é representada pelo seu valor central v_{ij} em torno do qual seus p elementos estão concentrados. O critério usado para melhorar a partição inicial é o critério da variância. Neste caso, mede-se a dissimilaridade entre os pontos em um grupo e seu valor central pela distância Euclidiana, d_{ik} , que, de acordo com Bezdek & Pal (1992), é dada por:

$$d_{ik} = \|x_k - v_i\| = \left[\sum_{j=1}^p (x_{kj} - v_{ij})^2 \right]^{1/2}. \quad (4)$$

A partição Fuzzy, de acordo com este critério, é feita minimizando o seguinte funcional ou função objetiva:

$$F(\mu, v) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m (d_{ik})^2, \quad (5)$$

o que resulta em

$$\mu_{ik} = \left[\frac{1}{\|x_k - v_i\|^2} \right]^{1/(m-1)} \left/ \sum_{j=1}^c \left[\frac{1}{\|x_k - v_j\|^2} \right]^{1/(m-1)} \right., \quad (6)$$

tal que

$$v_i = \frac{1}{\sum_{k=1}^n (\mu_{ik})^m} \sum_{k=1}^n (\mu_{ik})^m x_k, \quad \text{para } 1 \leq i \leq c. \quad (7)$$

O sistema descrito pelas equações (4), (5), (6) e (7) não pode ser resolvido analiticamente. A solução recai, então, na aplicação de técnicas de algoritmos iterativos que aproximam o mínimo do funcional a partir de uma dada posição. Um dos algoritmos mais conhecidos para solucionar tal problema é o algoritmo ISODATA, proposto por Bezdek (1980) onde, para cada $m \in (0, \infty)$, resolve-se iterativamente as condições necessárias dadas pelas

equações (1), (2) e (3), e ao mesmo tempo converge para um ótimo local da equação (5).

O algoritmo ISODATA pode ser condensado em 6 passos básicos:

PASSO 1: Seleciona-se o número ótimo de grupos c ($2 \leq c \leq n$) para o conjunto de dados disponível e o expoente m ($1 \leq m \leq \infty$), normalmente tomado como 2. O valor ótimo de c é calculado utilizando a seguinte medida de validade Fuzzy, proposta por Fukuyama & Sugeno (1989)

$$S(c) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^m (\|x_k - v_i\|^2 - \|v_i - \bar{x}\|^2), \quad (8)$$

onde x_k é o k -ésimo vetor de dados, v_i é o vetor valor central do i -ésimo grupo e \bar{x} é a média do dado. Os dois termos entre parênteses representam a variância dos dados dentro de cada grupo e a variância do próprio grupo, respectivamente. Assim, o número ótimo de grupos é determinado como o mínimo do funcional $S(c)$ quando c cresce; A Figura 1 indica a aplicação da medida de validade Fuzzy no dado real de um poço da Bacia de Campos utilizado neste trabalho, sendo definido o número ótimo de grupos como sendo 5.

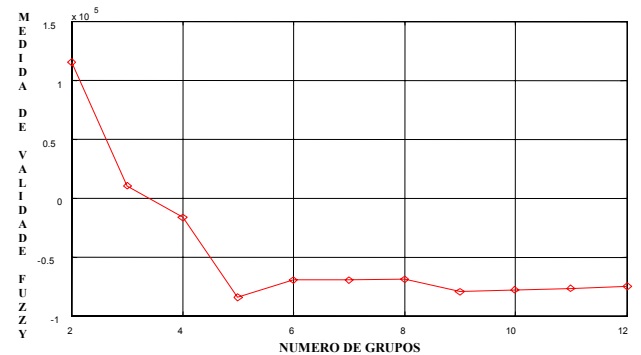


Figura 1: Gráfico do critério do mínimo valor da “validade Fuzzy”, utilizado nos dados reais deste trabalho, segundo a qual definiu-se o número ótimo de grupos como sendo 5.

PASSO 2: Inicializa-se a matriz de partição Fuzzy $U^{(0)}$, de forma aleatória, fazendo o índice de iteração $l = 0$;

PASSO 3: Faz-se $l = l + 1$;

PASSO 4: De posse dos valores de c , m e da matriz $\mu_{ik} = U^{(l-1)}$, calcula-se os centros de cada grupo usando a equação (7);

PASSO 5: Calcula-se a nova matriz de partição Fuzzy $U^{(l)}$ através da equação (6) e usando os valores dos centros dos grupos do PASSO 4.

PASSO 6: Escolhe-se uma norma matricial adequada e calcula:

$$\Delta = \|U^{(l)} - U^{(l-1)}\|, \text{ e um critério de parada } (\Delta \leq \epsilon).$$

Se $\Delta > \varepsilon$ vai para o PASSO 3;

Se $\Delta \leq \varepsilon \rightarrow$ FIM.

Aplicação

Após análise criteriosa, foram estabelecidas para o “Padrão-Folhelho” as seguintes variáveis: [A + MA] para o perfil GR, [MB + B] para o perfil ILD, [A + MA] para o perfil PHIN e [B + M + A] para os perfis RHOB e DT. Adicionalmente, vinculou-se a aceitação do “Padrão-Folhelho” a valores de frequências superiores a 2 e corte- α maior ou igual a 0,8, conforme sugerido por Das Gupta (1999).

O perfil sintético representa a convolução entre a “função resposta” de cada sonda de perfilagem e os valores reais de cada parâmetro a ser medido.

Para o perfil sintético foram definidas 4 eletrofácies: folhelho, arenito com óleo, arenito com água, arenito com gás. Foram simuladas para cada fácies litologias com composições e proporções mineralógicas diferentes, como exemplo, dentro da fácies folhelho, há folhelhos ricos em ilita e folhelhos ricos em montmorilonita.

Com o auxílio do software MATLAB™ R12 aplicou-se a regra fuzzy de acordo com os parâmetros definidos para o “padrão folhelho”, obtendo-se resultado coerente com a alocação dos folhelhos definidas no perfil sintético. Porém, devido à resolução limitada de cada sonda, as espessuras das camadas nem sempre coincidiram com as espessuras reais, diferindo em até 25% na resolução vertical das mesmas. A Figura 2 mostra um desenho esquemático com a resolução das ferramentas usadas na perfilagem.

A Figura 3 mostra gráficos Perfil *versus* Perfil oriundo dos dados sintéticos, destacando a distribuição dos próprios dados e o centro dos grupos.

A Figura 4 exibe a comparação entre o perfil litológico definido com dados sintéticos e o perfil identificando as zonas de folhelho através da aplicação da regra fuzzy. Note que o folhelho delgado (30cm) a 200cm de profundidade não é reconhecido devido ao problema de resolução. Como pode ser observado na Figura 2, a resolução da ferramenta que mede os raios gama naturais tem resolução de 30 cm aproximadamente prejudicando o reconhecimento.

Quanto aos dados reais, uma análise pormenorizada dos perfis GR e ILD, e das interpretações realizadas através de testemunhos, foi identificada a presença de algumas camadas de arenito com alto grau de radioatividade, confundindo-se com a eletrofácies folhelho no perfil GR. Para corrigir esta ambigüidade, um vínculo adicional foi imposto para a aceitação final dos pontos de folhelho: Resistividade < valor central da Fácies M do perfil ILD. Saliencia-se, no entanto, que, se este arenito radioativo estivesse saturado em água, também apresentaria resistividade baixa. Porém, o objetivo final do trabalho justifica este vínculo, pois o que se pretende é identificar rocha saturada em óleo, que apresenta resistividade alta, não tendo, então, prejuízo na identificação, já que esta porção seria efetivamente descartada nas etapas seguintes. Outra solução seria separar o perfil GR nos

espectros K-U-Th e eliminar-se a contribuição do urânio. Se ainda permanecesse radioatividade alta, poder-se-ia indicar este segmento como sendo folhelho.

A Figura 5 mostra gráficos Perfil *versus* Perfil, destacando a distribuição dos dados e o centro dos grupos. Nota-se no gráfico GR *versus* ILD, a dificuldade de definição dos grupos, devido à presença do arenito radioativo saturado em óleo.

A Figura 6(a) indica os perfis de poço utilizados e a interpretação litológica sugerida com base, também, em testemunhos. A Figura 6(b) indica o resultado obtido, aplicando-se a metodologia fuzzy, sendo a cor verde representativa das camadas de folhelho, objetivo desta etapa do trabalho. As análises visuais dos resultados comprovam a eficácia da metodologia fuzzy na identificação de folhelhos em perfis de poço. Os trechos em branco representam eletrofácies a serem futuramente identificadas na seqüência do trabalho.

Conclusões

Os resultados obtidos com a metodologia fuzzy foram coerentes na indicação de folhelhos nos dados sintéticos e com as análises visuais e de testemunhos feitas anteriormente para o poço real. Isto comprova a eficácia da metodologia fuzzy na identificação de folhelhos em perfis de poço, atendendo ao objetivo de tornar-se uma ferramenta com a qual se possa segmentar os perfis de poços em zonas que atendam padrões pré-estabelecidos. Na seqüência deste projeto, as zonas identificadas como folhelho serão descartadas. Sobre as regiões restantes, aplicar-se-á a metodologia fuzzy para identificar o padrão “rocha saturada em hidrocarboneto”. Nesta porção, separar-se-á óleo de gás. Todas estas etapas de aplicação serão realizadas tanto em perfis sintéticos, quanto em perfis reais. Depois de concluída, a técnica proposta neste trabalho será aplicada em vários perfis de poços do Campo de Namorado, Bacia de Campos, a fim de melhor testar a eficácia do método e aperfeiçoá-lo de acordo com as variáveis que venham a surgir, disponibilizando-o a fim de que se torne mais um fator que auxilie na avaliação de potenciais reservatórios de petróleo.

Referências

- Bezdek, J. C.**, 1980, A convergence theorem for the fuzzy ISODATA clustering algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2, p 1-8.
- Bezdek, J. C. and Pal, S. K.**, 1992, Fuzzy Models for Pattern Recognition Method. IEEE Press, New York, 539pp.
- Das Gupta, S. P.**, 2001, Application of a fuzzy pattern recognition method in borehole geophysics, Computers & Geosciences 27, p 85-89.
- Ellis, D. V.**, 1987, Well Logging for Earth Scientists. Elsevier, New York, p532.
- Fukuyama, Y. and Sugeno, M.**, 1989, A new method of choosing the number of clusters for fuzzy c-means method. Proc. 5th Fuzzy System Symposium, p 247-250.

Agradecimentos

Os autores agradecem a Agência Nacional do Petróleo (ANP) pelo suporte técnico e financeiro, que, através do programa de capacitação de recursos humanos PRH18, possibilitam a realização deste projeto.

Agradece-se, também, o apoio do estudante Maximiliano de Seixas Gonçalves pela consultoria na área de informática.

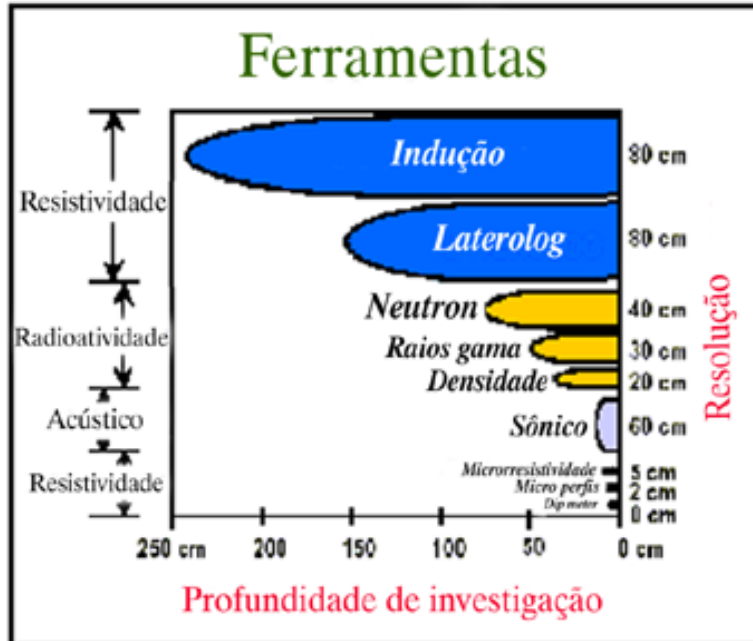


Figura 2: Figura representativa das resoluções dos diferentes tipos de ferramentas.

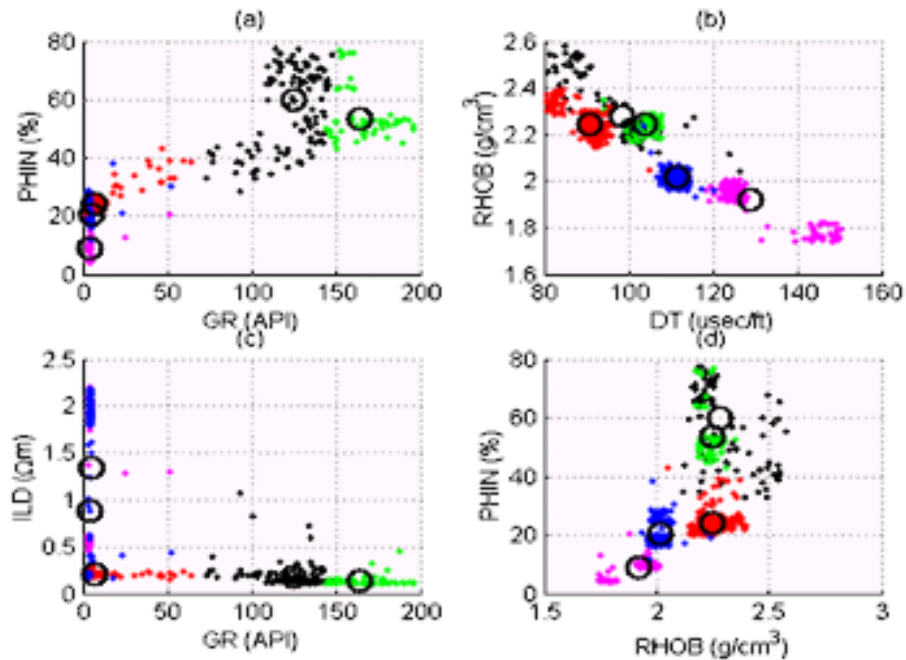


Figura 3: Figuras representativas perfil versus perfil, sintéticos, gerados pelo aplicativo MATLAB, indicando a distribuição dos pontos pelos grupos, sendo o círculo negro indicativo do centro dos grupos.

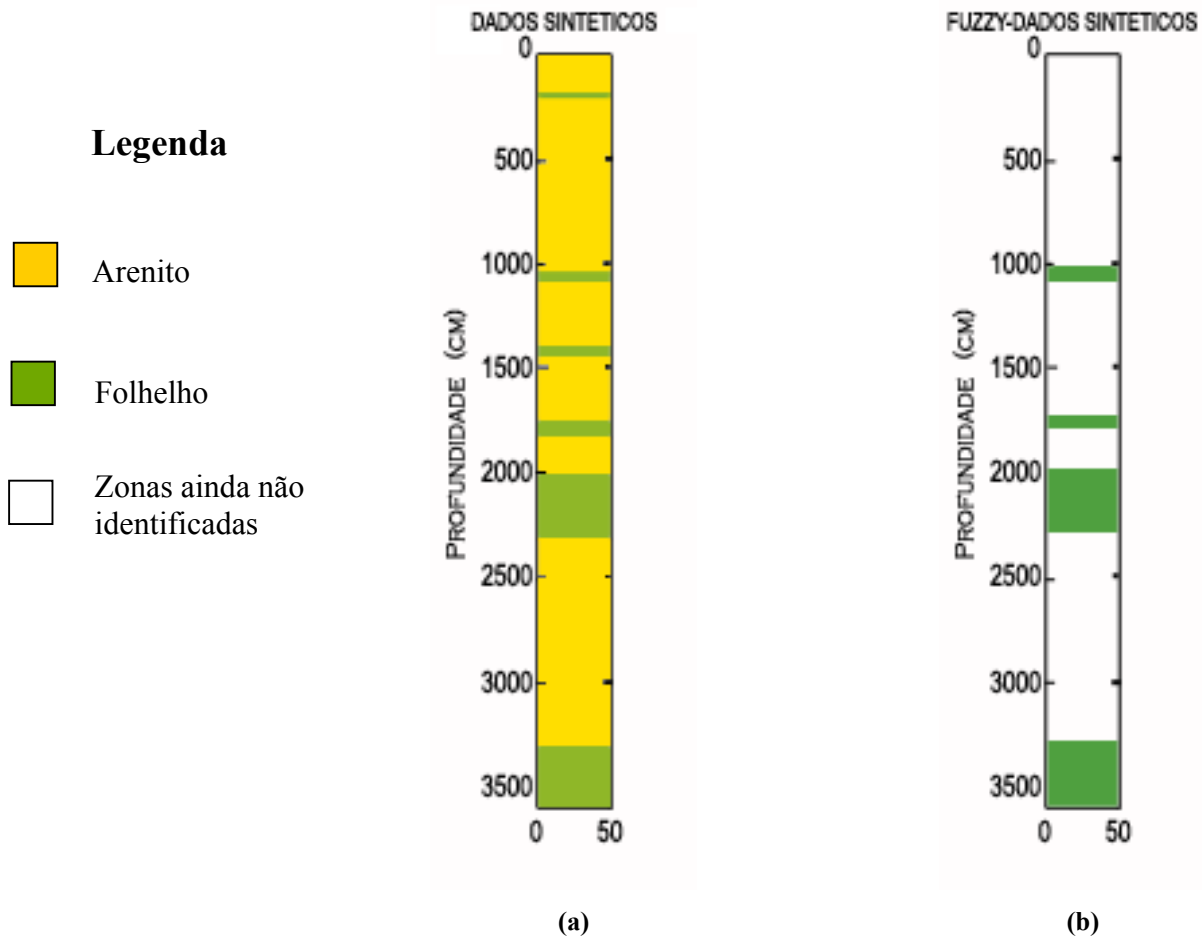


Figura 4: A Figura 4(a) mostra o perfil litológico gerador dos dados sintéticos. A Figura 4(b) mostra a coluna obtida ao aplicar-se à metodologia fuzzy aos dados sintéticos, identificando com a cor azul as zonas de folhelho.

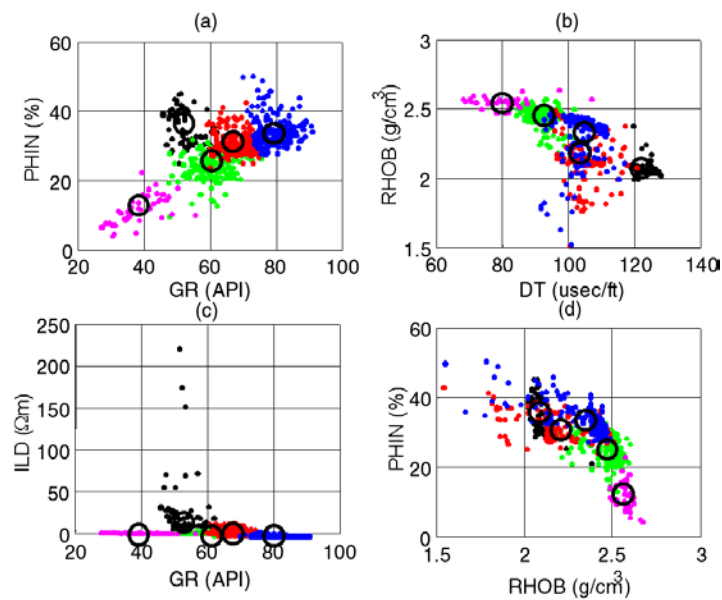


Figura 5: Figuras representativas perfil versus perfil, gerados pelo aplicativo MATLAB, indicando a distribuição dos pontos pelos grupos, sendo o círculo negro indicativo do centro dos grupos. Na Figura 5(c) percebe-se a dificuldade de definição destes dados nos grupos, relativo à presença do arenito radioativo.

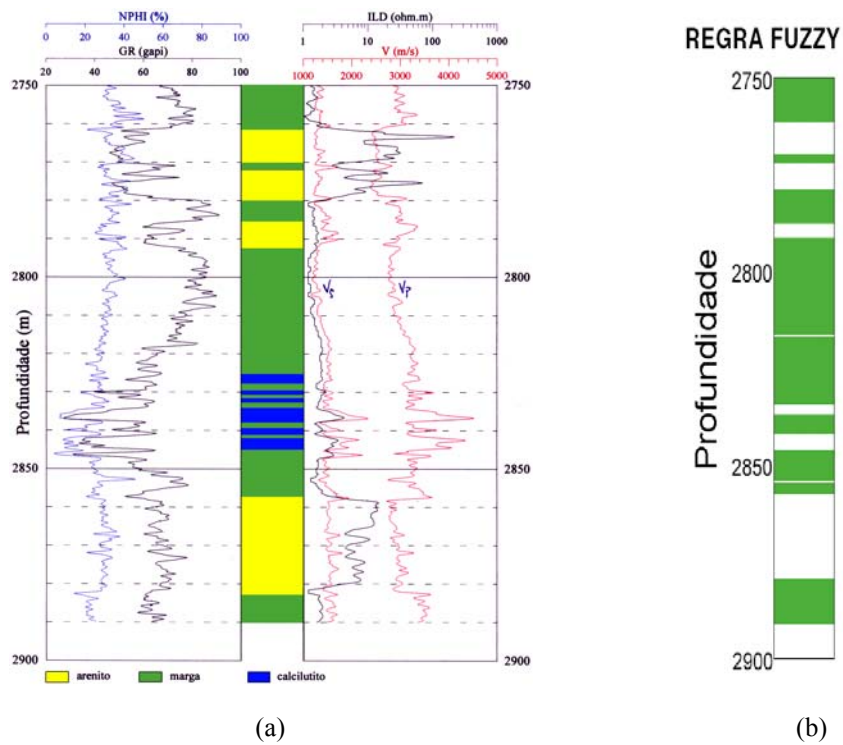


Figura 6: A Figura 6(a) mostra a coluna interpretada a partir dos dados dos perfis em conjunto com dados de testemunho, além de mostrar os perfis utilizados. A Figura 6(b) mostra a coluna obtida ao aplicar-se à metodologia fuzzy, identificando com a cor verde as zonas de folhelho.