



Identificação Litológica com Affinity Propagation

Nayara Safira Da Silva Caldas*, UFOPA

Carlos Eduardo Guerra, UFOPA

André Andrade, UFPA

Copyright 2016, SBGf - Sociedade Brasileira de Geofísica

Este texto foi preparado para a apresentação no VII Simpósio Brasileiro de Geofísica, Ouro Preto, 25 a 27 de outubro de 2016. Seu conteúdo foi revisado pelo Comitê Técnico do VII SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

Resumo

This work presents a methodology for solving the problem of the identification of lithologies at depth, directly from wireline logs. For this purpose, the *M-N* cross-plot is used as the formation evaluation technique for identifying the lithologies in the logged borehole, in terms of the physical properties of the main mineral. However, the visual interpretation of this graphic is limited by the large spread of points in the graph. The proposed method is algorithm for a computational interpretation of the *Vsh-M-N* data by means of the Affinity Propagation (AP) clustering technic. A minimum distance criterion, with respect to *M-N* fixed points, is applied to the set of exemplary points, acquired from AP, in order to associate the lithologies. Finally, the methodology was applied on a real *M-N* data, proving its effectiveness, both in reducing the number of groupings obtained by the clustering algorithm as well as in the lithological identification over highly-spread *M-N* data set.

Introdução

A avaliação de uma formação depende fundamentalmente da informação petrofísica e geológica produzida pela análise de testemunhos. A informação geológica obtida em um perfil de poço permite a correlação e / ou atualização estratigráfica do campo de petróleo e tem duas aplicações básicas para a avaliação de formação: a primeira trata do cálculo de porosidade, o que requer informações sobre os parâmetros de porosidade da matriz rochosa [1]. A segunda aplicação refere-se ao zoneamento do poço, ou a identificação das profundidades de topo e base de diferentes camadas [2].

Alguns fatores dificultam os procedimentos de testemunhagem tais como a inviabilidade técnica de realização da aquisição em poços horizontais, o alto custo das operações de testemunhagem e o factor de recuperação, que é a impossibilidade de recuperação de 100% da rocha retida no amostrador em uma operação de testemunhagem. Na ausência de tal informação, um grande número de técnicas foram desenvolvidas para a identificação de litologias [3-5]. Uma interpretação quantitativa destes perfis de poços é a interpretação do Gráfico *M-N* que, em princípio, proporcionaria uma rápida interpretação visual da litologia de qualquer ponto ao longo da profundidade de um poço. No entanto, a presença de ruído nos perfis de porosidade e modelo de rocha simplificado que desconsidera a argilosidade, torna a interpretação visual do Gráfico *M-N* complexa e ambígua.

Neste artigo, apresentamos um método que visa minimizar as limitações na interpretação do Gráfico *M-N*, e produzir a identificação automática de litologias através da utilização do algoritmo de agrupamento Affinity Propagation para processamento dos dados. O resultado é a simplificação dos dados a serem interpretados pela determinação de exemplares para os agrupamentos resultantes e, posteriormente, a classificação de litologias em profundidade.

Fundamentos teóricos

Gráfico *M-N*

O gráfico *M-N* [3] resulta de uma particular combinação dos três perfis de porosidade (sônico, densidade e porosidade neutrônica) para produzir a identificação litológica de um trecho de poço perfilado, em termos das propriedades físicas do mineral principal de cada litologia. O Gráfico *M-N* é construído com pares ordenados, definidos pelos parâmetros *M* e *N*, que são formulados de modo a torna-los relativamente independentes do efeito da porosidade sobre os registros dos perfis de porosidade. O modelo de rocha considerado para a definição dos parâmetros *M* e *N* desconsidera a presença da argila ($V_{sh} = 0$) na constituição da rocha reservatório. Burke [3] define os parâmetros *N* e *M* por:

$$N = \frac{\phi_{Nmf} - \phi_{Nb}}{\rho_b - \rho_{ma}} \quad (1)$$

$$M = \frac{\Delta t_{mf} - \Delta t_b}{\rho_b - \rho_{ma}} \times 0,01 \quad (2)$$

onde ρ_{ma} , ϕ_{Nmf} , Δt_{mf} e ρ_b , ϕ_{Nb} , Δt_b são os parâmetros da matriz e os dados dos perfis, respectivamente.

Na Eq.(2) é introduzido o fator de escala 0,01.

Interpretação do Gráfico M-N

O Gráfico M-N é construído por pontos formados por pares ordenados (N, M) calculados para os principais minerais constituintes das rochas sedimentares. Cada mineral produz um único ponto, denominados pontos fixos do Gráfico M-N (**Fig.1**). Admite-se que os pontos fixos apresentem valores dos parâmetros M e N suficientemente diferentes, tal que cada ponto fixo ocupe uma posição relativamente distante dos outros, de modo a permitir a classificação das litologias em função das propriedades físicas do seu mineral principal [1].

Os pares (N,M) para os pontos em profundidade de um poço perfurado são calculados a partir da substituição dos parâmetros de matriz (ρ_{ma} , ϕ_{ma} , Δt_{ma}) pelos dados dos perfis nas equações (1) e (2). Estes pontos se distribuem no plano N-M aproximadamente nas vizinhanças dos minerais principais de cada litologia. A identificação da litologia é produzida pela proximidade do ponto do perfil ao ponto fixo do mineral principal da sua litologia.

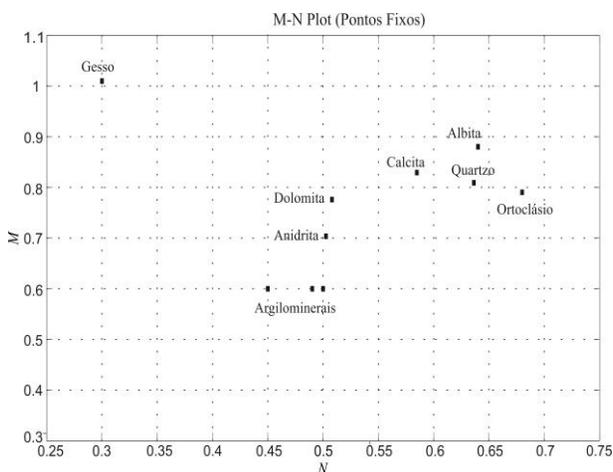


Figura 01 - Gráfico M-N. Os quadrados em preto representam os minerais mais comuns com seus respectivos valores de M-N.

GRÁFICO VSH-M-N

O Gráfico Vsh-M-N, possui eixos escalonados pelos parâmetros M e N e pela argilosidade (Vsh), calculada pelo perfil de raios gama natural. O Gráfico Vsh-M-N mostra uma distribuição espacial de pontos em função da constituição mineralógica e da argilosidade das rochas reservatório, mas são perdidas as relações de profundidade entre as diversas camadas atravessadas pelo poço.[1].

Cluster analysis

Os cross-plot de perfis de poços são sempre utilizados para mapeamento de fácies e litologia e frequentemente apresentam uma tendência para formar grupos [3, 4]. Esses grupos podem ser identificados e classificados por um processo de classificação que utiliza técnicas de avaliação de formação e análise de agrupamento.

A análise de agrupamento é uma espécie de processo não supervisionado ou uma técnica que divide um conjunto de objetos em grupos homogêneos [6-8]. Esta técnica pode ser usada para reduzir o tamanho de um conjunto de dados, reduzindo um grande número de objetos aos principais representantes do conjunto [9]. Estas técnicas de agrupamento consistem em reunir objetos com características comuns em grupos (clusters), com características que os diferem entre outros. Existem vários algoritmos que visam dividir um espaço de dados amostra em grupos cujos elementos têm algumas características semelhantes. A maior parte dos algoritmos utilizar um critério pré-estabelecido de dissimilaridade. Normalmente, a métrica utilizada é a distância euclidiana, mas outras métricas podem ser usadas, de preferência as que melhor refletem alguma característica comum dos dados.

Affinity propagation

A maioria dos algoritmos de agrupamento usam como parâmetro de entrada, um número pré-determinado k de clusters para particionar o espaço amostral de dados. O Affinity Propagation (AP) [10, 11] adota o princípio de que todos os pontos de dados pode ser eleitos como um "exemplar do agrupamento". O conjunto de dados forma uma configuração de rede em que os pontos representam os nós e as transmissões de mensagens ocorrem entre as arestas da rede [12]. O objetivo é otimizar a convergência de uma função de similaridade para a determinação de um conjunto de exemplares que serão os representantes do grupo ao qual estão vinculados.

AP tem basicamente uma entrada, os valores das similaridades $s(i, k)$ entre cada par de pontos $\{x_i, x_k\}$, em que estas semelhanças indicam como o ponto x_k

seria apto para representar o ponto x_i . Quando o objetivo é minimizar o quadrado do erro, cada similaridade é definido como o negativo do quadrado do erro (norma euclidiana)

$$s(i, k) = -\|x_i - x_k\|^2.$$

Além das similaridades, Affinity Propagation também tem como parâmetro de entrada valores $s(k, k)$ denominados "preferências" para cada ponto k . Sabendo, que em qualquer métrica a diagonal principal da matriz de similaridades é nula, o que indica que a maior similaridade ocorre de um ponto a ele mesmo, substituímos este valor por $s(k, k) \neq 0$ que indica preferências para esses pontos serem exemplares. Esses valores de preferência interferem nos resultados (número de clusters) encontrados pelo AP, altos e baixos valores de $s(k, k)$ resultam em grandes e baixos números de clusters encontrados, respectivamente.

A troca de mensagens entre os pontos podem ser de dois tipos, transmissão de responsabilidades e disponibilidades. As responsabilidades $r(i, k)$, mensagens enviadas de um ponto i para um possível exemplar k indicam o quão adequado seria o ponto k ser exemplar para o ponto i . As disponibilidades $a(i, k)$, enviadas do candidato a exemplar k para o ponto i , indicam evidências de quão adequado seria para o ponto i escolher o ponto k como seu exemplar.

As responsabilidades são definidas pela seguinte regra:

$$r(i, k) \leftarrow s(i, k) - \max_{k' \neq k} \{a(i, k') + s(i, k')\}.$$

As disponibilidades, que definem se um exemplar é um bom exemplar, são assim definidas:

$$\begin{cases} a(i, k) \leftarrow \min \left\{ 0, r(k, k) + \sum_{i' \in \{i, k\}} \max \{0, r(i', k)\} \right\}, \\ a(k, k) \leftarrow \sum_{i' \neq k} \max \{0, r(i', k)\}. \end{cases}$$

Assim, a definição de um exemplar se dá quando a combinação de responsabilidades e disponibilidades, segundo abaixo, assume valor máximo.

$$AP = \max \{a(i, k) + r(i, k)\}.$$

Metodologia/ Problema Investigado

Na avaliação de formação, em função da rarefação de poços testemunhados, utilizam-se os dados dos perfis geofísicos codificados em Gráfico M-N para a identificação de litologias. Em razão das simplificações adotadas no modelo de rocha do Gráfico M-N, como a ausência de argilosidade e a ocorrência de um único mineral na constituição da matriz, o que na maioria dos casos práticos não são contemplados pelas rochas reais, torna-se ambígua a identificação de litologias. Assim, os pontos definidos pelos pares ordenados (N, M) calculados para um intervalo perfilado apresentam um grande espalhamento, que os afasta dos pontos de referência e dificulta a interpretação visual do Gráfico M-N. Uma forma de atenuar as distorções geradas pela desconsideração da argilosidade é adotar o Gráfico Vsh-M-N e realizar a identificação litológica em termos das propriedades físicas do mineral principal e da argilosidade na constituição da rocha reservatório. Na prática a interpretação do Gráfico Vsh-M-N deve ser guiada pela informação geológica local de modo a serem eliminados os pontos de referência que representam minerais ausentes ou que apareçam apenas como traços na constituição das rochas presentes.

O problema da interpretação computacional do Gráfico Vsh-M-N pode ser estabelecido na forma de um problema de agrupamento (clustering), de modo a determinar um ponto do gráfico que seja representativo de um grupo ou conjunto de pontos com características semelhantes. Esta abordagem busca reduzir o número de pontos a serem classificados em função dos pontos fixos do Gráfico M-N. Diferente de um problema geral de agrupamentos, a interpretação computacional do Gráfico Vsh-M-N requer o estabelecimento da coerência geológica e da continuidade em profundidade de uma mesma camada.

A metodologia aqui apresentada pode ser dividida em duas etapas. A primeira etapa representa a interpretação computacional do Gráfico Vsh-M-N com a utilização do algoritmo Affinity Propagation. Considerando a disponibilidade dos perfis de porosidade para o cálculo dos parâmetros M e N e do perfil de raios gama natural para o cálculo da argilosidade (Vsh) preenche-se o Gráfico Vsh-M-N com os pontos do intervalo perfilado. Assim os pontos do Gráfico Vsh-M-N formam os dados de entrada no algoritmo Affinity Propagation. A operação natural do algoritmo Affinity Propagation tem como resposta a determinação dos denominados exemplares, ou particulares pontos do conjunto de dados que representam alguns pontos com características semelhantes. Assim, o resultado desta etapa é a identificação dos exemplares no Gráfico Vsh-M-N que são representativos das rochas reservatório e das rochas selantes. Com isto, tem-se uma considerável redução do volume de dados a serem tratados nos futuros processamentos. Tratando apenas com os exemplares, nesta etapa, procede-se à eliminação dos exemplares representativo das rochas selantes e rochas

reservatório com alta argilosidade, com a adoção de um valor de corte na argilosidade igual a 60%. Este procedimento visa eliminar todas as possíveis camadas com baixa permeabilidade, que poderiam, eventualmente, resultar em reservatórios de baixa qualidade.

A segunda etapa trata especificamente da identificação das litologias das rochas reservatório. Considerando apenas os exemplares que não foram eliminados pelo corte da argilosidade, realiza-se a aplicação do algoritmo Affinity Propagation para produzir a interpretação computacional do Gráfico M-N. Para a interpretação do Gráfico M-N, os dados de entrada para a construção da matriz de similaridades do algoritmo Affinity Propagation são tomados como a projeção dos exemplares remanescentes sobre o plano M-N para produzir a sua classificação ou a identificação de litologias em função dos pontos fixos do Gráfico M-N compatíveis com a geologia local. Ainda persiste a limitação do Gráfico M-N para a correta identificação litológica das rochas com matrizes constituídas por vários minerais em proporções volumétricas semelhantes. Na metodologia aqui apresentada, adota-se como a litologia de uma camada sedimentar, aquela correspondente ao mineral em maior proporção volumétrica na sua constituição. Assim, uma rocha com matriz composta proporcionalmente por 40% de calcita e 60% de quartzo terá a sua identificação litológica apontada como arenito.

Resultados

A metodologia foi empregada em dados sintéticos e reais, que satisfazem a equação geral dos perfis de porosidade. No entanto, neste trabalho consta apenas um caso real. O dado real apresentado aqui é de um poço do campo de namorado na bacia de Campos, Brasil.

Fig. 2 mostra um conjunto de perfis registrados na Bacia de Campos. Os testemunhos estão descritos na **Tabela 1** para os intervalos de profundidade selecionados (realçados em verde e cinza nas curvas dos perfis).

Tabela- 1 Descrição Geologica

Camada	Litologia	Composição Mineral
Verde	Arenito	Quartzo
Cinza	Folhelho Radioativo	Argilominerais

Fig. 3 apresenta o gráfico de M-N baseado nos perfis apresentados na **Fig. 2**. Os círculos em preto representam os pontos (N, M) e os quadrados pretos os principais minerais de pontos fixos. A interpretação

visual do Gráfico M-N sugere a existência de três grupos compactos e bem definidos. No entanto, a interpretação visual não necessariamente reflete o número de diferentes litologias in situ. Um problema comum na avaliação de formação é o deslocamento dos pontos (N, M) devido a argilosidade na direção dos argilominerais, impedindo uma classificação mais fiel.

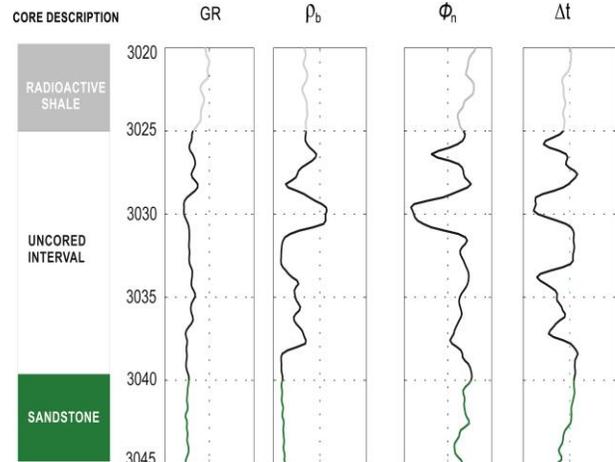


Figura 2- Perfis de dado real.

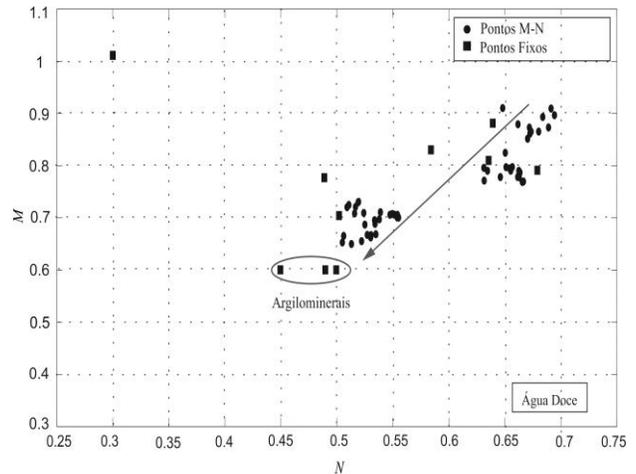


Figura 3- Gráfico M-N com três agrupamentos sugeridos pela interpretação visual.

A **Fig. 4** mostra o correspondente Gráfico Vsh-M-N. O eixo da argilosidade foi obtido com a utilização do perfil de raios gama natural [1]. Aqui a interpretação visual indica a presença de duas camadas, uma com alta argilosidade, podendo indicar um possível folhelho, a outra com baixo volume de argila, indicando uma possível camada reservatório.

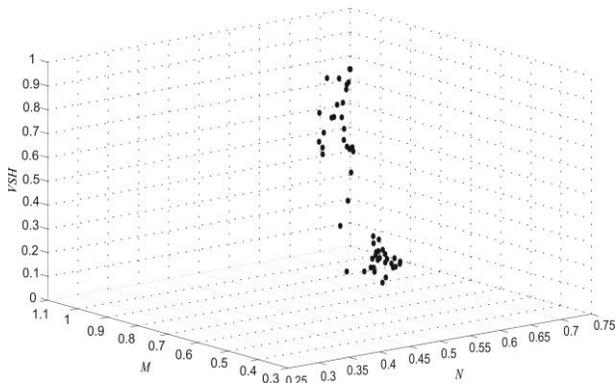


Figura 4- Gráfico Vsh-M-N para os perfis de dados reais.

A **Fig. 5** mostra o resultado da aplicação do algoritmo Affinity Propagation para a interpretação do Gráfico Vsh-M-N mostrado na **Fig. 4**. Em acordo com a interpretação visual do Gráfico Vsh-M-N, o algoritmo AP resultou na identificação de dois exemplares representativos das camadas reservatório (quadrados azul claro e amarelo). A **Fig. 5** permite apenas a determinação do número de camadas, mas não a sua classificação litológica. Assim, tem-se determinado pelo algoritmo Affinity Propagation o número de camadas reservatório.

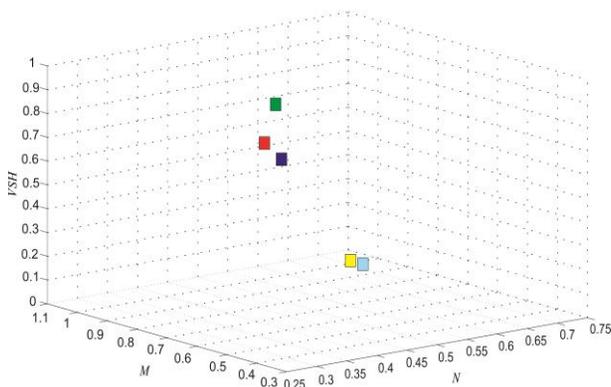


Figura 5- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.

A **Fig 6** apresenta os exemplares determinados pelo AP sobre o Gráfico M-N. A interpretação visual sugere duas camadas reservatório (azul claro/amarelo) e a ambiguidade referente aos 3 exemplares das rochas selantes (Gráfico Vsh-M-N) mas que poderiam ser interpretadas como Dolomita/Anidrita na análise do Gráfico M-N.

Mostra-se na **Fig. 7** os agrupamentos resultantes nas cores verde, vermelho e azul próximos aos argilominerais e, tendo analisado que estes pontos

possuem alta argilosidade, deduz-se assim, que esses pontos são referentes a rocha selante (folhelho). Os clusters em amarelo e azul claro estão aproximados do mineral principal quartzo, indicando a litologia referente à rocha reservatório arenito.

A identificação litológica das camadas atravessadas pelo poço perfurado é feita pela análise da distância euclidiana entre os exemplares e os pontos de minerais principais das litologias que são evidenciadas na bacia dispostos no Gráfico M-N.

A classificação final pode ser visualizada na **Fig.8 (a)**. A **Fig.8(b)** mostra o perfil de Raio Gama onde estão destacadas as camadas em profundidade. O primeiro intervalo, em profundidade, está associado aos minerais componentes do folhelho (argilominerais), o segundo intervalo é associado com o ponto fixo representante do mineral quartzo, principal componente do arenito.

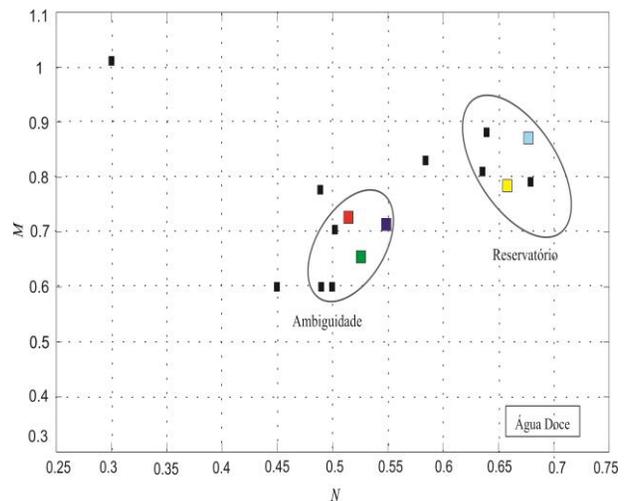


Figura 6- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.

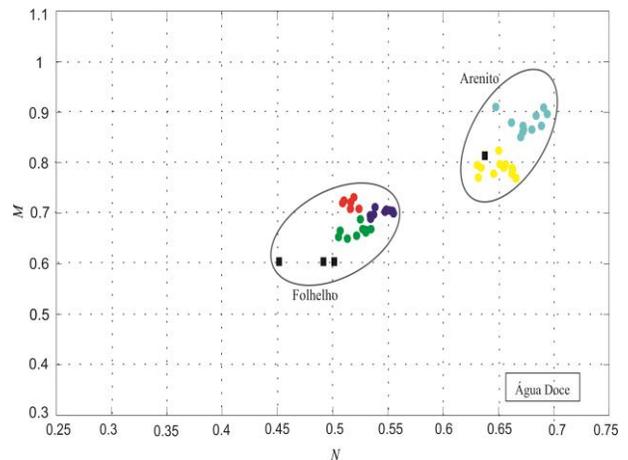


Figura 7- Representação dos agrupamentos pelo AP.

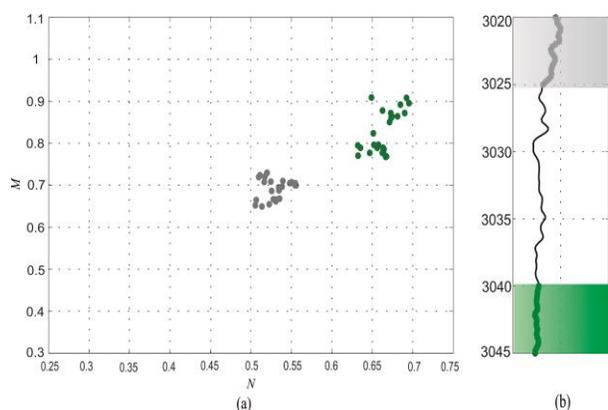


Figura 8- (a) Classificação final por associação de exemplar.(b) Mapeamento de litologias em profundidade.

Discussão e Conclusões

A identificação de litologias, no sentido estrito da geofísica de poço, é um problema bastante complexo de ser solucionado exclusivamente com os perfis geofísicos, uma vez que não é completamente válida a premissa de correspondência entre variação de propriedades físicas e litologias.

Uma interpretação quantitativa dos perfis é proporcionada pelo Gráfico M-N que, em princípio, poderia proporcionar uma rápida interpretação visual dos pontos ao longo das profundidades do poço. No entanto, a presença de ruído nos perfis geofísicos e o modelo de rocha simplificado que desconsidera a argilosidade, torna a interpretação visual do Gráfico M-N complexa e ambígua. Uma solução para este tipo de problema de classificação foi apresentada aqui um método não supervisionado que, independente da disponibilidade ou não dos testemunhos, é capaz de realizar a identificação e a distinção litológica das rochas atravessadas por um poço.

A representação de um cluster por seu exemplar, produzida pelo algoritmo Affinity Propagation na interpretação do Gráfico M-N resulta em uma redução significativa do conjunto de dados de conteúdo a ser processado, e atenua a ambiguidade relativamente a argilosidade presente no conjunto de dados M-N Plot ou dispersão devido ao ruído nos dados de poço.

As vantagens obtidas por meio do algoritmo Affinity Propagation é o fato de que não é necessário determinar a priori o número de agrupamentos, como um parâmetro de entrada, um requisito necessário na maioria dos métodos clássicos de agrupamento. Este requisito obriga à existência de algum conhecimento sobre dados a serem processados, o que, na maioria dos casos, não é possível. Outra vantagem de utilizar o AP em comparação com outro algoritmo de classificação é a substituição do centróide pelo exemplar, em que os centróides são representantes de

um cluster no gráfico, mas não são necessariamente pontos dos dados processados. Os exemplares, no entanto, são representantes dos clusters e ainda um dado relevante.

Em geral, a metodologia obteve resultados satisfatórios, e pode ser utilizada para a interpretação de dados de perfis de poços com o objetivo de definir os tipos de litologias atravessadas por poço.

Referências

- [1] C.A. Barros, A..In: SIMPOSIO BRASILEIRO DE GEOFÍSICA, Determination of water saturation by intelligent algorithm, in: SBGf (Ed.) SIMPOSIO BRASILEIRO DE GEOFÍSICA, SBGf, Belém, Para, Brazil, 2008.
- [2] D.V. Ellis, J.M. Singer, Well Logging for Earth Scientists, Springer, 2007.
- [3] J.A.C.J. Burke, R.L.; Schmidt, A.W., The Lithoposity cross plot: a method of determining rock characteristics for computation of log data., in: Society of Petroleum Engineers (Ed.) SPE ILLINOIS BASIN REGIONAL MEETING, Society of Petroleum Engineers, Illinois Basin.Evansville, Indiana., 1969, pp. p. 187- 198.
- [4] C. Clavier, D. Rust, MID-plot: A new lithology technique, The log analyst, 17 (1976) 16-24.
- [5] G.B. Asquith, C.R. Gibson, Basic well log analysis for geologists, American Association of Petroleum Geologists, 1982.
- [6] G. Gan, C. Ma, J. Wu, Data clustering: theory, algorithms, and applications, Society for Industrial and Applied Mathematics, 2007.
- [7] M. Steinbach, M., Karypis, G., & Kumar, V. 2000. A comparison of document clustering techniques.
- [8] P. Rai, S. Singh, A survey of clustering techniques, International Journal of Computer Applications, 7 (2010) 1-5.
- [9] H. Spath, Cluster analysis algorithms for data reduction and classification of objects, Ellis Horwood, Ltd., 1980.
- [10] B.J. Frey, D. Dueck, Clustering by passing messages between data points, Science, 315 (2007) 972-976.
- [11] B.a.D. Frey, D., Affnity Propagation | Frey Lab, in, 2007.
- [12] B.J. Frey, F.R. Kschischang, H.-A. Loeliger, N. Wiberg, Factor graphs and algorithms, in: Proceedings of the Annual Allerton Conference on Communication Control and Computing, UNIVERSITY OF ILLINOIS, 1997, pp. 666-680.