



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOFÍSICA
DISSERTAÇÃO DE MESTRADO

NAYARA SAFIRA DA SILVA CALDAS

IDENTIFICAÇÃO LITOLÓGICA COM *AFFINITY PROPAGATION*

Belém - Pará

2013

NAYARA SAFIRA DA SILVA CALDAS

IDENTIFICAÇÃO LITOLÓGICA COM *AFFINITY PROPAGATION*

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Geofísica do Instituto de Geociências da Universidade Federal do Pará - UFPA, em cumprimento às exigências para obtenção de Mestre em Geofísica.

Orientador: André José Neves Andrade

Coorientador: Carlos Eduardo Guerra

Belém - Pará

2013

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Caldas, Nayara Safira da Silva

Identificação litológica com *affinity propagation* / Nayara Safira da Silva Caldas; Orientador: André José Neves Andrade; Coorientador: Carlos Eduardo Guerra - 2013.

Dissertação (mestrado em geofísica) – Universidade Federal do Pará, Instituto de Geociências, Programa de Pós-Graduação em Geofísica, Belém, 2013.

1. Geofísica. 2. Informação geológica. 3. Identificação litológica. 4. Affinity Propagation. I. Andrade, André José Neves, *orient.* II. Guerra, Carlos Eduardo, *coorient.* III. Universidade Federal do Pará III. Título.

CDD 22º ed.: 550

NAYARA SAFIRA DA SILVA CALDAS

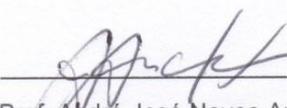
IDENTIFICAÇÃO LITOLÓGICA COM *AFFINITY PROPAGATION*

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Geofísica do Instituto de Geociências da Universidade Federal do Pará - UFPA, em cumprimento às exigências para obtenção de Mestre em Geofísica.

Data de Aprovação: 26 / 03 / 2013

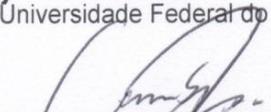
Conceito: Excelente

Banca Examinadora



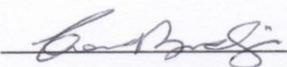
Prof. André José Neves Andrade – Orientador

Doutor em Geofísica
Universidade Federal do Pará



Prof. Carlos Eduardo Guerra – Coorientador

Doutor em Geofísica
Universidade Federal do Oeste do Pará



Prof.ª Carolina Barros da Silva – Membro

Doutora em Geofísica
Universidade Federal do Pará

AGRADECIMENTOS

Ao Prof.Dr. Carlos Eduardo Guerra pelo apoio e pela orientação na realização deste trabalho.

Ao Prof. Dr. André Andrade pela orientação.

Ao amigo Diogo Lopes pela inestimável ajuda e suporte técnico.

À minha amiga Renata Sena pelo suporte estético.

À PETROBRÁS (Rede de Geofísica), pelas oportunidades e suporte ofertado ao curso e a minha pesquisa.

RESUMO

Este trabalho apresenta uma metodologia para a solução do problema da extração de uma informação geológica, como a identificação de litologias em profundidade, diretamente dos perfis geofísicos de poços. Para isto, faz-se o uso do Gráfico Vsh-M-N, que é uma adaptação do Gráfico M-N, clássico da avaliação de formação, para a identificação litológica em termos das propriedades físicas do mineral principal e da argilosidade. A interpretação visual do Gráfico Vsh-M-N é limitada devido ao grande espalhamento dos pontos. Assim, apresenta-se uma forma de interpretação computacional para o Gráfico Vsh-M-N com a utilização do algoritmo Affinity Propagation, que em função das suas características de transporte da informação entre objetos semelhantes possibilita a interpretação dos pontos no Gráfico Vsh-M-N segundo as suas propriedades físicas e a sua continuidade em profundidade. Esta metodologia é apresentada com dados sintéticos e perfis de um poço testemunhado do Campo de Namorado, na Bacia de Campos. Brasil.

Palavras-chave: Geofísica de poço; Identificação litológica; Algoritmo Affinity Propagation.

ABSTRACT

This work presents a methodology for solving the problem of extracting geological information, such as identification of lithologies at depth, directly from wireline logs. For this propose, the Vsh-M-N 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 in each lithology. However, the visual interpretation of this graphic is limited by the large spread of the points in the graph. To evaluate a form of computational interpretation of the Vsh-M-N, the affinity propagation algorithm is used in reason of its characteristic transport of information among similar objects, which enables the interpretation of Vsh-M-N plot considering the physical properties and continuity in the depth. This methodology is presented using synthetic data and well logs from one borehole in the Namorado field. Campos basin. Brazil.

Keywords: Wireline logging; Lithological identification; Affinity Propagation Algorithm.

LISTA DE ILUSTRAÇÕES

Figura 01 - Gráfico M-N. Os quadrados em azul representam os minerais mais comuns com seus respectivos valores de M-N	20
Figura 02 - Gráfico Vsh-M-N mostrado as regiões ocupadas por diferentes rochas sedimentares.	21
Fonte: Adaptado de BARROS. (2011).	21
Figura 03 - Ilustração da troca de mensagens entre os dados segundo AP.	25
Figura 04 - Perfis de dados sintéticos (Exemplo 1).	32
Figura 05- Gráfico M-N resultante dos dados dos perfis sintéticos da Figura 3	33
Figura 06- Gráfico Vsh-M-N resultantes dos dados dos perfis sintéticos da Figura 3.	33
Figura 07- Representação dos exemplares no Gráfico Vsh-M-N.....	34
Figura 08- Disposição de exemplares no Gráfico M-N.....	35
Figura 09- Agrupamentos: Aplicação Affinity Propagation para clusterização dos dados.	36
Figura 10- Agrupamentos por Fuzzy C-means clustering. Os pontos vermelhos em destaque são os centróides da cada grupo.....	36
Figura 11- Gráfico M-N: Os centróides estão destacados em azul, os quadrados pretos são os pontos de minerais principais e os exemplares são representados por (*).	37
Figura 12- Zoneamento do poço em profundidade.	38
Figura 13- Perfis de dados sintéticos (exemplo 2).	39
Figura 14- Gráfico M-N resultante dos perfis de dados sintéticos da Figura 12.	40
Figura 15- Gráfico Vsh- M-N resultante dos perfis de dados sintéticos da Figura 13.	40
Figura 16- Representação dos exemplares no Gráfico Vsh-M-N.....	41
Figura 17- Disposição de exemplares e Pontos fixos referentes aos minerais de rochas sedimentares mais comuns.	42

Figura 18- Agrupamentos: Aplicação do Affinity Propagation para Clusterização dos dados.....	43
Figura 19- Agrupamentos por Fuzzy C-means clustering.	43
Figura 20- Os centróides estão destacados em círculos azuis, os quadrados pretos são os pontos de minerais principais e os exemplares em quadrados coloridos. Agrupamentos por Fuzzy C-means clustering. .	44
Figura 21- Zoneamento do poço em profundidade.	45
Figura 22- Perfis de dado real.....	46
Figura 23- Perfis de dado real. As curvas destacadas em preto representam o intervalo escolhido a ser processado na metodologia.....	47
Figura 24- Gráfico M-N resultantes dos perfis de dados reais representados na Figura 22.	48
Figura 25- Gráfico Vsh-M-N para os perfis de dados reais (Figura 23).....	48
Figura 26- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.....	49
Figura 27- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.....	50
Figura 28- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.....	51
Figura 29- Zoneamento do poço em profundidade. A representação das litologias em profundidade estão identificadas por arenito (verde) e folhelho (cinza).	51
Figura 30- Índices de validação internos: Coeficiente de partição e Classificação de entropia.	52
Figura 31- Índices de validação internos: Dunn e Dunn ajustado.	53

LISTA DE TABELA

<i>Tabela 1 - Valores de M e N para alguns minerais.</i>	20
---	-----------

SUMÁRIO

1 INTRODUÇÃO	11
2 AVALIAÇÃO DE FORMAÇÃO	13
2.1 Perfil de Raios Gama Natural.....	13
2.2 Perfis de Porosidade	14
2.2.1 Perfil de Densidade	15
2.2.2 Perfil de Porosidade Neutrônica	16
2.2.3 Perfil Sônico	16
2.3 GRÁFICO M-N	17
2.3.1 Parâmetro N	17
2.3.2 Parâmetro M.....	18
2.3.3 Interpretação do Gráfico M-N	19
3 MÉTODOS DE AGRUPAMENTO	21
3.1 FUZZY C-MEANS CLUSTERING	22
3.2 AFFINITY PROPAGATION	23
3.3 ÍNDICE DE VALIDAÇÃO.....	26
3.3.1 Coeficiente de Partição (PC)	27
3.3.2 Classificação de Entropia (CE)	27
3.3.3 Índice Dunn (DI).....	27
3.3.4 Índice Dunn ajustado (ADI).....	28
4 METODOLOGIA	29
5 RESULTADOS	31
5.1 DADOS SINTÉTICOS	31
5.1.1 Dado sintético: Exemplo 1.....	31
5.1.2 Dado sintético: Exemplo 2.....	38
5.2 DADO REAL.....	46
6 CONCLUSÕES	54
REFERÊNCIAS	55

1 INTRODUÇÃO

As operações de perfilagem de poço são complexas e particularmente, no caso das bacias submersas, essas complexidades são ainda maiores em função do grande número de poços horizontais ou poços desviados de alto ângulo. Neste caso, aumentam as dificuldades da avaliação de formação, pela ausência de informações geológicas e de laboratório (análise de testemunho), em função da ausência de afloramentos e da rarefação de poços testemunhados.

A avaliação de formação depende fundamentalmente de informações petrofísicas e geológicas produzidas pela análise de testemunhos. A ausência deste tipo de informação induz ao desenvolvimento de métodos, que possam atenuar as dificuldades e produzir um cálculo de reservas realista.

A informação geológica obtida em um poço permite a correlação ou atualização do perfil estratigráfico do campo petrolífero e possui ainda duas aplicações fundamentais para a avaliação de formação: A primeira trata do cálculo da porosidade, que necessita de informações sobre os parâmetros de porosidade da matriz rochosa (BARROS, 2011), ou de forma simplificada, das propriedades físicas do mineral principal. A segunda aplicação refere-se ao zoneamento do poço, ou a identificação das profundidades de topo e base de cada uma das diferentes camadas atravessadas pelo poço (ELLIS,2005).

O modo ideal para a realização da identificação litológica e delimitação das camadas é a análise geológica dos testemunhos retirados ao longo da profundidade de todo o poço. Três fatores impedem a realização cotidiana deste procedimento. O primeiro refere-se à inviabilidade técnica de realização da testemunhagem nos poços horizontais. O segundo fator é o alto custo da testemunhagem e o último trata-se do fator de recuperação, que é a impossibilidade de recuperação de 100% da rocha retida no amostrador em uma operação de testemunhagem.

Ao longo da história da avaliação de formação foram desenvolvidas algumas técnicas para a identificação das litologias, a partir da interpretação quantitativa dos perfis geofísicos e em termos de um modelo de rocha reservatório. O Gráfico M-N (BURKE, 1969) é um método muito utilizado no cotidiano da avaliação de formação para a identificação de litologias a partir das propriedades físicas do mineral principal. Este método combina os perfis de porosidade (densidade, sônico e

porosidade neutrônica) e um modelo de rocha (isenta de argilosidade) para gerar um gráfico composto por pontos definidos pelos parâmetros M e N, que por construção atenuam o efeito da porosidade sobre as medidas dos perfis de porosidade. O Gráfico M-N apresenta um conjunto de pontos (N,M) obtidos a partir destas três propriedades físicas medidas para os principais minerais na constituição das rochas sedimentares, que são tomados como pontos de referência para a identificação de litologias a partir de sua proximidade aos pontos (N,M) calculados diretamente dos perfis de porosidade. Este tipo de gráfico, em princípio, forneceria uma rápida interpretação visual da litologia de um ponto qualquer ao longo da profundidade de um poço. Entretanto, a presença de ruído nos perfis de porosidade e a simplificação do modelo de rocha, que desconsidera a argilosidade, torna a interpretação visual do Gráfico M-N complexa e ambígua.

Apresenta-se uma metodologia que busca atenuar as limitações do modelo de rocha adotado no Gráfico M-N e produzir a identificação automática de litologias composta pela integração de dois métodos. Em termos da codificação dos perfis em litologias é adotado o Gráfico Vsh-M-N, que propõe a inserção do perfil de raios gama natural, com um terceiro eixo ao Gráfico M-N, de modo a considerar o efeito da argilosidade sobre os perfis de porosidade (BARROS, 2011). Para a realização da identificação automática de litologias, adota-se o método numérico classificado como de mineração de dados (data mining) ou de classificação global, conhecido como Affinity Propagation (FREY; DUECK, 2007), que possui como característica de interesse a realização da troca de informação entre os objetos a serem classificados, em termos da similaridade ou afinidade entre os elementos de uma mesma classe. A afinidade ou similaridade é definida em termos de proximidade em profundidade, quando da interpretação computacional do Gráfico Vsh-M-N, de modo a prevalecer a continuidade em profundidade da litologia de uma mesma camada.

Apresenta-se e avalia-se esta metodologia com a utilização de dados sintéticos e perfis reais, registrados em poços do Campo de Namorado na Bacia de Campos. Brasil. Mostra-se ainda, uma comparação dos resultados da interpretação do Gráfico Vsh-M-N produzidos com o Affinity Propagation e com um método fuzzy de classificação conhecido como média C-fuzzy.

2 AVALIAÇÃO DE FORMAÇÃO

A geofísica de poço é o conjunto de operações e técnicas que envolvem a aquisição, o processamento e a interpretação das medidas de algumas propriedades físicas das rochas atravessadas por um poço. Tais medidas são adquiridas na operação de perfilagem de poço. A perfilagem (well logging), do termo francês *carottage électrique*, surgiu em 1927, como uma técnica de prospecção geofísica (ELLIS, 1987). Trata-se do registro de algumas propriedades físicas das rochas atravessadas pelo poço por um dispositivo de medida (ferramenta e sonda de perfilagem), que se desloca de forma contínua, na direção ascendente do poço.

Os perfis de poços são registros pontuais das variações de uma das propriedades medidas pelas ferramentas em função da profundidade. A análise desses perfis pode ser realizada de forma qualitativa ou de forma quantitativa. A primeira, é feita basicamente com evidências visuais das variações das propriedades físicas com a profundidade, podendo eventualmente, distinguir as diferentes camadas atravessadas pelo poço. Enquanto que a análise quantitativa, também conhecida como avaliação de formação, é realizada através de cálculos numéricos, nos quais pode-se determinar as propriedades petrofísicas (argilosidade, porosidade e a saturação fluida) das diversas camadas, e com isto realizar o cálculo das reservas ou do volume de hidrocarboneto. De um modo generalizado, a avaliação de formação pode ser conceituada como um conjunto de métodos que visam extrair informações dos perfis geofísicos, com o objetivo de identificar, quantificar e produzir um modelo geológico para uma acumulação de hidrocarbonetos em subsuperfície.

2.1 PERFIL DE RAIOS GAMA NATURAL

É o registro da variação da radiação gama natural ao longo da profundidade de um poço. A radioatividade natural é devido ao decaimento radioativo de isótopos instáveis com altos valores de meia-vida, tais como Potássio (K^{40}), o Tório (T_h^{232}) e o Urânio (U^{238}).

O perfil de raios gama natural é utilizado para classificar as rochas perfuradas em rochas reservatório ou rochas selantes devido às diferentes

concentrações de isótopos radioativos em cada uma delas. As rochas selantes, como os folhelhos, normalmente emitem uma maior quantidade de radiação gama que as rochas reservatório, como o arenito, por exemplo, em razão da presença de feldspatos (potássico) e de matéria orgânica.

O perfil de raios gama natural é utilizado como um indicador da argilosidade ou volume relativo de argila nas rochas sedimentares, na forma,

$$V_{sh} = \frac{RG - RG_{min}}{RG_{max} - RG_{min}} \quad (1)$$

Na equação 1, RG representa os valores registrados no perfil de raios gama natural. RG_{max} e RG_{min} representam, respectivamente, os valores máximos e mínimos da radiação gama, que foram interpretados no intervalo perfilado.

2.2 PERFIS DE POROSIDADE

É uma designação geral para o conjunto formado pelos perfis de densidade, porosidade neutrônica e sônico, que realizam o registro de propriedades físicas das rochas (densidade, porosidade neutrônica e tempo de trânsito) que, segundo o modelo de rocha, podem ser expressos por uma combinação linear da propriedade física medida individualmente em cada material presente na constituição da rocha e os seus volumes relativos (CRAIN, 2005).

Considerando um modelo de uma rocha reservatório constituída por matriz (mineral), fluido (água doce) e argila, pode-se escrever a equação geral para um perfil de porosidade qualquer, na forma

$$p = \phi p_w + V_{sh} p_{sh} + (1 - \phi - V_{sh}) p_m \quad (2)$$

Na equação 2, p é um perfil de porosidade. p_w é o valor da propriedade medida na água doce. ϕ é a porosidade e p_{sh} é o valor da propriedade medida na argila, V_{sh} é a argilosidade e p_m é o valor da propriedade medida no mineral principal (BARROS, 2011).

2.2.1 Perfil de Densidade

O perfil de densidade (ρ) é o registro contínuo das variações das densidades das rochas atravessadas por um poço. No caso de rochas porosas, a medição realizada pelo perfil inclui tanto a densidade da matriz da rocha, como a do fluido contido no espaço poroso.

Sabe-se que a atenuação de um feixe de raios gama que atravessa a matéria, pode ser relacionada com a densidade dos elétrons do material, se a interação predominante entre os raios gama e os átomos do material for o efeito Compton (ELLIS, 2005). No ambiente do poço, uma medida da atenuação dos raios gama pode ser usada para determinar a densidade da rocha próxima à parede do poço. A fonte de raios gama normalmente utilizada é o Césio (Cs^{137}), o qual emite raios gama com uma energia de 662 KeV, bem abaixo do limite para a ocorrência do efeito de produção de par. Este isótopo tem uma meia vida de cerca de trinta anos, o que dá uma estabilidade na intensidade do feixe energético, durante um período de tempo considerável. O poço é bombardeado por um fluxo de raios gama produzido pela fonte radioativa, que se encontra no interior da ferramenta. Um cintilômetro, montado no corpo ferramental e protegido da radiação direta da fonte, realiza o registro do fluxo de raios gama difundido pela formação. O sinal observado é relacionado com a densidade da formação, uma vez que ele é representativo do número de raios gama que foram desviados de sua trajetória original no interior da formação. Assim, quanto menor for o fluxo de raios gama detectados, maior é o número de elétrons (densidade eletrônica) responsáveis pela mudança na trajetória dos raios gama pelo efeito Compton, o que implica em um maior número de átomos por unidade de volume.

Como a ferramenta de densidade não responde diretamente à densidade (massa por unidade de volume), mas sim ao número de elétrons por unidade de volume, ou seja, a densidade eletrônica. Existe uma calibração que é realizada internamente pela ferramenta, que considera o calcário ($\rho_{cal} = 2,71 \text{ g/cm}^3$) saturado com água doce como rocha de referência para produzir o registro do valor real da densidade da rocha.

2.2.2 Perfil de Porosidade Neutrônica

A ferramenta de porosidade neutrônica, que registra as medidas do perfil de porosidade neutrônica (ϕ_N), foi o primeiro dispositivo nuclear descido em um poço (Nery, 2000). Baseado no fato que os nêutrons são partículas sem carga elétrica e com massa similar a do núcleo de hidrogênio, elas interagem com as rochas atingindo os núcleos dos elementos dos diversos materiais que as constituem, na forma elástica ou inelasticamente. Como o hidrogênio é o principal atenuador da energia de um fluxo de nêutrons e se faz presente na constituição da rocha na forma de água ou hidrocarboneto, a medida da atenuação de um fluxo de nêutrons está diretamente relacionado ao volume de hidrogênio ou a porosidade.

A ferramenta neutrônica é basicamente construída por uma fonte e dois detectores de nêutrons. A fonte emite nêutrons rápidos, os quais penetram nas camadas próximas da parede do poço. Através das múltiplas colisões elásticas, os nêutrons perdem parte da energia com que foram injetados, esta perda de energia está sujeita ao tamanho da massa do núcleo do átomo alvo. A atenuação de energia é máxima quando o núcleo alvo tem massa igual a do nêutron, caso do núcleo de hidrogênio.

O problema da ferramenta de porosidade neutrônica é sensível a todas as fontes de hidrogênio numa formação, não apenas a contida nos espaços porosos, mas também presente na argila, uma vez que o hidrogênio associado aos minerais de argila é visto pela ferramenta da mesma maneira como o hidrogênio nos fluidos.

2.2.3 Perfil Sônico

No início dos anos 50, os experimentos sobre amostra de rocha de Wyllie e outros, na Gulf Oil Co, constataram que a velocidade de propagação de uma onda acústica podia ser representada por uma combinação linear da velocidade de propagação da onda e o volume relativo de cada constituinte da rocha (CRAIN, 1989). A ferramenta sônica é composta por um ou mais conjuntos formados por um transmissor e dois receptores acústicos. Um receptor está mais próximo e outro mais afastado em relação ao transmissor, de modo a atenuar o efeito da lama sobre o tempo de propagação das ondas compressãois (onda P) na rocha em função da disposição de fonte e receptor no corpo da ferramenta (em linha). O

perfil sônico avalia a velocidade de propagação de uma onda compressional realizando o registro das variações do tempo de transito em função da profundidade. O tempo de trânsito (Δt) é conceituado como o tempo de propagação de uma onda refratada na parede do poço na distancia de um pé (30,48 cm), que por construção é a separação entre os dois receptores.

2.3 GRÁFICO M-N

Gráfico M-N (BURKE et al., 1969) 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 tornar-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.

2.3.1 Parâmetro N

O parâmetro N é definido como a inclinação da reta de porosidades determinada em um gráfico com os eixos definidos pelos perfis de densidade e porosidade neutrônica. A reta de porosidade é o segmento de reta que une o ponto de matriz ($\phi = 0$) ao ponto de água ($\phi = 1$). Cada matriz ou mineral principal de uma litologia define uma única reta de porosidade. Considere os perfis de densidade e porosidade neutrônica escritos na forma de suas equações gerais,

$$\begin{aligned}\rho_b &= \phi\rho_{mf} + (1-\phi)\rho_{ma} \\ \phi_{Nb} &= \phi\phi_{mf} + (1-\phi)\phi_{ma}\end{aligned}\tag{3}$$

onde os índices b , mf e ma identificam a propriedade referida na formação, filtrado de lama e matriz, respectivamente. Considerando o sistema de equações lineares (equação 3), para um mesmo valor de porosidade(ϕ) pode-se escrever a

relação do perfil de porosidade neutrônica em função do perfil de densidade, na forma:

$$\phi_{Nb} = \frac{\phi_{Nmf} - \phi_{Nma}}{\rho_{mf} - \rho_{ma}} \rho_b + \left(1 + \frac{\rho_{ma}}{\rho_{mf} - \rho_{ma}} \right) \phi_{Nma} - \frac{\phi_{Nmf} \rho_{ma}}{\rho_{mf} - \rho_{ma}}. \quad (4)$$

Assim, por definição, o parâmetro N é tomado como o coeficiente angular da reta $\phi_{Nb} - \rho_b$.

$$N = \frac{\phi_{Nmf} - \phi_{Nma}}{\rho_{mf} - \rho_{ma}}. \quad (5)$$

Na relação linear (equação 4), o termo independente ou a intersecção da reta com o eixo das ordenadas não tem significado físico em termos das propriedades registradas nos perfis.

2.3.2 Parâmetro M

O parâmetro M é definido como a inclinação da reta de porosidades determinada em um gráfico com os eixos definidos pelos perfis de densidade e sônico. Considere os perfis de densidade e sônico escritos na forma de suas equações gerais,

$$\begin{aligned} \rho_b &= \phi \rho_{mf} + (1 - \phi) \rho_{ma} . \\ \Delta t_b &= \phi \Delta t_{mf} + (1 - \phi) \Delta t_{ma} . \end{aligned} \quad (6)$$

onde os índices acima são similares aos descritos na definição do parâmetro N. Considerando o sistema de equações lineares (equação 6), para um mesmo valor de porosidade(ϕ) pode-se escrever a relação do perfil sônico em função do perfil de densidade, na forma:

$$\Delta t_b = \frac{\Delta t_{mf} - \Delta t_{ma}}{\rho_{mf} - \rho_{ma}} \rho_b + \left(1 + \frac{\rho_{ma}}{\rho_{mf} - \rho_{ma}} \right) \Delta t_{ma} - \frac{\Delta t_{mf} \rho_{ma}}{\rho_{mf} - \rho_{ma}}. \quad (7)$$

Assim, por definição, M é caracterizado como sendo o coeficiente angular da reta $\Delta t_b - \rho_b$.

$$M = \frac{\Delta t_{mf} - \Delta t_{ma}}{\rho_{mf} - \rho_{ma}} \times 0,01 . \quad (8)$$

Na equação 8 realiza-se a adequação de escala com o fator 0,01. Na relação linear (equação 7) o termo independente ou a intersecção da reta com o eixo das ordenadas não tem significado físico em termos das propriedades registradas nos perfis.

2.3.3 Interpretação do Gráfico M-N

O Gráfico M-N (Figura 01), 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 (Tabela 1), denominados pontos fixos do Gráfico M-N. 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 outro, de modo a permitir a classificação das litologias em função das propriedades físicas do seu mineral principal (Barros, 2011).

Os pares (N,M) para os pontos em profundidade de um poço perfilado são calculados a partir da substituição dos parâmetros de matriz ($\rho_{ma}, \phi_{ma}, \Delta t_{ma}$) pelos dados dos perfis nas equações (9) e (10). 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.

$$N = \frac{\phi_{Nmf} - \phi_{Nb}}{\rho_b - \rho_{ma}} . \quad (9)$$

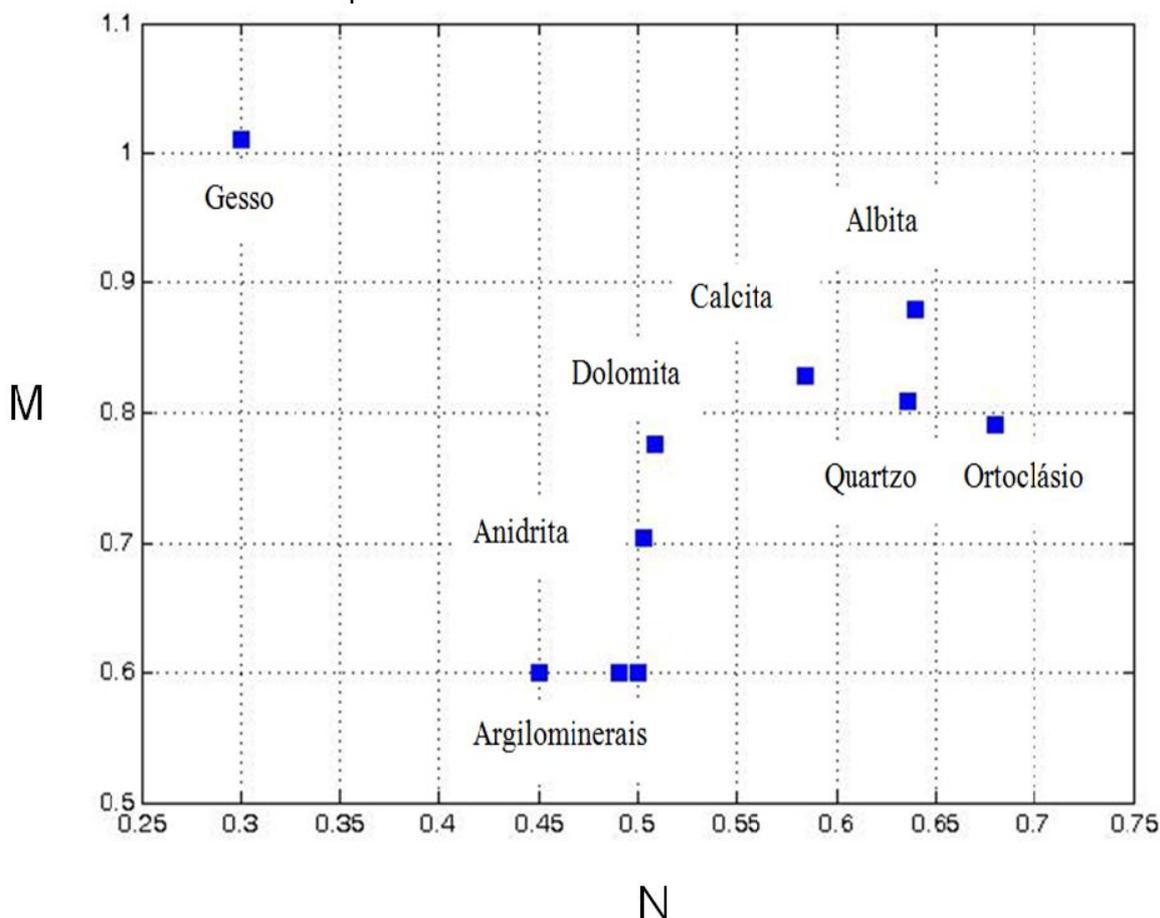
$$M = \frac{\Delta t_{mf} - \Delta t_b}{\rho_b - \rho_{ma}} \times 0,01 . \quad (10)$$

Tabela 1 - Valores de M e N para alguns minerais.

Mineral	ρ	Φ_N	Δt	M	N	Composição
Quartzo	2,65	-0,05	55,5	0,81	0,64	SiO_2
Calcita	2,71	0,00	47,0	0,83	0,59	$CaCO_3$
Dolomita	2,86	0,05	43,6	0,78	0,508	$CaMg(CO_3)_2$
Ortoclásio	2,55	-0,05	66,5	0,79	0,68	$KAlSi_3O_8$
Albita	2,62	-0,04	46,4	0,88	0,64	$NaAlSi_3O_8$
Anidrita	2,96	0,02	51,8	0,70	0,50	$CaSO_4$
Gesso	2,32	0,604	55,7	1,01	0,30	$CaSO_4 \cdot 2H_2O$
Argilomineral						
Caulinita	2,42	0,36	103,8	0,60	0,45	-
Ilita	2,53	0,25	97,2	0,60	0,49	-
Esmectita	2,12	0,44	121,8	0,60	0,50	-

Fonte: Adaptado de Burke, Campbell JR e Smith. (1969).

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

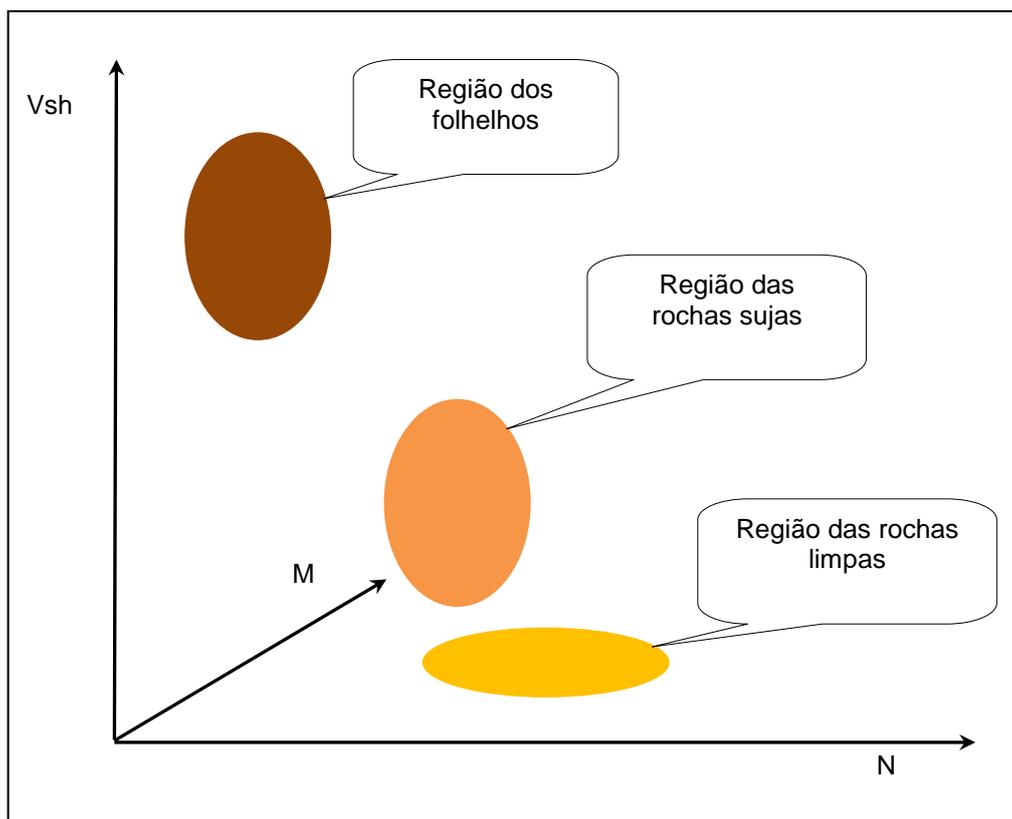


Fonte: Da autora

2.4 GRÁFICO VSH-M-N

O Gráfico Vsh-M-N, apresentado na Figura 02, 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 (BARROS, 2011).

Figura 02 - Gráfico Vsh-M-N mostrando as regiões ocupadas por diferentes rochas sedimentares.



Fonte: Adaptado de BARROS. (2011).

3 MÉTODOS DE AGRUPAMENTO

A análise de agrupamentos é uma ferramenta útil para análise de dados em diferentes situações. Essa técnica pode ser usada para reduzir a dimensão de um conjunto de dados, reduzindo uma ampla gama de objetos à informação do centro do seu conjunto. Técnicas de agrupamentos consistem em reunir objetos com características comuns em grupos (Clusters) ou classes. Um cluster é um conjunto de objetos com características em comum e com características que os diferem de outros clusters. Tendo em vista que clustering é uma técnica de agrupamento de dados que pode ser implementada em sistemas de aprendizado sendo útil para extrair características escondidas dos dados e desenvolver as hipóteses a respeito de sua natureza (LINDEN, 2009).

Existem diversos algoritmos que visam dividir um espaço amostral de dados em grupos cujos elementos possuem algumas características similares. A maioria dos algoritmos usa o critério de dissimilaridade pré-estabelecido, normalmente, a métrica utilizada é a distância euclidiana, mas podem ser usadas outras métricas, de preferência as que retratem melhor alguma característica comum dos dados com os quais se esteja trabalhando.

Aplicações que envolvem análise de agrupamentos são muito utilizadas em distintos campos da ciência e dos negócios como: Bioinformática, análise de dados espaciais, mineração de dados, geração e prova de hipóteses, predição baseada em agrupamento, para posterior classificação (VILLANUEVA; ZUBEN, 2008).

3.1 MEDIA C-FUZZY (FCM)

A teoria clássica dos conjuntos, um subconjunto A de um conjunto S é definido como um mapeamento dos elementos de S nos elementos do conjunto $\{0,1\}$. Esse mapeamento é usualmente expresso por pares ordenados, em que o primeiro elemento do par é representativo de um dos elementos do conjunto S , e o segundo, um elemento do conjunto $\{0,1\}$. Esses valores 1 e 0 representam, respectivamente, a pertinência e a não pertinência de um elemento de S em A (ou a verdade e a falsidade da afirmação de que um elemento de S pertence a A (AGUIAR ; OLIVEIRA JUNIOR, 2007).

Seja A um conjunto no universo X , a pertinência ou não de $x \in X$ em A pode ser caracterizada, pela lógica clássica por:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in X \\ 0 & \text{se } x \notin X \end{cases} \quad (11)$$

Diferente da teoria clássica dos conjuntos que conhecemos, na qual um elemento é classificado como pertencente ou não a um conjunto, ou seja, assume valor 1 ou 0, na lógica Fuzzy um elemento é classificado como pertencente a um conjunto segundo graus de pertinência que variam no intervalo $[0,1]$, onde 0 representa a total não pertinência e 1 representa a total pertinência.

O algoritmo da média C-fuzzy é o um método de agrupamento baseado na teoria dos conjuntos fuzzy desenvolvido por Dunn (1973) e melhorado por Bezdek em 1981. O algoritmo da média C-fuzzy é um método que organiza um determinado conjunto de objetos em grupos (clusters) tal que ocorra a minimização da função:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m < \infty . \quad (12)$$

em que m é qualquer número maior que 1, u_{ij} é o grau de pertinência de x_i no cluster j , x_i é o i -ésimo dado, c_j é o centro do cluster e $\|\cdot\|$ é qualquer norma que expresse similaridades entre o centro e os dados processados.

3.2 AFFINITY PROPAGATION (AP)

Os principais algoritmos de agrupamento utilizam como parâmetro de entrada um número k pré-determinado de cluster para particionar o espaço amostral de dados. O Affinity Propagation parte do princípio de que todos os pontos ou dados podem ser eleitos um "exemplar". O conjunto de dados forma a configuração de uma rede onde os pontos representam um nó e ocorre a transmissão de mensagens entre as arestas da rede que visam otimizar a

convergência para a eleição de um conjunto de exemplares que serão considerados representantes do grupo ao qual estão vinculado.

Affinity Propagation tem por entradas uma coleção de valores de similaridades entre dois pontos (dados), onde a similaridade $s(i, k)$ indica quão adequado o ponto com índice k seria para ser um exemplar para o ponto de índice i . Quando o objetivo é minimizar o quadrado do erro, cada similaridade é definida como o negativo do quadrado do erro (distância euclidiana): Para os pontos x_i e x_k , $s(i, k) = -\|x_i - x_k\|^2$. Esse método pode ser aplicado em casos mais gerais, onde a similaridade é derivada de pares de imagens, microarrays, sentenças em inglês, e pares de cidades (FREY ; DUECK, 2007).

A matriz de similaridades, depende do número de dados em questão, isto é, seja n dados, a matriz de similaridades tem ordem n^2 . A similaridade esta relacionada a alguma métrica d sobre os dados, conceito esse que incorpora a noção geométrica geral de distância. Por definição:

Seja um conjunto M , uma métrica em M é uma função $d : M \times M \rightarrow \mathbb{R}$ tal que as seguintes propriedades são satisfeitas:

- (1) $d(x, y) \geq 0, \forall x, y \in M$.
- (2) $d(x, y) = d(y, x), \forall x, y \in M$.
- (3) $d(x, z) \leq d(x, y) + d(y, z), \forall x, y \text{ e } z \in M$.
- (4) $d(x, y) = 0 \Leftrightarrow x = y$.

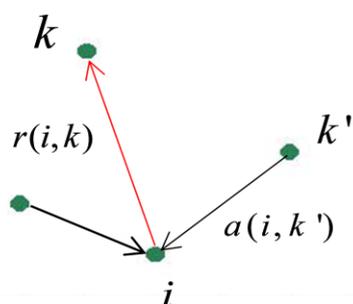
Além das similaridades, Affinity Propagation também tem como parâmetro de entrada valores $s(k, k)$ denominados "preferências" para cada ponto k . Sabendo, que em qualquer métrica a diagonal principal da matriz de similaridades é nula, o que indica que a maior similaridade ocorre de um ponto a ele mesmo, 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.

Um bom resultado é obtido adotando para preferências o valor médio das similaridades.

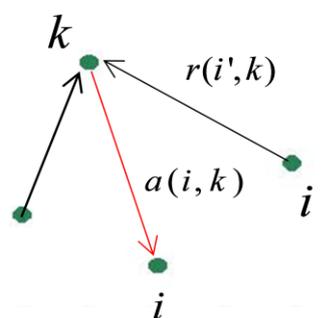
A troca de mensagens entre os pontos podem ser de dois tipos, transmissão de responsabilidades e disponibilidades. As responsabilidades $r(i, k)$, mensagens enviadas de um ponto i para um possível exemplar k indicam o quão adequado seria o ponto k ser exemplar para o ponto i . As disponibilidades $a(i, k)$, enviadas do candidato a exemplar k para o ponto i , indicam evidencias de quão adequado seria para o ponto i escolher o ponto k como seu exemplar. A troca de mensagens pode ser visualizada na Figura 03.

Figura 03 - Ilustração da troca de mensagens entre os dados segundo AP.

Enviando Responsabilidades



Enviando disponibilidades



Fonte: Da autora

As responsabilidades são definidas pela seguinte regra:

$$r(i, k) \leftarrow s(i, k) - \max_{k' \neq k} \{a(i, k') + s(i, k')\}. \quad (13)$$

onde $a(i, k) = 0$, para a primeira interação. Para $i = k$, a responsabilidade $r(k, k)$, auto responsabilidade, reflete a evidencia de que k é um exemplar.

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

$$\begin{aligned}
 a(i,k) &\leftarrow \min \left\{ 0, r(k,k) + \sum_{i' \in \{i,k\}} \max \{0, r(i',k)\} \right\}. \\
 a(k,k) &\leftarrow \sum_{i' \neq k} \max \{0, r(i',k)\}.
 \end{aligned}
 \tag{14}$$

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

$$AP = \max \{a(i,k) + r(i,k)\}. \tag{15}$$

3.3 ÍNDICES DE VALIDAÇÃO

O índice de validação tem por objetivo avaliar qualidade de grupos resultantes de um processo de agrupamento. Essa avaliação serve para comparar diferentes algoritmos de agrupamento, partições, e fazer comparação de dois grupos (clusters).

Os algoritmos de validação podem assumir alguns aspectos, dentre eles: determinar a tendência de agrupamentos de dados (verificar se os dados são ou não aleatórios); Comparar os resultados de uma análise de agrupamento com resultados conhecidos (tendo um conhecimento prévio dos dados, verificar que índices correspondem); Comparar diversos algoritmos de agrupamentos ou determinar o valor mais apropriado do número de grupos (clusters).

Medidas numéricas que são aplicadas para avaliar os vários aspectos da validação de agrupamentos são classificadas em três grupos: Índices Externos, usados para avaliar o agrupamento gerado de acordo com uma estrutura pré-especificada, imposta ao conjunto de dados, exemplos: Índice Rand ajustado (adjusted Rand) e índice de Jaccard; Índices Internos, usados para medir a qualidade de um agrupamento com base apenas nos dados originais (instâncias ou matriz de similaridade), exemplos: Índice Davies-Bouldin, Índice Dunn, Silhuetas; Índices Relativos, usados para comparar diversos agrupamentos para decidir qual deles é o melhor em algum aspecto. Em geral, pode ser qualquer um dos índices acima definidos (SOUTO, 2002).

Seja c o número de clusters, u_{ij} o valor de pertinência do dado j no cluster i . Podemos definir os índices aplicados nesta metodologia:

3.3.1 Coeficiente de Partição (PC)

Segundo este índice, o número ótimo de clusters ocorre quando o índice atinge valor máximo.

$$PC(c) = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N (u_{ij})^2. \quad (16)$$

3.3.2 Classificação de Entropia (CE)

O número ótimo de clusters é dado quando se atinge o valor mínimo para os valores de índices obtidos, dados pela equação:

$$CE(c) = -\frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N u_{ij} \log(u_{ij}). \quad (17)$$

3.3.3 Índice Dunn (DI)

O índice Dunn é utilizado para identificação de agrupamentos compactos e bem definidos, cujo valor ótimo se dá quando o índice atinge valor máximo. Temos:

$$DI(c) = \min_{i \in c} \left\{ \min_{j \in c, j \neq i} \left\{ \frac{\delta(A_i, A_j)}{\max_{k \in c} \{\Delta(A_k)\}} \right\} \right\}. \quad (18)$$

onde

$$\delta(A_i, A_j) = \min \{d(x_i, x_j) \mid x_i \in A_i \text{ e } x_j \in A_j\}.$$

$$\Delta(A_k) = \max \{d(x_i, x_j) \mid x_i, x_j \in A_k\}.$$

d é uma função de distância e A_i é o conjunto de pontos pertencentes ao cluster i .

$\delta(A_i, A_j)$ é uma função de dissimilaridade entre agrupamentos e $\Delta(A_k)$ é uma medida de dispersão de dados internos do cluster.

3.3.4 Índice Dunn ajustado (ADI)

Para este índice, como no caso anterior, o valor ótimo de cluster é apontado pelo valor máximo para o índice.

Seja v_j o centro do cluster j , temos:

$$ADI(c) = \min_{i \in c} \left\{ \min_{j \in c, j \neq i} \left\{ \frac{\delta(A_i, A_j)}{\max_{k \in c} \{\Delta(A_k)\}} \right\} \right\}. \quad (19)$$

onde

$$\delta(A_i, A_j) = \min \left\{ |d(x_j, v_j) - d(x_i, v_j)| \mid x_i \in A_i \text{ e } x_j \in A_j \right\}.$$

$$\Delta(A_k) = \max \left\{ d(x_i, x_j) \mid x_i, x_j \in A_k \right\}.$$

d é uma função de distância e A_i é o conjunto de pontos pertencentes ao cluster i .

4 METODOLOGIA

A identificação das litologias ou dos tipos de rocha atravessados por um poço é realizada pelo geólogo, no local da perfuração, que aplica técnicas visuais e químicas diretamente sobre o testemunho, possibilitando além da identificação litológica, o posicionamento em profundidade de cada camada.

Na situação prática da avaliação de formação com a utilização dos perfis geofísicos e em função da rarefação de poços testemunhados pode ser utilizado o 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, 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 brutal 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.

5 RESULTADOS

A metodologia é empregada em dados sintéticos e reais, que satisfazem a equação geral dos perfis de porosidade. Os dados reais são de poços do campo de namorado na bacia de Campos, Brasil.

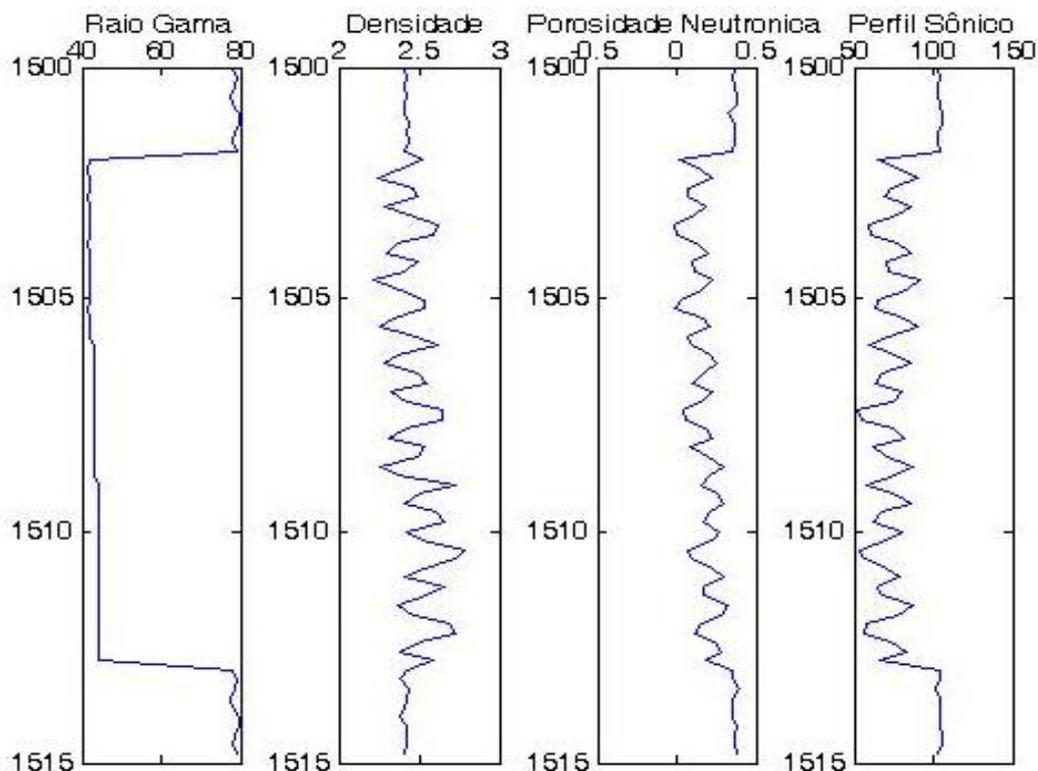
5.1 DADOS SINTÉTICOS

Com a utilização dos dados sintéticos investiga-se a capacidade de identificação litológica do algoritmo Affinity Propagation na interpretação do Gráfico M-N em dois casos. No primeiro, investiga-se apenas o efeito da argilosidade sobre os parâmetros M-N. No segundo caso, busca-se representar um exemplo mais realista considerando a argilosidade e a mistura de minerais na constituição da matriz.

5.1.1 Dado sintético: Exemplo 1

Neste caso, adota-se uma sequência geológica com três camadas reservatório monomineral: Arenito, Calcário e Dolomita. O interesse aqui é investigar o efeito da argilosidade sobre os parâmetros M-N. A Figura 04 mostra os perfis sintéticos para a sequência de rochas reservatório limitadas no topo e na base por camadas de folhelho. Na primeira trilha aparece o perfil de raio gama natural e nas trilhas subsequentes os perfis de densidade, porosidade neutrônica e sônico. A interpretação visual qualitativa das quatro curvas possibilita apenas a identificação de uma rocha reservatório pelo perfil de raio gama natural. Nenhum dos quatro perfis mostrados na Figura 04 indica a presença de três camadas reservatório.

Figura 04 - Perfis de dados sintéticos (Exemplo 1).

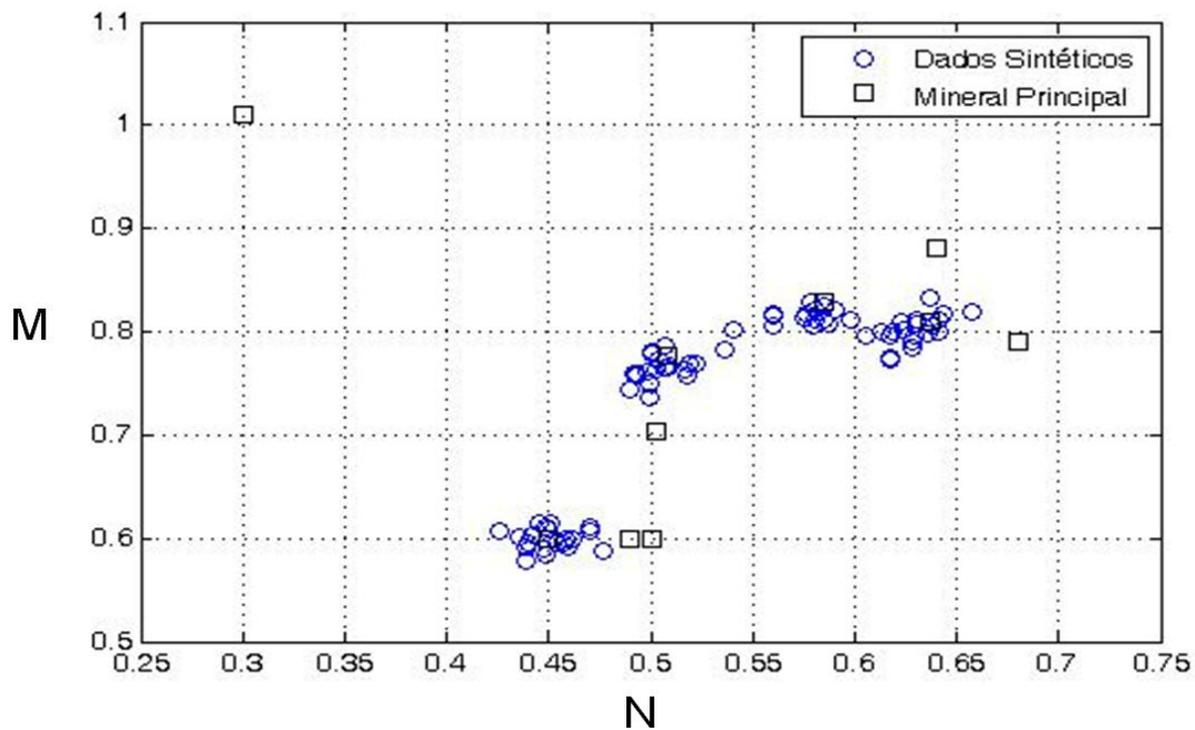


Fonte: Da autora

A Figura 05 apresenta o Gráfico M-N com os pontos representativos dos perfis mostrados na Figura 04 representados por círculos azuis. Os pontos fixos são representados pelos quadrados pretos. A interpretação visual da Figura 05 sugere a existência de mais de uma camada reservatório. No entanto, o efeito da argilosidade representado pelo deslocamento dos pontos na direção dos argilominerais impede uma classificação mais confiável.

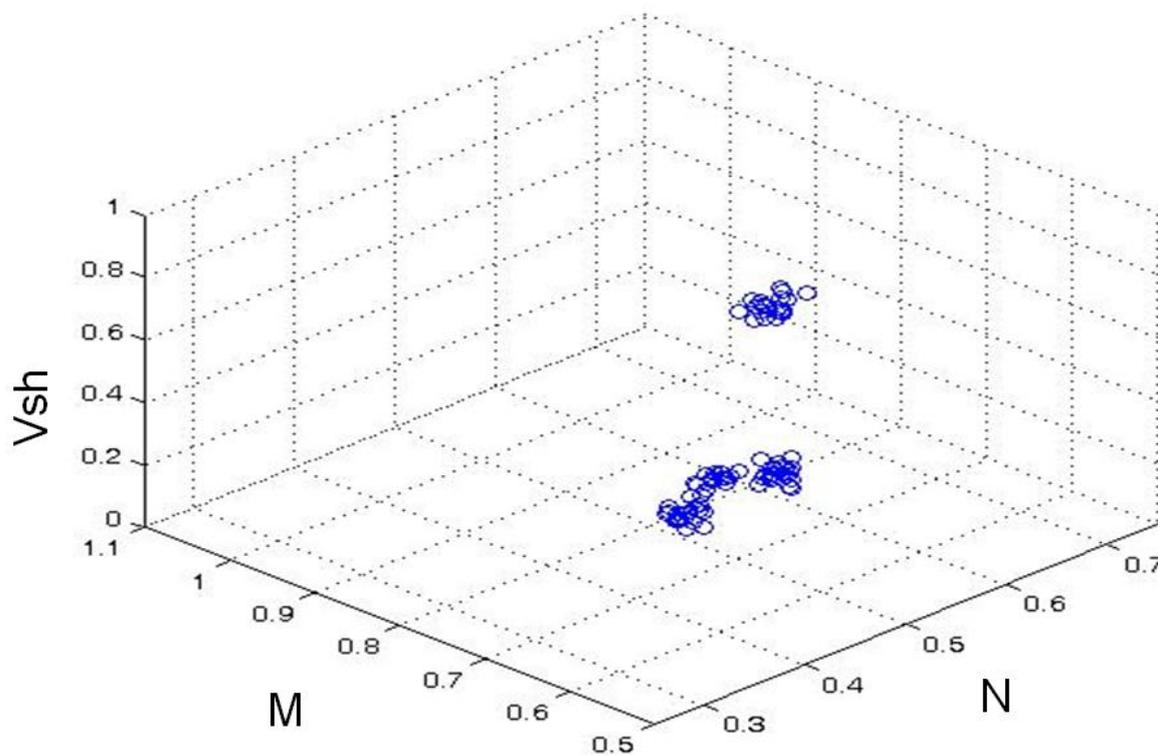
A Figura 06 mostra o correspondente Gráfico Vsh-M-N. O eixo da argilosidade foi obtido com a utilização do perfil de raios gama natural através da equação (1). Aqui a interpretação visual da presença de três camadas reservatório é mais confiável.

Figura 05- Gráfico M-N resultante dos dados dos perfis sintéticos da Figura 3



Fonte: Da autora

Figura 06- Gráfico Vsh-M-N resultantes dos dados dos perfis sintéticos da Figura 3.

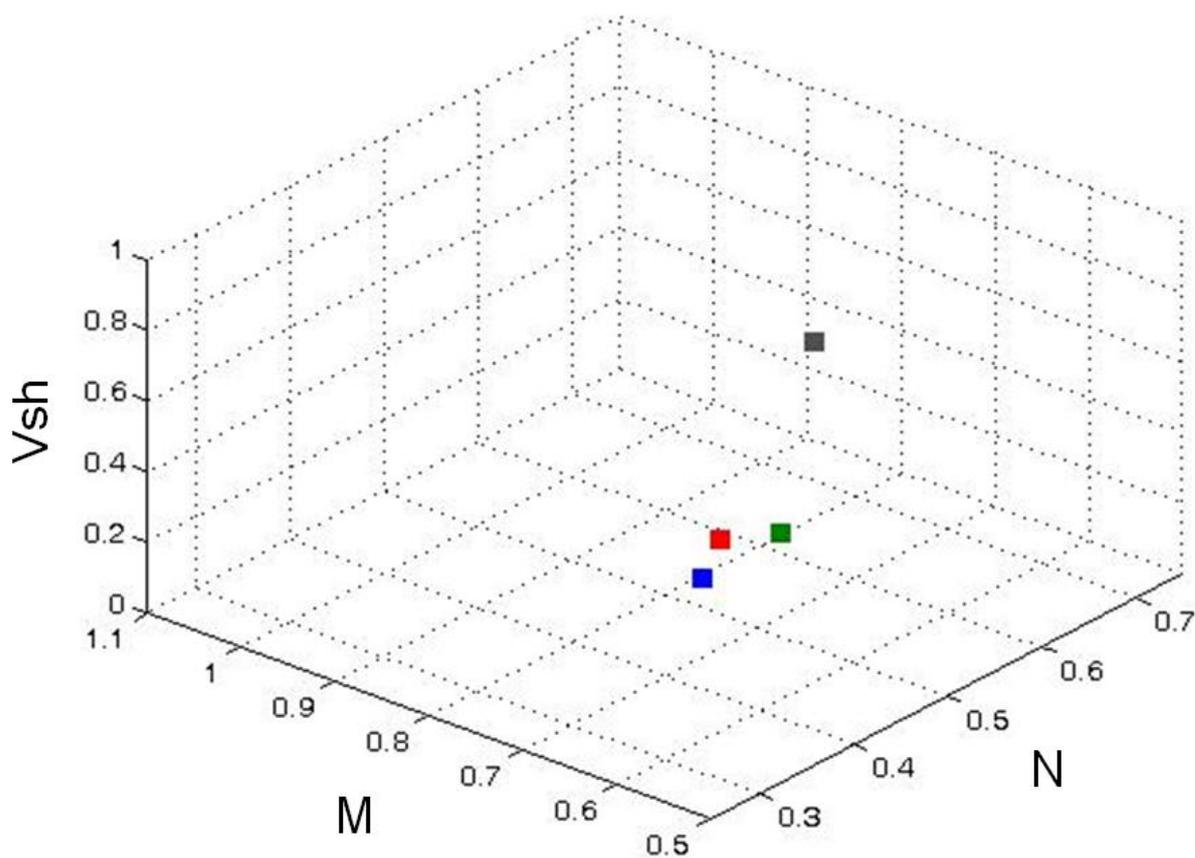


Fonte: Da autora

A Figura 07 mostra o resultado da aplicação do algoritmo Affinity Propagation para a interpretação do Gráfico Vsh-M-N mostrado na Figura 06. Em acordo com a interpretação visual do Gráfico Vsh-M-N mostrado na Figura 06, o algoritmo Affinity Propagation resultou na identificação de três exemplares representativos das camadas reservatório. Estes exemplares são apresentados pelos quadrados azul, vermelho e verde. A Figura 07 ou o Gráfico Vsh-M-N 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 igual ao adotado no modelo sintético.

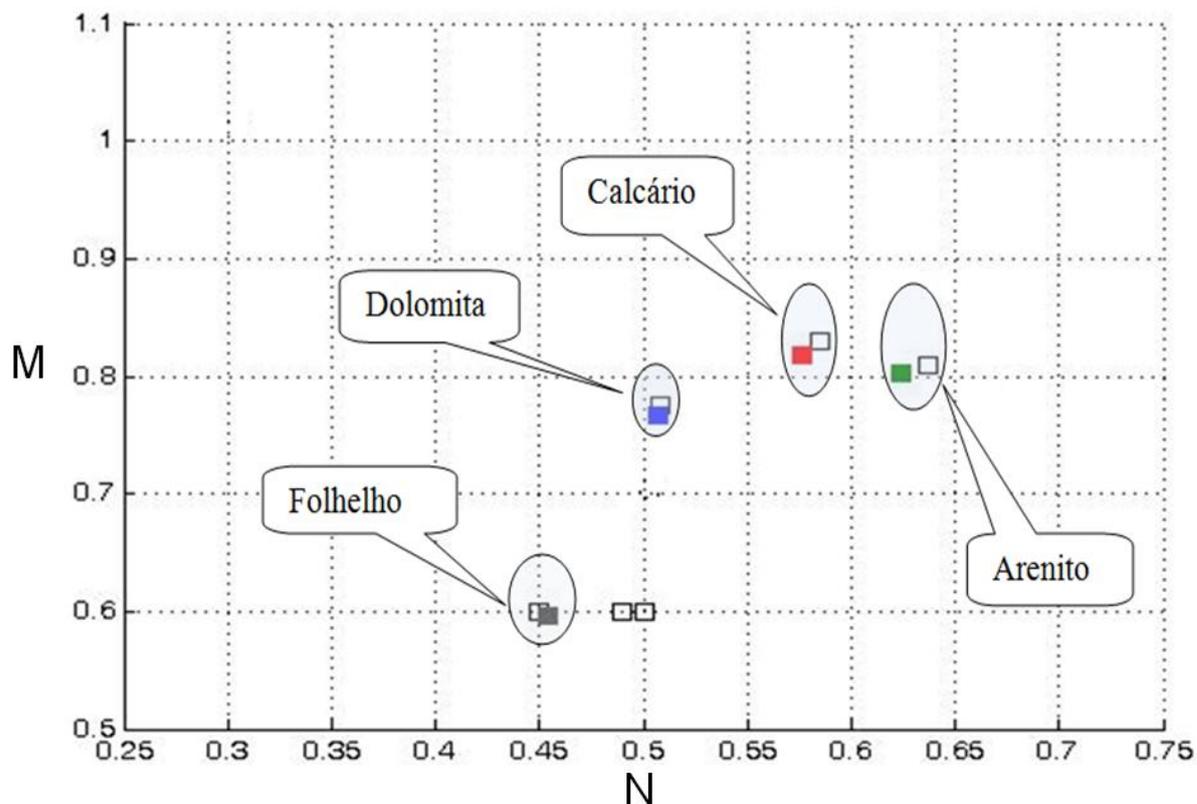
A Figura 08 mostra o resultado da classificação litológica das camadas reservatório. O exemplar azul é classificado como dolomita, e todos os pontos do Gráfico M-N que são representados por este exemplar são classificados como dolomita. Raciocínio semelhante é aplicado para os demais exemplares.

Figura 07- Representação dos exemplares no Gráfico Vsh-M-N



Fonte: Da autora

Figura 08- Disposição de exemplares no Gráfico M-N.

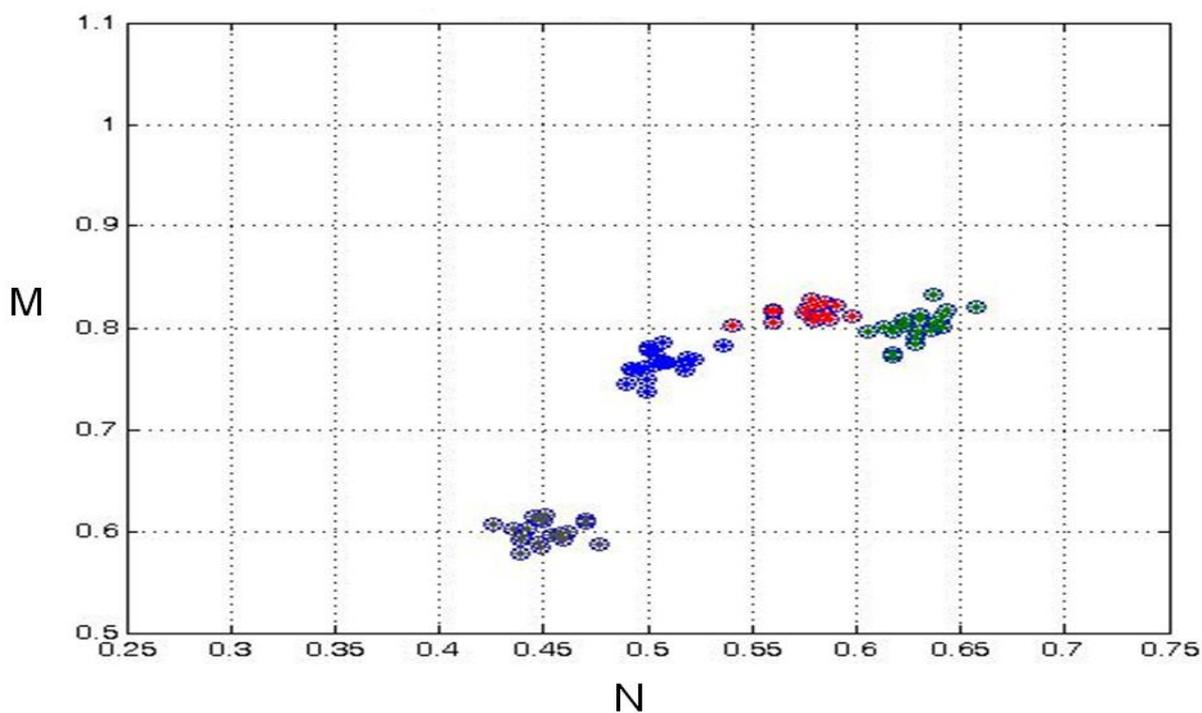


Fonte: Da autora

A Figura 09 apresenta o resultado final da identificação litológica das camadas atravessadas pelo poço perfurado no Exemplo 1. Mostra-se na Figura 09 os pontos classificados como dolomita na cor azul. Os pontos classificados como calcário na cor vermelha e aqueles classificados como arenito na cor verde. Para efeito da identificação litológica de todo o trecho perfurado apresentou-se em todas as figuras o exemplar e os pontos classificados como folhelho.

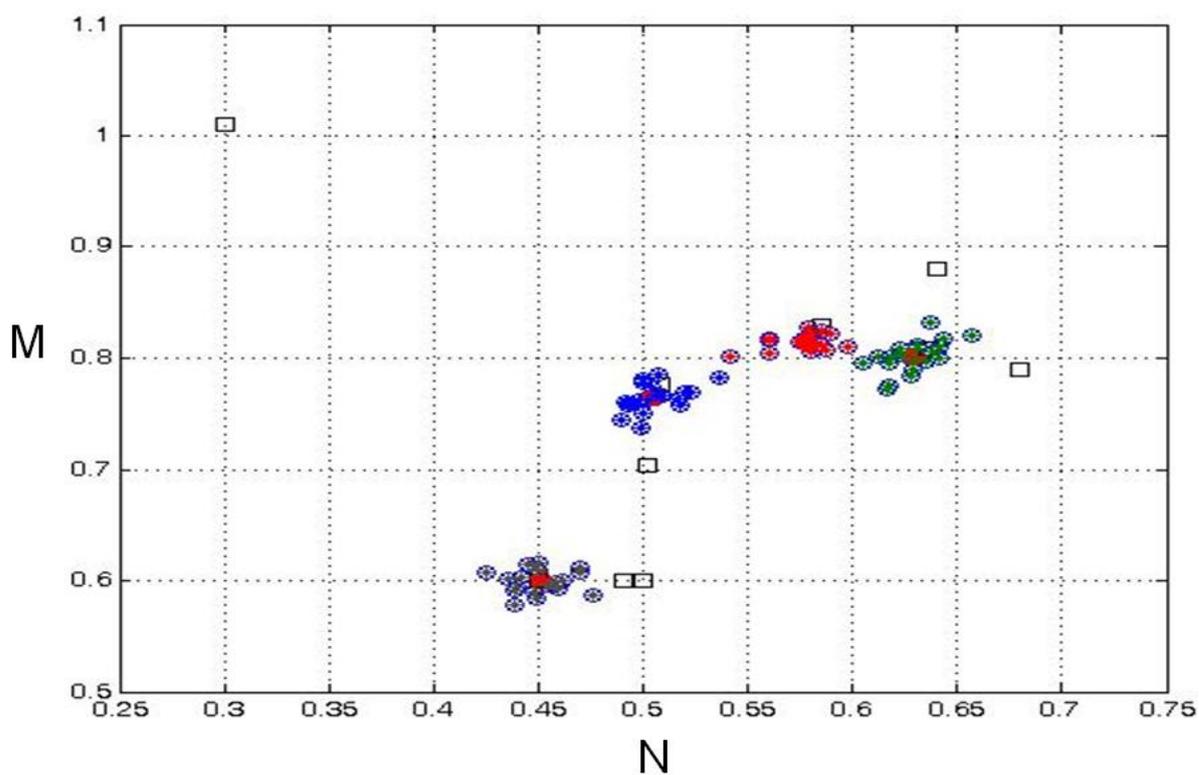
A Figura 10 mostra o resultado de classificação feita pelo fuzzy C-means clustering, algoritmo de clusterização que exige a pré-determinação de um número k de clusters como entrada, já utilizado na análise de perfis de poços (CLIMACO, 2009). Observa-se a concordância de resultados entre o FCM e o Affinity Propagation, para este caso. Para uma análise mais compacta reduzimos os dados aos centróides dos agrupamentos resultantes pela aplicação do FCM e aos exemplares resultantes da clusterização com AP (Figura 11).

Figura 09- Agrupamentos: Aplicação Affinity Propagation para clusterização dos dados.



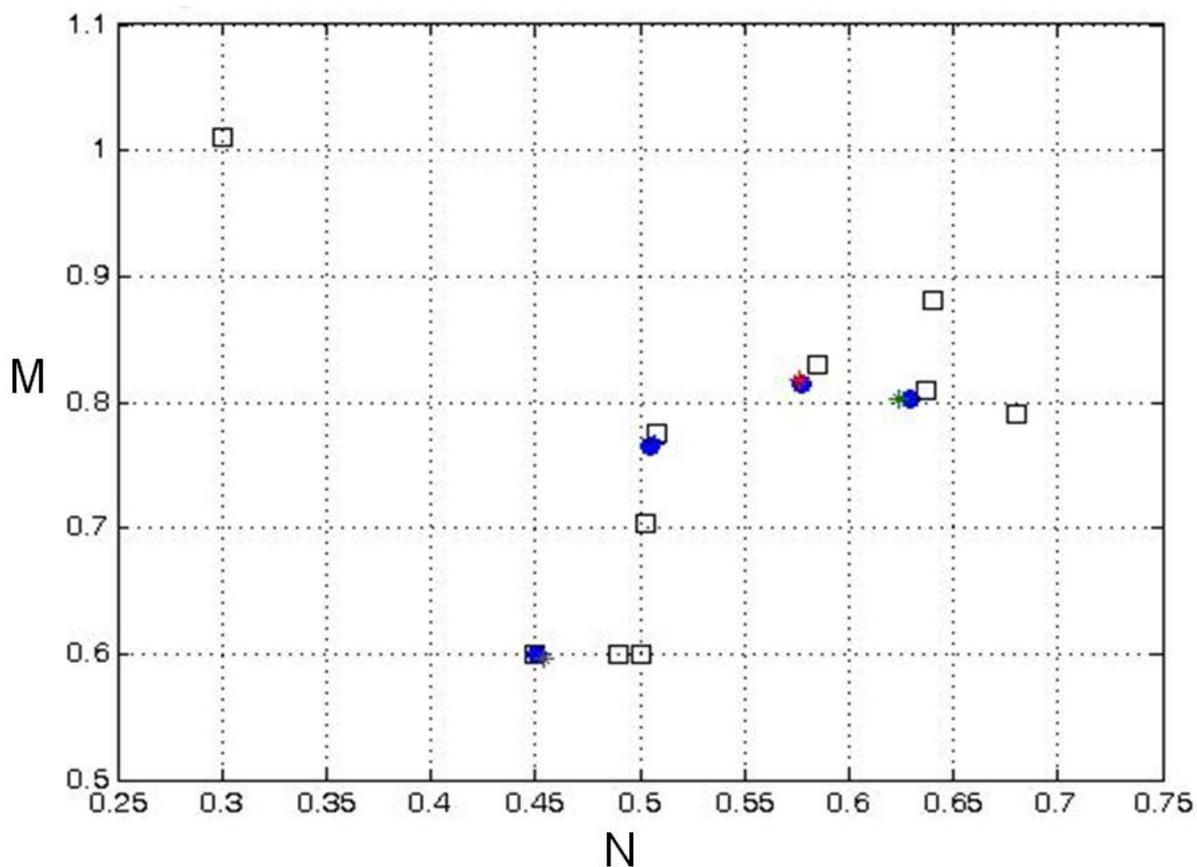
Fonte: Da autora

Figura 10- Agrupamentos por Fuzzy C-means clustering. Os pontos vermelhos em destaque são os centróides da cada grupo.



Fonte: Da autora

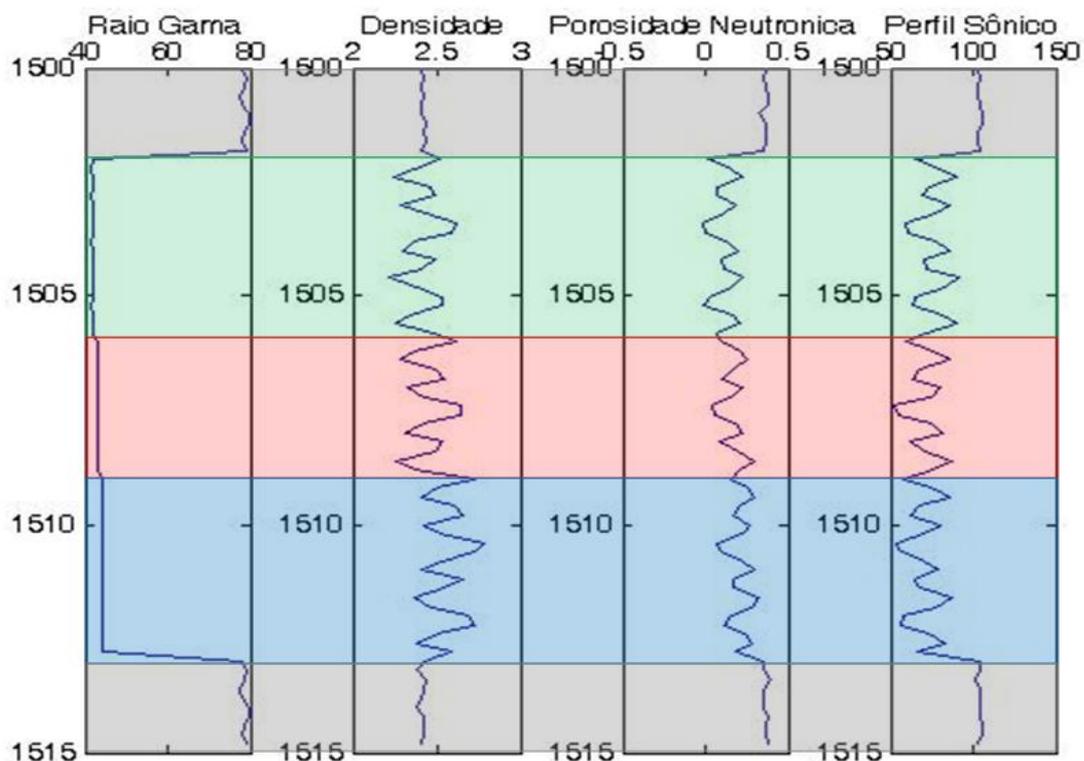
Figura 11- Gráfico M-N: Os centróides estão destacados em azul, os quadrados pretos são os pontos de minerais principais e os exemplares são representados por (*).



Fonte: Da autora

A Figura 12 mostra a representação em profundidade da identificação litológica produzida pelo algoritmo Affinity Propagation interpretando computacionalmente o Gráficos Vsh-M-N e o Gráfico M-N para o intervalo perfilado mostrado no Exemplo 1. A representação das litologias em profundidades estão identificadas por cores, onde arenito (verde), calcário (vermelho), dolomita (azul) e folhelho (cinza) foram classificados.

Figura 12- Zoneamento do poço em profundidade.



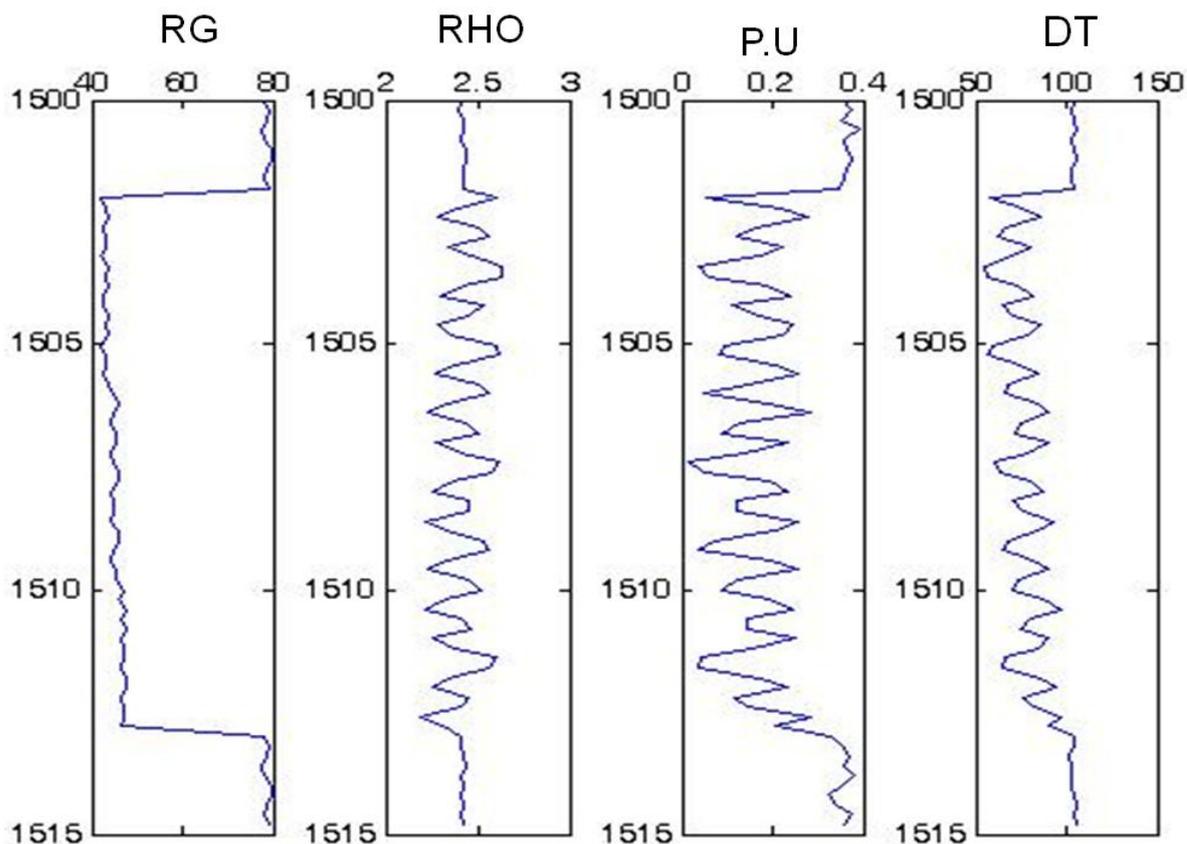
Fonte: Da autora

5.1.2 Dado sintético: Exemplo 2

Neste exemplo, busca-se um modelo de rocha mais realístico considerando além da argilosidade a mistura de minerais na composição da matriz.

A Figura 13 mostra os perfis sintéticos registrados em um cenário geológico semelhante ao da Figura 04. A primeira camada reservatório é um calcário exclusivamente composta por calcita. A segunda é um calcarenito com matriz composta por 60% de quartzo e 40% de calcita. A última camada reservatório é um arenito com matriz de 100% quartzo. A argilosidade cresce com a profundidade.

Figura 13- Perfis de dados sintéticos (exemplo 2).

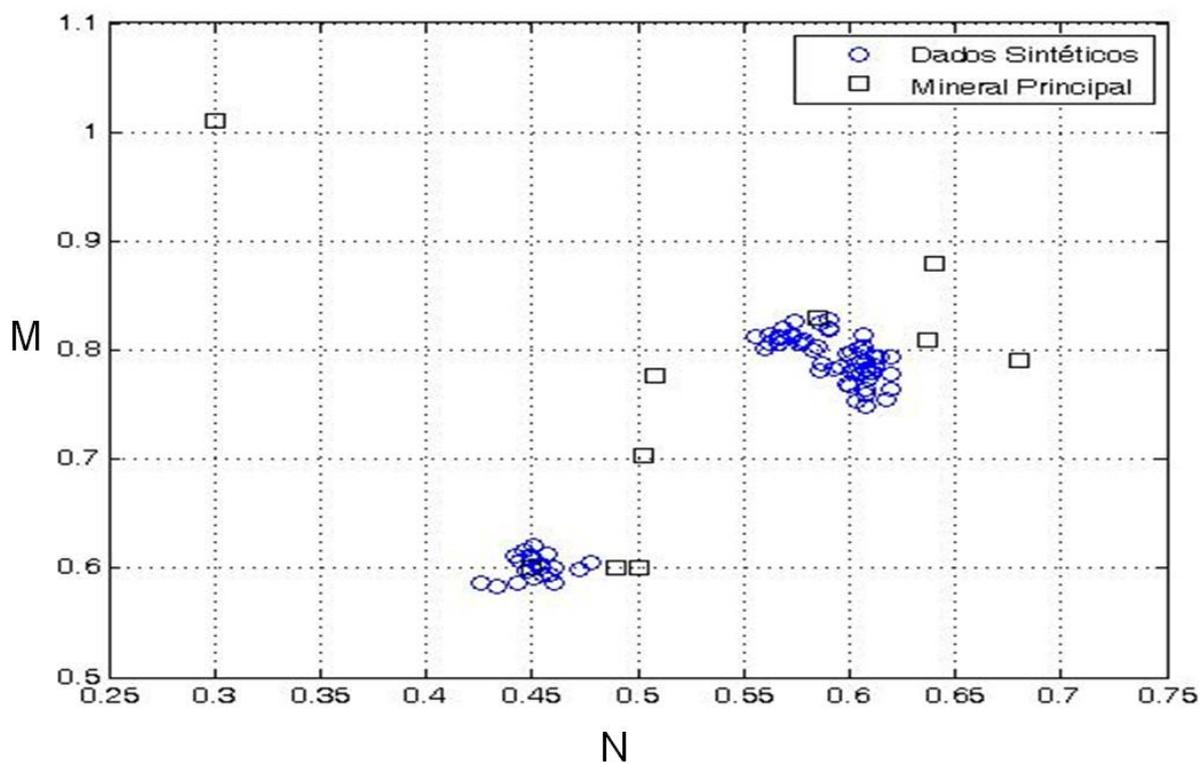


Fonte: Da autora

A Figura 14 apresenta o Gráfico M-N com os pontos representativos dos perfis mostrados na Figura 13 representados por círculos azuis. Os pontos fixos são representados pelos quadrados pretos. A interpretação visual do Gráfico M-N mostrado na Figura 14 sugere a existência de mais de uma camada reservatório. No entanto, o efeito da argilosidade e mistura de minerais na matriz produzem o deslocamento dos pontos em relação aos pontos fixos prejudicando a realização de uma classificação visual confiável.

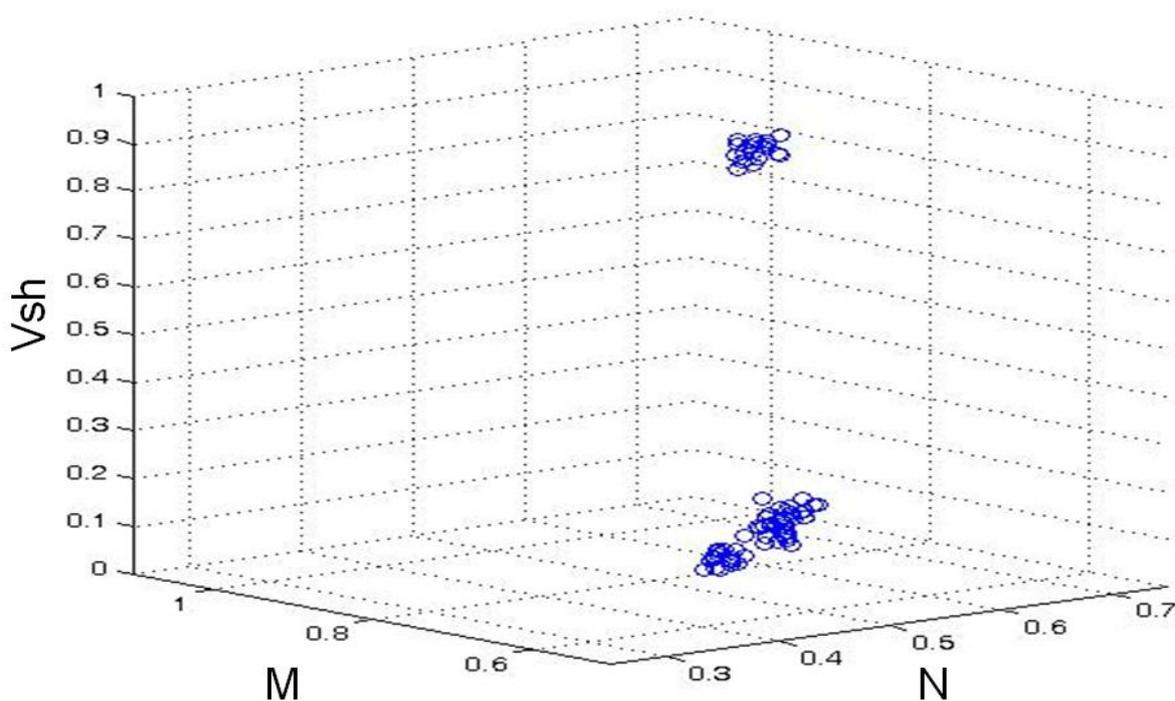
A Figura 15 mostra o correspondente Gráfico Vsh-M-N. O eixo da argilosidade foi obtido com a utilização do perfil de raios gama natural através da equação (1). O Gráfico Vsh-M-N permite a interpretação visual da presença de três camadas reservatório.

Figura 14- Gráfico M-N resultante dos perfis de dados sintéticos da Figura 12.



Fonte: Da autora

Figura 15- Gráfico Vsh- M-N resultante dos perfis de dados sintéticos da Figura 13.

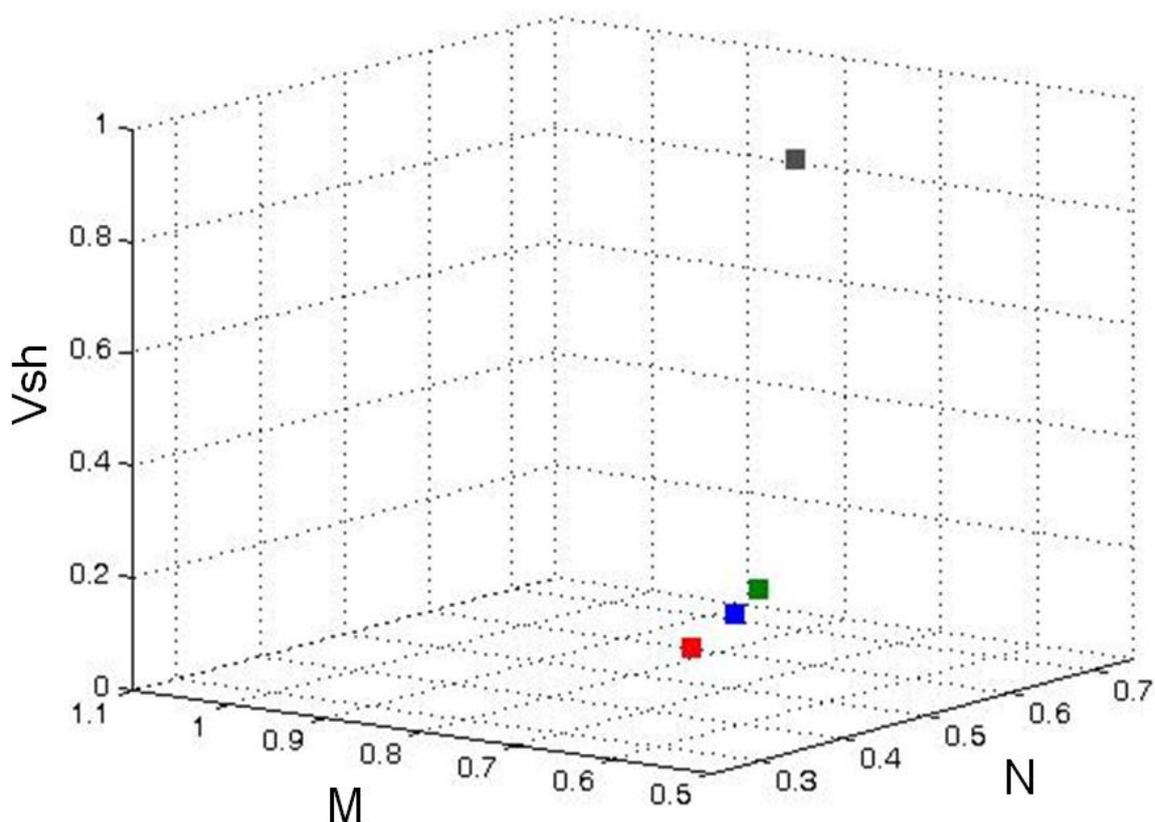


Fonte: Da autora

A Figura 16 mostra o resultado da aplicação do algoritmo Affinity Propagation para a interpretação do Gráfico Vsh-M-N mostrado na Figura 15. Em acordo com a interpretação visual do Gráfico Vsh-M-N mostrado na Figura 15, o algoritmo Affinity Propagation resultou na identificação de três exemplares representativos das camadas reservatório. Estes exemplares são apresentados pelos quadrados azul, vermelho e verde. A Figura 15 ou o Gráfico Vsh-M-N 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 igual ao adotado no modelo sintético.

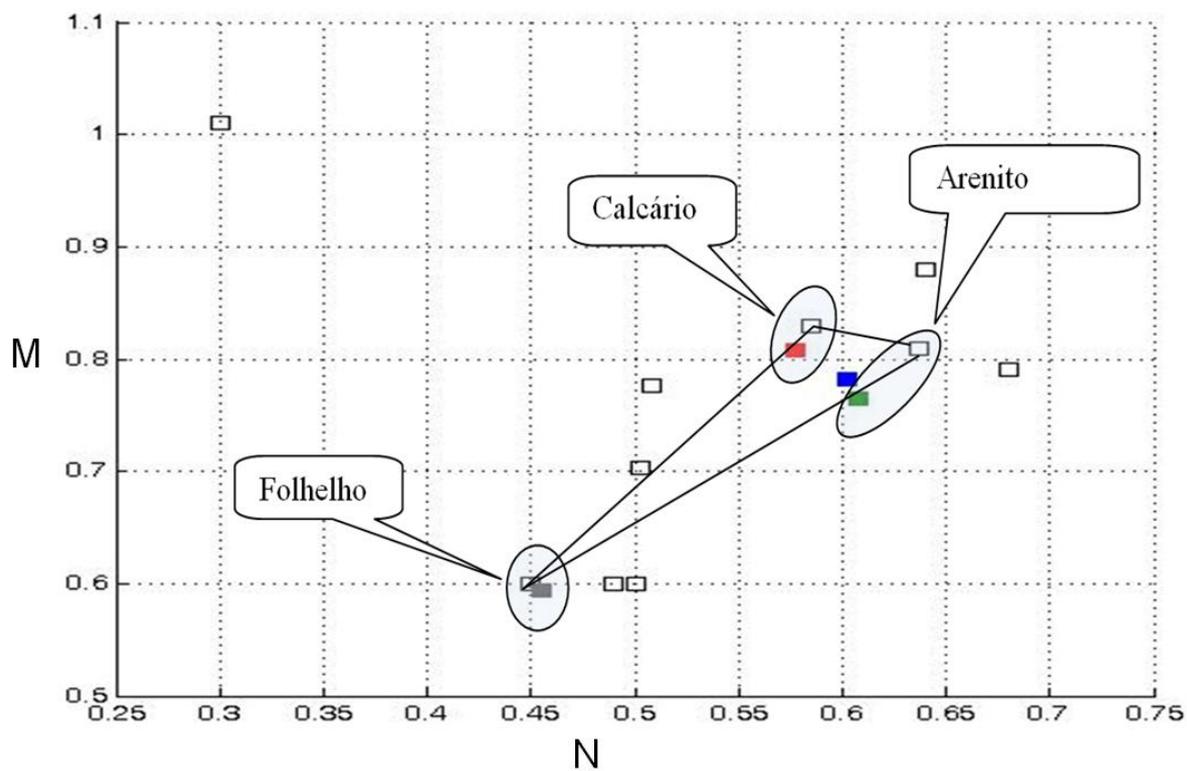
A Figura 17 mostra o resultado da classificação litológica das camadas reservatório. O exemplar azul é classificado como arenito, devido a predominância do mineral quartzo e a identificação litológica abordada ser monomineral, e todos os pontos do Gráfico M-N que são representados por este exemplar são classificados como arenito. Raciocínio semelhante é aplicado para os demais exemplares.

Figura 16- Representação dos exemplares no Gráfico Vsh-M-N.



Fonte: Da autora

Figura 17- Disposição de exemplares e Pontos fixos referentes aos minerais de rochas sedimentares mais comuns.

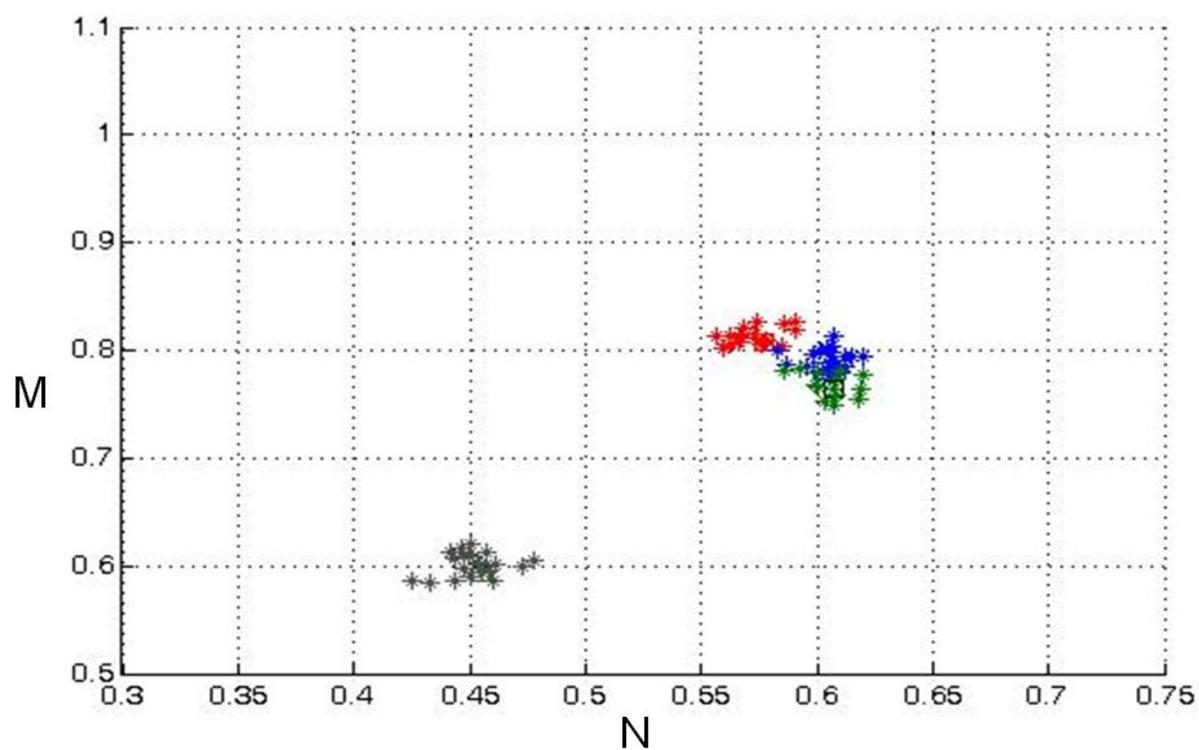


Fonte: Da autora

A Figura 18 apresenta o resultado final da identificação litológica das camadas atravessadas pelo poço perfurado no Exemplo 2. Mostra-se na Figura 18 os pontos classificados como calcarenito na cor azul. Os pontos classificados como calcário na cor vermelha e aqueles classificados como arenito na cor verde. Para efeito da identificação litológica de todo o trecho perfurado apresentou-se em todas as figuras o exemplar e os pontos classificados como folhelho.

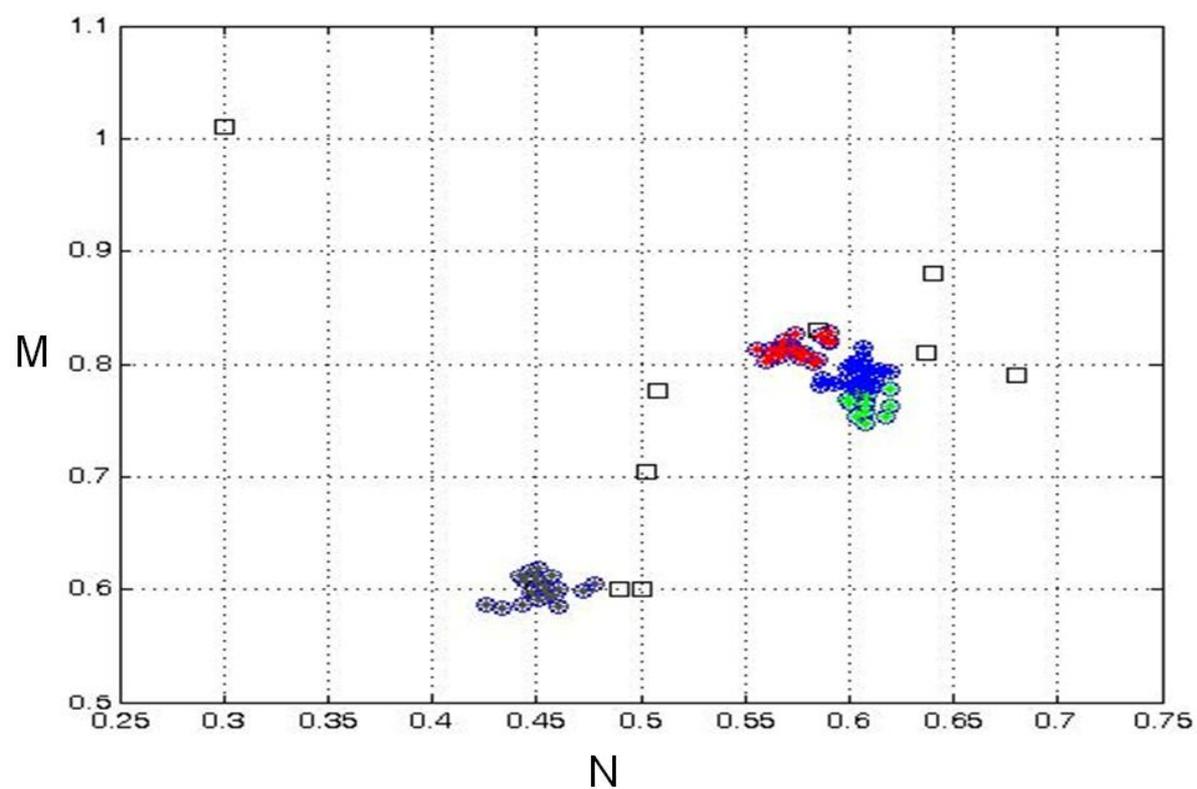
A Figura 19 mostra o resultado de classificação feita pelo fuzzy C-means clustering. Observa-se a concordância de resultados entre o FCM e o Affinity Propagation, para este caso. Para uma análise mais compacta reduzimos os dados aos centróides dos agrupamentos resultantes pela aplicação do FCM e aos exemplares resultantes da clusterização com AP (Figura 20).

Figura 18- Agrupamentos: Aplicação do Affinity Propagation para Clusterização dos dados.



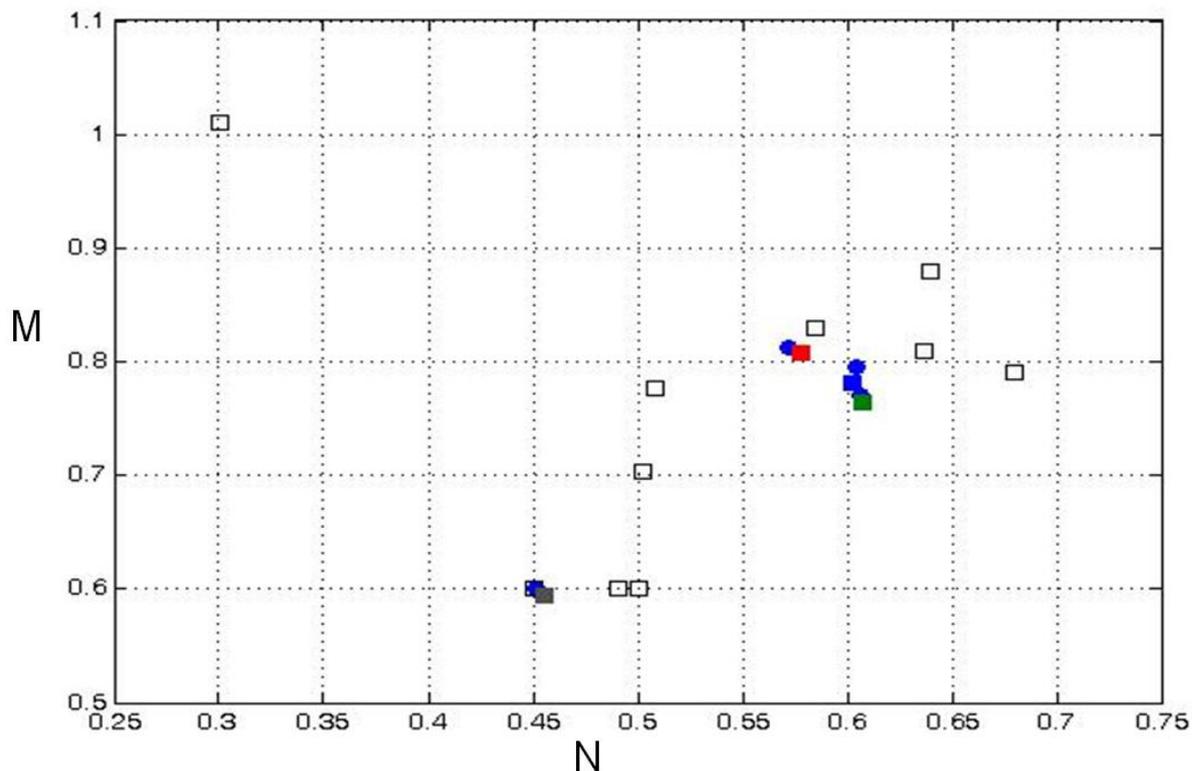
Fonte: Da autora

Figura 19- Agrupamentos por Fuzzy C-means clustering.



Fonte: Da autora

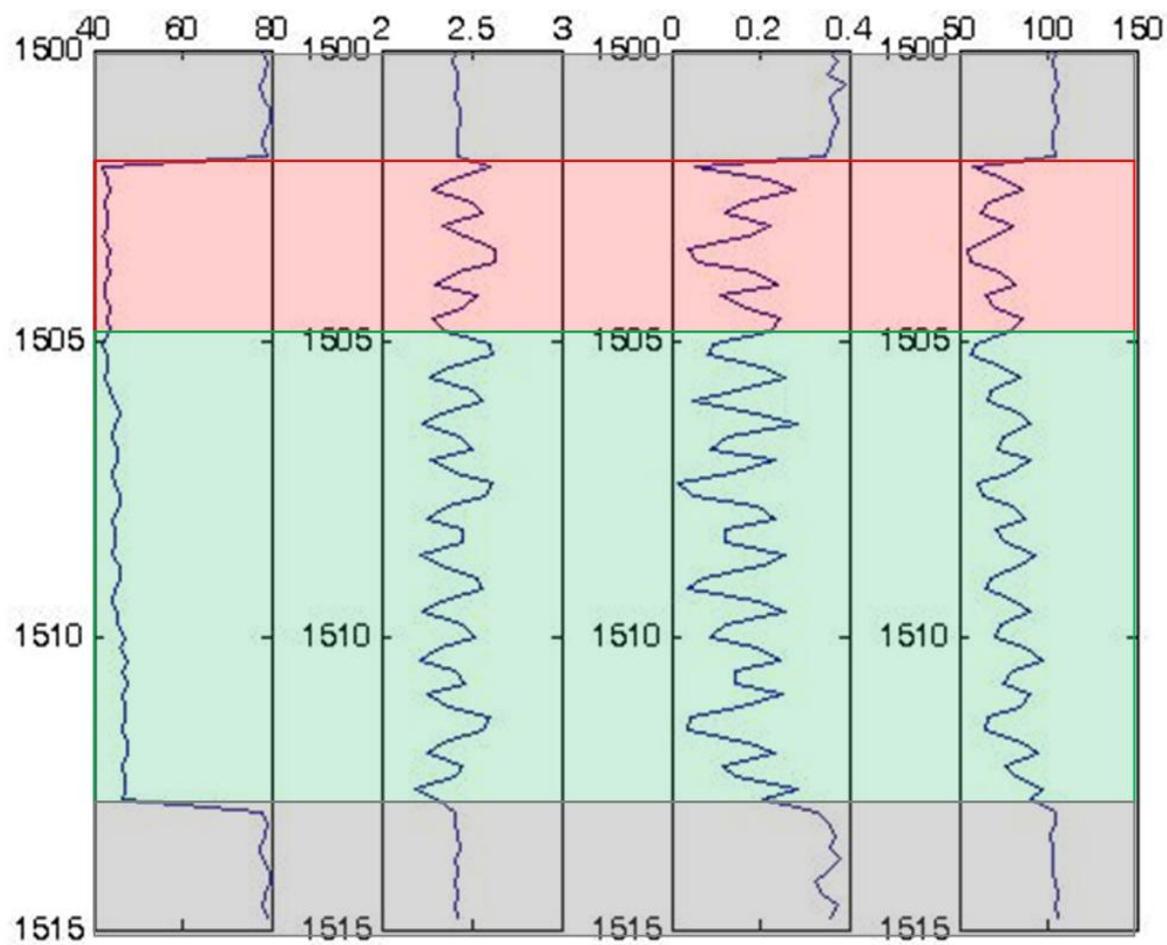
Figura 20- Os centróides estão destacados em círculos azuis, os quadrados pretos são os pontos de minerais principais e os exemplares em quadrados coloridos. agrupamentos por Fuzzy C-means clustering.



Fonte: Da autora

A Figura 21 mostra a representação em profundidade da identificação litológica produzida pelo algoritmo Affinity Propagation interpretando computacionalmente o Gráfico Vsh-M-N e o Gráfico M-N para o intervalo perfilado mostrado no Exemplo 2. A representação das litologias em profundidade estão identificadas por cores, onde arenito (verde), calcário (vermelho) e folhelho (cinza) foram classificados.

Figura 21- Zoneamento do poço em profundidade.



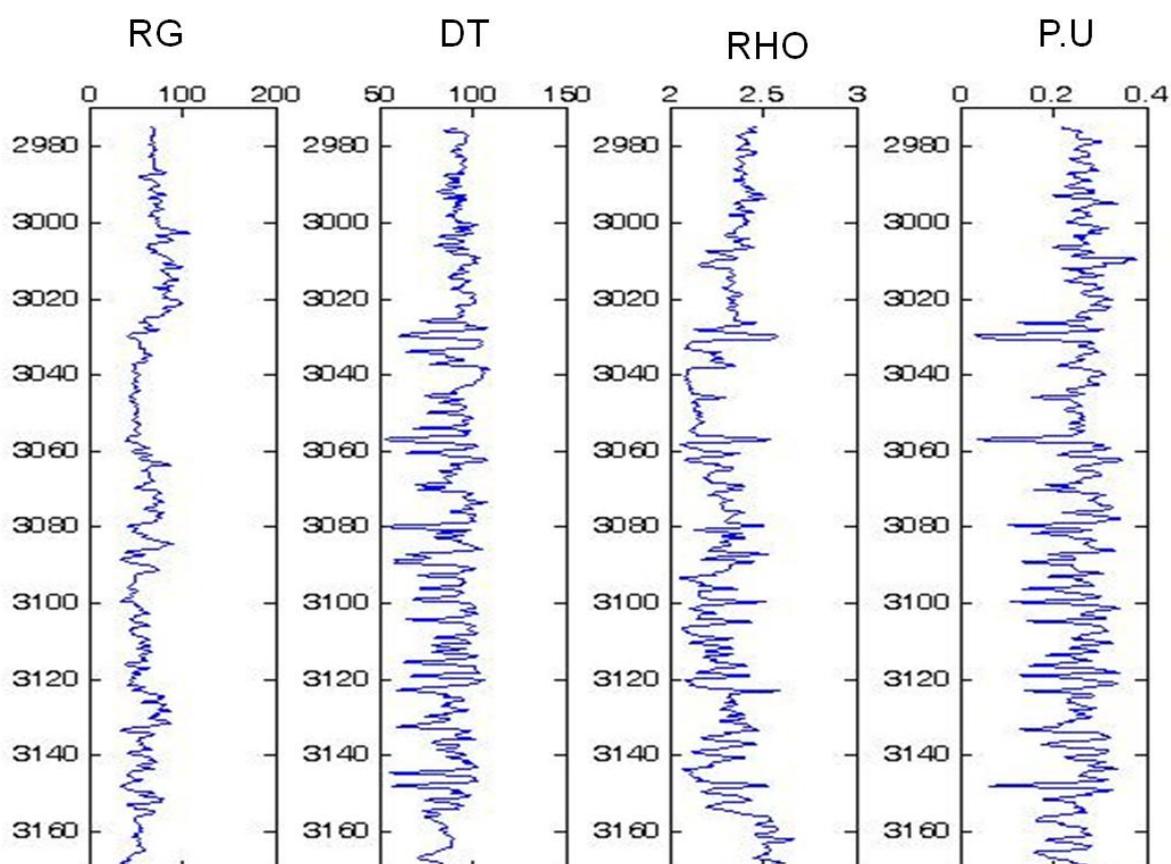
Fonte: Da autora

5.2 DADO REAL

Após a aplicação da metodologia nos dados sintéticos investiga-se a capacidade de identificação litológica do algoritmo Affinity Propagation na interpretação do Gráfico M-N em um caso real.

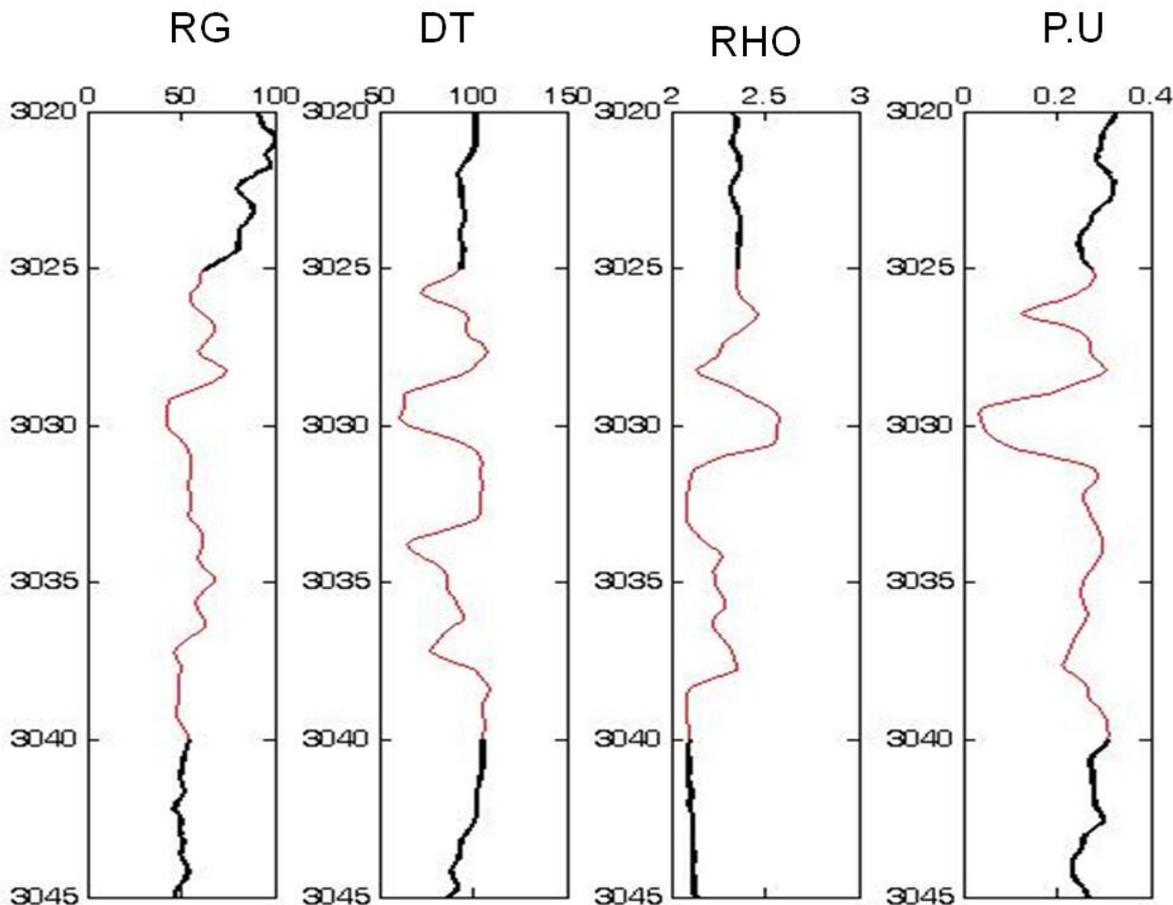
A Figura 22 mostra os perfis de dados reais registrados em um poço do cenário geológico da bacia de campos Namorado. Com o intuito de comprovar a acurácia dos resultados, foram selecionados intervalos testemunhados em profundidade. O intervalo de profundidades selecionado pode ser observado na Figura 23. Neste caso, busca-se interpretar de forma mais realista considerando a argilosidade e a mistura de minerais na constituição da matriz.

Figura 22- Perfis de dado real



Fonte: Da autora

Figura 23- Perfis de dado real. As curvas destacadas em preto representam o intervalo escolhido a ser processado na metodologia.

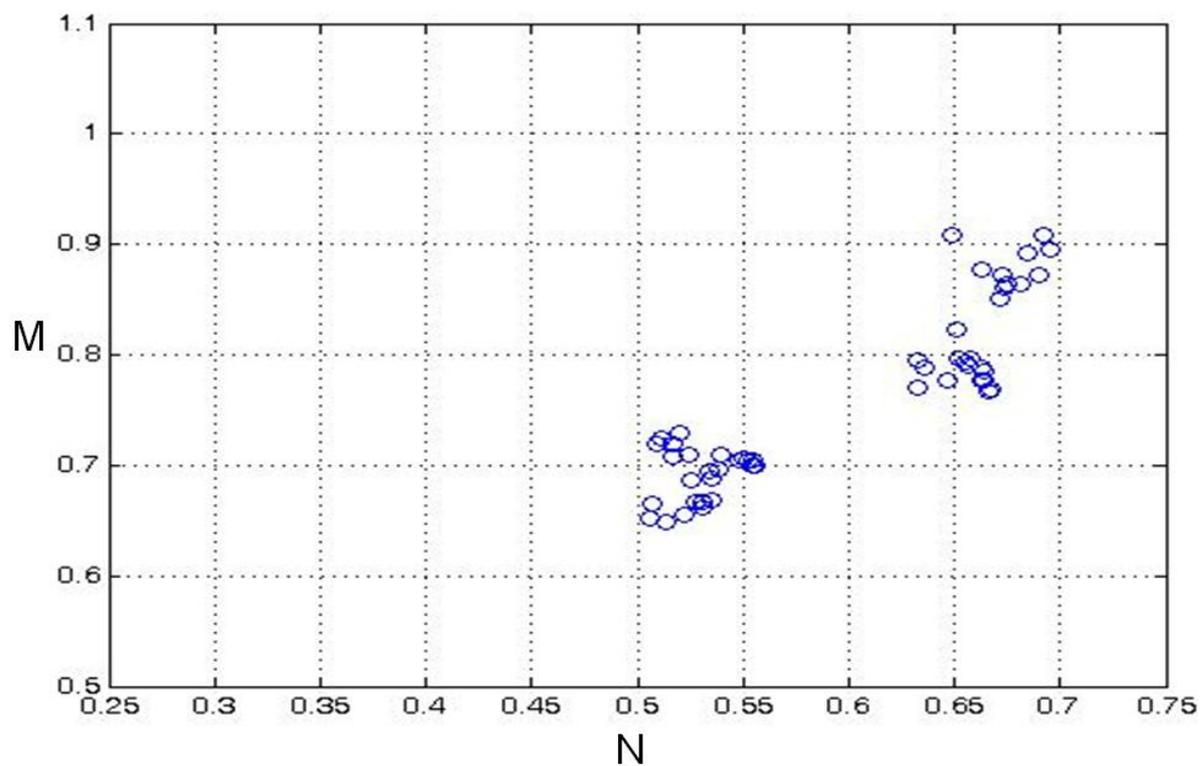


Fonte: Da autora

A Figura 24 apresenta o Gráfico M-N com os pontos representativos dos perfis mostrados na Figura 23 representados por círculos azuis. A interpretação visual da Figura 24 sugere a existência de três agrupamentos bem compactos e definidos. No entanto, não necessariamente a interpretação visual reflete o número de diferentes litologias *in situ*, visto que é comum o deslocamento dos pontos em função da argilosidade na direção dos argilominerais impedindo uma classificação mais confiável. Como solução, busca-se uma interpretação geológica para os resultados numéricos obtidos pelo algoritmo.

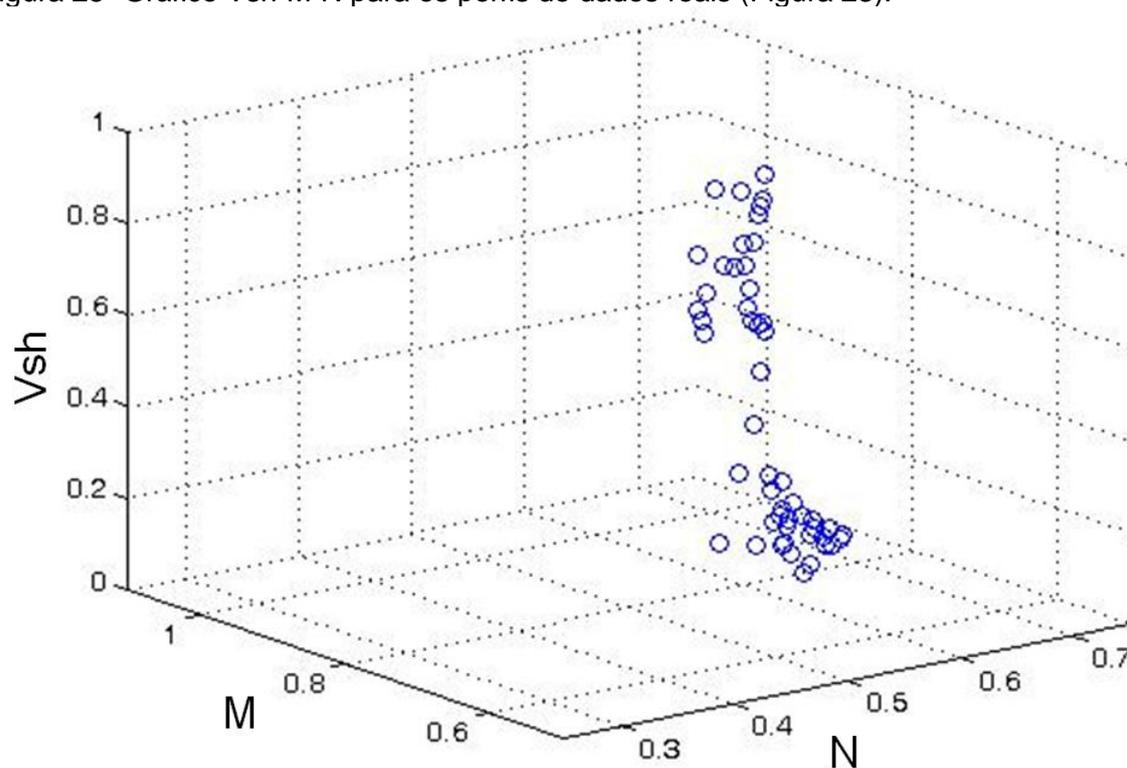
A Figura 25 mostra o correspondente Gráfico Vsh-M-N. O eixo da argilosidade foi obtido com a utilização do perfil de raios gama natural através da equação (1). Aqui a interpretação visual indica a presença de duas camadas, uma com alta argilosidade, podendo indicar um possível folhelho, a outra com baixo volume de argila, indicando uma possível camada reservatório.

Figura 24- Gráfico M-N resultantes dos perfis de dados reais representados na Figura 22.



Fonte: Da autora

Figura 25- Gráfico Vsh-M-N para os perfis de dados reais (Figura 23).

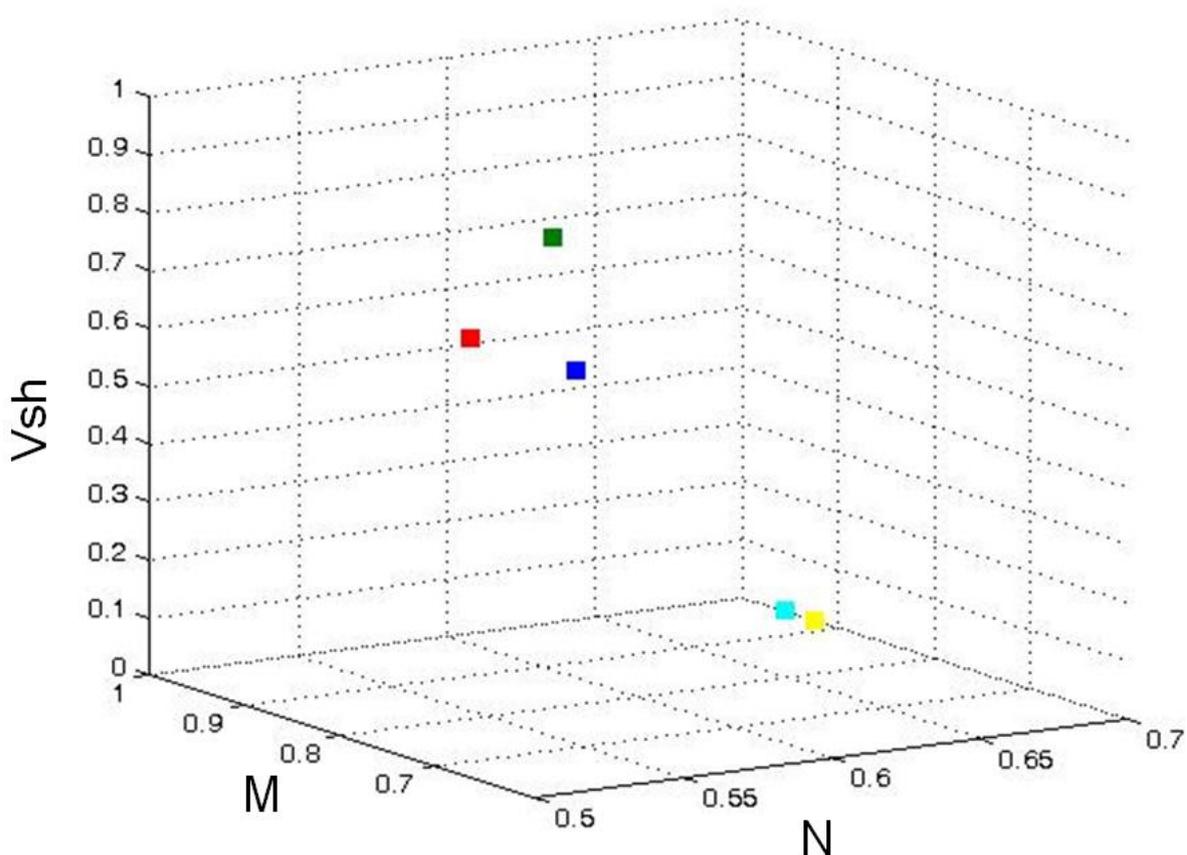


Fonte: Da autora

A Figura 26 mostra o resultado da aplicação do algoritmo Affinity Propagation para a interpretação do Gráfico Vsh-M-N mostrado na Figura 25. Em acordo com a interpretação visual do Gráfico Vsh-M-N mostrado na Figura 25, o algoritmo Affinity Propagation resultou na identificação de dois exemplares representativos das camadas reservatório. Estes exemplares são apresentados pelos quadrados azul claro e amarelo. A Figura 25 ou o Gráfico Vsh-M-N 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.

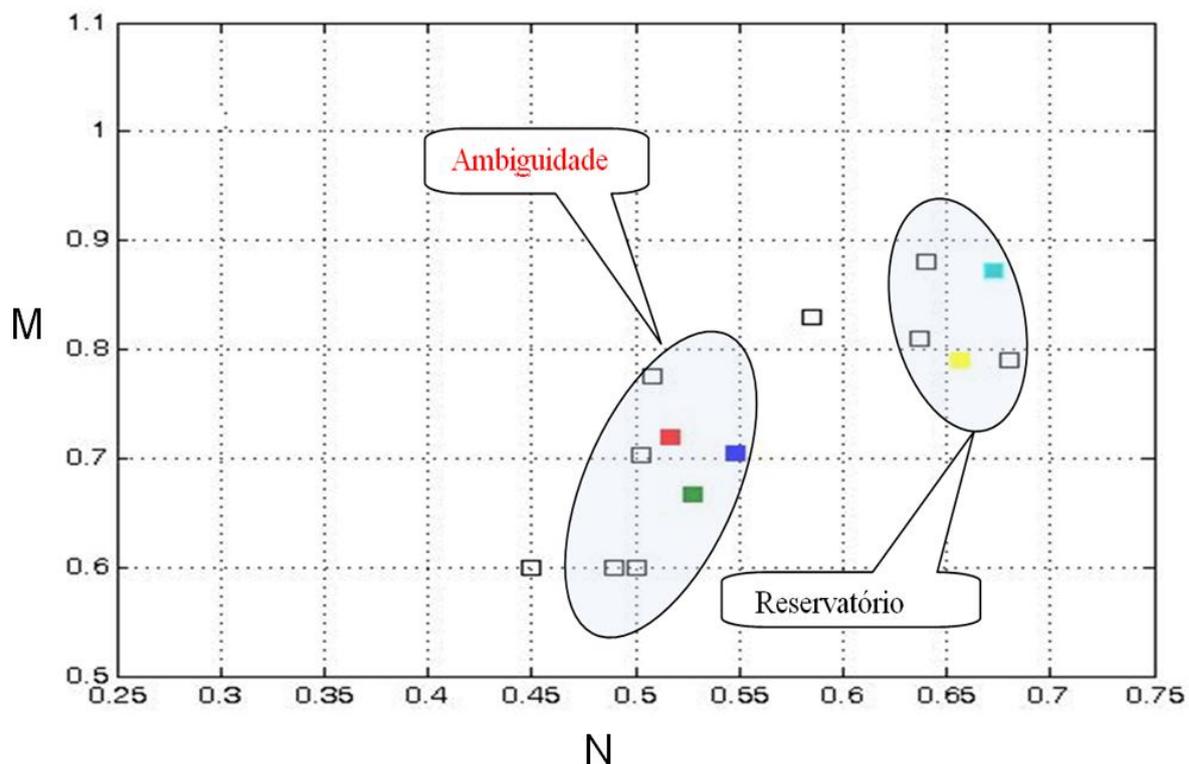
A Figura 27 mostra o resultado da classificação litológica das camadas reservatório. O exemplar azul é classificado como calcarenito, e todos os pontos do Gráfico M-N que são representados por este exemplar são classificados como calcarenito. Raciocínio semelhante é aplicado para os demais exemplares.

Figura 26- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.



Fonte: Da autora

Figura 27- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.

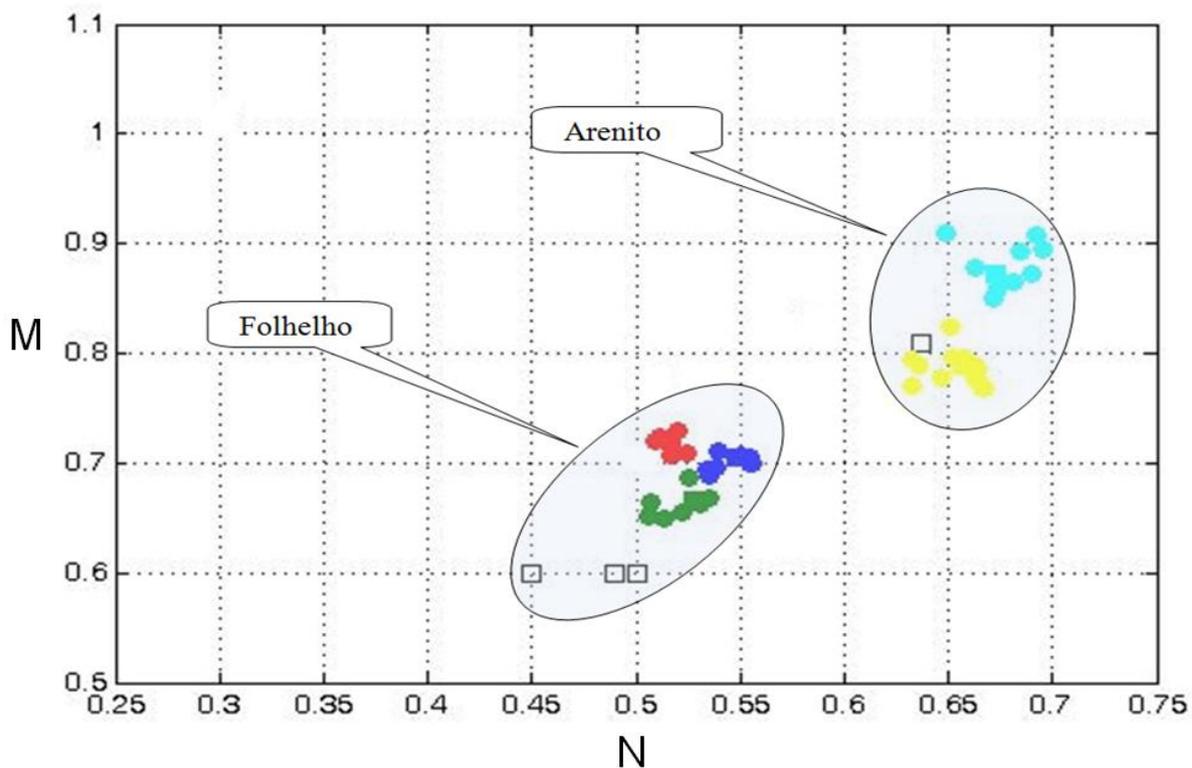


Fonte: Da autora

A Figura 28 apresenta o resultado dos agrupamentos resultantes pelo algoritmo AP. 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. Mostra-se na Figura 28 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 a rocha reservatório arenito.

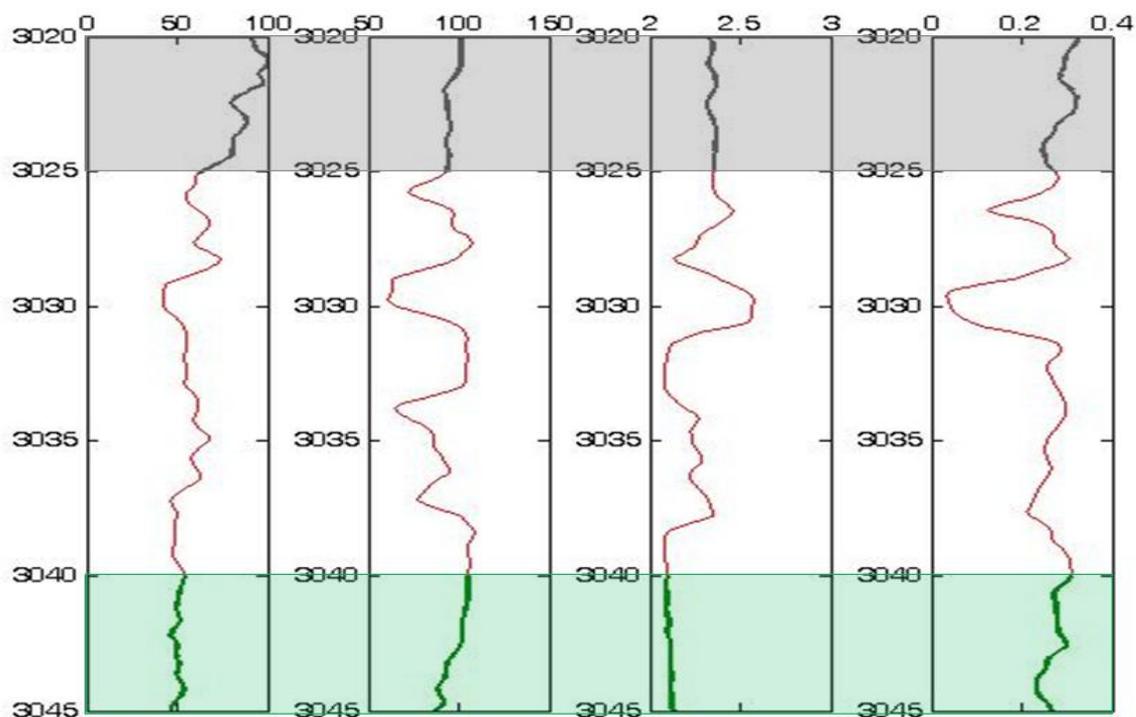
A Figura 29 mostra a representação em profundidade da identificação litológica produzida pelo algoritmo Affinity Propagation interpretando computacionalmente o Gráficos Vsh-M-N e o Gráfico M-N para o intervalo perfurado mostrado na Figura 23.

Figura 28- Representação dos exemplares a partir dos dados do gráfico Vsh-M-N.



Fonte: Da autora

Figura 29- Zoneamento do poço em profundidade. A representação das litologias em profundidade estão identificadas por arenito (verde) e folhelho (cinza).

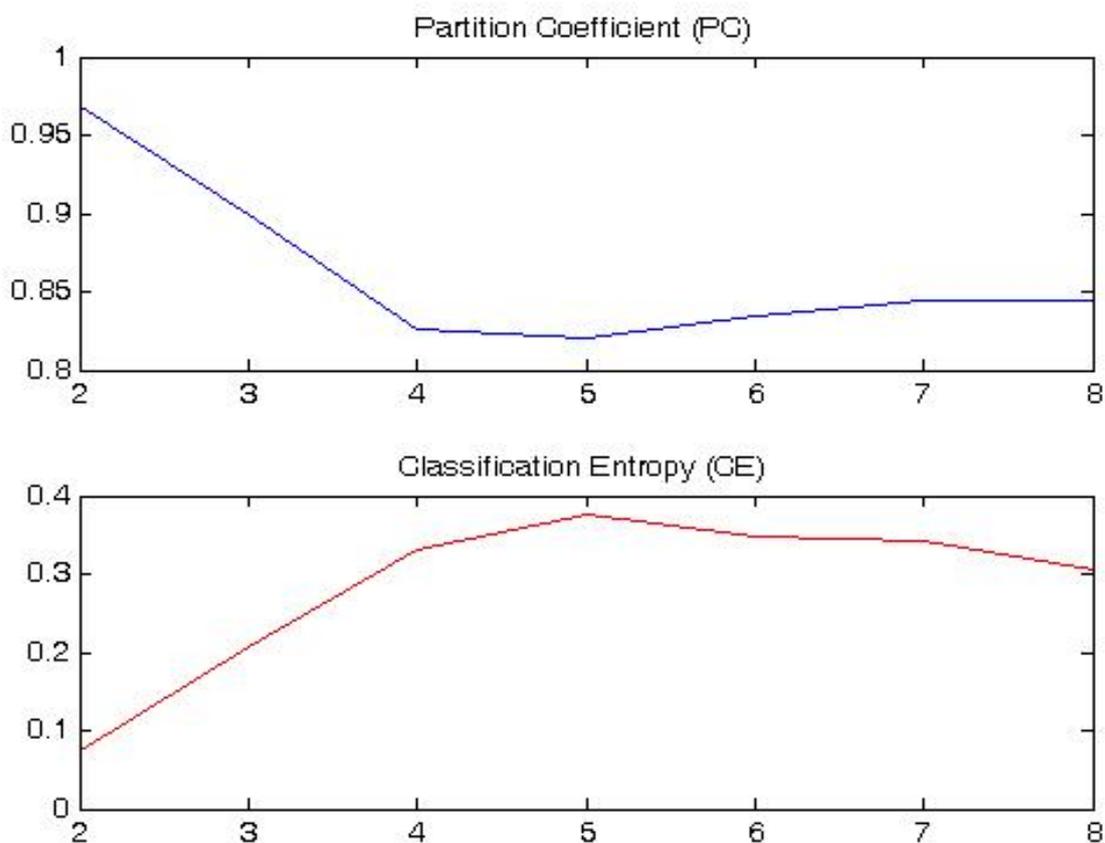


Fonte: Da autora

Para validar os resultados obtidos pela classificação litológica, usa-se os índices de validação sobre os dados do Gráfico M-N.

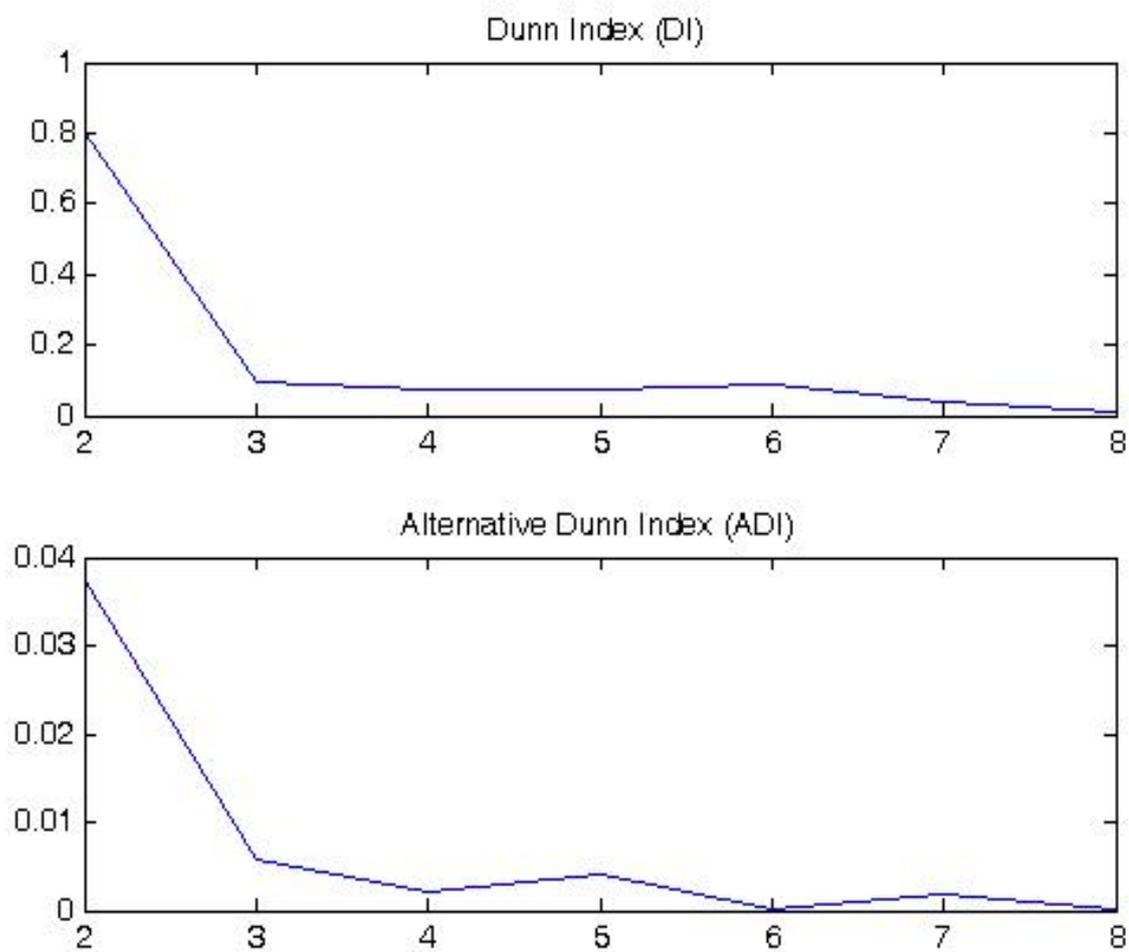
A Figura 30 apresenta os valores de índice coeficiente de partição e classificação de entropia correspondentes a diferentes números de cluster em um intervalo de [2,8]. O índice PC indica o melhor resultado quando assume valor máximo. Já o índice de classificação de entropia (CE) indica uma melhor resposta quando assume um valor mínimo. Para este caso, os melhores resultados apontam para um valor 2 de grupos. A Figura 31 apresenta os valores de índices Dunn em um intervalo de [2,8]. Ambos os índices indicam o melhor resultado quando assume valor máximo. Nesse caso, apontam para o número de grupos igual a dois, os mais alto valor para o índice. De modo geral, os índices confirmam a identificação de litologias obtidas pela metodologia.

Figura 30- Índices de validação internos: Coeficiente de partição e Classificação de entropia.



Fonte: Da autora

Figura 31- Índices de validação internos: Dunn e Dunn ajustado.



Fonte: Da autora

6 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. A contribuição deste trabalho é a de apresentar um método não supervisionado que, independente da disponibilidade ou não dos testemunhos, que é capaz de realizar a identificação e a distinção litológica das rochas atravessadas por um poço.

A representação de um agrupamento de dados pelo seu exemplar produzido pelo algoritmo Affinity Propagation na interpretação do Gráfico Vsh-M-N resulta em uma significativa simplificação e redução do conjunto de dados a serem processados, além de atenuar a ambiguidade em relação a argilidade presente no Gráfico M-N.

O benefício obtido ao utilizar o algoritmo Affinity Propagation está no fato de que não se faz necessário estabelecer *a priori* o número de agrupamentos, como parâmetro de entrada, uma exigência necessária na maioria dos métodos clássicos de agrupamento, como a média C-fuzzy (FCM). Esta obrigatoriedade força a existência de um conhecimento prévio dos dados a serem processados, que na maioria dos casos práticos da avaliação de formação não é possível.

Uma outra vantagem da utilização algoritmo Affinity Propagation em comparação a média C-fuzzy está na substituição de centróides por exemplares, onde centróides são apenas pontos representantes de um grupo no gráfico, mas não são, necessariamente, pontos dos dados processados. Os exemplares, no entanto, são representantes dos agrupamentos e ainda dados relevantes.

A análise dos dados, pelos índices de validação, mostrou-se em concordância com os resultados obtidos na identificação das litologias em profundidade.

REFERÊNCIAS

AGUIAR ;OLIVEIRA JUNIOR, **Lógica Difusa – Aspectos Práticos e Aplicações**; Editora Interciência, 2007.1 v.

BURKE, J. A.; CAMPBELL JR, R.L.; SCHIMIDT, A.W. The Litho-porosity cross plot: a method of determining rock characteristics for computation of log data. In: SPE ILLINOIS BASIN REGIONAL MEETING, 1969, Evansville, Indiana. **Illinois Basin**.Evansville, Indiana: Society of Petroleum Engineers, 1969. p. 187- 198.

BARROS, C.; ANDRADE, A.Determination of water saturation by intelligent algorithm.In: SIMPOSIO BRASILEIRO DE GEOFÍSICA, 3., 2008, Belém. **Resumosexpandidos**.Belém:SBGf, 2008.1 CD-ROM.

CHIU, S. Identificação do modelo fuzzy com base na estimativa de Cluster. **Journal of Intelligent and Fuzzy Systems**, [s.l.], v. 2, n. 3, p.267-278, 3 set. 1994.

CRAIN, E. R. **The Log analysis handbook**: quantitative log analysis methods, Tulsa: Pennwell Books, 1986. 1 v.

ELLIS, D. V. **Well logging for earth scientists**. New York: Elsevier, 1987.

FREY,B.J.; DUECK,D. Clustering by passing messages between data points.**Science**, v.315, p.972-951, fev.2007.

KOHONEN, T.**Self-organization and associative memory**. Berlin, Springer-Verlag, 1989.312 p.

LINDEN, R. Técnicas de agrupamento.**Revista de Sistema de Informação da FSMA**, [s.l.], n. 3, p. 18-36.2009.

MENDES, J.C. **Elementos de estratigrafia**, Rio de Janeiro:T.A. Queiroz Editor, 1984. v. 12

MOORE, R.C.**Meaning of facies**, 1949. Geol. Soc. Am., Mem., 39, p. 1-34.

NIKRAVESH, M. Soft computed based computational intelligent for reservoir characterization.**Expert Systems with Applications**,v. 26, p.19-38.2004.

REINECK, H. E; SINGH, I. B. **Depositional sedimentary environments**: with reference to terrigenousclastics. 2. ed. Springer: Verlag, Berlim, Heidelberg, Nova York, 1986. 549p.

RODRIGUES, C. C. **Determinação de Litologias em Perfis através de Média-C-Fuzzy**. 2007. 40 f. Trabalho de Conclusão de Curso (Graduação em Geofísica) – Universidade Federal do Pará, Instituto de Geociências, Belém, 2009.

SOUTO, M. **Validação de Agrupamentos**. DIMAp/UFRN, 2002. Disponível em: www.dimap.ufrn.br/~marcilio/AM/validacao-cluster.ppt.

WELLER, J.M. Stratigraphic facies differentiation and nomenclature. **Am.Ass.Petrol.Geologists, Bull.**, 40, p.17-50.1958.