

6º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

Visualização e Classificação Automática de Petrofácies Sedimentares

AUTORES:

Juliana Tonoli Cevolani¹, Leonardo Costa de Oliveira¹, Leonardo Goliatt², Egberto Pereira³

INSTITUIÇÃO:

1 Centro Universitário Norte do Espírito Santo (CEUNES-UFES)

2 Universidade Federal de Juiz de Fora (UFJF)

3 Universidade do Estado do Rio de Janeiro (UERJ)

Este Trabalho foi preparado para apresentação no 6º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás- 6º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás-ABPG, no período de 09 a 13 de outubro de 2011, em Florianópolis-SC. Esse Trabalho foi selecionado pelo Comitê Científico do evento para apresentação, seguindo as informações contidas no documento submetido pelo(s) autor(es). O conteúdo do Trabalho, como apresentado, não foi revisado pela ABPG. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões da Associação Brasileira de P&D em Petróleo e Gás. O(s) autor(es) tem conhecimento e aprovação de que este Trabalho seja publicado nos Anais do 6ºPDPETRO.

VISUALIZAÇÃO E CLASSIFICAÇÃO AUTOMÁTICA DE PETROFÁCIES SEDIMENTARES

Abstract

This work aimed to develop a computational method based on clusters models to determine petrofacies amount in sedimentary thin-sections. The data from boreholes drilled in the Parana Basin was processed and analyzed through a computational method implemented to infer the different petrogenetic associations in the studied samples. The data mining technique have proved useful in the reservoirs heterogeneities interpretation accelerating the process of classification and sharply reducing the labor expended by geologist/analyst when using the conventional method of petrofacies classification.

Introdução

O estudo de reservatórios de hidrocarbonetos é estrategicamente importante para a caracterização e definição da produtividade e o caráter comercial de um campo de óleo e gás. Para uma rocha ser considerada um bom reservatório, a mesma deve possuir as seguintes características: (i) principalmente, uma extensão considerável, (ii) boa porosidade (em torno de 15% a 20%), (iii) uma apreciável permeabilidade e (iv) uma boa eficiência de recuperação. Essas características, denominadas petrofísicas, são as preponderantes na definição da qualidade do reservatório. As alterações decorrentes da diagênese, que acometem os sedimentos após estes serem depositados, controlam fortemente as características petrofísicas dos reservatórios. As fácies sedimentares são um conjunto de características que individualizam um determinado grupo de rochas, tais como: cor, granulometria, estrutura sedimentar, geometria deposicional, presença de fósseis, paleocorrente, entre outros [1]. A quantificação das petrofácies sedimentares em um reservatório permite a inferência do grau de heterogeneidade do mesmo.

Os modelos preditivos bem como o mapeamento da distribuição de heterogeneidades e qualidade em reservatórios são de fundamental importância para exploração e otimização da produção de campos de petróleo. No entanto, as etapas metodológicas que envolvem todo o processo de amostragem, posterior geração de dados e futuras interpretações dos mesmos, por vezes é muito longa e nem toda informação é aproveitada, em razão do excesso de dados que o geólogo tem em mãos. Dessa maneira, à medida que a quantidade de dados cresce, torna-se necessária a mudança do uso de métodos diretos realizados manualmente, para análises indiretas e automáticas usando ferramentas mais complexas e sofisticadas que otimizem a análise dos dados em conjunto.

Nesse contexto, as técnicas de Mineração de Dados (MD) [2] surgem como uma ferramenta interessante para auxiliar na identificação de petrofácies sedimentares. Estas técnicas são frequentemente utilizadas para ganhar conhecimento sobre os dados e se baseiam na busca por padrões que ajudam na explicação de determinadas características dos dados de interesse. Os modelos diagenéticos podem ser aperfeiçoados através das técnicas de agrupamento, pelas quais são procurados um conjunto de elementos que juntos se adequam; técnicas de agrupamento, que classificam exemplos que tinham passado despercebidos; ou ainda por associações entre as características dos dados analisados, os quais resultam em um modelo expresso como um código computacional em software adequado, usado para classificação, predição, estimação ou tarefas similares.

Este trabalho objetiva o desenvolvimento de um modelo computacional para a caracterização de petrofácies que auxilie nas etapas posteriores da análise de bacias sedimentares, permitindo, sobretudo, elucidar algumas relações e interações na diagênese de rochas siliciclásticas.

Metodologia

A metodologia usada neste trabalho envolve dois procedimentos principais: (1) visualização dos agrupamentos das lâminas em duas dimensões, com o objetivo de sugerir ao geólogo as possíveis

petrofácies e facilitar o entendimento da relação entre as lâminas, e (2) elaboração de um procedimento computacional para determinação do número de petrofácies e posterior classificação das lâminas nas petrofácies determinadas. Os dados resultantes da análise das lâminas são interpretados como entidades matemáticas e a forma com que se agrupam é tratada por meio de algoritmos de redução de dimensionalidade e de agrupamentos de dados, para facilitar a visualização e classificação das lâminas. O processo de obtenção dos dados é descrito em seguida.

Os arenitos, que compõem a base de dados para as considerações geológicas e computacionais, são da seção devoniana da Bacia do Paraná, registradas nas regiões de Tibagi e Ponta Grossa, no estado do Paraná. Estes arenitos foram amostrados a partir de um conjunto de furos de sondagem (posicionados entre os paralelos 24°00' S e 25°30' S e entre os meridianos 50°45' W e 49°30' W) disponibilizado pelo projeto Arcabouço Estratigráfico refinado das Formações Ponta Grossa e Irati das Bacias do Paraná e Chaco Paraná (Paleosul). Os dados deste trabalho são de dois destes furos (Figura 1). Os testemunhos encontram-se disponíveis no Laboratório Geral de Processamento de Amostras da Faculdade de Geologia da Universidade do Estado do Rio de Janeiro (LGPA-FGEL-UERJ).

A análise quantitativa das amostras foi feita através da contagem (em microscópio petrográfico de luz transmitida) de 300 pontos em cada lâmina petrográfica com espaçamento de 0.3 mm. Em cada lâmina foram contabilizadas as porcentagens de 22 constituintes: quartzo, feldspatos totais, muscovita, minerais opacos, turmalina, zircão, rutilo, glauconita, clorita, pseudo matriz, litoclasto, bioclasto, crescimento secundário de quartzo, caolinita, illita/smectita, pirita, siderita, cimento carbonático, cimento silicoso, cimento ferruginoso, porosidade intergranular e porosidade intragranular [3]. Cada lâmina extraída do furo de sondagem é uma amostra com 22 atributos, que correspondem à quantidade de constituintes contabilizados. Para uma melhor avaliação da história deposicional da seção estratigráfica estudada, as lâminas de um mesmo furo de sondagem com características diagenéticas semelhantes são agrupadas em petrofácies.

As petrofácies foram classificadas pelo método convencional (individualização manual), usando-se como critério diferenciador a porcentagem dos minerais detríticos das rochas estudadas, bem como as características texturais dos mesmos. A determinação manual é um processo intensivo que requer uma quantidade de tempo considerável do especialista. De acordo com os resultados obtidos com a análise quantitativa dos dados petrográficos foi possível individualizar os arenitos (pelo método convencional) em 4 petrofácies distintas: PT-1, PT-2, PT-3 e PT-4 [4]. As petrofácies denominadas PT-1 e PT-3 estão presentes nos poços em discussão neste trabalho e intercalam-se ao longo dos furos. Para o Poço 1 foram analisadas 12 lâminas classificadas de DEPA-

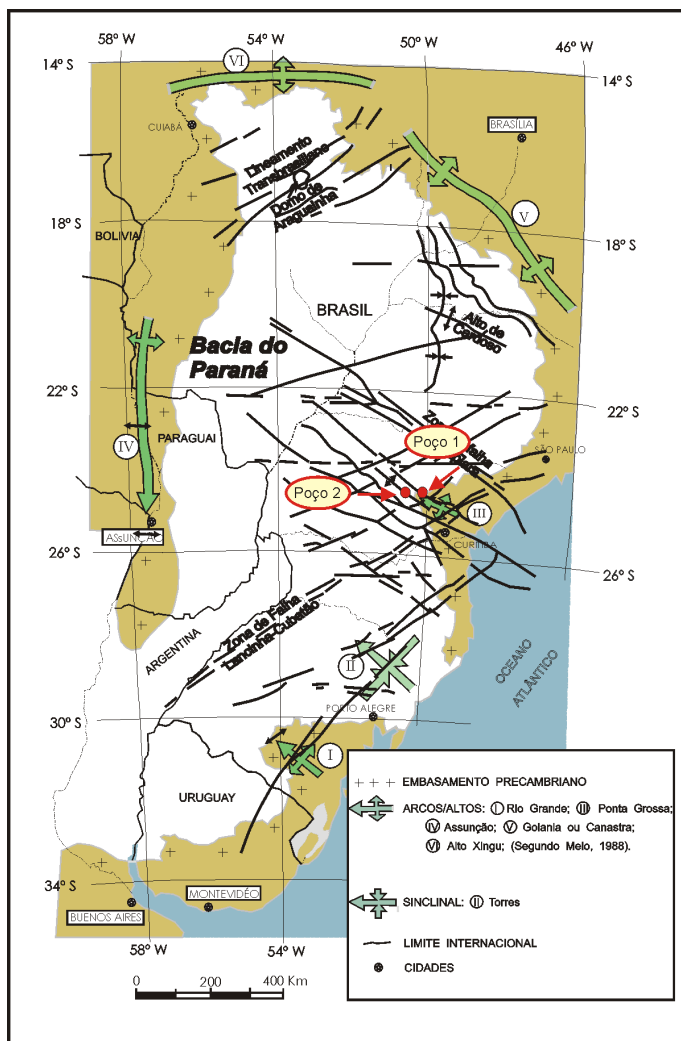


Figura 1: Mapa de localização da área de estudo e dos furos estudados (modificado de [3]).

83 até DEPA-94. A lâmina DEPA-93 foi identificada como pertencente à petrofácies denominada PT-3; as demais lâminas foram agrupadas à petrofácies chamada PT-1. Do Poço 2 foram analisadas 4 lâminas, DEPA-100 a DEPA-103, onde a lâmina DEPA-102 foi classificada na petrofácies PT-3 e o restante na petrofácies PT-1.

Para a visualização dos agrupamentos foi usada a técnica de redução de dimensionalidade baseada em análise de componentes principais. A análise de componentes principais (ACP) é uma técnica estatística para projetar dados n -dimensionais em um espaço de baixa dimensão. Na representação de um problema, cada variável tem a mesma importância estatística. As componentes principais são variáveis geradas através de uma transformação linear sobre as variáveis originais, onde cada componente principal traz uma informação estatística diferente das outras, e com importância estatística decrescente. As componentes principais são ortogonais entre si, e dessa maneira podem ser analisadas separadamente, servindo para visualizar o conjunto de amostras apenas pelo gráfico das duas primeiras componentes principais, que detêm maior parte da informação estatística. A ACP auxilia na elaboração de hipóteses gerais a partir dos dados coletados e na separação da informação importante da redundante e aleatória. Para a visualização dos dados foi usada a técnica de escalonamento multidimensional (EM). O EM [7] consiste no emprego de técnicas para análise de proximidade em conjunto de dados com o objetivo de revelar sua estrutura topológica. Esta técnica toma medidas de distâncias (ou dissimilaridades) entre objetos e retorna um conjunto de pontos (em duas ou três dimensões) de modo que as suas distâncias sejam aproximadamente as mesmas. Como resultado, os dados das lâminas com 22 dimensões (22 constituintes) são reduzidos para duas que detêm maior parte da informação estatística, onde os agrupamentos podem ser visualizados. Este procedimento permite visualizar aproximadamente como as lâminas naturalmente se agrupam, assim como a formação de diferentes agrupamentos. Cada agrupamento é constituído de lâminas com características petrológicas semelhantes, e sugere ao geólogo/analista uma ideia da formação das petrofácies.

Embora auxilie na classificação, esta metodologia pode levar a erros quando os agrupamentos encontram-se muito próximos, e neste caso um mesmo conjunto de lâminas pode ser agrupado de diferentes formas. Para contornar este problema, foi desenvolvido um procedimento para a detecção automática do número de agrupamentos e a posterior classificação das lâminas nos agrupamentos encontrados. O método proposto tem o intuito de individualizar cada petrofácies por meio de um procedimento computacional baseado em Análise de Agrupamentos (AA). Através deste método computacional é possível reunir objetos em grupos, de tal forma que os objetos que estão em um mesmo grupo são mais similares entre si do que objetos que estão em outro grupo, segundo uma medida de similaridade. Quanto ao método de formação do agrupamento, os algoritmos de AA podem ser classificados em hierárquicos (que constroem uma hierarquia representada graficamente por árvores hierárquicas ou dendrogramas) e não hierárquicos (ou de partição). Com base no objetivo, foi usado o algoritmo de partição em torno de medóides (PAM, *Partitioning Around Medoids*) [5] para formar os agrupamentos. O PAM, apesar de ter a desvantagem de não ser eficiente em grandes volumes de dados, é uma versão mais robusta do algoritmo das k -médias, um dos mais empregados em tarefas de agrupamento. Ele se baseia na busca por objetos representativos, chamados medóides. Os medóides são os objetos cuja a média de dissimilaridade para todos os outros objetos é mínima, sendo a dissimilaridade a distância euclidiana entre o objeto representativo (medóide) e os demais objetos do conjunto de dados. A quantidade de medóides encontrada corresponderá à quantidade de agrupamentos formados. Cada agrupamento constitui então uma petrofácies no poço, em função dos dados dos constituintes diagenéticos.

Para avaliar os agrupamentos em função do número de petrofácies, usou-se a técnica de análise de silhueta, que fornece para cada objeto uma medida de posicionamento dentro de um agrupamento. Para cada observação i , a largura de silhueta $s(i)$ é definida como

$$s(i) = (b(i) - a(i)) / \max(a(i), b(i))$$

onde $a(i)$ é a dissimilaridade média entre i e todos os outros objetos dentro do mesmo agrupamento ao qual i pertence. Considerando um outro agrupamento C e $d(i,C)$ a dissimilaridade média entre i e todas as observações de C , o menor destes $d(i,C)$ é $b(i) = \min_C d(i,C)$, e pode ser visto como a

dissimilaridade entre i e seu agrupamento vizinho, ou seja o agrupamento mais próximo ao qual i não pertence. Este então seria a segunda melhor escolha para o objeto i . Se i é a única observação dentro do seu agrupamento então define-se $s(i)=0$. Observações com valores de $s(i)$ próximos de 1 são considerados bem agrupados, enquanto valores de $s(i)$ próximos de zero indicam que a observação recai entre dois agrupamentos. As observações com $s(i)<0$ indicam que o objeto i pode estar no agrupamento errado. A medida de silhueta de um agrupamento é a média das silhuetas de seus objetos e mostra quais objetos estão bem classificados no interior de seus agrupamentos e quais estão meramente em uma posição intermediária entre eles.

Para se obter o número de petrofácies (agrupamentos) adequada para cada poço procedeu-se da seguinte maneira: (i) o algoritmo PAM foi executado para 2, 3, 4 e 5 agrupamentos (ii) calculou-se o coeficiente de silhueta global (média das medidas de silhueta de cada objeto no agrupamento) (iii) os valores dos coeficientes de silhueta foram arranjados em uma tabela em função do número de agrupamentos (iv) o número de petrofácies adequado é a quantidade de agrupamentos que retorna a maior medida de silhueta.

Resultados e Discussão

O método computacional foi implementado em linguagem R [6]. Foram usados os pacotes de agrupamentos de dados e de análise estatística. Com base na informação fornecida pelo geólogo, buscou-se através do método computacional identificar as petrofácies e confrontá-las com os dados gerados na classificação manual.

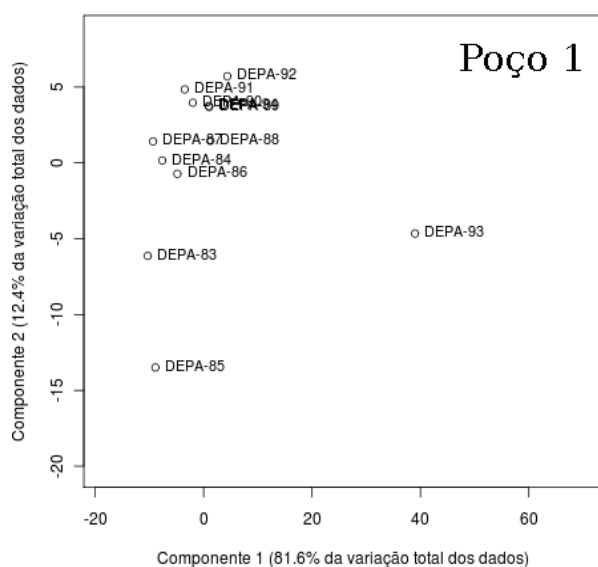


Figura 2: Visualização dos agrupamentos para as lâminas correspondentes ao Poço 1.

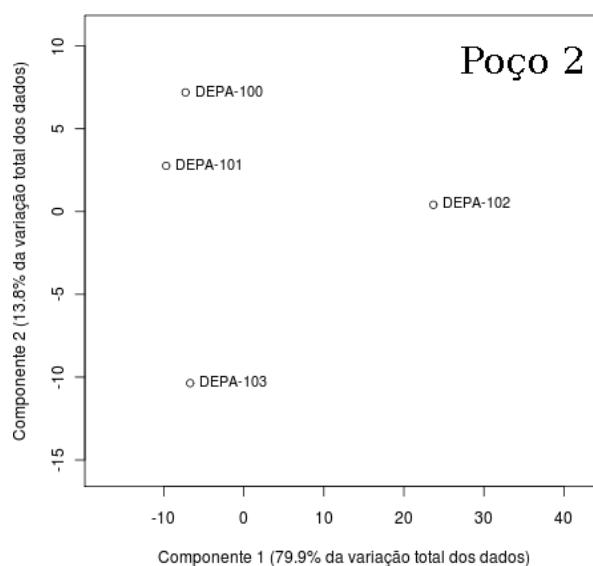


Figura 3: Visualização dos agrupamentos para as lâminas correspondentes ao Poço 2

A Figura 2 mostra o resultado da análise de componentes principais para o Poço 1. A análise revelou que as duas primeiras componentes descrevem 94% da variação total dos dados, onde a primeira componente principal descreve 81,6% da variação total e 12,4% é descrita pela segunda componente principal. Analisando a disposição dos dados em função das duas componentes principais percebe-se um agrupamento nítido das lâminas DEPA-92, DEPA-91, DEPA-90, DEPA-89, DEPA-94, DEPA-88, DEPA-87, DEPA-84 e DEPA-86. Este agrupamento de lâminas de características semelhantes sugere que estas pertencem à uma mesma petrofácies. Porém, outras formas de agrupamento podem ser encontradas na visualização da distribuição das lâminas do Poço 1. A

visualização da disposição das lâminas para o Poço 2 é mostrada na Figura 3. A primeira e segunda componentes descrevem respectivamente 79,9% e 13,8% da variação total dos dados. Duas formas distintas de agrupamento podem ser vistas para este poço, ambas constituindo duas petrofácies: (i) um agrupamento formado pelas lâminas DEPA-100 e DEPA-101 e outro formado pelas lâminas DEPA-102 e DEPA-103 ou (ii) um agrupamento formado por DEPA-100, DEPA-101 e DEPA-103 e outro com somente uma lâmina, DEPA-102. Neste caso torna-se necessária a intervenção do geólogo, que deve usar seu conhecimento para realizar a classificação, de acordo com os constituintes presentes em cada amostra.

Embora este procedimento permita visualizar de maneira rápida o agrupamento das lâminas, não quantifica quantos agrupamentos podem existir, e não faz uma designação das lâminas aos agrupamentos. Os algoritmos de agrupamento, juntamente com a técnica de silhueta permitem realizar esta tarefa, e fornecer estimativas mais precisas da quantidade de agrupamentos e da maneira como são formados.

As medidas de silhueta em função do número de agrupamentos para os poços analisados são mostrados na Tabela 1. As maiores medidas silhueta aparecem em negrito. O Poço 2, por possuir somente 4 lâminas, permite somente o agrupamento em 2 ou 3 petrofácies, uma vez que, por hipótese, pelo menos uma das petrofácies deve conter 2 lâminas. Esta hipótese foi inserida dentro do código computacional para evitar que o algoritmo separasse n lâminas em n petrofácies, ou seja, cada petrofácies com somente uma lâmina. O número ideal de agrupamentos é aquele que retorna a maior medida de silhueta. Os resultados obtidos para as medidas de silhueta sugerem duas petrofácies para ambos os poços. A separação dos dados em 3 ou mais petrofácies resultaram em baixos valores de silhueta, evidenciando que para estes poços a separação em duas petrofácies é a mais adequada.

Poço \ N° de Petrofácies	2	3	4	5
Poço 1	0,66	0,26	0,25	0,20
Poço 2	0,40	0,12	—	—

Tabela 1: Medidas de silhueta para os Poços 1 e 2 considerando 2, 3, 4 e 5 agrupamentos (petrofácies).

O objetivo do método computacional aqui proposto é individualizar agrupamentos “intra-poço” e facilitar as decisões do geólogo na análise das petrofácies. Buscou-se então identificar as petrofácies e confrontá-las com os dados gerados na classificação manual. As Figuras 4 e 5 mostram o agrupamento das lâminas dos Poços 1 e 2 em duas petrofácies. A classificação resultante da análise manual aparece entre parênteses, e a classificação obtida com o método computacional desenvolvido é mostrada pelos agrupamentos indicados por quadrados (□) e círculos (○). Os resultados mostram que o método desenvolvido foi capaz de quantificar o mesmo número de petrofácies encontrado na classificação manual. Além disso, classificou corretamente todas as lâminas, designando cada uma à sua respectiva petrofácies.

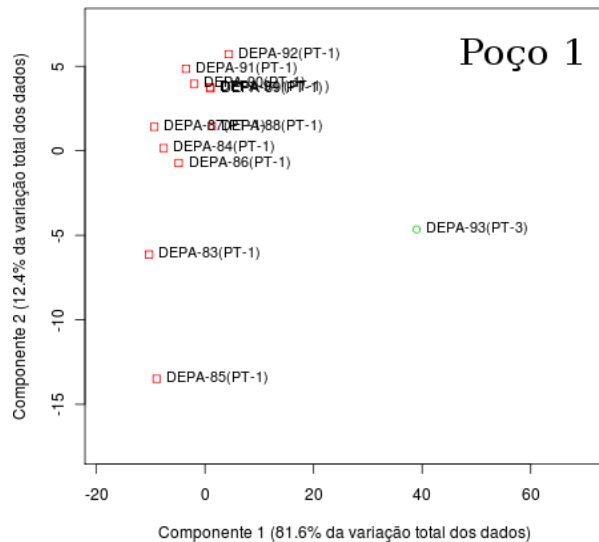


Figura 4: Agrupamentos (petrofácies) encontrados a partir do método computacional para o poço 1. Os símbolos ○ (círculo) e □ (quadrado) correspondem a dois diferentes agrupamentos. Entre parênteses são mostrados os agrupamentos encontrados a partir da classificação manual.

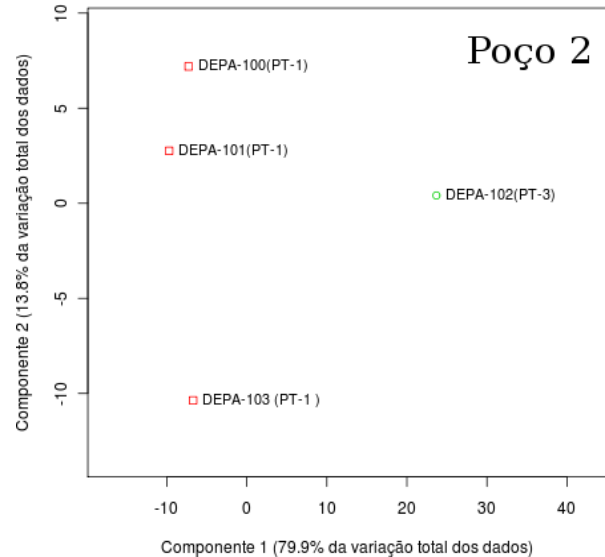


Figura 5: Agrupamentos (petrofácies) encontrados a partir do método computacional para o poço 2. Os símbolos ○ (círculo) e □ (quadrado) correspondem a dois diferentes agrupamentos. Entre parênteses são mostrados os agrupamentos encontrados a partir da classificação manual.

Comparando os resultados do método convencional (individualização manual) com os do método computacional é possível confirmar a eficiência da metodologia empregada, pois como visto, a quantidade de petrofácies identificadas nos dois métodos foram idênticas para os dois poços, bem como a distribuição das lâminas para cada petrofácies.

O método usado anteriormente, baseado somente no escalonamento multidimensional, pode conduzir a diversas interpretações, pois depende do conhecimento do analista sobre os dados que estão sendo tratados. Sem o conhecimento geológico dos dados do Poço 1, por exemplo, poderíamos classificar as lâminas DEPA-83, DEPA-93 e DEPA-85 em um mesmo agrupamento o que não seria coerente com a análise dos constituintes das lâminas do ponto de vista geológico. Assim, tal procedimento serve como uma análise prévia dos resultados, mas que deve ser complementada pela indicação no número de petrofácies e o posterior agrupamento das lâminas, caso não seja utilizado o método convencional (individualização manual). Dessa forma, o uso da Análise de Agrupamentos juntamente com a análise de silhueta se mostrou um método eficiente que permite otimizar o processo de análise de lâminas e determinação de petrofácies, pois, tais tarefas, quando realizadas de maneira convencional, exigem do geólogo/analista, além de grande experiência em petrografia sedimentar, tempo para adquirir conhecimento geológico suficiente sobre os dados de interesse, e mais tempo para contagem dos constituintes de cada lâmina, normalização dos dados e, por fim, a classificação das lâminas em diferentes petrofácies. Assim, empregando-se o método computacional o geólogo economiza tempo e trabalho na individualização das petrofácies, e consegue chegar a resultados similares aos que seriam obtidos quando do uso do método convencional.

Conclusões

O código computacional desenvolvido facilitou a visualização dos agrupamentos das petrofácies dessas rochas, permitindo que se obtenham de maneira rápida informações sobre a similaridade entre as amostras em um mesmo poço. As técnicas empregadas mostraram-se úteis na interpretação das heterogeneidades dos reservatórios, podendo, portanto, otimizar o processo de exploração e posterior produção de campos de petróleo em rochas análogas às aqui estudadas,

acelerando o processo de classificação e diminuindo o trabalho despendido pelo geólogo/analista ao usar o método convencional. O próximo passo no desenvolvimento deste trabalho é desenvolver um modelo computacional para a identificação da mesma petrofácies em diferentes poços e assistir as etapas posteriores da análise de bacias sedimentares.

Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Espírito Santo (FAPES) pela concessão de bolsa de Iniciação Científica para desenvolvimento do estudo e à Universidade do Estado do Rio de Janeiro (UERJ) por conceder acesso às lâminas utilizadas para análise e obtenção dos dados.

Referências Bibliográficas

- [1] J. C. Della Fávera. *Fundamentos de Estratigrafia Moderna*. EdUERJ, 2001.
- [2] D. Hand and A. Manilla. *Principles of Data Mining*. MIT, 2001.
- [3] L.C. Oliveira, E. Pereira. Aplicação de parâmetros diagenéticos para a caracterização do arcabouço estratigráfico do Devoniano da Bacia do Paraná. *In 5º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás, 2009*.
- [4] Oliveira, L. C. 2009. Estudo das relações entre o arcabouço estratigráfico e as alterações diagenéticas observadas na seção Devoniana da bacia do Paraná. Dissertação de mestrado. Faculdade de Geologia – UERJ. Brasil, 100 p.
- [5] L. Kaufman, P.J. Rousseeuw. *Finding Groups in Data An Introduction to Cluster Analysis*. Wiley Interscience, New York, 1990.
- [6] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0.
- [7] T. F. Cox, M. A. A. Cox. *Multidimensional Scaling*. Chapman and Hall, 1994.