

ANÁLISE DE DADOS MULTIVARIADOS: UMA EFICIENTE FERRAMENTA PARA DESCRIÇÃO E CARACTERIZAÇÃO DE RESERVATÓRIOS

MULTIVARIATE DATA ANALYSIS: AN EFFICIENT TOOL IN RESERVOIR DESCRIPTION AND CHARACTERIZATION

Olinto G. Souza Jr.¹

RESUMO — Os campos de hidrocarbonetos, principalmente aqueles em adiantado estágio de desenvolvimento, possuem grande quantidade de dados originados de perfis, testes de formação, testemunhos, descrição de amostras de calha e histórico de produção. O grande desafio enfrentado pelos grupos multidisciplinares de gerenciamento de reservatórios é manusear com eficiência este fabuloso volume de dados, além de promover a transferência destas informações para diferentes escalas. Este conjunto de dados, por apresentar um caráter multivariado, e por estar sujeito a variabilidades aleatórias, pode ser estudado por meio de modelos probabilísticos. Este enfoque tem sido rotineiramente aplicado em estudos de reservatórios realizados nos campos de hidrocarbonetos da Bacia Potiguar. O objetivo deste trabalho é fazer uma breve retrospectiva das aplicações técnicas da análise de dados multivariados nos estudos de reservatórios da Bacia Potiguar.

(Originais recebidos em 11.10.91).

ABSTRACT — *Hydrocarbon fields, especially those in a late phase of development, make available a large quantity of data in the form of logs, DSTs, core samples, well cutting descriptions, and production records. A major challenge to multidisciplinary groups working with reservoir management is the efficient handling of this remarkable volume of data and the translation of the information into different scales. Because of their multivariate character and the fact that they are subject to random variabilities, such sets of data can be studied using probabilistic models. This approach has been routinely applied in reservoir studies of Potiguar Basin hydrocarbon fields. The present paper offers a brief overview of the technical applications of multivariate data analysis in these studies.*

(Expanded abstract available at the end of the paper).

1 – INTRODUÇÃO

A prospecção de jazidas de hidrocarbonetos, assim como a de outros bens minerais, é uma atividade de alto risco devido ao elevado grau de incerteza, onde as chances de sucesso são consideravelmente reduzidas. Na indústria do petróleo, este risco torna-se ainda mais significativo, já que os investimentos nesta área são muito superiores aos de outras atividades minerais.

Nas atividades exploratórias, a filosofia utilizada parte da premissa de que as informações geológicas são as mais eficientes para a descoberta de jazidas, apesar destas informações terem um caráter fundamentalmente estatístico. Esta natureza estatística dos dados geológicos está relacionada às informações incompletas, limitadas a pontos de controle, que são os dados dos poços e os dados geofísicos, o que leva a

uma grande incerteza sobre as variações, em superfície, do fenômeno em estudo. Os modelos probabilísticos reconhecem esta incerteza e fornecem ferramentas para, entre outras coisas, a estimativa de valores em locais não amostrados, desde que sejam feitas suposições sobre as características estatísticas do fenômeno. Neste caso, os dados geológicos passam a ser vistos como realizações de processos aleatórios. Esta suposição, na verdade, entra em conflito com a realidade, pois o processo que dá origem a um reservatório é puramente determinístico, sendo o resultado de complexas interações físicas e químicas (sedimentação, compactação, diagênese, etc); no entanto, o conhecimento é tão reduzido que esta complexidade parece ter um comportamento aleatório. Resumindo, pode-se dizer que a relação funcional entre a geologia e a ocorrência de hidrocarbonetos não é uma relação determinística.

1 - Setor de Tratamento de Dados Geológicos (SERTRAG), Divisão Regional de Geologia (DIRGEO), Distrito de Exploração da Bacia Potiguar (DEBAR), Av. Interventor Mário Câmara, 2783, Nazaré, CEP 59070, Natal, RN, Brasil.

Esta incerteza a respeito do comportamento do fenômeno entre os pontos de controle revela-se em todas as escalas. Por exemplo, a arquitetura de reservatórios heterogêneos no espaço compreendido entre dois pontos é, na maioria das vezes, desconhecidas, assim como não é possível saber exatamente o que se passa entre os pontos amostrados por meio de perfis ao longo de um poço, pois é possível a existência de camadas mais delgadas do que a resolução vertical das ferramentas de perfilagem.

O uso rotineiro de modelos probabilísticos torna-se possível graças ao desenvolvimento, em paralelo, de requintados modelos sedimentológicos e da ciência da computação. Este enfoque tem sido aplicado nos estudos de campos de hidrocarbonetos da Bacia Potiguar, seguindo uma tendência que se observa em todo o mundo.

Na PETROBRÁS, durante a década de 70, foram realizados diversos trabalhos de caráter exploracionista (Bettini *et al.* 1976 e Bettini e Silva, 1979), nos quais se utilizaram modelos probabilísticos. No entanto, somente na segunda metade da década de 80 é que começaram a surgir os trabalhos ligados à lavra de hidrocarbonetos. Atualmente, existe um enorme potencial para aplicação de técnicas estatísticas, pois a empresa já conta com um grande número de campos em adiantada fase de desenvolvimento, alguns dos quais necessitando de detalhadas descrições de seus reservatórios para tornarem-se economicamente viáveis. Porém, é importante que se volte a realizar trabalhos com suporte estatístico na área de exploração, principalmente, levando-se em conta que algumas bacias brasileiras estão em fase madura de investigação exploratória.

Operacionalmente, a aplicação destas técnicas no âmbito do Departamento de Exploração - DEPEX tornou-se extremamente facilitada graças a aquisição de *software* versáteis e poderosos que, entre outras qualidades, permitem a aplicação de modernas técnicas estatísticas. Sarzenski e Toledo (1990) oferecem excelentes exemplos de estudos quantitativos de reservatórios, e de como utilizar estes programas.

2 -- PRINCIPAIS MÉTODOS MATEMÁTICOS UTILIZADOS NA PREDIÇÃO GEOLÓGICA.

Os estudos quantitativos para caracterização de reservatórios envolvem duas etapas principais: a primeira consiste na aplicação de estatística clássica, principalmente, análise de dados multivariados para a determinação de litofácies, eletrofácies, unidade de fluxo e propriedades petrofísicas, enquanto que, na segunda etapa, utiliza-se a geoestatística para estudos de estimativa (krigagem) e estudos de variabilidade (simulação condicional) dos parâmetros obtidos na fase anterior.

Na Bacia Potiguar, técnicas de análise de dados multivariados têm sido utilizadas na determinação de eletrofácies e unidades de fluxo, a partir das litofácies reconhecidas nos testemunhos e das medidas petrofí-

sicas feitas em laboratório (porosidade e permeabilidade).

As eletrofácies, que podem ser definidas como um ou mais tipos litológicos que mostram respostas similares frente aos perfis, são determinadas por meio de um processo conhecido por "correlação rocha *versus* perfil". O mapeamento das eletrofácies fornece uma idéia da variabilidade lateral do reservatório, possibilitando um planejamento mais efetivo para as locações de desenvolvimento.

Um outro conceito, ligado às propriedades petrofísicas da rocha e que, portanto, fornece informações sobre a drenagem do reservatório, é o de unidades de fluxo, definidas como sendo as porções do reservatório que apresentam médias similares para aquelas propriedades que afetam a movimentação dos fluidos em subsuperfície, como a permeabilidade. A determinação destas unidades de fluxo pode ser feita usando-se uma análise de regressão entre variáveis petrofísicas medidas em laboratório e variáveis de perfil.

A seguir, serão comentadas e ilustradas, através de exemplos, as principais técnicas de análise de dados multivariados que tem sido utilizadas em estudos quantitativos de reservatórios.

2.1 -- Análise de Componentes Principais

É um método de projeção com redução de dimensões de um espaço vetorial, podendo portanto ser utilizado para reduzir um número de variáveis que caracterizam uma certa observação. No caso bidimensional, esta projeção é efetuada sobre os eixos de uma elipse de rotação cujo eixo maior, denominado de primeiro componente principal, explica uma fração de variância total que, por sua vez, é maior do que aquela explicada por qualquer uma das variáveis tomadas isoladamente (fig. 1). Elek (1988) utilizou-se desta técnica para fazer correlações entre poços, reduzindo os vários perfis disponíveis, a uma única curva que continha a maior parte da variância dos dados dos originais.

2.2 -- Análise de Grupamento

Esta técnica tem por objetivo reunir observações ou variáveis em grupos, de modo que os elementos de um mesmo grupo tenham um alto grau de associação natural entre si, e que os grupos sejam relativamente distintos uns dos outros. Na prática, monta-se uma matriz de associação composta por coeficiente de similaridade ou de distância, plotando-os num dendograma.

A análise de grupamento pode ser utilizada na determinação de eletrofácies, reunindo em um mesmo grupo aquelas litofácies, de preferência com alguma associação genética, que se comportem de modo semelhante frente às curvas dos perfis (fig. 2). De forma similar, esta técnica pode ser utilizada na determinação de litofácies, onde as variáveis poderão ser, entre outras, a textura e as estruturas sedimenta-

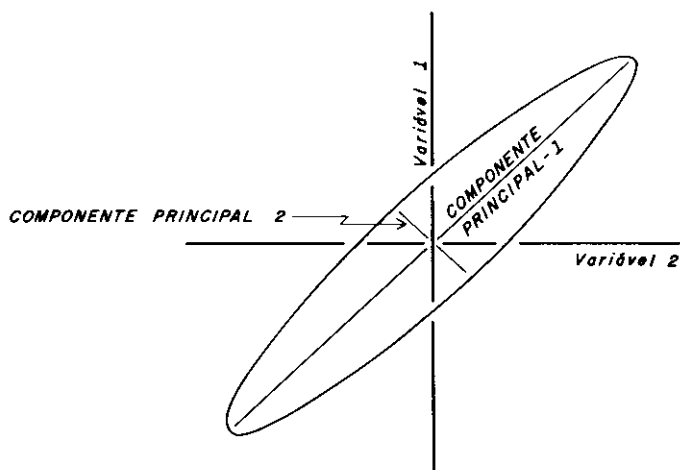


Fig. 1 - Análise de componentes principais.
Fig. 1 - Principal component analysis.

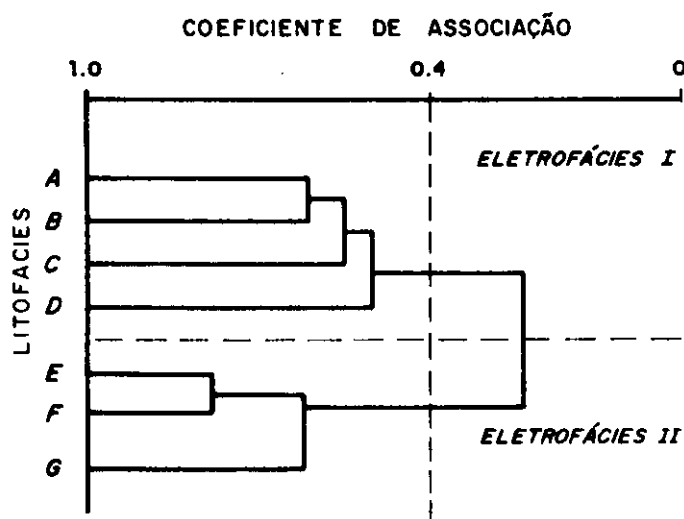


Fig. 2 - Análise de grupamento.
Fig. 2 - Cluster analysis.

res.

2.3 – Análise Discriminante

É um método que objetiva determinar uma função linear que discrimine os grupos (definidos *a priori*, por meio da análise de agrupamento ou não), de modo que a probabilidade de classificação errônea de um elemento em qualquer um dos grupos, seja minimizada. Este objetivo é atingido usando-se uma combinação linear das variáveis discriminantes que maximiza a diferença entre os grupos e minimiza a variabilidade interna de cada grupo (fig. 3). Sakurai e Melvin (1988) utilizaram-se a análise discriminante para prever, a partir de perfis, a ocorrência de litofácies em poços não testemunhados.

2.4 – Análise de Regressão

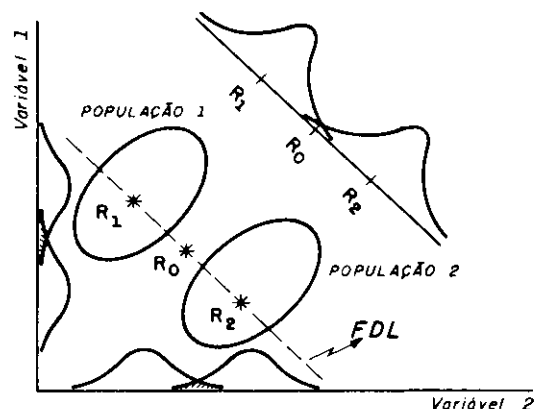
A análise de regressão é utilizada para se prever o valor de uma dada variável, chamada de dependente, a partir de outras variáveis, tidas como independentes. Genericamente, um modelo de regressão linear múltipla tem a seguinte expressão:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$$

onde:

- y = variável dependente ou variável resposta;
- X_1, \dots, X_n = variáveis independentes ou preditivas;
- β_0 = termo independente da equação de regressão;
- β_1, \dots, β_n = coeficientes das variáveis independentes e;
- ϵ = erro aleatório que representa a variação de y não explicada pelo modelo.

A aplicação mais comum para essa técnica em estudos de reservatório tem sido a previsão de valores de permeabilidade em intervalos não testemunhados, onde se dispõe apenas das variáveis fornecidas pelos perfis.



R_0 = Índice discriminante

R_1 = Centro geométrico da população 1

R_2 = Centro geométrico da população 2

FDL = Função discriminante linear

Fig. 3 - Análise discriminante.
Fig. 3 - Discriminant analysis.

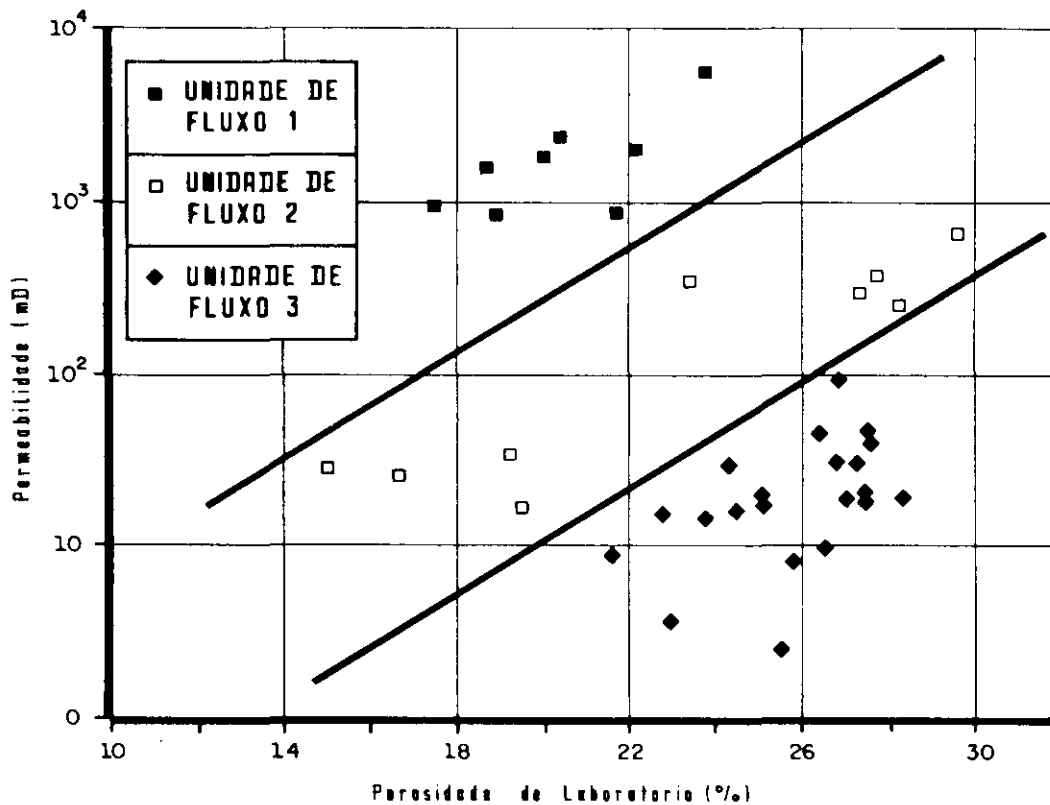


Fig. 4 - Gráfico permeabilidade versus porosidade com discriminação das unidades de fluxo definidas para a Zona C1 do Campo de Livramento (In: Souto Filho et al. 1989).

Fig. 4 - Permeability versus porosity cross-plot showing flow units defined for Zone C1, Livramento Field (In: Souto Filho et al. 1989).

3 – CASOS ESTUDADOS

Os estudos de reservatórios nos campos da Bacia Potiguar têm, sistematicamente, utilizado algumas das técnicas comentadas anteriormente, com a obtenção de resultados satisfatórios em alguns casos. A seguir, serão discutidos alguns destes trabalhos.

Souza Jr. (1988) caracterizou, por meio da análise discriminante, as eletrofácies A, B e C em reservatórios da Formação Açu e, posteriormente, estas eletrofácies foram subdivididas em unidades de fluxo que, por sua vez, foram utilizadas em exercícios de simulação condicional.

Souto Filho et al. (1989) aplicaram as técnicas de análise de grupamento, análise discriminante e análise de regressão linear para caracterização de eletrofácies e unidades de fluxo nos reservatórios da Formação Pendência. A partir de litofácies reconhecidas nos testemunhos, foram determinadas três eletrofácies, sendo uma delas constituída por depósitos de lobos e canais do sistema deltaico que deu origem ao reservatório. Posteriormente, subdividiu-se esta eletrofácies em duas unidades de fluxo (fig. 4), já que os depósitos de canais apresentavam valores de permoporosidade bastante superiores aos dos lobos deltaicos. A permeabilidade para cada unidade de fluxo foi prevista por meio de equações de regressão do tipo:

$$\log k = b_1 * GR + b_2 * PHIN + b_3 * RHOC + b_4 * SW - b_5 * VSH - b_0$$

onde:

- $\log k$ = logaritmo natural da permeabilidade absoluta;
- GR = raios gama;
- PHIN = porosidade neutrônica;
- RHOC = densidade;
- SW = saturação de água;
- VSH = volume de argila.

Souza Jr. (1991), estudando os reservatórios da Formação Açu na área do Campo de Canto do Amaro, dividiu os ciclos de grandecrescência ascendente em duas unidades básicas: uma inferior (reservatórios), constituída por arenitos grosseiros a médios, e outra superior (não-reservatórios), formada por arenitos finos e folhelhos. Para a caracterização destas unidades, por meio de perfis, gerou-se uma função discriminante linear que resultou em duas eletrofácies. A próxima etapa foi a utilização das medidas de permeabilidade da eletrofácies constituída por rochas-reservatório e efetuar uma análise de regressão múltipla com variáveis de perfis. Deste modo, tornou-se viável a identificação do tipo litológico, assim como a previsão do valor da permeabilidade a cada profundidade do perfil (fig. 5). A partir da análise visual de indícios de óleo, foram estabelecidos valores de corte que deram origem às unidades de fluxo, como é apresentada na seção esquemática da figura 6.

Poletto (1991) também utilizou regressão múltipla

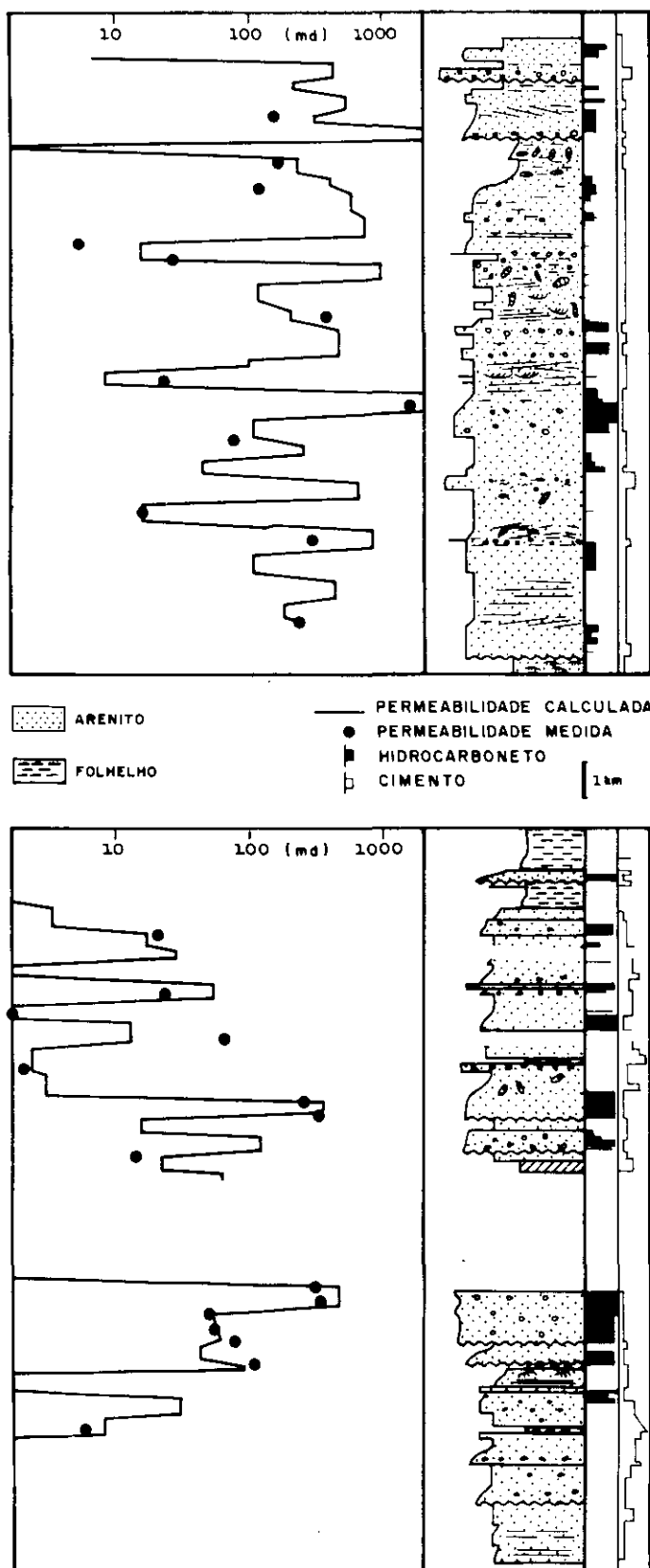


Fig. 5 - Comparação entre a permeabilidade medida (K) e a permeabilidade calculada (KC) através de equação de regressão multivariada (In: Relatório do projeto Canto do Amaro, inédito).

Fig. 5 - Measured permeability (K) versus permeability calculated (KC) according to multivariate regression equation (Canto do Amaro project report, unpublished).

pla para prever valores de permeabilidade a partir de perfis elétricos neste mesmo campo.

Teixeira e Souza Jr. (1991), fazendo uso da análise de grupamento, conseguiram classificar as águas da Formação Açú, na Bacia Potiguar, em três grupos. As variáveis utilizadas foram parâmetros físico-químicos da água de formação, medidos em laboratório.

4 - CONCLUSÕES

Os dados geológicos, por apresentarem um caráter multivariado, além de variabilidades aleatórias, podem ser eficientemente estudados através de modelos probabilísticos.

As principais técnicas usadas para a determinação de eletrofácies e unidades de fluxo, a partir de testemunhos e perfis, são: análise de componentes principais, análise de grupamento, análise discriminante e análise de regressão.

Os exemplos citados mostram que, graças ao surgimento de *software* estatísticos de fácil utilização, a análise de dados geológicos, através de métodos matemáticos, tem tomado um grande impulso nos últimos anos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLEN, J. R. Prediction of permeability from logs by multiple regression, In: SOCIETY OF PROFESSIONAL WELL LOG ANALYSTS, 6, London, 1992.
- BETTINI, C., FUSCHILO, B. L., BELTRAMI, C. V. et al. *Uso da análise de regressão múltipla para prever a posição de um novo campo de petróleo em áreas de produção conhecida*. Rio de Janeiro: PETROBRÁS/DEXPRO/DIVEX/CPDG, 1976.
- BETTINI, C., SILVA, R. R. *Mapeamento da probabilidade condicional de ocorrência de petróleo no Membro Água Grande, Formação Itaparica, Bacia do Recôncavo*. Rio de Janeiro: PETROBRÁS/DEPEX/DITREX, 1979. 10 p. Relatório interno.
- DAVIS, J. C. *Statistics and data analysis in geology*. New York: John Wiley & Sons, 1973. 550 p.
- ELEK, I. Some applications of principal component analysis. *Geobyte*, Tulsa, ok, v. 3, n. 2, p. 25-34, 1988.
- POLETO, C. A. *Análise quantitativa dos reservatórios através de perfis elétricos no Campo de Canto do Amaro*. [s.l. :s.n.], 1991. (STOG/91).
- SAKURAI, S., MELVIN, J. Facies discrimination and permeability estimation from well logs for the Endicon field. In: SOCIETY OF PROFESSIONAL WELL LOG ANALYSTS. ANNUAL LOGGING SYMPOSIUM, 29, [s.l.], 1991.
- SARZENSKI, D. J., TOLEDO, J. B. *Correlação rocha-perfil : conceitos e aplicações em reservatórios heterogêneos*. Rio de Janeiro: PETROBRÁS/DEPEX, 1990. 394 p. Relatório interno.
- SOUTO FILHO, J. D., AQUINO, P. C. A., SOUZA Jr., O. G. *Avaliação do Campo de Livramento, Zona C (análise integrada de perfis)*. Natal: PETROBRÁS/DEBAR/DIRGEO, 1989. Relatório interno.
- SOUZA Jr., O. G. *Simulação condicional de unidades de fluxo*

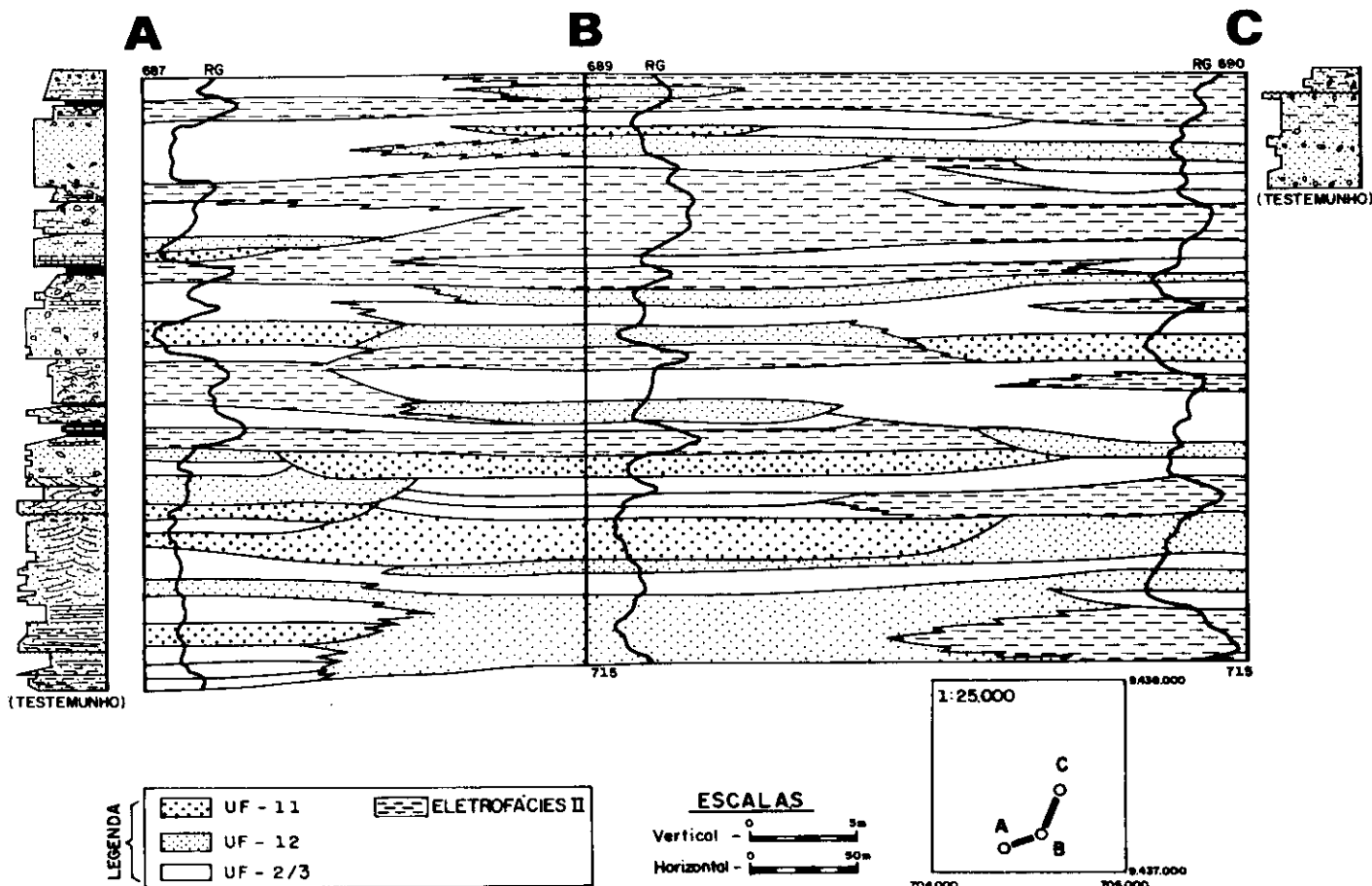


Fig. 6 - Seção esquemática da distribuição lateral das unidades de fluxo e da eletrofácies II. Campo de Canto do Amaro Zona 400 (In: Relatório do projeto Canto do Amaro, inédito).

Fig. 6 - Schematic cross-section of lateral distribution of flow units and electrofacies II, Canto do Amaro Field, Zone 400 (Canto do Amaro project report, unpublished).

na área II do projeto-piloto de injeção de vapor na Formação Açú (Ksup), Campo de Estreito, Bacia Potiguar, Brasil. Ouro Preto: Universidade Federal de Ouro Preto, 1988. 168 p. Tese de mestrado.

SOUZA Jr., O. G. *Análise quantitativa dos reservatórios da Zona 400 dos campos de Canto do Amaro e Alto da Pedra, Bacia Potiguar*. Natal: PETROBRÁS/DEBAR/DIRGEO, 1991. Relatório interno.

TEIXEIRA, I. E. M., SOUZA Jr, O. G. *Estudos de salinidade da Formação Açú*. [s.l. : s.n.], 1991. (STOG/91).

EXPANDED ABSTRACT

Hydrocarbon fields make available a large quantity of fundamentally statistical data obtained from incomplete well and geophysical information that is related to certain limited points of control. The resultant gaps in information leave uncertainties regarding the subsurface variability of properties important to reservoir characterization studies. The major challenge for multidisciplinary groups assigned to reservoir management is how to efficiently handle this large volume of data, derived mainly from well cuttings, DSTs, production records, logs, and core samples, among other sources. Since these data have a multivariate character and are subject to random variabilities, they may be analyzed using probabilistic models that recognize and quantify associated uncertainties. The use of such models has been made viable thanks to great strides in modern computer and graphic techniques, in addition to enhanced sedimentological models.

Quantitative studies in reservoir characterization consist of two main stages (the second of which is not explored in this paper): (1) the application of multivariate techniques (principal component, discriminant, cluster, and regression analyses) to arrive at electrofacies and flow units and (2) the use of geostatistical techniques such as kriging and conditional simulation to regionalize the variables determined in the first stage. Reservoir characterization studies in the Potiguar Basin have systematically relied on this methodology, in most cases with satisfactory results. Due to the extreme complexity of these reservoirs and to the viscosity of Potiguar oil, this research has been a constant challenge to geologists and production engineers. The use of linear discriminant functions has made it possible to identify the most homogeneous intervals of the main reservoir facies, and permeability values have been calculated based on a linear regression equation.